Naturwissenschaftlich-Technische Fakultät

Universität Siegen

Towards a simultaneous measurement of the top-quark mass and decay width with 13 TeV ATLAS data using $bb4\ell$ Monte-Carlo predictions

Dissertation
zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Katharina Voß

Eingereicht bei der Naturwissenschaftlich-Technischen Fakultät
der Universität Siegen

2026

1. Gutachter:	Prof. Dr. Markus Cristinziani
2. Gutachterin:	PD Dr. Carmen Diez Pardos

Tag der Promotion: 01.06.2026

Abstract

This thesis presents a sensitivity study for a simultaneous measurement of two parameters of the heaviest known elementary particle: the top-quark mass and its decay width. The analysis is based on the full Run 2 dataset, corresponding to an integrated luminosity of $140\,\text{fb}^{-1}$ of proton–proton collisions at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$, collected by the ATLAS experiment at the Large Hadron Collider between 2015 and 2018. The study focuses on the $WWbb$ production process in the dilepton decay channel and exploits the sensitivity of the invariant mass distribution of a charged lepton and a b-jet to both the top-quark mass and decay width. This sensitivity can be only reliably described using next-to-leading-order predictions for the full $WWbb$ final state in the dilepton channel. A central aspect of this work is the validation of a corresponding Monte-Carlo simulation within the ATLAS Collaboration, including a comprehensive evaluation of the associated modelling uncertainties. This simulation constitutes a valuable input not only for the measurement presented here, but also for future ATLAS analyses targeting the $WWbb$ final state. In the context of this measurement, the use of these improved predictions is shown to significantly enhance the sensitivity to the top-quark decay width, reducing the expected uncertainty by approximately a factor of two compared to the currently most precise measurement.

Zusammenfassung

Diese Dissertation stellt eine Sensitivitätsstudie für die simultane Messung zweier Parameter des schwersten bekannten Elementarteilchens dar: der Masse des Top-Quarks und seiner Zerfallsbreite. Die Analyse basiert auf dem vollständigen Run-2-Datensatz, der einer integrierten Luminosität von $140\,\text{fb}^{-1}$ Proton-Proton-Kollisionen bei einer Schwerpunktsenergie von $\sqrt{s} = 13\,\text{TeV}$ entspricht, aufgezeichnet vom ATLAS-Experiment am Large Hadron Collider zwischen 2015 und 2018. Die Studie konzentriert sich auf den $WWbb$-Produktionsprozess im vollleptonischem Zerfallskanal und nutzt die Sensitivität der invarianten Massenverteilung eines geladenen Leptons und eines b-Jets gegenüber sowohl der Top-Quark-Masse als auch der Zerfallsbreite. Diese Sensitivität kann nur zuverlässig durch Vorhersagen für den vollständigen $WWbb$-Endzustand im vollleptonischem Zerfallskanal beschrieben werden. Ein zentraler Aspekt dieser Arbeit ist die Validierung einer entsprechenden Monte-Carlo-Simulation innerhalb der ATLAS-Kollaboration, einschließlich einer umfassenden Ausarbeitung der zugehörigen Unsicherheiten. Diese Simulation ist nicht nur für die hier präsentierte Messung wichtig, sondern auch für zukünftige ATLAS-Analysen, die sich auf den $WWbb$-Endzustand konzentrieren. Im Kontext dieser Messung wird gezeigt, dass die Verwendung dieser verbesserten Vorhersagen die Sensitivität gegenüber der Zerfallsbreite des Top-Quarks deutlich erhöht und die erwartete Unsicherheit im Vergleich zur derzeit präzisesten Messung um etwa einen Faktor zwei reduziert.

Contents

1 Introduction **1**

2 From fixed-order calculations to the modelling of high-multiplicity final states **3**
2.1 The Standard Model Lagrangian . . . 3
2.2 Fixed-order calculation . . . 10
2.3 Parton shower . . . 18
2.4 The POWHEG method . . . 21
2.5 Predicting $WWbb$ with NLO+PS accuracy . . . 23

3 Modelling uncertainties in the ATLAS $bb4\ell$ Monte-Carlo sample **31**
3.1 The nominal ATLAS $bb4\ell$ Monte-Carlo sample . . . 32
3.2 Uncertainties in the hard process generation . . . 45
3.3 Uncertainties in the parton-shower matching . . . 54
3.4 Uncertainties in the parton showering . . . 59
3.5 Conclusions . . . 70

4 From proton–proton collisions to reconstructed physics objects **73**
4.1 The CERN accelerator complex . . . 73
4.2 Luminosity and pile-up . . . 75
4.3 The ATLAS detector and low-level object reconstruction . . . 77
4.3.1 Inner detector . . . 79
4.3.2 Track reconstruction . . . 81
4.3.3 Primary vertex reconstruction . . . 82
4.3.4 Calorimeters . . . 83
4.3.5 Topological energy cell cluster reconstruction . . . 87
4.3.6 Muon spectrometer . . . 87
4.3.7 Muon track reconstruction . . . 89
4.3.8 Trigger and data acquisition . . . 90
4.4 Physics object reconstruction . . . 91
4.4.1 Jets . . . 91
4.4.2 Heavy-flavour jets . . . 94
4.4.3 Electrons . . . 96
4.4.4 Muons . . . 99
4.4.5 Missing transverse momentum . . . 100

5 Previous top-quark mass and decay width measurements **103**
5.1 Top-quark mass measurements . . . 105
5.2 Top-quark decay width measurements . . . 113
5.3 Lessons learned from previous measurements . . . 117

6 Data and Monte-Carlo simulation **119**
6.1 Data events . . . 120
6.2 Monte-Carlo events . . . 120
6.2.1 Signal process . . . 121
6.2.2 Background process . . . 123

7 Analysis strategy **127**
7.1 Object definition and event selection . . . 127
7.2 Event reconstruction in the signal region . . . 130
7.3 Controlling the additional heavy-flavour contribution . . . 135
7.4 Controlling jet energy scale uncertainties . . . 139

8 Systematic uncertainties **141**
8.1 Experimental uncertainties . . . 141
8.1.1 Jet uncertainties . . . 142
8.1.2 Flavour tagging uncertainties . . . 149
8.1.3 Lepton uncertainties . . . 150
8.1.4 Missing transverse momentum uncertainties . . . 151
8.2 Modelling uncertainties . . . 151
8.2.1 Signal modelling uncertainties . . . 151
8.2.2 Background normalisation . . . 154

9 Fit strategy **155**
9.1 The likelihood function . . . 155
9.2 Parameter optimisation and uncertainty estimation . . . 159
9.3 Parameterising the top-quark mass and decay width sensitivity . . . 162

10 Estimating the top-quark mass and decay width sensitivity **165**
10.1 Profile likelihood fit using the Asimov data set . . . 165
10.2 The blinded fit . . . 174
10.3 Future developments . . . 179

11 Conclusions **181**

A Testing alternative parameterisations of the top-quark mass and decay width dependence **183**

B Further studies of the top-quark mass and decay width measurement **187**
B.1 Simultaneous or independent extraction of the top-quark mass and decay width 187
B.2 Influence of jet pre-recommendation uncertainties . . . 190
B.3 Compatibility in validation regions . . . 193

C High-p_{T} jet tagging using a Boosted Decision Tree **197**
C.1 Monte-Carlo samples and object definitions 199
C.2 Track Classification Tool 201
C.3 Structure of the Classified Track Tagger 206
C.4 Results 208
C.5 Conclusion 212

Acknowledgements **215**

Bibliography **217**

CHAPTER 1

Introduction

The curiosity of generations of scientist trying to understand the fundamental building blocks of matter and their interactions has resulted in the Standard Model (SM) of particle physics, a theoretical framework that describes most experimental data with remarkable accuracy. The experimental program to discover all fundamental particles predicted by the SM, and to probe both the energy and precision frontiers for deviations from its predictions, has led to a series of particle physics experiments. The highest-energy particle collider ever built is thereby the Large Hadron Collider (LHC) at CERN. With proton–proton collision provided by the LHC, in 2012, the final missing particle of the SM, the Higgs boson, was discovered by the ATLAS and CMS collaborations. Despite its success, the SM is widely regarded as incomplete. There is compelling evidence for physics beyond the Standard Model (BSM), particularly from astroparticle physics observations. For example, the SM does not provide a viable candidate for dark matter, which constitutes about 25 % of the mass–energy content of the universe, whereas ordinary matter described by the SM accounts for only about 5 %. This discrepancy strongly motivates continued searches for new physics. The data collected by the LHC enable searches for BSM physics in two complementary ways: by searching for the direct production of BSM particles in exploring the highest-energy collisions ever achieved in laboratory conditions, and by performing high-precision measurements that can reveal small but significant deviations from SM predictions. Such deviations may arise from higher-order corrections involving virtual BSM particles. This thesis focuses on the latter approach. It presents a study of two fundamental parameters of the heaviest known elementary particle in the SM: the top-quark mass and decay width. The analysis exploits the large dataset recorded by the ATLAS detector during 2015–2018 (Run 2), corresponding to approximately 115 million events top-quark pair production events. The combination of this large dataset, excellent detector performance, and advanced object reconstruction techniques enables precision measurements of these quantities. To fully benefit from the experimental precision, high requirements on the accuracy of the theoretical predictions are placed. In particular, the analysis presented in this thesis relies on predictions at next-to-leading order (NLO) accuracy in quantum chromodynamics (QCD) for the final state $b\bar{b}\ell^-\bar{\nu}_\ell\ell^+\nu_\ell$, which

is expected for a fully leptonically decaying top-quark pair. A Monte-Carlo implementation providing such predictions has been developed by the theory community [1, 2]. In order to use these predictions effectively in an experimental analysis, besides the nominal theory predictions, the associated modelling uncertainties must be carefully evaluated. This thesis presents a detailed study of such a Monte-Carlo dataset, with a particular focus on the assessment of its modelling uncertainties. Using these simulations, which were made available within the ATLAS Collaboration in the course of this thesis, the achievable sensitivity to the top-quark mass and decay width is evaluated.

In this thesis, after an introduction to the SM, theoretical methods for obtaining predictions for high-multiplicity final states are discussed, with particular emphasis on the $W^+W^-b\bar{b}$ production process, in Section 2. The study of the ATLAS $bb4\ell$ Monte-Carlo sample describing the full $b\bar{b}\ell^-\bar{\nu}_\ell\ell^+\nu_\ell$ final state at NLO accuracy in QCD, in particular the evaluation of the associated modelling uncertainties, is presented in Section 3. An overview of the experimental setup — including the LHC, the ATLAS detector, and physics object reconstruction — is given in Section 4. A summary of previous measurements of the top-quark mass and decay width is provided in Section 5. The analysis strategy developed in this thesis is then described, including the dataset and simulated samples used (Section 6), the overall analysis approach (Section 7), the considered systematic uncertainties (Section 8), and the employed fitting strategy (Section 9). The results for the sensitivity estimation to the top-quark mass and decay width are presented in Section 10. Finally, Section 11 summarizes the findings and provides an outlook.

CHAPTER 2

From fixed-order calculations to the modelling of high-multiplicity final states

The work presented in this thesis aims at measuring two fundamental parameters of a theoretical framework that describes the constituents of matter and their interactions. In this chapter, the conceptual development of this theoretical framework, the Standard Model, is outlined, along with the methods by which it enables precise predictions that can be directly compared to experimental data. Mendeleev's periodic table of chemical elements, whose indeed periodic structure suggested an underlying substructure can be regarded as a starting point to understand the building blocks of atoms, the nuclei, protons and neutrons held together by the strong force, and the electrons, bound to the nuclei through the electromagnetic force. The radioactive decay of neutrons into protons can be described through weak interactions, accompanied by an electron and a neutrino, another fundamental particle in the theory. A multitude of discovered particles, which are bound by the strong force, as the proton and the neutron, suggested that indeed these particles are not fundamental themselves, but are build from fundamental particles, the quarks. Experimental evidence suggested that quarks and leptons exists in three different families. This thesis focuses on a quark of the third generation, the top quark.

2.1 The Standard Model Lagrangian

The Standard Model (SM) of particle physics is a relativistic quantum field theory that successfully describes the electromagnetic, weak, and strong interactions of elementary particles. It is in remarkable agreement with experimental data, such that — within current experimental and theoretical uncertainties — no significant deviations from its predictions have been observed in laboratory experiments. The dynamics and interactions of the SM are encoded in the SM Lagrangian $\mathcal{L}_{SM}$, which is built upon the requirement of Lorentz invariance to ensure consistency with special relativity. Further, the interactions between fundamental particles

arise from the requirement of local gauge invariance under the symmetry group $\mathrm{SU}(3)_C\times$ $\mathrm{SU}(2)_L\times$ $\mathrm{U}(1)_Y$. These gauge symmetries are associated with the conservation of colour charge, weak isospin, and weak hypercharge Y, respectively. A starting point for constructing the SM Lagrangian is the Dirac Lagrangian for a free spin-$\frac{1}{2}$ fermion field:

$$\mathcal{L}_{\text{Dirac}} = \bar{q}(x)(i\gamma_\mu\partial^\mu - m)q(x), \tag{2.1}$$

where $q(x)$ denotes a fermion field of mass m, and $\bar{q} = q^\dagger\gamma^0$. The matrices γ^μ are the Dirac γ-matrices. In the following, a brief overview is provided of how imposing local gauge invariance on this Lagrangian leads to the structure of the SM Lagrangian, following mainly [3].

Quantum Electrodynamics (QED) The electromagnetic interactions can be described by requiring invariance of the Lagrangian in Equation 2.1 under a local U(1) symmetry, under which the fermion fields transform as

$$q(x) \xrightarrow{\mathrm{U}(1)} e^{i\alpha(x)}q(x). \tag{2.2}$$

While the mass term $m\bar{q}q$ in Equation 2.1 remains invariant under this local transformation, the kinetic term involving the derivative is not:

$$\bar{q}(x)\partial^\mu q(x) \xrightarrow{\mathrm{U}(1)} \bar{q}(x)\partial^\mu q(x) + i(\partial^\mu\alpha)\bar{q}(x)q(x). \tag{2.3}$$

Gauge invariance can be restored by introducing a gauge field $A_\mu(x)$ and replacing the derivative ∂_μ by the covariant derivative

$$D_\mu = \partial_\mu + ieA_\mu(x), \tag{2.4}$$

with e the electric charge of the fermion field q. The Lagrangian then becomes

$$\mathcal{L} = \bar{q}(x)\left(i\gamma_\mu D^\mu - m\right)q(x), \tag{2.5}$$

which is invariant under local U(1) transformations provided that the gauge field transforms as

$$A_\mu(x) \xrightarrow{\mathrm{U}(1)} A_\mu(x) - \frac{1}{e}\partial_\mu\alpha(x). \tag{2.6}$$

To identify the vector field A_μ with the photon mediating the electromagnetic interaction, a kinetic term for a propagating photon needs to be added. Since the electromagnetic interaction has infinite range, the photon must be massless. A mass term would explicitly break gauge invariance, as discussed later in this section. To ensure that this kinetic term for the photon field also satisfies U(1) gauge invariance, the electromagnetic field tensor $F_{\mu\nu}$ is derived from the commutator of the covariant derivate as

$$\left[D_\mu, D_\nu\right] = D_\mu D_\nu - D_\nu D_\mu = ieF_{\mu\nu} = ie(\partial_\mu A_\nu - \partial_\nu A_\mu). \tag{2.7}$$

The full QED Lagrangian is thus given by

$$\mathcal{L}_{\text{QED}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \bar{q}(x)(i\gamma_\mu D^\mu - m)q(x) = \bar{q}(x)(i\gamma_\mu(\partial^\mu + ieA^\mu) - m)q(x). \tag{2.8}$$

Quantum Chromodynamics (QCD) Following the construction of QED from local U(1) gauge invariance, QCD is obtained by requiring invariance of the Lagrangian under local SU(3) transformations. This raises two questions: why the SU(3) group, and what are the degrees of freedom it acts on? Experimental observations indicate that quarks carry an additional quantum number, referred to as colour charge, conventionally labelled red, green and blue. Evidence for this structure is provided, for instance, by measurements of the ratio of the cross-section for e^+e^- annihilation into hadrons to that into muon pairs ($\mu^+\mu^-$) [4], which can be explained by assuming that three distinct colour states of each quark flavour exists: e.g. there is not the up-quark u, but instead the three states u_r, u_g and u_b. The SU(3) gauge symmetry is therefore introduced to act on the three-dimensional colour space of quarks. A local SU(3) acts on the colour degrees of freedom as

$$q(x) \xrightarrow{\text{SU(3)}} U(x)q(x) = e^{i\alpha^a(x)T^a}q(x), \tag{2.9}$$

where $q = (q_r, q_g, q_b)$ and T^a are the eight generators of the SU(3) Lie group. In analogy to QED, gauge invariance of the kinetic term of the quark field in the free Dirac Lagrangian is ensured by introducing the covariant derivative

$$D_\mu = \partial_\mu - ig_s G^a_\mu T^a, \tag{2.10}$$

with the coupling strength g_s and the eight gauge fields $G^a_\mu(x)$, identified with the gluons. In contrast to the Abelian U(1) case, SU(3) is a non-Abelian group, with generators satisfying

$$[T^a, T^b] = if_{abc}T^c, \tag{2.11}$$

where f_{abc} are the structure constants. Requiring the gauge invariance of the covariant derivate, the SU(3) transformation of the gauge fields is given by

$$G^a_\mu \xrightarrow{\text{SU(3)}} G^a_\mu + \frac{1}{g_s}\partial_\mu\alpha^a + f_{abc}G^b_\mu\alpha^c. \tag{2.12}$$

The kinetic term of the gluon fields is again obtained from the commutator of covariant derivatives and the field strength tensor is evaluated to

$$G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f_{abc}G^b_\mu G^c_\nu. \tag{2.13}$$

The last term in this equation arises due to SU(3) being a non-Abelian symmetry group and has immediate consequence on the properties of QCD: it describes gluon self-interactions with three- and four-gluon vertices. Gluons carry colour charge and thus interact themselves under exchange of QCD gauge bosons (the gluons). One important property of QCD is asymptotic freedom, meaning that the scale-dependent strong coupling constant $\alpha_S^2(Q) = g_s^2(Q)/(4\pi)$ vanishes at scales $Q \to \infty$, while for low energy scales, $\alpha_S^2(Q)$ becomes large, leading to the confinement of coloured particles.

Weak interactions While quantum electrodynamics and quantum chromodynamics successfully describe electromagnetic and strong interactions, respectively, several experimental observations indicated the existence of an additional interaction. One example is the comparatively long life-times of the charged pions ($\pi^- \to \mu^- \bar{\nu}_\mu$) of around 10^{-8} s, which is much larger than typical timescales of strong (10^{-23} s) and electromagnetic ($10^{-16}\,s$, e.g. $\pi^0 \to \gamma\gamma$) decays. This suggests a different interaction mechanism with a weaker coupling strength. An early theoretical description of weak interactions was proposed by Fermi in the context of β-decay. Since neutrinos do not couple to photons or gluons, this process cannot be mediated by electromagnetic or strong interactions. Fermi's model introduced a four-fermion interaction with a structure analogous to QED, involving vector–vector currents of the form $\bar{\nu}_e \gamma^\mu e$. While this approach successfully described certain features of β-decay, it failed to account for the experimentally observed violation of parity symmetry in weak interactions, meaning that the interaction is not invariant under space-inversion. By analysing for example the preferred emission direction of electrons and consequently of neutrinos emitted in the β-decay of polarized nuclei, it was concluded that the weak interaction only couples to left-handed fermions and right-handed anti-fermions. This behaviour is incorporated when replacing $\gamma^\mu \to \frac{1}{2}\gamma^\mu(1-\gamma^5)$. This vector-axial (V minus A or $V-A$) structure of the weak interaction can be directly probed by measuring angular distributions in electron–neutrino scattering experiments. At this point, weak interactions couple fermions within the same weak isospin doublet: leptons and quarks in the same family, e.g. (e, ν_e), (u, d) and (c, s) doublets. A problem arises however when considering e.g. kaon decays, $K^-(s\bar{u}) \to \mu^- + \bar{\nu}_\mu$, which should be forbidden. This issue is resolved by introducing weak eigenstates that differ from the mass eigenstates. At this time, only the u, d and s-quarks were experimentally discovered, but the existence of the c-quark was postulated,[1] and the mixing of the quark-mass eigenstates with two quark generations can be parameterised with one parameter: the Cabibbo angle θ_C. The d and s mass eigenstates now participate in the weak interactions only in linear superpositions, e.g. $d' = d\cos\theta_C + s\sin\theta_C$. Similarly to the postulation of the existence of the c-quark, after the discovery of the b quark, the existence of the top-quark was predicted, the weak isospin partner of the b quark. Now, the mixing of mass eigenstates to form the weak eigenstates is described by a 3×3-matrix, the Cabibbo-Kobayashi-Maskawa (CKM) matrix. The CKM matrix can be parameterised by three real parameters and an imaginary phase, providing a source for CP-violation in the SM,[2] which was indeed discovered in the decay of neutral kaons.[3] In addition to the charged weak interactions, neutral weak interactions were observed in $\nu_\mu e^- \to \nu_\mu e^-$ scattering processes. Measurements of the corresponding angular differential cross-sections demonstrated that neutral weak interactions cannot be described by a pure $V-A$ structure, in contrast to charged current interactions, providing a hint about the structure of the unified electroweak theory.

[1] The c-quark existence was postulated by Glashow, Iliopoulos and Maiani to explain the small $K_L^0 \to \mu^+\mu^-$ branching fraction, which can be explained through the GIM suppression mechanism describing the negative interference of $K^0 \to \mu^+\mu^-$ box-diagrams with intermediate u and c-quarks.

[2] A general weak interaction $ab \to cd$ is CP-invariant, if $V_{ca}V_{db}^* = V_{ca}V_{db}$ with V_{ij} corresponding to the elements of the CKM-matrix. This is the case for a real CKM matrix, but violated in the SM due to the complex phase needed in the 3×3 matrix.

[3] Footnote is printed on the next page.

Electroweak interactions The first attempt to describe the weak interaction in terms of a symmetry group, was to use the $SU(2)_L$ symmetry group. Thereby, left-handed fermions are arranged into weak isospin doublets, for example

$$\chi_L = \begin{pmatrix} \nu \\ e^- \end{pmatrix}_L , \tag{2.14}$$

which transform non-trivially under $SU(2)_L$, while the right-handed fermions, such as e_R, remain invariant. Introducing the corresponding gauge fields W_μ^a $(a = 1, 2, 3)$, the interaction term in the Lagrangian is given by

$$-ig\bar{\chi}_L \gamma_\mu T^i W_i^\mu \chi_L, \tag{2.15}$$

where T^i are the generators of $\mathrm{SU}(2)_L$, and g the coupling strength of the weak interaction. The charged weak gauge bosons $W^\pm$ can be derived with the linear combination of W^a corresponding to the charge lowering and raising combination of T^i,

$$W_\mu^\pm = \sqrt{\frac{1}{2}}(W_\mu^1 \pm iW_\mu^2). \tag{2.16}$$

However, a theory based solely on $\mathrm{SU}(2)_L$ cannot account for the observed structure of neutral current interactions. In particular, it would predict purely $V - A$ neutral currents, in contradiction with experimental results. This issue is resolved by unifying the weak and electromagnetic interactions within the gauge group $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$. The additional $\mathrm{U}(1)_Y$ symmetry introduces the hypercharge gauge field B_μ, with interaction term

$$-ig'\bar{\psi}\gamma_\mu \frac{Y}{2}\psi B^\mu, \tag{2.17}$$

where g' is the corresponding coupling strength and Y the weak hypercharge. The physical neutral gauge bosons arise from a mixing of the fields W_μ^3 and B_μ, given by

$$A_\mu = B_\mu \cos\theta_W + W_\mu^3 \sin\theta_W \tag{2.18}$$

$$Z_\mu = -B_\mu \sin\theta_W + W_\mu^3 \cos\theta_W \tag{2.19}$$

where θ_W denotes the Weinberg angle. The field A_μ is identified with the photon, while Z_μ corresponds to the neutral weak boson. A problem arises when trying to introduce mass terms

[2] The neutral kaons $K^0(d\bar{s})$ and $\bar{K}^0(s\bar{d})$ — with s and d denoting the mass eigenstates — are transformed into each other by the CP transformation $\mathrm{CP}|K^0\rangle = |\bar{K}^0\rangle$, where the CP operation combines the charge conjugation (particles transformed into antiparticles and vice versa) with the parity operation. The CP eigenstates K_S^0 (CP=+1) and K_L^0 (CP=−1) can be constructed through linear superposition of the K^0 and $\bar{K}^0$. In this case, the K_S^0 would always decay into two pions (CP=+1) and K_L^0 into three pions (CP=−1). After a sufficient decay length, a neutral kaon beam includes almost only K_L^0 due to the larger decay phase space of the K_S^0. Thereby it was found, that with a small branching fraction of 10^{-3} also the K_L^0 decay into two pions: a CP-violating decay. Where does the weak interaction enter this specific analysis of the K^0 system? It is responsible for the mixing of the K^0 and $\bar{K}^0$ states, as the flavour-changing charged current is only mediated by the weak interaction.

in the Lagrangian for the weak vector bosons, which are required to account for the short range of weak interactions, as well as for fermions. Explicit fermion, $m\bar{\psi}\psi = m(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R)$, and gauge boson mass terms are not invariant under the $SU(2)_L \times U(1)_Y$ symmetry group.

Brout–Englert–Higgs mechanism The resolution to the problem of generating gauge-invariant mass terms is provided by the Brout–Englert–Higgs mechanism [5–7]. Thereby, a complex scalar field $\phi(x)$ is introduced, transforming as a doublet under $\mathrm{SU}(2)_L$ with hypercharge $Y = 1$, such that

$$\phi(x) \xrightarrow{\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y} e^{i\alpha_i(x)T^i/2} e^{i\beta(x)} \phi(x). \tag{2.20}$$

The corresponding Lagrangian is given by

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi) \tag{2.21}$$

with the Higgs potential

$$V(\phi) = \mu^2 \phi^\dagger \phi + \frac{1}{2}\lambda(\phi^\dagger \phi)^2, \tag{2.22}$$

and the covariant derivative

$$D_\mu = \partial_\mu + igW_\mu^i T^i + ig' \frac{Y}{2} B_\mu. \tag{2.23}$$

The scalar field can be written as the $\mathrm{SU}(2)_L$ doublet,

$$\phi = \sqrt{\frac{1}{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \tag{2.24}$$

For $\mu^2 < 0$ and $\lambda > 0$, the Higgs potential develops a continuum of degenerate minima at

$$\phi^\dagger \phi = \frac{1}{2}(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2) = \frac{v^2}{2}, \qquad v = \sqrt{-\frac{\mu^2}{\lambda}}. \tag{2.25}$$

Choosing a particular vacuum ground state ϕ_0,

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \tag{2.26}$$

spontaneously breaks the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetry down to $\mathrm{U}(1)_{\text{em}}$. Expanding the field around ϕ_0 gives

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{2.27}$$

where $h(x)$ denotes the physical Higgs field. The three remaining scalar degrees of freedom are absorbed as longitudinal polarization modes of the massive gauge bosons. The chosen vacuum is invariant under the electromagnetic charge operator $Q = T_3 + Y/2$ with $T_3 = \pm\frac{1}{2}$ the third component of the weak isospin, ensuring that $\mathrm{U}(1)_{\text{em}}$ remains unbroken, and the

photon stays massless. Expanding the Higgs potential around the vacuum yields

$$V(\phi(x)) = -\frac{\lambda v^4}{4} + \frac{1}{2}(2\lambda v^2)h^2 + \lambda v h^3 + \frac{\lambda}{4}h^4, \tag{2.28}$$

which contains the Higgs mass term with $m_H^2 = 2\lambda v^2 = -\mu^2$, as well as the description of Higgs-boson self interactions. Gauge boson masses arise from the kinetic term $(D_\mu\phi)^\dagger(D^\mu\phi)$ after symmetry breaking,

$$(D^\mu\phi)^\dagger(D^\mu\phi) = \frac{1}{2}(\partial_\mu h)^2 + m_W^2 W_\mu^+ W^{\mu-} + \frac{1}{2}m_Z^2 Z_\mu Z^\mu \tag{2.29}$$

$$+ \left(2\frac{m_W^2}{v}W_\mu^+ W^{\mu-} + \frac{m_Z^2}{v}Z_\mu Z^\mu\right)h \tag{2.30}$$

$$+ \left(\frac{m_W^2}{v^2}W_\mu^+ W^{\mu-} + \frac{m_Z^2}{2v^2}Z_\mu Z^\mu\right)h^2 \tag{2.31}$$

with $m_W = \frac{gv}{2}$ and $m_Z = \frac{\sqrt{g^2+g'^2}}{2}v$. Additionally, the interactions of the massive gauge bosons with the Higgs boson are encoded in these terms.

The final missing ingredients for the Standard Model Lagrangian are mass terms for the fermions, which are invariant under $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$. This is achieved by generating fermion masses through Yukawa interactions. For example, the Yukawa term for electrons is given by

$$\mathcal{L}_{\text{Yukawa}} = -y_e(\bar{\nu}_e, \bar{e})_L \begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} e_R + \text{h.c.}, \tag{2.32}$$

where h.c. stands for Hermitian conjugate. After spontaneous electroweak symmetry breaking, this term reads

$$\mathcal{L}_{\text{Yukawa}} = -\frac{y_e}{\sqrt{2}}\left(v(\bar{e}_L e_R + \bar{e}_R e_L) + (\bar{e}_L e_R + \bar{e}_R e_L)h\right), \tag{2.33}$$

including the electron mass term $m_e = y_e v/\sqrt{2}$ as well as the electron–Higgs boson interaction with the coupling strength m_e/v. Quark masses are generated analogously, with up-type quarks coupling to the conjugate Higgs doublet.[3] Neutrinos are massless in the Standard Model, although experimental evidence from neutrino oscillations indicates that they do possess non-zero masses. While the Yukawa sector relates fermion masses to their couplings to the Higgs field, the Yukawa couplings themselves remain free parameters of the theory. One of the central goals of the measurement strategy presented in this thesis is the determination of the top-quark mass parameter m_t.[4]

[3] The conjugate Higgs doublet is given by $\phi_c = i\sigma^2\phi^*$.

[4] Details on the definition of the top-quark mass are given in the next section.

2.2 Fixed-order calculation

As described in the previous section, the interactions of elementary particles in the SM are encoded in the SM Lagrangian (density) $\mathcal{L}_{\text{SM}}$, from which the corresponding Hamiltonian (density) $\mathcal{H}_{\text{SM}}$ can be derived. Theoretical predictions such as scattering cross-sections are based on computing transition amplitudes between an initial state in the far past and a final state in the far future: $|{}_{\text{out}}\langle p_1 p_2 \cdots | k_A k_B \rangle_{\text{in}}|^2$, where $|k_A, k_B\rangle$ describes the incoming particles, and the final state may contain several particles. The transition amplitude between these states can be computed by evolving them to a common reference time using the Hamiltonian,

$$ {}_{\text{out}}\langle p_1 p_2 \cdots | k_A k_B \rangle_{\text{in}} = \lim_{T\to\infty} \langle \underbrace{p_1 p_2 \cdots}_{+T} | \underbrace{k_A k_B}_{-T} \rangle = \lim_{T\to\infty} \langle p_1 p_2 \cdots | e^{-iH(2T)} | k_A k_B \rangle \,. \tag{2.34} $$

It is intuitive how a perturbative series arises: including more terms from the expansion of the exponential function, improves the accuracy of the transition amplitude calculation. The individual contributions can be represented diagrammatically by Feynman diagrams, which can be translated into mathematical expressions (resulting from the terms above) via Feynman rules. Physical observables are defined by summing over all final states that are experimentally indistinguishable — for example, an n-particle final state and an $n+1$-particle final state, in which the additional particle is either collinear w.r.t. another external particle or has vanishing momentum. The Kinoshita–Lee–Nauenberg (KLN) theorem [8, 9] states that all contributions at the same order in the coupling constants need to be considered in the calculation of physical observables to obtain finite results. For the final-state of interest in this thesis, the $WWbb$ final state in fully leptonic decay ($b\bar{b}\ell^+\ell'^-\nu_\ell\bar{\nu}_{\ell'}$), the leading, non-trivial, order (LO) in the expansion of Equation 2.34, is given by diagrams of the form shown in Figure 2.1, referred to as Born contribution. The NLO corrections include virtual corrections, as shown in Figure 2.2, as well as real corrections, leading to an $(n+1)$-particle final state, as illustrated in Figure 2.3.

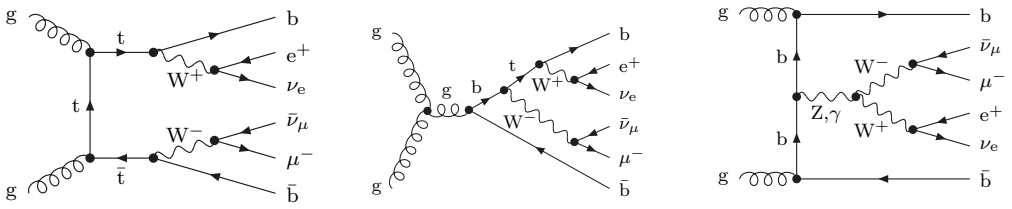


Figure 2.1: Exemplary LO Feynman diagrams of $b\bar{b}e^+\nu_e\mu^-\bar{\nu}_\mu$ production [10].

The full NLO differential cross-section can be written as

$$ \sigma^{\text{NLO}} = \int \mathrm{d}\Phi_n |\mathcal{M}_{\text{Born}}|^2 + 2\text{Re}\left(\mathcal{M}_{\text{Born}}\mathcal{M}^*_{\text{virtual}}\right) + \int \mathrm{d}\Phi_{n+1} |\mathcal{M}_{\text{real}}|^2, \tag{2.35} $$

where $\mathrm{d}\Phi_n$ denotes the n-particle Lorentz-invariant phase space and $\mathcal{M}$ the invariant matrix element, which corresponds to the non-trivial part of the transition amplitude between

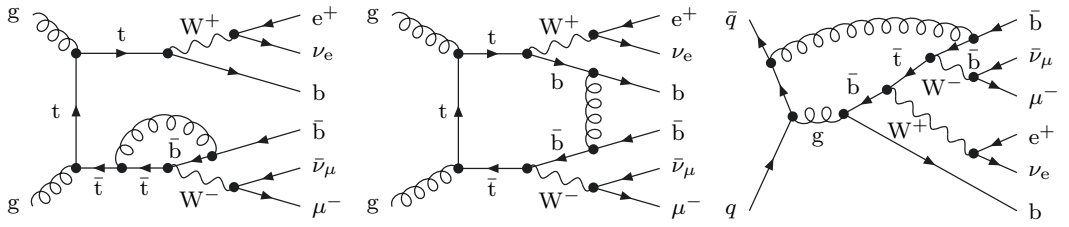


Figure 2.2: Exemplary Feynman diagrams of NLO virtual corrections to $b\bar{b}e^+\nu_e\mu^-\bar{\nu}_\mu$ production [10].

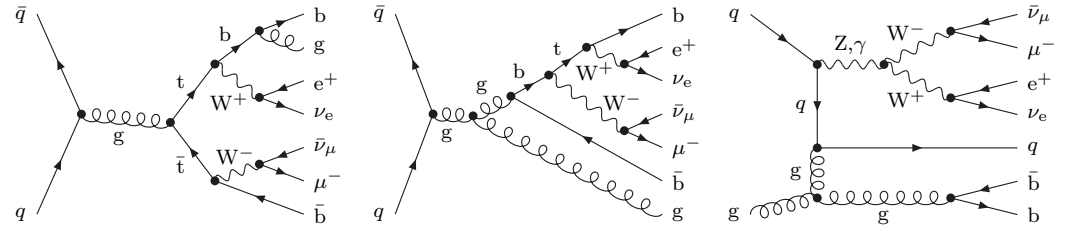


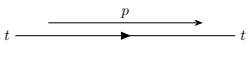
Figure 2.3: Exemplary Feynman diagrams of NLO real corrections to $b\bar{b}e^+\nu_e\mu^-\bar{\nu}_\mu$ production [10].

initial and final states. Details of the derivation are absorbed in these terms, as e.g. the representation of initial and final state as wave-packets instead of momentum eigenstates introduces flux factors and energy–momentum conservation leads to δ distributions in the momenta p_i, $i \in \{a, b, 1, \ldots, n\}$. The discussion so far has assumed quarks or gluons as particles in the initial state. However, only colourless hadrons are observed experimentally, and at the LHC production processes with protons in the initial state are of interest. The seemingly hopeless situation is resolved by the factorisation theorem, which separates the short-distance physics described by the partonic cross-section calculation $\hat{\sigma}(\Phi_{ab\to X})$ from the long-distance kinematics associated with the proton structure. The latter is encoded in the parton distribution functions (PDFs), which describe the probability of finding a parton a carrying the longitudinal momentum fraction x_a of the proton inside this proton. The production cross-section for the process $p_1p_2 \to X$ can thus be written as

$$\sigma_{p_1p_2\to X} = \sum_{a,b\in\{q,g\}} \int \mathrm{d}x_a \int \mathrm{d}x_b f_a^{p_1}(x_a, \mu_F^2) f_b^{p_2}(x_b, \mu_F^2) \int d\Phi_{ab\to X} \frac{d\hat{\sigma}(\Phi_{ab\to X}, \mu_F^2)}{\mathrm{d}\Phi_{ab\to X}}. \quad (2.36)$$

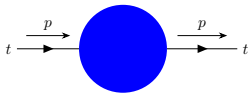
The factorisation scale μ_F is introduced in this procedure to separate long- and short-distance kinematics. In case the observable $\sigma_{p_1p_2\to X}$ could be calculated to all orders, the final result is independent of this auxiliary scale. In practice, a residual dependence on μ_F remains at fixed order and is used to estimate the uncertainty due to missing higher-order corrections. The PDFs themselves are not calculable from first principles in perturbative QCD, but their scale dependence is governed by perturbatively calculable evolution equations. Starting from non-perturbative input extracted from experimental data at a given scale, the PDFs are evolved to alternative scales using the probabilities for parton branchings, such as $g \to q\bar{q}$.

Top-quark mass in matrix-element calculations In Section 2.1, mass terms for fermions and bosons in the SM Lagrangian were described. The (bare) top-quark propagator depends on the (bare) top-quark mass m_0 in the Lagrangian as



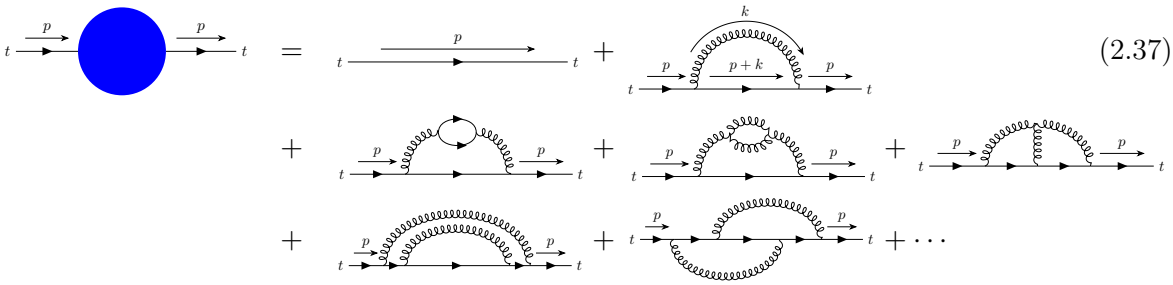
$$\frac{i(\not{p}+m_0)}{p^2-m_0^2+i\epsilon}=\frac{i}{\not{p}-m_0+i\epsilon}.$$

with $\not{p}=\gamma^\mu p_\mu$. Including higher-order QCD corrections leads to the dressed propagator,



$$\frac{i}{\not{p}-m_0-\Sigma(p,m_0,\mu_R)}$$

with the renormalisation scale μ_R. These higher-order perturbative corrections, include the following diagrams (see e.g. [11]):



(2.37)

The quark self-energy $\Sigma(p,m_0,\mu_R)$ is known to $\mathcal{O}(\alpha_S^4)$ [12, 13] and the diagrams are ultraviolet divergent due to loop diagrams with high-energy internal momenta. This behaviour can already be illustrated with a simpler scalar loop integral,

$$\int\frac{\mathrm{d}^4k}{16\pi^2}\frac{1}{k^2}\frac{1}{(k+p)^2}, \tag{2.38}$$

which exhibits a logarithmic divergence for $k\to\infty$. To describe these divergences, a regularisation procedure is required. A commonly used scheme is dimensional regularisation, where the space-time dimension in the momentum integral is continued to $4-2\epsilon$. A generic loop integral then takes the form

$$g^2\int\frac{\mathrm{d}^4k}{k^4}\cdots\xrightarrow{\text{dim.reg.}}g^2\mu_R^{2\epsilon}\int\frac{\mathrm{d}^{4-2\epsilon}k}{k^4}\cdots=ig^2\mu_R^{2\epsilon}\left(\frac{C_{-1}}{\epsilon}+C_0+C_1\epsilon+\mathcal{O}(\epsilon^2)\right). \tag{2.39}$$

Thereby, the renormalisation scale μ_R was introduced to maintain the correct mass dimension of the integral.[5] The ultraviolet divergence is now encoded in the limit $\epsilon\to0$. Rewriting

[5] Since the action $\int\mathrm{d}^{4-2\epsilon}x\mathcal{L}$ needs to be dimensionless, the mass dimension of the Lagrangian in dimensional regularisation needs to be $4-2\epsilon$. By requiring that the kinematic terms fullfill this condition, the mass dimensions of the fields are determined and in the interaction terms the coupling constant has to be rescaled to $g_0\to g_0\mu_R^\epsilon$, see also [14].

$\mu_R^{2\epsilon} = \exp[\epsilon \log \mu_R^2]$ and expanding the exponential in a Taylor series in ϵ leads to

$$\begin{aligned} \mu_R^{2\epsilon}\left(\frac{C_{-1}}{\epsilon} + C_0 + C_1\epsilon + \mathcal{O}(\epsilon^2)\right) &= (1 + \epsilon \log \mu_R^2 + \mathcal{O}(\epsilon^2))\left(\frac{C_{-1}}{\epsilon} + C_0 + C_1\epsilon + \mathcal{O}(\epsilon^2)\right) \\ &= \frac{C_{-1}}{\epsilon} + C_0 + C_{-1}\log \mu_R^2 + \mathcal{O}(\epsilon). \end{aligned} \tag{2.40}$$

In the full calculation a term including a logarithm of a typical energy scale of the problem $\log Q^2$ is present, leading to a dimensionless logarithm $C_{-1} \log \frac{\mu_R^2}{Q^2}$. The quark self-energy can be schematically written as

$$\Sigma(p, m_0, \mu_R) \sim m_0 \frac{\alpha_S(\mu_R)}{\pi}\left(\frac{1}{\epsilon} + \ln(4\pi e^{-\gamma_E}) + A^{\text{fin}}(m_t/\mu_R)\right) + \mathcal{O}(\alpha_S^2) \tag{2.41}$$

with the finite term A^{fin} as $\epsilon \to 0$. In the top-quark mass renormalisation the divergent terms in $\Sigma(p, m_0, \mu_R)$ are absorbed into the renormalised mass definition.[6] The mass renormalisation scheme is defined by which terms of the quark self-energy are included in the renormalised mass definition. The pole mass scheme is defined such that top-quark propagator has a pole at $p^2 = m_t$, meaning $\not{p} - m_0 - \Sigma(p, m_0, \mu_R) = 0$ for $p^2 = m_t$. In contrast, in the $\overline{\text{MS}}$ scheme, only the $1/\epsilon$-pole and a conventional logarithmic term are included:

$$m_t^{\text{pole}} = m_0\left(1 + \frac{\alpha_S(\mu_R)}{\pi}\left(\frac{1}{\epsilon} + \ln(4\pi e^{-\gamma_E}) + A^{\text{fin}}(m_t/\mu_R)\right)\right) + \mathcal{O}(\alpha_S^2) \tag{2.42}$$

$$\bar{m}_t(\mu_R) = m_0\left(1 + \frac{\alpha_S(\mu_R)}{\pi}\left(\frac{1}{\epsilon} + \ln(4\pi e^{-\gamma_E})\right)\right) + \mathcal{O}(\alpha_S^2) \tag{2.43}$$

The conversion between different mass schemes can be performed order-by-order in perturbation theory.[7]

Testing the Standard Model with top-quark mass measurements Through the electroweak unification and the Higgs mechanism in the SM, described in Section 2.1, several parameters in the SM are directly related. These relations can be tested in electroweak precision measurements. One important example is the prediction of the W-boson mass [16], given by

$$M_W^2 = \frac{M_Z^2}{2}\left(1 + \sqrt{1 - \frac{\sqrt{8}\pi\alpha(1 + \Delta r)}{G_F M_Z^2}}\right), \tag{2.44}$$

[6] See e.g. the review [15].

[7] While in the $\overline{\text{MS}}$ scheme corrections above the scale μ_R are included in the renormalised mass definition, perturbative corrections from all scales are included in the pole-mass definition. This leads to the so-called infrared renormalon ambiguity, originating from low-energy fermion-bubble insertions into the gluon line of the $t \to tg \to t$ diagram. These contributions first decrease with the order of the calculation n, but after a turn-over point, start to grow, leading to divergent patterns. In the full calculation, these contributions cancel with non-perturbative corrections. But, as these are absorbed into the mass definition in the pole-mass scheme, this leads to the mentioned ambiguity. See the discussion in [15].

where M_Z denotes the Z-boson mass, α the fine-structure constant, G_F the Fermi constant and Δr encodes radiative corrections to the tree-level relations. These corrections are crucial for the electro-weak precision fits. In particular, loop contributions involving the top quark are sizeable, introducing a strong dependence on the top-quark mass. By combining multiple electroweak precision observables, global fits can simultaneously constrain the top-quark mass and the W-boson mass. The resulting predictions can then be compared to their direct experimental measurements, as illustrated in Figure 2.4. The inclusion of the measured Higgs-boson mass significantly improves the sensitivity of this consistency test. It should be mentioned, however, that the top-quark mass parameter entering the electroweak precision fits must be defined within a well-specified renormalisation scheme. It is noted in [16], that more precise measurements of the top-quark mass parameter based purely on kinematic methods, the so-called MC mass parameter, are of limited use in the context of electroweak precision fits. The interpretability of the top-quark mass extraction from the decay product kinematics is further discussed in Section 5.1.

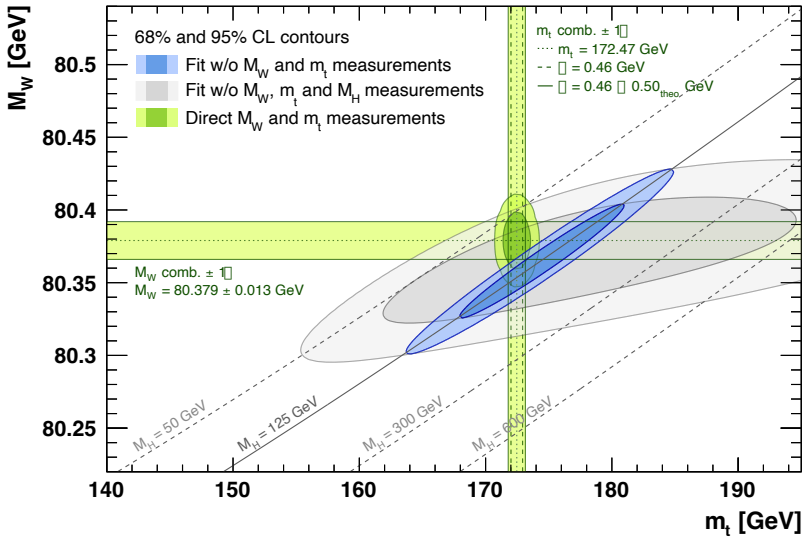


Figure 2.4: Contours at 68 % and 95 % confidence level obtained from scans of M_W versus m_t in the electroweak precision fit either including (blue) or excluding the direct measurement of M_H (grey). The predicted M_W and m_t values are compared to the results from direct measurements, shown with the horizontal and vertical green areas, as well the corresponding two-dimensional 1σ and 2σ ellipses [17].

Besides its role in electroweak precision fits, the top-quark mass parameter is also of central importance for the extrapolation of the SM to very high energy scales, up to the Planck scale of $\sim 10^{19}$ GeV [18–22]. The nature of the electroweak vacuum is determined by the shape of the Higgs potential introduced in Equation 2.22, which depends on the Higgs self-coupling λ. When higher-order corrections are taken into account, the Higgs self-coupling must be renormalised, thereby acquiring a scale dependence. This scale-dependence is thereby significantly influenced by loop corrections involving the top quark. In particular, these contributions can lead to negative values of λ for high scales, such that the Higgs potential may develop either a deeper minimum at large field values or become unbounded from below. Depending on the values

of the Higgs-boson mass and the top-quark mass — again, in a well-defined renormalisation scheme — the electroweak vacuum can be classified as stable, metastable, or instable. A metastable potential corresponds to the case where the electroweak vacuum is not the global minimum of the potential, but the probability for quantum tunnelling to the deeper minima of the potential is sufficiently suppressed such that the vacuum lifetime exceeds the age of the Universe. This interpretation relies on the assumption that no additional contributions from physics beyond the Standard Model modify the running of λ up to the Planck scale. Current measurements of the Higgs-boson and top-quark masses indicate that the Standard Model lies close to the boundary between stability and metastability. Future improvements in the precision of the m_t and m_H measurements, for instance at the High-Luminosity LHC, will help to probe this boundary region further, as illustrated in Figure 2.5.

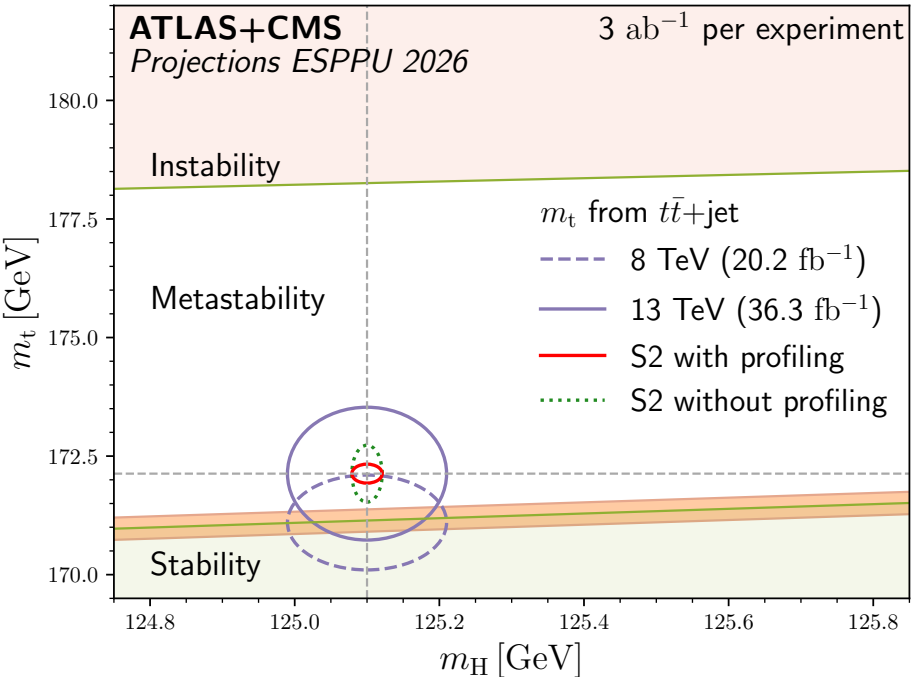


Figure 2.5: The 1σ (m_t, m_H)-ellipses using the current top-quark pole mass measurements from ATLAS at $\sqrt{s} = 8\,\text{TeV}$ [23] and CMS at $\sqrt{s} = 13\,\text{TeV}$ [24] and the expected sensitivity at the High Luminosity LHC are compared to the (m_t, m_H)-parameter regions, which indicate a stable, metastable or instable electroweak vacuum [25].

Top-quark decay width in matrix element calculations The total decay width of an unstable particle t is defined as

$$\Gamma_t = \frac{1}{2m_t} \sum_X \int d\Pi_X |\mathcal{M}(t \to X)|^2, \tag{2.45}$$

where $\mathcal{M}(t \to X)$ denotes the matrix element for the decay into a final state X, and $d\Pi_X$ represents the corresponding phase-space element.

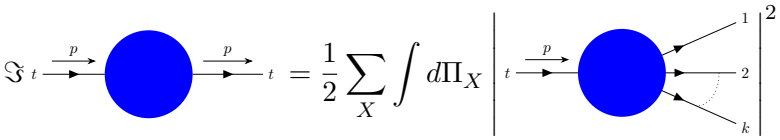
For the top quark, the decay is overwhelmingly dominated by the channel $t \to Wb$, such that $\Gamma_t \simeq \Gamma(t \to Wb)$. At LO in QCD, the partial width is given by [26]

$$\Gamma_t^{\mathrm{LO}}(t \to Wb) = \frac{G_F m_t^3}{8\sqrt{2}\pi} |V_{tb}|^2 \left(1 - \frac{m_W^2}{m_t^2}\right)^2 \left(1 + 2\frac{m_W^2}{m_t^2}\right), \tag{2.46}$$

where terms of order m_b^2/m_t^2 are neglected and $|V_{tb}| \simeq 1$ denotes the CKM matrix element. This expression highlights the strong dependence of the decay width on the top-quark mass, $\Gamma_t \propto m_t^3$. How the top-quark decay width enters in cross-section predictions can be seen with the optical theorem,[8] which relates the imaginary part of the $\mathcal{M}(t \to t)$ matrix element calculation to the top-quark decay width as

$$\Im\mathcal{M}(t \to t) = \sum_X \int d\Pi_X |\mathcal{M}(t \to X)|^2 \tag{2.47}$$

or pictorially:



(2.48)

In fact, in the derivation of the self-energy ($\Sigma(p^2) = -\mathcal{M}(t \to t)(p2)$) for the definition of different renormalised top-quark masses, we considered stable top-quarks. Instead, top-quarks decay — actually very fast — before top-quarks can hadronise in the order of 10^{-25} s. Compared to the top-quark mass, the top-quark decay width can still be considered small, as $\Gamma_t \ll m_t$ and thus formulae for long-lived particles can be applied. The top-quark mass pole mass m_t is defined through the real part of the quark self-energy diagram. The imaginary part of the self energy of an unstable particle is reflected in the propagator

$$\frac{i}{p^2 - m_0^2 - \Sigma(p^2)} \to \frac{i}{p^2 - m_t^2 - i\Im\Sigma(p^2)} = \frac{i}{p^2 - m_t^2 + i m_t \Gamma_t} \tag{2.49}$$

and leads to cross-sections for the production and decay of unstable particles including a characteristic Breit–Wigner shape of

$$\sigma(X \to t \to Y) \sim \left|\frac{1}{s - m_t^2 + i m_t \Gamma_t}\right|^2 = \frac{1}{(s - m_t^2)^2 + m_t^2 \Gamma_t^2} \tag{2.50}$$

in the kinematic region $p^2 \sim m_t^2$. An NNLO QCD and NLO EW correction accuracy could be reached in the theoretical prediction of the total top-quark decay width [29–31].

[8] The discussion follows [27] and Chapter 7.3 of [28].

Top-quark decay width sensitivity to Beyond the Standard Model theories The top-quark decay width parameter is sensitive to BSM contributions, in case they either enter loop corrections to the top-quark decay width calculation or describe new top-quark decay channels into BSM particles. While this constitutes a broad class of BSM theories, several examples for BSM top-quark decays are discussed below, following [32, 33].

- **Top-quark decay to charged Higgs:** Several BSM scenarios include Two Higgs Doublet Models (2HDM), which extend the SM Higgs sector by introducing a second Higgs doublet. After electroweak symmetry breaking, this results in two neutral scalars (h^0, H^0), one neutral pseudoscalar (A^0), and a pair of charged Higgs bosons ($H^\pm$). The impact of this extended Higgs sector on the top-quark decay width is twofold: it modifies loop corrections to the SM top-quark decay width and opens a new decay channel $t \to H^+ b$, provided that $m_{H^+} < m_t$. Both effects depend on the parameter $\tan\beta$, defined as the ratio of the vacuum expectation values of the two Higgs doublets. Electroweak supersymmetric (SUSY) loop corrections can modify the decay width by up to $\mathcal{O}(10\%)$ [34]. The leading-order partial width for the decay into a charged Higgs boson is given by

$$\Gamma(t \to bH^+) = \frac{G_F m_t^3}{8\sqrt{2}\pi} \left(1 - \frac{m_{H^+}^2}{m_t^2}\right)^2 \left[\left(\frac{m_b}{m_t}\right)^2 \tan^2\beta + \cot^2\beta\right], \tag{2.51}$$

  which is minimal at $\tan\beta \approx 6$–8. However, for small or large values of $\tan\beta$, the corrections to Γ_t can be substantial, reaching even 100 % when including higher-order SUSY corrections [35].

- **Top-quark decay to stop:** Supersymmetric extensions of the SM predict a supersymmetric partner for each SM particle. In the Minimal Supersymmetric Standard Model (MSSM), the scalar partner of the top quark (stop) can be sufficiently light to allow the decay $t \to \tilde{t}\tilde{\chi}^0$, where $\tilde{\chi}^0$ denotes a neutralino. This channel is kinematically allowed only if the stop mass is smaller than the top-quark mass. In such scenarios, the total top-quark decay width can be modified by up to $\sim 20\%$. However, current experimental searches impose strong constraints on the relevant stop mass parameter space [36–39].

- **Flavour Changing Neutral Currents top-quark decay:** In the SM, flavour-changing neutral current (FCNC) top-quark decays, such as $t \to c\gamma$, are absent at tree level and highly suppressed at loop level. The corresponding branching fractions, $\mathrm{BR}(t \to Xc)$ with $X \in \{g, Z, \gamma, H\}$, are predicted to lie in the range 10^{-12} to 10^{-15} [40]. In contrast, many BSM scenarios allow FCNC processes at tree level, leading to substantial enhancements. Branching fractions as large as $\mathcal{O}(10^{-4}$–$10^{-5})$ are predicted in some models, several orders of magnitude above the SM expectations [41].

2.3 Parton shower

While fixed order calculations describe a production process for all kinematic configurations of the initial and final states exactly at a given order in perturbation theory, the calculable final-state multiplicity is limited due to the rapidly increasing complexity of the calculations with the number of final state particles.[9] To simulate realistic events that can be compared to data, which typically exhibit high final-state multiplicities of $\mathcal{O}(100) - \mathcal{O}(1000)$ [42], the process-independent factorisation of soft-collinear emissions is iteratively applied. Specifically, the cross-section for producing $n+1$ final-state particles can be approximated in terms of the cross-section for producing n final-state particles and a universal branching probability:

$$\mathrm{d}\sigma_{n+1} \simeq \mathrm{d}\sigma_n \cdot \sum_j K_{\tilde{\imath}\to ij}\mathrm{d}\Phi_{+1}. \tag{2.52}$$

Here, the radiation functions $K_{\tilde{\imath}\to ij}$ describe the emission of a parton j in the branching $\tilde{\imath} \to ij$. This factorisation is exact in the soft and collinear limit and is the basis of parton-shower algorithms. Graphically, this is represented in Figure 2.6.

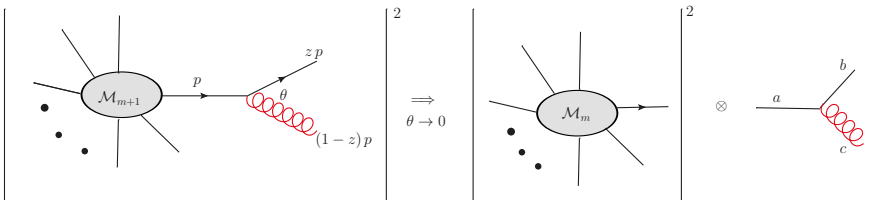


Figure 2.6: Graphic representation of collinear factorisation of the matrix-element squared [43].

The collinear limit In the limit, in which two massless final-state particles i and j become collinear, the squared matrix element $|\mathcal{M}_{n+1}|^2$ factorises as

$$|\mathcal{M}_{n+1}|^2 \to \frac{8\pi\alpha_s}{2p_i \cdot p_j} P_{\tilde{\imath}\to ij}|\mathcal{M}_n|^2, \tag{2.53}$$

where $P_{\tilde{\imath}\to ij}$ are the Altarelli–Parisi splitting functions, which only depend on the type of parton splitting and the momentum sharing fraction.

The soft limit The soft limit refers to the emission of a gluon j with vanishing four-momentum. In this limit, the squared matrix element factorises as [44]

$$|\mathcal{M}_{n+1}|^2 \to 8\pi\alpha_s \sum_{i<k} \mathcal{C}_{ik} \frac{2p_i \cdot p_k}{(2p_i \cdot p_j)(2p_k \cdot p_j)} |\mathcal{M}_n|^2, \tag{2.54}$$

[9] The number of Feynman diagrams required to compute a matrix-element squared grows factorially with the number of final-state particles [42].

where $\mathcal{C}_{ik}$ depends on the colour charges of the emitting partons i and k in interfering matrix elements. When terms suppressed by $1/N_c^2$ are neglected, interference contributions between emissions from different colour lines vanish. In this leading-colour approximation ($N_c \to \infty$), a soft gluon emission can be treated as an independent emission from a single colour dipole, i.e., a colour-connected pair of partons. An important consequence of including interference effects beyond the leading-colour approximation is colour coherence, which gives rise to the angular ordering of soft gluon emissions. To illustrate this, consider soft gluon emission from a $q\bar{q}$ pair produced by a virtual photon. The differential cross-section for the resulting $q\bar{q}g$ final state can be decomposed into contributions associated with the quark and anti-quark, W_q and $W_{\bar{q}}$. After integrating over the azimuthal angle ϕ of the soft gluon, these terms lead to:

$$\int \frac{\mathrm{d}\phi}{2\pi} W_{q/\bar{q}} = \begin{cases} \frac{1}{1-\cos\theta_{g(q/\bar{q})}} & \text{if } \theta_{g(q/\bar{q})} < \theta_{q\bar{q}} \\ 0 & \text{otherwise,} \end{cases} \tag{2.55}$$

where $\theta_{g(q/\bar{q})}$ is the angle between the soft gluon and the quark q or anti-quark $\bar{q}$ and $\theta_{q\bar{q}}$ is the angle between the quark and anti-quark. Colour coherence in this example thus leads to a suppression of soft gluons outside a cone with opening angle $\theta_{q\bar{q}}$ around the quark and anti-quark direction, while the soft gluon emission inside the cone can be treated as independent. Intuitively, a soft gluon at large angles cannot resolve the individual colour charges of the quark and the anti-quark. Instead, it couples to their combined colour charge, which vanishes for a colour-singlet system such as a virtual photon. More generally, soft emissions probe the total colour charge of partons that are unresolved at the emission angle, leading to angular ordering in parton showers. In addition to the factorisation of the squared matrix element, the phase space also factorises in the soft and collinear limits. The term $d\Phi_{+1}$ in Equation 2.52 therefore represents the phase space associated with the additional emission.

Iterating the factorisation described above yields an all-orders description of additional QCD radiation. This description is exact only in the limit where these emissions are either soft, i.e. have infinitesimal energy, or collinear to their emitters, i.e. their three-momentum is aligned with that of the emitter. The differential cross-section in Equation 2.52 describes the inclusive production of one additional parton in the soft or collinear limit at fixed order. To obtain an exclusive description of high multiplicity final states — i.e. events with a well-defined number of resolved partons above a certain resolution scale — it is necessary to include the probability that no emission occurs above a given scale. This is achieved through the Sudakov form factor, and the procedure is applied iteratively. An intuitive understanding for the Sudakov form factor can be obtained from the analogy in [45], which considers the decay of a radioactive isotope with a time-dependent decay width $\Gamma(t)$, and consequently a time-dependent half-life $\tau(t)$, where $\Gamma(t) = 1/\tau(t)$. The probability that the isotope has not decayed between the time 0 and t is given by

$$P^{\text{nodec}}(t, 0) = \exp\left(-\int_0^t \Gamma(t')\mathrm{d}t'\right). \tag{2.56}$$

The probability density for the decay to occur at time t is then obtained by differentiating the decay probability $P^{\text{dec}}(t,0) = 1 - P^{\text{nodec}}(t,0)$ with respect to t:

$$\frac{\mathrm{d}P^{\text{dec}}(t,0)}{\mathrm{d}t} = \Gamma(t)P^{\text{nodec}}(t,0). \tag{2.57}$$

In the context of parton showers, the decay rate $\Gamma(t)$ is replaced by the splitting probability density. The Sudakov form factor $\Delta(T,t)$ can thus be interpreted as the probability for no branching to occur between scales t and T. The cross-section after one step of the parton shower evolution from an n-particle state can then be written as

$$\mathrm{d}\sigma_{n+1} = \mathrm{d}\sigma_n \left(\underbrace{\Delta(t_{\text{min}})}_{\text{no emission above } t_{\text{min}}} + \int_{t>t_{\text{min}}} \underbrace{\frac{\mathrm{d}t}{t}\frac{\alpha_s(t)}{2\pi}\int \mathrm{d}z K_{\tilde{\imath}\to ij}(z,t)}_{\text{emission at scale } t} \underbrace{\Delta(t)}_{\text{no emission above } t} \right), \tag{2.58}$$

where the Sudakov form factor is defined as

$$\Delta(t) = \exp\left[-\int_{t'>t} \frac{\mathrm{d}t'}{t'} \int \mathrm{d}z \frac{\alpha_S(t')}{2\pi} K_{\tilde{\imath}\to ij} \right]. \tag{2.59}$$

The evolution is terminated at a cutoff scale t_{min}, below which non-perturbative effects become important.

From the previous discussion, it might appear that the parton shower approach is simply a convenient tool for describing high-multiplicity final states. However, a reliable description of the soft-collinear phase-space region is not possible by using fixed-order matrix-element calculations alone due to infrared divergences. These divergences are encoded in Equation 2.53 and in Equation 2.54 in the respective limits of $p_i p_j = E_i E_j (1-\cos\theta) \xrightarrow{\theta\to 0} 0$ (for massless particles) and $p_j \to 0$. Through the all-order resummation of the leading logarithmic terms associated with these divergences[10] in the soft-collinear region through the Sudakov factor, this divergence is appropriately regulated. This is illustrated in Figure 2.7, which shows a generic observable that is sensitive to soft and collinear radiation. An example is the transverse momentum (p_{T}) of a gluon in a $q\bar{q}g$ final state. When computed at fixed-order NLO, the prediction exhibits a divergence in the soft-collinear region. Also, the fixed-order description of the soft-collinear region (LO+$\text{PS}_{\text{NLO\_FO}}$) — corresponding to Equation 2.52 — shows this divergence. In contrast, in both the LO+PS and the NLO+PS predictions the divergence is regulated through the parton shower approach, leading to finite and similar predictions in the soft-collinear region. Differences between the two approaches arise in the hard-emission region, in which the improved fixed-order accuracy of the NLO+PS prediction leads to a more reliable description, resulting in visible deviations from the LO+PS result.

While important components of the Monte-Carlo event generators Pythia and Herwig, including the underlying event modelling and hadronisation, have not been discussed here, they will be addressed in Section 3.4, along with the associated modelling uncertainties in the ATLAS $bb4\ell$ Monte Carlo sample.

[10] Logarithmic divergences appear in the phase-space integral of the real radiation through terms like $\mathrm{d}t/t$ [42].

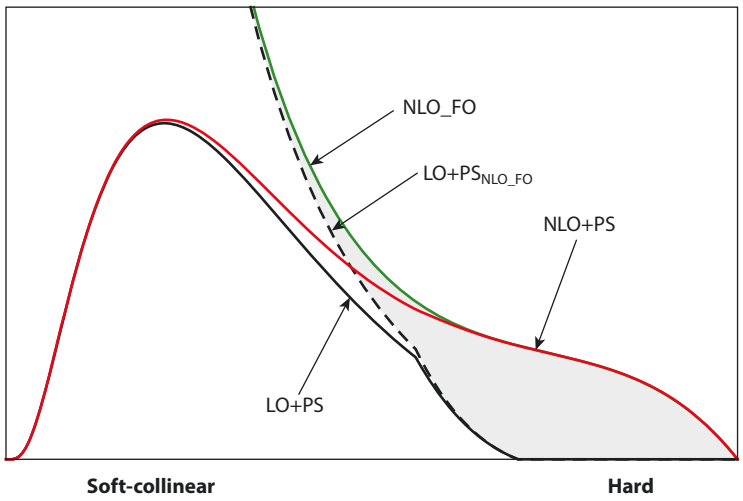


Figure 2.7: Comparison of exact fixed-order NLO QCD (NLO_FO) predictions, as well as the fixed order result in the soft-collinear approximation (LO+PS$_{NLO_FO}$) corresponding to Equation 2.52 to the NLO+PS and LO+PS results for a generic distribution, which is sensitive to soft-collinear radiation [15].

2.4 The Powheg method

When combining fixed-order NLO QCD calculations with parton showers, care must be taken to avoid double counting of real emission contributions. While the parton shower provides an approximate description of QCD radiations that is formally accurate only in the soft and collinear limits, the real-emission contribution in a NLO calculation describes the production of one additional parton at tree level across the full phase space. As a result, emissions in the soft-collinear region are described by both approaches, leading to an overlap and thus double counting if they are naively combined. Moreover, an important difference between fixed-order calculations and parton showers is that fixed-order results are inclusive with respect to additional radiation, whereas parton showers are exclusive above a certain emission scale. As described in the previous Section, parton showers explicitly veto emissions above a specific resolution scale, while fixed-order calculations treat such emissions as possible but unobserved contributions.

A consistent matching procedure must therefore both remove the double counting of real emissions and combine the inclusive description of fixed-order calculations with the exclusive formulation of the parton shower. One such method is for instance provided by the POWHEG approach.

The Powheg formula The POWHEG method [46, 47] achieves NLO accuracy while preserving the leading-logarithmic accuracy of the parton shower by generating the hardest emission first, using the full NLO QCD matrix elements, and subsequently letting the parton shower account for softer emissions. In this way, double counting is avoided while retaining the

strengths of both descriptions. Unlike pure fixed-order calculations, which yield inclusive predictions, POWHEG generates events which are exclusive with respect to radiation above the transverse-momentum scale of the real emission. This is achieved through the introduction of a Sudakov form factor, which resums leading logarithmic contributions and ensures a no-emission probability above the scale of the hardest emission. When a parton shower is subsequently applied, it is restricted to generate only emissions which are softer than the one already generated by POWHEG. The POWHEG method is realized through the generation of hard-process events according to the POWHEG formula:

$$\mathrm{d}\sigma^{\mathrm{PWHG}} = \bar{B}(\Phi_n)\mathrm{d}\Phi_n \left(\Delta_{\mathrm{PWHG}}(\Phi_n, p_{\mathrm{T}}^{\mathrm{min}}) + \sum_{\alpha} \frac{d\Phi_{\mathrm{rad}}^{\alpha} R^{\alpha}(\Phi_{n+1}) \Delta_{\mathrm{PWHG}}(\Phi_n, k_T(\Phi_{n+1}))}{B(\Phi_n)} \right). \tag{2.60}$$

The term $\bar{B}(\Phi_n)$ represents the NLO-accurate differential cross-section evaluated at Born-level kinematics:

$$\bar{B}(\Phi_n) = B(\Phi_n) + V(\Phi_n) + \sum_{\alpha} \int d\Phi_{\mathrm{rad}}^{\alpha} \left[R^{\alpha}(\Phi_{n+1}) - \mathcal{C}^{\alpha}(\Phi_{n+1}) \right]. \tag{2.61}$$

Here, $B(\Phi_n)$ denotes the Born contribution, including parton distribution functions, flux factors, and sums over spin and colour degrees of freedom. The real-emission contribution $R = \sum_{\alpha} R^{\alpha}$ contains soft and collinear divergences when integrated over the full phase space. These divergences are cancelled by corresponding singularities in the virtual NLO corrections, yielding finite physical predictions at NLO. For a Monte-Carlo implementation of the NLO calculation, this cancellation is achieved by introducing subtraction terms $\int \mathrm{d}\Phi_{\mathrm{rad}}^{\alpha} \mathcal{C}^{\alpha}(\Phi_{n+1})$, which reproduce the functional form of the soft-collinear singularities. These terms are added to the virtual contribution, yielding the finite term $V(\Phi_n)$ and subtracted from the real contribution. The integral over the subtracted real contribution $R^{\alpha}(\Phi_{n+1}) - \mathcal{C}^{\alpha}(\Phi_{n+1})$ is thus finite. The subscript α refers to the different phase space regions used to isolate the singularities, reflecting the partitioning of the real emission amplitude into soft and collinear regions. A similar concept was already introduced in the previous section, where the $n+1$ final-state differential cross-section was constructed from the n final-state differential cross-section with an additional soft or collinear radiation. The POWHEG Sudakov factor,

$$\Delta_{\mathrm{PWHG}}(\Phi_n, p_{\mathrm{T}}) = \exp\left(- \sum_{\alpha} \int_{k_t(\Phi_{n+1}) > p_{\mathrm{T}}} \frac{d\Phi_{\mathrm{rad}}^{\alpha} R^{\alpha}(\Phi_{n+1})}{B(\Phi_n)} \right), \tag{2.62}$$

describes the probability of no radiation occurring above a given transverse momentum scale p_{T}. Comparing this expression to the Sudakov form factor used in parton showers in Equation 2.59, the ratio $R(\Phi_{n+1})/B(\Phi_n)$ replaces the splitting kernel $\frac{\alpha}{2\pi}\frac{1}{t}K_{\tilde{i}\to ij}$ used in the shower approximation. Indeed, this is exactly how the splitting kernel $K_{\tilde{i}\to ij}$ are derived, i.e. reproducing the ratio of $\mathrm{d}\sigma_{n+1}/\mathrm{d}\sigma_n$ in the soft and collinear limit. Finally, the full POWHEG formula in Equation 2.60 closely resembles that of the parton-shower expression for a single emission (Equation 2.58): both include a no-emission probability for emissions above a certain cut-off scale $p_{\mathrm{T}}^{\mathrm{min}}$ and generate exactly one additional emission at a scale p_{T}, while vetoing harder emissions. It has been shown [46, 47], that the POWHEG formula leads

to NLO accurate observable description. Thereby, the fixed order NLO result is reproduced and the additionally included terms are of higher order in the strong coupling α_s.[11]

Interface to the parton shower If a POWHEG event is generated according to the first term on the right-hand side of Equation 2.60, no resolved real emission is produced and the final-state particle content corresponds to the Born configuration. In this case, these final state particles have been evolved until $p_{\mathrm{T}}^{\mathrm{min}}$ with the POWHEG Sudakov. When interfaced with a parton shower, only emissions with transverse momentum below the minimum POWHEG scale $p_{\mathrm{T}}^{\mathrm{min}}$ are allowed. If instead the POWHEG event is generated according to the second term on the right-hand side of Equation 2.60, a real emission is included, and the event corresponds to an $(n+1)$-particle final state. Thereby, the POWHEG formula includes the no-emission probability for emissions harder than the generated real-emission scale $k_{\mathrm{T}}^{\mathrm{Pwhg}}$. To avoid double counting, the subsequent parton shower must therefore only generate emissions with $p_{\mathrm{T}} < k_{\mathrm{T}}^{\mathrm{Pwhg}}$.

For parton showers ordered in transverse momentum, such as PYTHIA 8, this matching is implemented using a vetoed shower. Since the definition of transverse momentum in the shower typically differs from that used in POWHEG, the shower is started at the maximal phase-space scale. Emissions are then generated according to the shower algorithm, but any emission with a transverse momentum — recomputed using the POWHEG definition — exceeding $k_{\mathrm{T}}^{\mathrm{Pwhg}}$ is vetoed. For angular-ordered parton showers, such as HERWIG, a modified matching procedure is required. In these angular-ordered parton showers, emissions are generated starting from wide-angle radiation, generally not those emissions with the highest transverse momentum. If the parton shower were started directly from the $(n+1)$-particle configuration provided by POWHEG, the large-angle soft emissions, which HERWIG would have generated first, when starting from the Born configurations, would be missed. Therefore, a truncated shower is introduced. This modified evolution begins from a reconstructed Born-level configuration and evolves the system from the maximum allowed angular scale $\tilde{q}$ down to the scale $\tilde{q}_{\mathrm{Pwhg}}$, which corresponds to the angular scale associated with the POWHEG real emission. After this step, the real-emission final-state kinematics generated by POWHEG are restored, and the standard parton-shower evolution continues. During this evolution, a transverse momentum veto is imposed to ensure that all parton-shower-generated emissions satisfy $p_{\mathrm{T}} < k_{\mathrm{T}}^{\mathrm{Pwhg}}$.

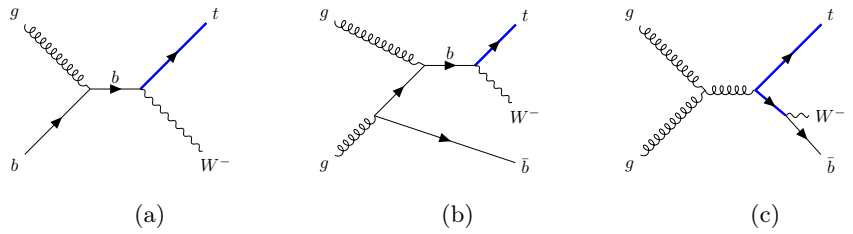
2.5 Predicting $WWbb$ with NLO+PS accuracy

In this section, the available NLO+PS implementations for the description of the $WWbb$ production process with varying theoretical accuracy are presented. Starting from the production of on-shell $t\bar{t}$ pairs, which, together with the separately simulated tWb production process, describes the $WWbb$ final state, the full NLO QCD calculation of the $WWbb$ production process with dileptonic decay, as implemented in the $bb4\ell$ POWHEG generator and used in this thesis, is discussed.

[11] These higher-order terms arise from the expansion of the POWHEG Sudakov form factor in Equation 2.62, but do not spoil the NLO accuracy of the method.

The $t\bar{t}$ production process Predictions for the production of a $t\bar{t}$ final state at hadron colliders, which can be matched to a parton shower[12] are well established at NLO QCD accuracy [58–62], and more recently also available at NNLO QCD accuracy [63, 64] via the MINNLO method [65, 66]. Since a $t\bar{t}$ final state is simulated, the top quarks in the ME calculations are on-shell, meaning that their invariant mass is fixed to the nominal input top-quark mass used in the calculation. For a precise experimental measurement, the accurate modelling of the kinematics of the top-quark decay products is crucial. The most precise description is achieved when simulating the full $2 \to 6$ process, such as $pp \to b\bar{b}\ell^-\bar{\nu}_\ell\ell'^+\nu_{\ell'}$, which will be discussed in more detail later. But, due to the much larger top-quark mass compared to the top-quark width, $m_t \gg \Gamma_t$, the narrow-width approximation (NWA) can be and is used in the calculations described above. In the NWA, i.e. $\Gamma_t \to 0$, the top-quark production and decay subprocess can be factorised, in the sense that there is no interference between the two subprocesses. The NWA is improved in the common $t\bar{t}$ NLO+PS generators, by including spin correlations and a Breit–Wigner smearing of the top-quark mass through the FLMW approach [67]. This is implemented, for example, in the POWHEG heavy-flavour hadroproduction process (`hvq`) [59] or in an automated way in MADSPIN [68]. Differences in implementation of the FLMW method — particularly in how final-state momenta are reshuffled to reflect off-shell top quarks — lead to differences in the predicted distributions. These two implementations are used to define the ATLAS top-quark line-shape uncertainty in $t\bar{t}$ MC samples. Although, the top-quark decay is formally described at LO QCD accuracy, matrix element corrections in the PYTHIA 8 parton shower provide a reasonable approximation to the NLO QCD decay description [2]. A NLO+PS generator, that includes both top-quark production and top-quark decay at NLO QCD accuracy within the NWA is available in [69].[13]

The tW production process Generators predicting tW production at NLO QCD, matched to a parton shower, are available in the five-flavour number scheme (5FNS) [75–77], i.e. treating the b quark as massless. The LO contributions and exemplary real NLO correction to the tW production process are illustrated in Figure 2.8. At NLO, diagrams corresponding to processes



(a) (b) (c)

Figure 2.8: Exemplary Feynman-diagrams of: (a) LO tW production; (b) single-resonant NLO real correction; (c) double-resonant NLO real correction.

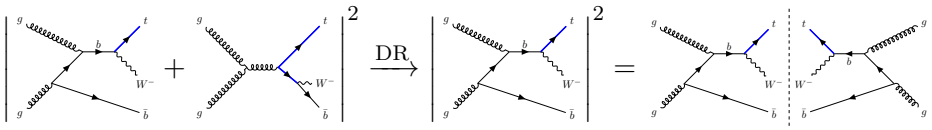
with two resonant top quarks (Figure 2.8(c)) must also be considered. These lead to the same

[12] Fixed-order $t\bar{t}$ on-shell predictions are available at NNLO QCD [48–51], NLO EW [52–56] and NNLO QCD + NLO EW [57] accuracy.

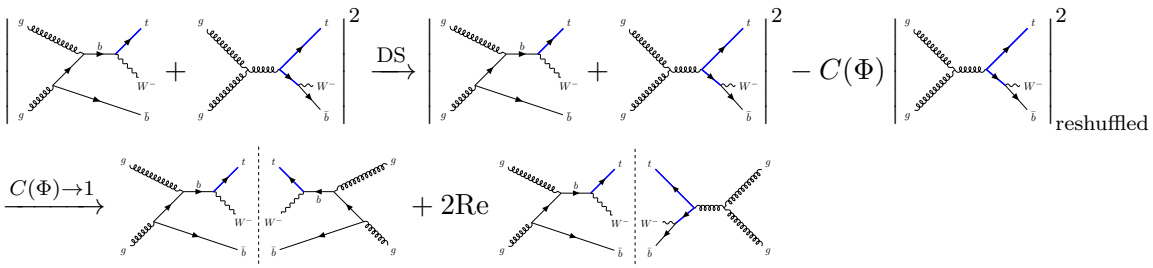
[13] Fixed order $t\bar{t}$ predictions are available in the NWA with NNLO QCD accuracy both in top-quark production and decay [70–74].

tWb final state as tW production with an additional real b-quark emission (Figure 2.8(b)). However, since the LO tW production cross-section is approximately 15 times smaller than the LO cross-section for double-resonant top-quark production [76], defining tW production at NLO as a standalone single-top process becomes ambiguous, since the $t\bar{t}$ contributions dominate the total NLO cross-section. At matrix-element squared level, two terms involving double-resonant production contribute: the squared amplitude of double-resonant diagrams, and the interference terms between single-resonant and double-resonant diagrams. To isolate the associated single-top tW production process at NLO, two methods can be applied to subtract the double-resonant contributions [76]:

- **Diagram Removal (DR)**: All diagrams involving double-resonant contributions are excluded from the calculation. This also removes the interference terms between the $t\bar{t}$ and tW production processes. Pictorially:



- **Diagram Subtraction (DS)**: All diagrams are included in the NLO calculation, and a subtraction term is introduced at cross-section level to locally remove the double-resonant contribution. This local subtraction term is based on the matrix-element squared of the double-resonant diagrams, with reshuffling of the final-state four-momenta to ensure that these are consistent with on-shell $t\bar{t}$ production. This is in fact not guaranteed, since off-shell effects in the intermediate anti-top quark resonance are taken into account through the inclusion of the full top-quark propagator. Further, a phase space dependent factor $C(\Phi)$ is introduced, approaching unity as $m_{Wb} \to m_t$. Pictorially:



Both methods have limitations: the DR scheme is not gauge invariant, while the DS differential cross-sections are not guaranteed to be positive definite across the whole phase space. Regions in phase space, in which large differences are observed between the DR and DS differential cross-sections, are particularly sensitive to interference effects between $t\bar{t}$ and tW production.[14]

[14] One example of an ATLAS measurement, which targets a phase space region particularly sensitive to the $t\bar{t}/tW$ interference modelling is performed in [78], where the $m_{\ell b}^{\text{minimax}}$ distribution of the $WWbb$ production process (simulated as stack of $t\bar{t} + tW$) was unfolded.

The $WWbb$ production process The issues discussed in the theoretical accuracy in the modelling of the top-quark decay products in on-shell $t\bar{t}$ production, as well as the ambiguities associated with defining NLO QCD corrections to tW production, can be resolved by including all Feynman diagrams for the full $2 \to 6$ process in the ME calculation. In the dileptonic decay channel, this corresponds for example to $pp \to b\bar{b}\ell^-\bar{\nu}_\ell\ell'^+\nu_{\ell'}$. While these calculations will be referred to as $WWbb$ production cross-sections in the following, they consistently account for off-shell effects of the intermediate W bosons and for production processes without an intermediate $WWbb$ state, since the complete process is computed explicitly. Fixed-order NLO QCD predictions for $WWbb$ production are available in both the 4FNS [79, 80] and 5FNS [10, 81–83] for the dileptonic channel. Predictions for the single-lepton channel are available in the 5FNS [84]. The 4FNS, in which b quarks are treated as massive and excluded from the initial state, enables a consistent treatment of both, the $t\bar{t}$ and tW topologies. In the 5FNS, initial-state collinear singularities from $g \to b\bar{b}$ splitting need to be suppressed by using adequate kinematic cuts, to yield finite results. This is not an issue in the previously discussed NLO QCD tW calculations in the 5FNS, since the corresponding infrared singularities in the virtual corrections, leading to a tW, i.e. WWb final state, cancel the divergence in the real corrections, i.e. $WWbb$ with $g \to b\bar{b}$ splitting. In the 4FNS $WWbb$ calculation, Feynman diagrams with a single top-quark propagator are described via topologies involving a $g \to b\bar{b}$ splitting in the initial state, ensuring a more accurate modelling of the spectator b-quark kinematics through the fixed-order matrix-element calculation. Fixed-order NLO QCD predictions of $WWbb$ production in both the dileptonic and single-lepton decay channel in the 4FNS, matched to a parton shower, are available via the $bb4\ell$ process [1, 2]. Implemented with the POWHEG method, these predictions thus include top-quark off-shell effects, as interference of double, single and non-resonant contributions, and NLO QCD corrections to the top-quark decay.

The $bb4\ell$ process in Powheg After having discussed general aspects of the NLO QCD calculation of the $WWbb$ final state and the basic concepts of the POWHEG method in Section 2.4, additional details about the implementation of the $bb4\ell$ process [1, 2] are presented in the following. The $bb4\ell$ predictions used in this thesis are obtained for the dilepton channel. Here, only matrix elements leading to different flavour leptons in the final state are included: $pp \to b\bar{b}\ell^-\bar{\nu}_\ell\ell'^+\nu_{\ell'}$ with $\ell \neq \ell'$. This choice avoids the need to include additional matrix elements in the fixed-order calculation that would be required for same-flavour final states, such as $pp \to ZZb\bar{b} \to \ell^+\ell^-\nu_\ell\bar{\nu}_\ell b\bar{b}$. Since the contribution of these additional topologies and their interference with the included topologies is expected to be small, the $bb4\ell$ MC sample used in this thesis is produced by also allowing same-flavour lepton final states, but still using the original ME calculation. This significantly increases the applicability of the MC sample in experimental data analyses. While the $WWbb$ ME calculation is performed in the 4FNS, the $bb4\ell$ implementation uses an interface to 5FNS PDFs. The different treatment of quark masses in the ME and PDFs is accounted for by applying a correction factor and, consistent with a 4FNS ME calculation, initial state b quarks from the PDFs are not considered. The ME calculation yields differential cross-sections in terms of the two initial-state and the final-state particles of the Born or real-emission configuration (e.g. $gg \to WWbb + X$), obtained by consistently including all double-, single- and non-resonant Feynman diagrams. Two main issues arise when applying the original POWHEG method, as implemented in POWHEG BOX [85], to processes involving intermediate resonances:

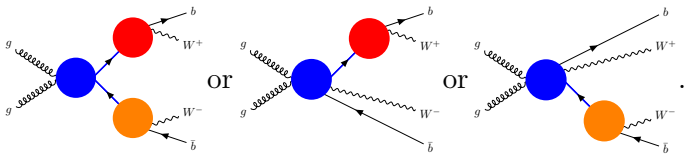
- **Mapping of the real to the underlying Born configuration** The mapping of real-emission kinematics to the corresponding underlying Born kinematics in a given singular region (soft or collinear) is crucial for both the construction of the NLO QCD subtraction terms and the computation of the R/B ratio in the POWHEG Sudakov factor. In the original POWHEG BOX implementation [85], these real-to-Born mappings are not designed to preserve the virtuality of intermediate resonances. This can lead to convergence issues in the calculation. Further, the same real-to-Born mapping is also used in the R^{α}/B ratios in the POWHEG Sudakov factor. Consider the example discussed in [1], where the real-emission configuration involves final state particles $W^+W^-b\bar{b}g$, with the two b-partons produced close together and the gluon originating from a $b \to bg$ splitting. In this case, the real contributions in the collinear regions corresponding to the $b \to bg$ (R^{α_b}) and $\bar{b} \to \bar{b}g$ ($R^{\alpha_{\bar{b}}}$) splittings dominate the total real-emission cross-section and are of comparable size, provided the emitted gluon is similarly close to both the b and the $\bar{b}$ parton. Now, the situation is considered in which the invariant masses of the W^+bg and the $W^-\bar{b}$ systems are both close to the top-quark mass. The real-to-Born mapping in region α_b will lead to a W^+b invariant mass near the top-quark mass, preserving the resonance virtuality. In contrast, the $\alpha_{\bar{b}}$ mapping will result in a $W^-\bar{b}$ invariant mass far from the top-quark mass, corresponding to a highly off-shell resonance. This leads to a large $R^{\alpha_{\bar{b}}}/B$ ratio in the POWHEG Sudakov due to the suppressed (off-shell) B contribution. Such a mismatch violates the collinear approximation on which the Sudakov factor's construction relies.

- **Preserving the invariant mass of intermediate resonances in the parton shower** Parton shower programs must implement recoil strategies to ensure energy-momentum conservation. These recoil strategies preserve the invariant mass of intermediate resonances — provided that the information about existing resonances is passed to the parton shower. This information is shared via the Les Houches Event (LHE) [86, 87][15] format. However, in the case of general POWHEG-generated events, the resonance structure is not explicitly communicated through the LHE format. Only the properties of the initial-state particles (e.g. gg) and final-state particles (e.g. $WWbbg$), are included. As a result, parton-shower programs are unaware of intermediate resonances such as top quarks or W bosons during their evolution, which can lead to distortions of resonance virtualities.

To address these issues, the POWHEG BOX RES framework [89] introduces a new real-to-Born-mapping, and as such, a partitioning of the real cross-section into singular regions. The new mapping restricts collinear singularities to occur only between two partons originating from the same intermediate resonance or between two partons that are directly associated with the production process. Moreover, it preserves the invariant mass of the intermediate resonance when mapping from the real to the underlying Born configuration. To implement this, a method is required to assign the most probable intermediate resonance structure to a given

[15] The LHE format is a standardized file structure used in high-energy physics Monte-Carlo generators to store event-level information, including the momenta, the PDG IDs [88], colour charges, and other properties of all initial- and final-state particles in an event. It also includes global event-level metadata such as the partonic centre-of-mass energy. Events stored in this format are typically generated from fixed-order calculations, potentially modified by terms of higher order in α_S, and are then passed to a parton-shower program for further evolution.

$WWbb$ final state. This assignment will in the following be referred to as resonance history. Once determined, the resonance history can also be written out in the LHE format, enabling parton-shower programs to apply resonance-aware recoil strategies that avoid distorting the resonance's invariant mass. In the $bb4\ell$ process [2], the resonance history assigned to a given $WWbb$ final state corresponds to one out of three possibilities: $t\bar{t}$, $tW^-\bar{b}$ or $\bar{t}W^+b$.[16] Graphically, this corresponds to



(2.63)

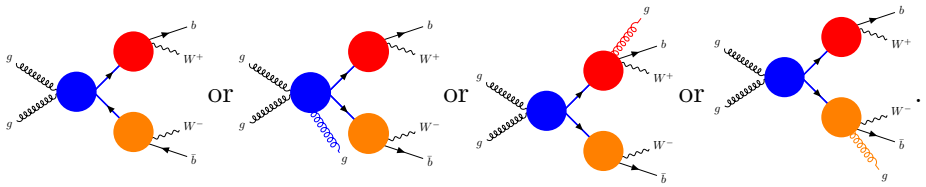
The assignment is performed using resonance projectors based on LO QCD squared matrix element, constructed with the Feynman diagrams involving only double-resonant, $|\mathcal{A}^{\text{LO}}_{t\bar{t}}|^2$, or only single-resonant topologies, $|\mathcal{A}^{\text{LO}}_{tW^-\bar{b}}|^2$ or $|\mathcal{A}^{\text{LO}}_{\bar{t}W^+b}|^2$. Thus, all interference effects between the $t\bar{t}$ and tWb contributions, as well as processes not involving top quarks (e.g. $Z^* \to WW$), are neglected in the construction of the resonance history projectors. It should be emphasized, however, that this affects only the projection of the full $WWbb$ final-state onto a given resonance history, while the underlying $WWbb$ ME calculation itself includes all off-shell and interference effects exactly at NLO QCD. Consequently, the partitioning of the real amplitude R is no longer performed only w.r.t. the singular regions α, but also w.r.t. the resonance histories h. Moreover, the possible singular regions $\Lambda(h)$ now depend on the resonance history. The partitioning of the real-emission amplitude is thus modified as:

$$R = \sum_{\alpha\in\Lambda} R^\alpha \to R = \sum_h \sum_{\alpha\in\Lambda(h)} R^{\alpha,h}. \tag{2.64}$$

The POWHEG formula from Equation 2.60, now modified to account for intermediate resonances, takes the form

$$\mathrm{d}\sigma = \sum_h \bar{B}_h(\Phi_n)\mathrm{d}\Phi_n \left(\Delta_h(\Phi_n, p_{\mathrm{T}}^{\min}) + \sum_{\alpha\in\Lambda(h)} \frac{d\Phi^\alpha_{\text{rad}} R^{h,\alpha}(\Phi_{n+1})\Delta_h(\Phi_n, k_T(\Phi_{n+1}))}{B^h(\Phi_n)} \right). \tag{2.65}$$

Pictorially, focussing on the $t\bar{t}$ resonance history, this corresponds to



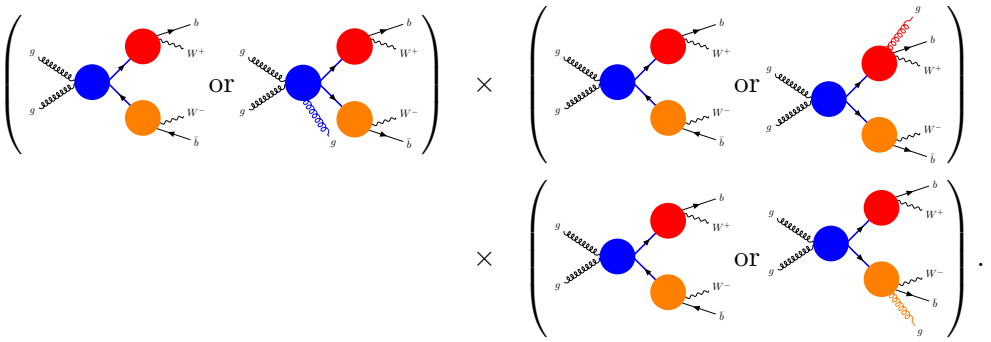
(2.66)

[16] In the first publicly available version of the $bb4\ell$ process [1], only the resonance histories corresponding to $t\bar{t}$ and $(Z^* \to WW)bb$ were implemented. Events with tWb-like final-state kinematics were projected onto a $t\bar{t}$ topology.

In the $bb4\ell$ process, three categories of singular regions can be identified: emissions from the production process, from the top-quark decay, and from the anti-top quark decay. In general, real emissions from the production process in POWHEG are expected to have a larger transverse momentum w.r.t. emissions from the decay subprocesses, where the b-quark mass regulates the soft-collinear divergence. As a result, using the standard POWHEG highest-bid algorithm, which generates emissions in each singular region and selects the one with the highest transverse momentum, would predominately lead to generated emissions associated with the production process. To preserve the NLO accuracy in the description of the top-quark decay, the `allrad` scheme [69] is employed. This allows for the generation of up to three emissions per event, one from each of the three singular region categories. This is implemented through the modified POWHEG BOX RES formula,

$$\mathrm{d}\sigma = \sum_h \bar{B}_h(\Phi_n)\mathrm{d}\Phi_n \prod_{s\in S(h)} \left(\Delta_s(\Phi_n, p_\mathrm{T}^\mathrm{min}) + \sum_{\alpha\in\Lambda(s)} \frac{d\Phi_\mathrm{rad}^\alpha R^{h,\alpha}(\Phi_{n+1})\Delta_s(\Phi_n, k_T(\Phi_{n+1}))}{B^h(\Phi_n)} \right), \tag{2.67}$$

where $S(h)$ denotes the set of singular region categories applicable to the resonance history h. Graphically, again focussing on the $t\bar{t}$ resonance history, this corresponds to



(2.68)

Expanding Equation 2.67 yields the same terms as in Equation 2.65, along with additional contributions at higher orders in the strong coupling constant. The main difference to Equation 2.65 is that emissions from the production processes, the top-quark decay and the anti-top quark decay are now generated independently. Which of these subprocesses can actually generate emissions depends on the resonance history, i.e. $s \in S(h)$. For instance, in a $tW^-\bar{b}$ topology, emissions can only occur from the production process and the top-quark decay, as there is no anti-top quark present in this projection. The emissions from the independent subprocesses are then merged into a single hard-scattering event, allowing for up to three emissions per event. Importantly, each POWHEG Sudakov factor Δ_s only evolves the partons associated with its specific subprocess s. For example, in the case of a $t\bar{t}$ resonance history with emissions generated at scales $k_\mathrm{T}^\mathrm{prod}$, $k_\mathrm{T}^{t\text{-dec}}$ and $k_\mathrm{T}^{\bar{t}\text{-dec}}$ in the production process, top-quark and anti-top-quark decay, respectively, all emissions above $k_\mathrm{T}^\mathrm{prod}$ are vetoed by Equation 2.67 for partons associated with the production process. Correspondingly, for partons associated with the top and anti-top-quark decay all emissions are vetoed above $k_\mathrm{T}^{t\text{-dec}}$ and $k_\mathrm{T}^{\bar{t}\text{-dec}}$. This modified procedure for generating POWHEG real emissions also has implications for the matching of POWHEG events with the parton shower, modifying the standard procedure described in Section 2.4, as detailed in Section 3.3.

CHAPTER 3

Modelling uncertainties in the ATLAS $bb4\ell$ Monte-Carlo sample

The abundance of top-quark pairs produced at the LHC makes precision measurements of top-quark properties possible. Exploiting the large available data set, not only the most populated regions in phase space, but also the tails of kinematic distributions can be studied. With improved experimental precision, the requirements on the accuracy of the theoretical predictions are increased. Important progress in the modelling of top-quark processes has been achieved with the matrix-element (ME) calculation of the full off-shell $t\bar{t} + tW$ production in different lepton flavour dileptonic decays, which can be consistently matched to a parton shower for a realistic modelling of high-multiplicity final states and is implemented in the $bb4\ell$ generator described in Section 2.5.

Another important improvement in top-quark pair modelling was achieved by matching next-to-next-to-leading order (NNLO) QCD predictions of on-shell $t\bar{t}$ production with a parton shower [63, 64]. While this is an essential correction in many considered phase spaces of $t\bar{t}$ analyses, in which often a NNLO per-event-reweighting based on e.g. the top-quark p_{T} is applied to improve the agreement of NLO+PS predictions with data [57, 90, 91], analyses which are sensitive to the correct off-shell treatment of the top-quark and the exact inclusion of the $t\bar{t}/tW$ interference at NLO will profit from the improved signal modelling provided by the $bb4\ell$ simulation. The employment of both samples, the NNLO QCD on-shell $t\bar{t}$ production and the NLO QCD off-shell $t\bar{t} + tW$ production interfaced with a parton shower is thus of utmost importance. Which of the two Monte-Carlo (MC) samples will be used in a particular measurement, will thereby depend on the targeted phase-space region.

However, in order to use a Monte-Carlo sample in an experimental analysis as a nominal prediction, an estimation of all relevant modelling uncertainties needs to be possible. These uncertainties are estimated by variations of the Monte-Carlo generator settings, which can not be determined by first principles. Since often the modelling uncertainties are among the most dominant systematic uncertainties in top-quark measurements, they need to be well

understood and studied. In this chapter, modelling systematics, which should be considered by ATLAS data analyses, when using the $bb4\ell$ sample as their nominal signal MC sample, are studied. The strategy for defining these systematic uncertainties builds on the prescription used for the $t\bar{t}$ POWHEG+PYTHIA 8 sample, which was scrutinized in several ATLAS top-quark measurements. The impacts of these uncertainties on the $bb4\ell$ and in the $t\bar{t}$ predictions are carefully compared and in case of differences further analysed.

This chapter is structured as follows: First, the nominal $bb4\ell$ MC sample modelling the off-shell $WWbb$ production process is compared to on-shell $t\bar{t}+tW$ generators in Section 3.1. After discussing systematic uncertainties related to the hard-process generation in Section 3.2, systematic uncertainties concerning the matching of $bb4\ell$ hard-process events and the parton shower will be discussed in Section 3.3. Systematic uncertainties related to the parton-shower modelling are studied in Section 3.4. Finally, in Section 3.5 the conclusions are presented.

3.1 The nominal ATLAS $bb4\ell$ Monte-Carlo sample

As described in Section 2.5, the modelling of the $b\bar{b}\ell^+\ell^-\bar{\nu}_\ell\nu_\ell$ final state is improved when using $bb4\ell$ instead of $t\bar{t}+tW$ predictions by including off-shell effects, the top-quark decay and the $t\bar{t}/tW$ interference description at NLO QCD. While it was already noted in [2], that the NLO top-quark decay description is well approximated by the PYTHIA 8 matrix-element corrections,[1] in this section, the nominal ATLAS $bb4\ell$ and $t\bar{t}+tW$ predictions are compared with a focus on their respective top-quark lineshape and $t\bar{t}/tW$ interference modelling.

The generation of the hard-scatter events of the nominal ATLAS $bb4\ell$, $t\bar{t}$ and tW MC samples, which are later interfaced to a parton shower, is performed with the POWHEG method at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$. While the $bb4\ell$ process [1, 2] is implemented in the POWHEG BOX RES framework[2] [89], the $t\bar{t}$ (`hvq` [59]) and the tW (`ST_wtch_DR` [75]) processes are available in the POWHEG BOX v2 framework [85].[3] In the POWHEG generation of these MC samples, the NNPDF3.0NLO [93] ($\alpha_S(M_Z) = 0.118$) PDF set was used, the top-quark mass was set to $m_{\text{top}} = 172.5\,\text{GeV}$ and the POWHEG h_{damp} parameter to $1.5m_{\text{top}}$. The $t\bar{t}$ and tW predictions are obtained with a five-flavour number scheme (5FNS) calculation, and the $bb4\ell$ calculation is performed within the 4FNS. Thereby, matching factors are applied to match the 4FNS $bb4\ell$ matrix element calculation to the 5FNS PDF set [2].

[1] The effect of matrix-element corrections in PYTHIA 8 was also studied using the nominal ATLAS $bb4\ell$ and $t\bar{t}$ MC samples. When these corrections are disabled in the PYTHIA 8 parton shower, the resulting $t\bar{t}$ distributions exhibit worse agreement with the $bb4\ell$ MC sample compared to the default $t\bar{t}$ simulation, which includes matrix element corrections. Further, their impact is smaller in the $bb4\ell$ simulation, since the hardest emission in the top-quark decay is already generated by POWHEG and is therefore unaffected by the PYTHIA 8 matrix-element corrections. In this case, the corrections only influence subsequent emissions from the b-quark from the top-quark decay.

[2] The $bb4\ell$ ATLAS MC sample studied in this thesis was produced using the $bb4\ell$ process version described in [2] and available in [92].

[3] The differences between the POWHEG BOX v2 and the POWHEG BOX RES frameworks are described in Section 2.5.

In the following, alternative $t\bar{t}+tW$ predictions are additionally studied, which are used in ATLAS measurements to estimate the uncertainty due to the approximations made in the MC generation of the $t\bar{t}+tW$ MC sample regarding the top-quark lineshape and the $t\bar{t}/tW$ interference treatment. These modelling uncertainties do not need to be considered when using the $bb4\ell$ MC sample. These predictions are build with the following MC samples:

- **Top-quark lineshape uncertainty:** An additional $t\bar{t}$ `hvq` MC sample is studied, in which the top-quark decay and the off-shell smearing is not done internally within the POWHEG BOX v2 framework, but instead the undecayed, on-shell top-quark pair events produced with the `hvq` process, are decayed with MADSPIN [67, 68] and then interfaced to PYTHIA 8. In the generation of this sample, a special MADSPIN configuration was used to correct for the different treatment of the b quark in MADSPIN and in PYTHIA (the former considers it as massless, the latter as massive), which was described in [90]. This configuration provided by the authors of MADSPIN, allows MADSPIN to use massless b quarks for the production, as a 5FNS PDF is used for the calculation of the LO spin-correlated $t\bar{t}$ production amplitudes, but massive b quarks in the decay of the top-quark, thereby aligning the b quark treatment in the top-quark decay of MADSPIN and PYTHIA.
- **$t\bar{t}/tW$ interference uncertainty:** In the nominal ATLAS tW sample, the interference with $t\bar{t}$ production is handled according to the diagram removal (DR) scheme [76, 94]. An alternative tW Monte Carlo sample was produced, in which the diagram subtraction (DS) scheme is used instead, applying the same simulation parameter settings as in the nominal tW DR Monte Carlo samples.

For a direct comparison of the $bb4\ell$ and $t\bar{t}+tW$ predictions, an additional $t\bar{t}$ MC sample is studied in this section, in which the PYTHIA 8 recoil scheme is set to recoil-to-top, as in the nominal ATLAS $bb4\ell$ sample. In contrast, the nominal $t\bar{t}$ MC sample uses the recoil-to-colour scheme. Further details are given in Section 3.4.

All POWHEG-generated Les Houches events (LHEs) are interfaced with PYTHIA 8 for modelling of the parton shower and hadronisation, using the A14 tune [95]. The PYTHIA 8 versions of the considered samples differ between the $bb4\ell$ (8.312),[4] the nominal `hvq` $t\bar{t}$ ATLAS sample (8.230), the `hvq` +MADSPIN +PYTHIA 8 $t\bar{t}$ sample (8.309) and the tW DR (8.309) and DS (8.307) samples.[5] For all processes the EVTGEN program [96] was used to simulate the decay of the bottom and charm hadrons.

[4] In the context of this study, it was found that the recoil-to-top setting was not correctly applied when showering $bb4\ell$ LHEs, due to the special $t \to Wbg$ vertices present in these LHEs. A patch provided by the PYTHIA 8 authors was used for the PYTHIA 8.312 version in ATLAS to enable the recoil-to-top option also for $bb4\ell$ events. This patch is included in the official PYTHIA 8 releases starting from PYTHIA version 8.313.

[5] It was verified that for the considered distributions in this study, the change in PYTHIA 8 versions has a small impact by comparing $t\bar{t}$ `hvq` +PYTHIA 8 distributions generated with the PYTHIA versions 8.312, 8.230 and 8.307.

All distributions are normalised to their NLO+PS cross-section, which is calculated during event generation.[6] The presented study is performed at NLO+PS level and includes both, pure simulation based comparisons between different MC distributions, allowing for flexible observable definitions and binning choices only limited by MC statistics, as well as comparisons to unfolded ATLAS and CMS data. Given the numerous results of unfolded pp-collision data for the $t\bar{t}$ production process by the ATLAS and CMS collaborations, comparisons to these results are presented in the following. The validity of comparing $bb4\ell$ and $t\bar{t}+tW$ MC predictions to $t\bar{t}$ unfolded data is verified by explicitly comparing the $t\bar{t}$ and $t\bar{t}+tW$ predictions. Since the single-resonant contribution from the tW process is found to have a negligible impact on the considered distributions, the comparison of $bb4\ell$ predictions to unfolded $t\bar{t}$ data is concluded to be valid. This verification will not be stated again explicitly in the following discussions. However, comparisons between the $t\bar{t}$ and $t\bar{t}+tW$ distributions are always shown in the corresponding figures. The creation of histograms was performed using the Rivet software [98].

Pure Monte Carlo comparisons: Description of event selection and observable definition
For the purpose of this study, a Rivet 3 routine targeting typical top-quark observables was prepared. For all distributions shown in this chapter, this Rivet routine was used, if not specified otherwise. The lepton objects at particle level are defined as the dressed leptons, which are built from a prompt bare lepton (lepton in the event record) and surrounding prompt photons in a cone of $\Delta R = 0.1$. The resulting dressed leptons need to satisfy $p_{\mathrm{T}} > 20\,\mathrm{GeV}$ and $|\eta| < 2.5$. Electron and muons from the decay of a prompt τ-lepton are not considered as prompt leptons in this Rivet routine implementation. Jets are defined by means of the anti-k_t clustering algorithm [99] with the radius parameter $R = 0.4$ using the FastJet implementation [100]. The particle-level jets need to satisfy selection criteria on their transverse momenta and pseudorapidities: $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.5$. In the jet-clustering algorithm, all stable particles with a mean lifetime greater than $0.3 \times 10^{-10}\,\mathrm{s}$ are considered, except for dressed prompt leptons and prompt neutrinos. Non-prompt muons and non-prompt neutrinos are also included in the jet clustering, which can originate for example from a semi-leptonic B-hadron decay. A jet is identified as a b-jet, if a B-hadron with $p_{\mathrm{T}} > 5\,\mathrm{GeV}$ is associated to the jet through the ghost-matching technique[7] [101]. A light jet is defined as not being ghost-associated neither to a B-hadron nor a D-hadron. The event selection requires exactly two dressed leptons with opposite electric charge, at least two neutrinos and at least two

[6] A correction factor was applied to the $t\bar{t}$ and tW NLO QCD ME+PS total cross-sections in the comparison presented in this section due to inconsistent settings of the leptonic top-quark decay branching ratio `tdec/elbranching 0.1802` ($\mathrm{BR}(W \to \ell\nu)$) in the `hvq` and `ST_wtch_DR` processes and the W-boson decay width (Γ_W^{NLO}) setting `GammaW 2.085` in the $bb4\ell$ simulation. The top-quark decay branching ratio and the W-boson decay width are related through the formula

$$\mathrm{BR}(W \to \ell\nu) = \Gamma(W \to \ell\nu)/\Gamma_W^{\mathrm{NLO}} = \left[(G_F m_W^3)/(6\pi\sqrt{2})\right]/\Gamma_W^{\mathrm{NLO}} = 0.2272/\Gamma_W^{\mathrm{NLO}},$$

with the mass of the W-boson m_W, the Fermi constant G_F and the LO partial decay width $\Gamma(W \to \ell\nu)$. This would imply a decay branching ratio `tdec/elbranching 0.108968`. Therefore, the tW NLO+PS predictions are scaled by a factor $(0.108968/0.1082) = 1.0071$ and the $t\bar{t}$ NLO+PS predictions by a factor of $(0.108968/0.1082)^2 = 1.01425$ [97].

[7] In the Rivet routine implementation, the `Rivet::Jet` class method `bTagged()` is utilized for this purpose.

b-tagged jets. For normalised distributions, the normalisation is performed with respect to the fiducial cross-section, defined by these selection criteria. The particle-level top-quark reconstruction is done by first reconstructing the decay products of the particle-level W-bosons by using the two leading neutrinos from the event record. The pairing of the neutrino and the charged lepton that originate from the same W-boson decay is found by minimizing $|m_{\ell\nu} - 80.4\,\text{GeV}| + |m_{\ell'\nu'} - 80.4\,\text{GeV}|$. Then, the particle level top-quark is constructed, by identifying those Wb-pairs which minimize the difference between m_{Wb} and $m_{\text{top}} = 172.5\,\text{GeV}$, using the two leading b-jets in the event. Since also the single resonant and non-resonant[8] contributions are included, the reconstructed top-quark refers to a particular combination of W-boson and b-jet, but does not imply that these particles originate from a top-quark decay. For brevity, in this chapter, still the expressions "reconstructed top quark" and "reconstructed top-quark pair" are used.

Comparison of predictions for on-shell and off-shell $WWbb$ production In the following, the nominal ATLAS $bb4\ell$ and $t\bar{t}+tW$ MC samples are compared. The figures used for this comparison follow the same structure as in Figure 3.1: The first ratio panel shows the ratio of the nominal $t\bar{t}+tW$ DR prediction to the nominal $bb4\ell$ ATLAS prediction. Since $bb4\ell$ uses the recoil-to-top Pythia 8 scheme, an alternative $t\bar{t}+tW$ prediction with the same recoil setting applied in the $t\bar{t}$ MC generation ($t\bar{t}$ (rT)+tW DR) is also displayed. Modelling uncertainties, relevant only for $t\bar{t}+tW$, are shown as a grey band. They are derived from the deviations of the $t\bar{t}$ MadSpin + tW DR and $t\bar{t}+tW$ DS variations with respect to the nominal $t\bar{t}+tW$ DR prediction, as listed in Table 3.1, added in quadrature and symmetrised (as detailed in the second ratio panel).

$t\bar{t}+tW$ uncertainty estimation of	Comparison of nominal $t\bar{t}+tW$ DR prediction to
$t\bar{t}/tW$ interference	$t\bar{t}+tW$ DS
top-quark lineshape	$t\bar{t}$ MadSpin + tW DR

Table 3.1: Summary of modelling uncertainties in the $t\bar{t}+tW$ prediction, which are removed when using instead the $bb4\ell$ prediction.

In the first panel, this uncertainty band is applied to the $t\bar{t}$ (rT)+tW DR prediction to assess whether it covers the observed differences to $bb4\ell$, while avoiding effects solely due to different Pythia 8 recoil settings. Additionally, $bb4\ell$ predictions including inverse-width corrections are shown. These corrections are expected to improve the agreement in the fiducial cross-sections [2] of $bb4\ell$ and $t\bar{t}+tW$, and are discussed in Section 3.2. The third ratio panel compares the tW DR contribution to the full $t\bar{t}+tW$ DR prediction and, analogously, the fraction of $bb4\ell$ events assigned to a tW resonance history to the full $bb4\ell$ prediction. For absolute distributions this corresponds to the respective fractions, while for normalised distributions it probes shape differences. This highlights phase-space regions particularly sensitive to single-resonant contributions and $t\bar{t}/tW$ interference modelling. Finally, the fourth ratio panel

[8] The latter only in the case of the $bb4\ell$ simulation.

compares LHE-level — i.e. final-state kinematics generated by POWHEG before applying the parton shower — and showered NLO+PS predictions for both, $bb4\ell$ and $t\bar{t}+tW$, identifying phase-space regions most sensitive to parton-shower effects and highlights differences in the impact of the parton shower on $bb4\ell$ and $t\bar{t}+tW$ predictions.

In Figure 3.1, this comparison is shown for the

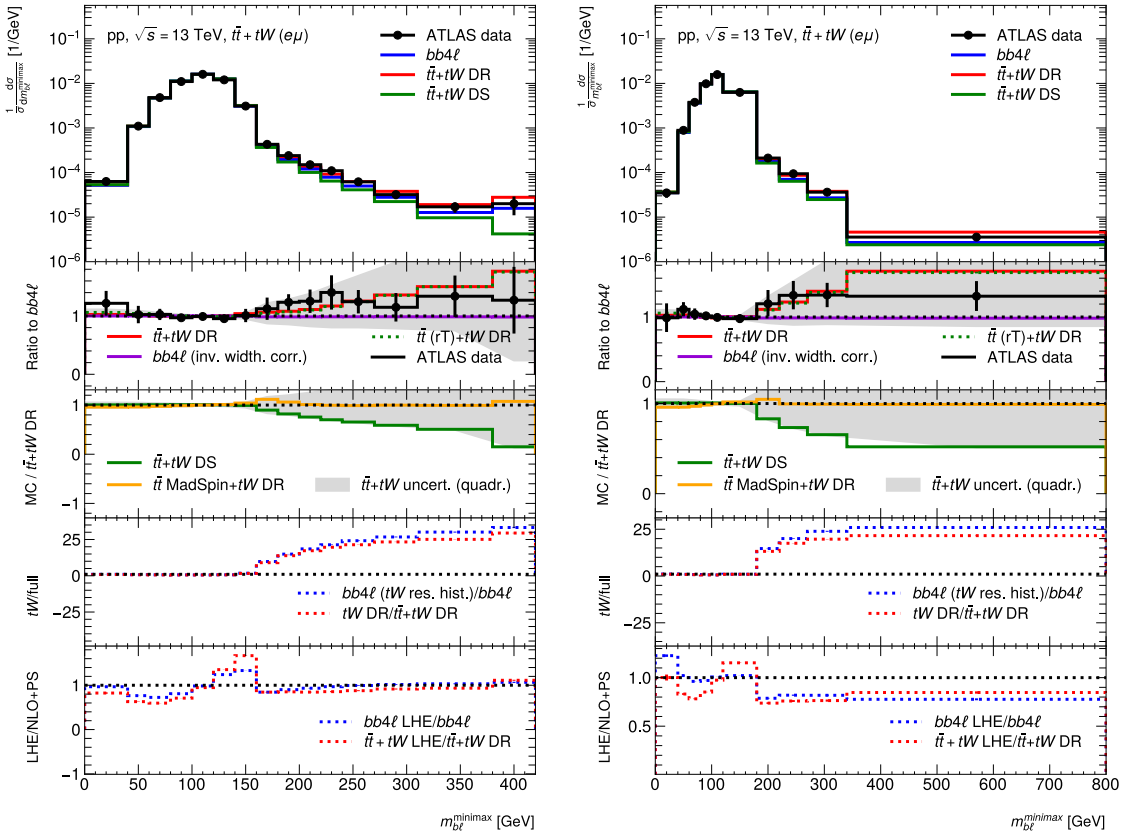
$$m_{\ell b}^{\text{minimax}} = \min\left(\max(m_{b_1\ell_1}, m_{b_2\ell_2}), \max(m_{b_1\ell_2}, m_{b_2\ell_1})\right) \tag{3.1}$$

distribution, including a comparison to unfolded $t\bar{t}+tW$ ATLAS data in the dilepton channel using (a) $36.1\,\text{fb}^{-1}$ [78] and (b) $140\,\text{fb}^{-1}$ [102] of pp collision data at $\sqrt{s}=13\,\text{TeV}$.[9] While (a) considers $t\bar{t}+tW$ decays into all dilepton final states ($ee, \mu\mu, e\mu$), the analysis presented in (b) is restricted to the $e\mu$-decay channel. For the $m_{\ell b}^{\text{minimax}}$ distribution, both measurements require at particle-level exactly two b-jets in the event with $p_{\text{T}} > 25\,\text{GeV}$ and $|\eta| < 2.5$. In addition, the measurement in (a) vetoes additional soft b-jets with transverse momenta of $p_{\text{T}} > 5\,\text{GeV}$. The unfolded data from these analyses provide an excellent opportunity to study $bb4\ell$ MC predictions, as they consider the full $t\bar{t}+tW$ final state. In contrast, most unfolded measurements concentrate on the double-resonant $t\bar{t}$ production process. The $m_{\ell b}^{\text{minimax}}$ distribution is studied, as the high-$m_{\ell b}^{\text{minimax}}$ tail is sensitive to the $t\bar{t}+tW$ interference modelling. This is clearly visible in Figure 3.1, where the tW contribution grows significantly and the DR/DS uncertainty in the $t\bar{t}+tW$ prediction becomes large. This uncertainty is found to cover the difference between the nominal $t\bar{t}+tW$ DR and the $bb4\ell$ prediction, while a general trend towards better agreement between the $bb4\ell$ and $t\bar{t}+tW$ DS predictions is observed. In both simulations, the inclusion the parton shower shifts the $m_{\ell b}^{\text{minimax}}$ distribution towards smaller values in the bulk of the distribution compared to the corresponding LHE-level prediction, as shown in the lowest ratio panel. Interestingly, while parton-shower effects have only a limited impact on the high-$m_{\ell b}^{\text{minimax}}$ tail in (a), a more pronounced effect is observed in (b). One possible explanation for this difference is the additional veto on soft b-jets applied in the particle-level event selection of (a), since the tail of the $m_{\ell b}^{\text{minimax}}$ distribution has been found to be populated by events with additional heavy-flavour jets.

In Figure 3.2(a), the invariant mass of the reconstructed top-quark pair is shown. Sizeable differences between the $bb4\ell$ and $t\bar{t}+tW$ predictions are observed in the threshold region, $m_{t\bar{t}} \lesssim 2m_{\text{top}}$, highlighting the challenges associated with modelling this phase space region accurately. The lineshape uncertainty in the $t\bar{t}$ sample is sizeable in this low $m_{t\bar{t}}$-region and partly covers the discrepancies observed between the on-shell and off-shell generators. However, the comparison of the nominal $t\bar{t}$ prediction and the corresponding $t\bar{t}$ sample in which the top-quark decays are simulated with MADSPIN must be interpreted with care. The description of the threshold region modelled with the $t\bar{t}$ POWHEG + MADSPIN MC sample is currently under scrutiny in ATLAS.[10] For both the $bb4\ell$ and $t\bar{t}+tW$ predictions, the relative

[9] Both measurements also include comparisons of the unfolded data to a $bb4\ell$ MC prediction. In those studies, an older version of the $bb4\ell$ generator software was used, in which e.g. only $t\bar{t}$ and Z-resonance history projectors were implemented.

[10] As described in Section 3 of [68], for events with unstable particles produced in the threshold region, the reshuffling of final-state particle momenta, required to ensure energy-momentum conservation when introducing off-shell effects, can be ill-defined. In such cases, the current decay kinematics are rejected, and



(a) (b)

Figure 3.1: Comparison of the nominal $bb4\ell$ and $t\bar{t}+tW$ predictions for the $m_{\ell b}^{\text{minimax}}$ distribution with unfolded $t\bar{t}+tW$ data from the ATLAS measurements (a) [78] and (b) [102]. In addition to the nominal predictions, $bb4\ell$ predictions with inverse width corrections applied ($bb4\ell$ inv. width corr.) are shown, together with $t\bar{t}+tW$ predictions employing different modelling choices: alternative top-quark line-shape modelling ($t\bar{t}$ MadSpin + tW DR), different treatments of $t\bar{t}/tW$ interference ($t\bar{t}+tW$ DS), and a modified $t\bar{t}$ Pythia 8 recoil setting using recoil-to-top ($t\bar{t}$ (rT) + tW DR). The grey uncertainty band is constructed by adding the $t\bar{t}+tW$ line-shape and interference-modelling uncertainties in quadrature and applying the resulting variation symmetrically in the upward and downward directions. In the first ratio panel, this uncertainty is applied to the $t\bar{t}$ (rT) + tW DR prediction, since the nominal $bb4\ell$ MC sample is also produced using the Pythia 8 recoil-to-top setting. This uncertainty band corresponds to modelling uncertainties that do not apply to the $bb4\ell$ MC prediction and therefore can lead to a significant reduction of the total modelling uncertainty. The third ratio panel compares the tW DR prediction to the $t\bar{t}+tW$ DR prediction, as well as the $bb4\ell$ prediction obtained only from events, which are assigned a tW resonance history, to the full $bb4\ell$ prediction. The fourth ratio panel illustrates the impact of the parton shower by comparing distributions obtained at LHE level for $bb4\ell$ and $t\bar{t}+tW$ to the corresponding NLO+PS predictions.

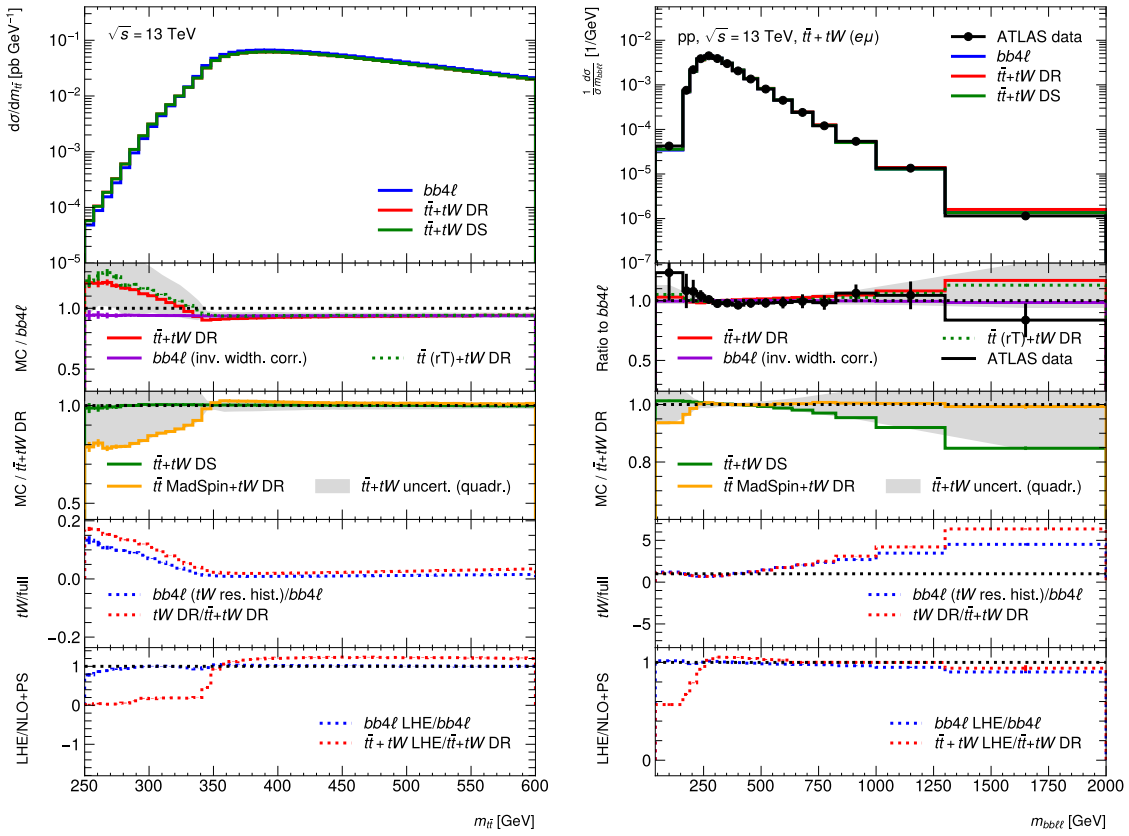
tW contribution is enhanced in this threshold region, since $m_{WWbb} < 2m_{\text{top}}$ is kinematically allowed for the single-resonant diagrams already at LO, even when only on-shell top-quarks are included in the matrix-element calculation. The shape of the tW contributions relative to the full $t\bar{t} + tW$ or $bb4\ell$ prediction is found to be very similar in the two simulations. The impact of the parton showering in the threshold region is significantly smaller in the $bb4\ell$ prediction compared to the $t\bar{t} + tW$ MC simulation. This difference can be attributed to the inclusion of NLO top-quark decays in the $bb4\ell$ simulation, where the hardest emission in the top-quark decay is already generated at the LHE level using the POWHEG method.[11]

In Figure 3.2(b), a closely related distribution is studied: the invariant mass of the two charged leptons and the leading and subleading b-jets. This distribution was unfolded in the ATLAS $t\bar{t} + tW$ analysis presented in [102], which was previously discussed in the context of Figure 3.1(b). In contrast to the particle-level event selection used for Figure 3.1(b), where exactly two b-jets are required, the selection for Figure 3.2(b) requires at least two b-jets. Qualitatively similar features to those observed in Figure 3.2(a) are present, but in this case the region $m_{bb\ell\ell} > 500\,\text{GeV}$ is also studied, providing additional sensitivity to the treatment of $t\bar{t}/tW$ interference effects. In this region, the unfolded data agree better with the $bb4\ell$ and $t\bar{t} + tW$ DS predictions than with the corresponding $t\bar{t} + tW$ DR prediction. A mild tension between the MC predictions and the unfolded data is observed in the threshold region. However, this is covered by the total uncertainties in the unfolded data points. This phase-space region has been the subject of several recent measurements investigating $t\bar{t}$ bound-state effects or incorporating such effects for a more accurate description of the $m_{t\bar{t}}$ spectrum by ATLAS [104] and CMS [105, 106].

The transverse momentum distribution of the leading lepton is shown in Figure 3.3(a). A sizeable normalisation difference is observed between the nominal $bb4\ell$ and $t\bar{t} + tW$ DR samples. This difference is largely resolved once inverse width corrections are applied to the $bb4\ell$ prediction, which results to an approximately flat rescaling of the distribution of about 6 %. A similar impact of the inverse width correction on the total cross-section has been reported in [2]. While small shape differences remain between the nominal $t\bar{t} + tW$ DR prediction and the inverse width corrected $bb4\ell$ prediction, the overall agreement is reasonable within the $t\bar{t} + tW$ modelling uncertainties. In particular, the high-p_{T} region is sensitive to the modelling of $t\bar{t}/tW$ interference, as reflected by the increasing importance of DR/DS uncertainty as well as the increasing tW contribution in this phase-space region. When comparing the LHE-level to the showered distributions in the last ratio panel, a normalisation difference is observed in the case of the $t\bar{t} + tW$ predictions, but not for the $bb4\ell$ predictions. A possible explanation is, that in the $bb4\ell$ simulation, the hardest emission from the top-quark decay is already present at LHE level, which can reduce the transverse momenta of the resulting b-jets. As a consequence, some events may fail the $p_{\text{T}} > 25\,\text{GeV}$ b-jet selection

new decay kinematics are generated for the on-shell $t\bar{t}$ hard process event. For this reason, in a recent ATLAS $t\bar{t}$ data analysis [103], which used a reconstructed $m_{t\bar{t}}$ distribution as sensitive observable, this uncertainty was instead evaluated by reweighting the parton-level invariant mass distribution of the top quark m_t and anti-top quark $m_{\bar{t}}$, separately, in the nominal $t\bar{t}$ MC sample to the corresponding distribution obtained with the $t\bar{t}$ MADSPIN MC sample, instead of directly comparing the $m_{t\bar{t}}$ spectra of the two generators.

[11] Using the modified `allrad` POWHEG method within the POWHEG BOX RES framework.



(a) (b)

Figure 3.2: Same comparison as in Figure 3.1, but for (a) the invariant mass of the reconstructed top-quark pair and for (b) the invariant mass distribution of the $bb\ell\ell$-system including the comparison of the latter to unfolded ATLAS $t\bar{t}+tW$ data [102].

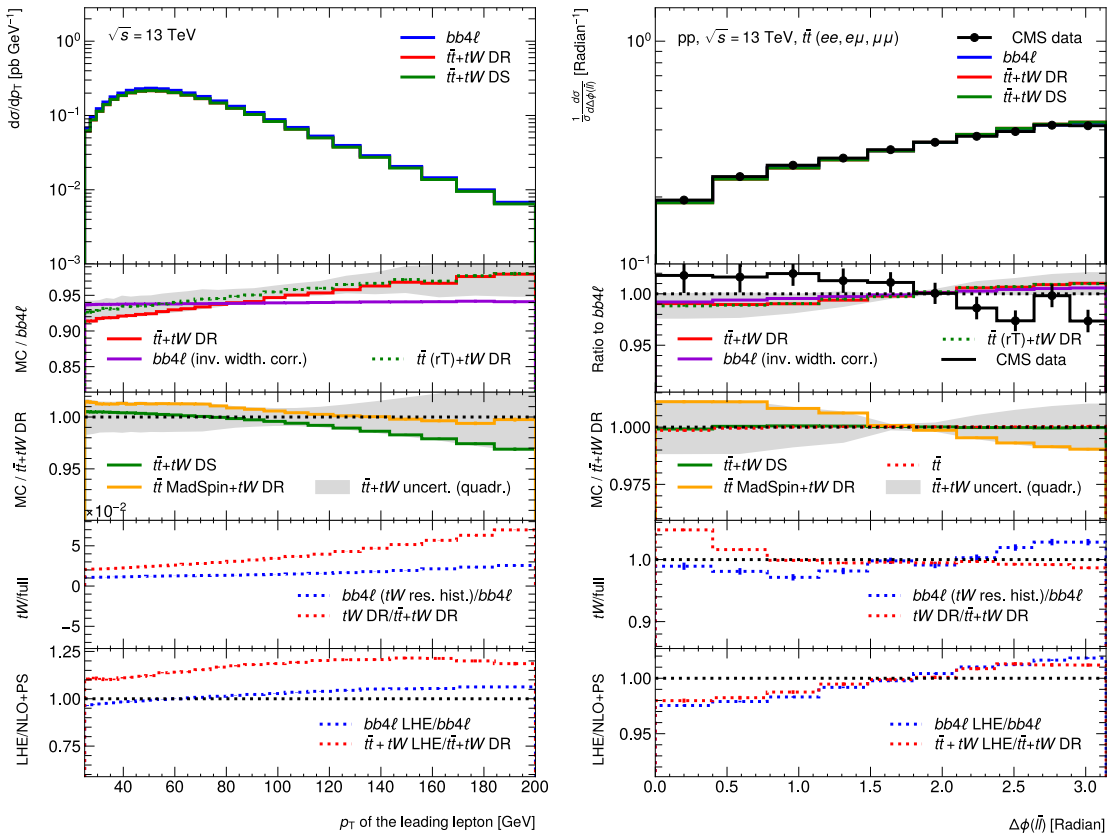
requirement imposed already at the LHE level, while this effect is seen in the $t\bar{t}+tW$ only in the showered predictions.

In Figure 3.3(b), the unfolded distribution of the difference in azimuthal angle, $\Delta\phi$, between the two charged leptons from the $t\bar{t}$ decay is shown, as measured by the CMS Collaboration using 39.5 fb^{-1} of pp collision data at $\sqrt{s} = 13\,\text{TeV}$ [107]. The truth-level event selection requires exactly two leptons with $p_{\text{T}} > 20\,\text{GeV}$ and $|\eta| < 2.4$, as well as at least two b-jets with $p_{\text{T}} > 30\,\text{GeV}$ and $|\eta| < 2.4$.[12] This observable is sensitive to the modelling of spin correlations, which are described at NLO QCD in the $bb4\ell$ MC sample. In contrast, the production spin correlations[13] in the $t\bar{t}$ and $t\bar{t}$ MadSpin ATLAS MC samples are modelled using LO spin-correlated matrix elements following the FMLW [67] method, but with different

[12] Additional requirements include overlap removal, at least two truth neutrinos and an invariant mass of the charged lepton pair $m_{\ell\ell} > 20\,\text{GeV}$. For more details, refer to [107, 108].

[13] While spin correlation observables in the top-quark decay subprocess can be described also when considering the production and the decay process in the event generation separately, this is not the case for production angular correlations.

implementations. These implementation differences result in a noticeable variation in the $t\bar{t}+tW$ prediction. When compared to the unfolded CMS data, the $bb4\ell$ prediction is in slightly better agreement than the nominal $t\bar{t}+tW$ DR distribution. As shown in the lower ratio panel, parton-shower effects are important in this distribution and are found to be similar for the $bb4\ell$ and $t\bar{t}+tW$ simulations. This is in contrast to the observations made with distributions that focus on b-jet kinematics, in which the presence of the first emission from the top-quark decay at LHE level in the $bb4\ell$ simulation leads to more pronounced differences in the impact of the parton shower between the $bb4\ell$ and $t\bar{t}+tW$ predictions.

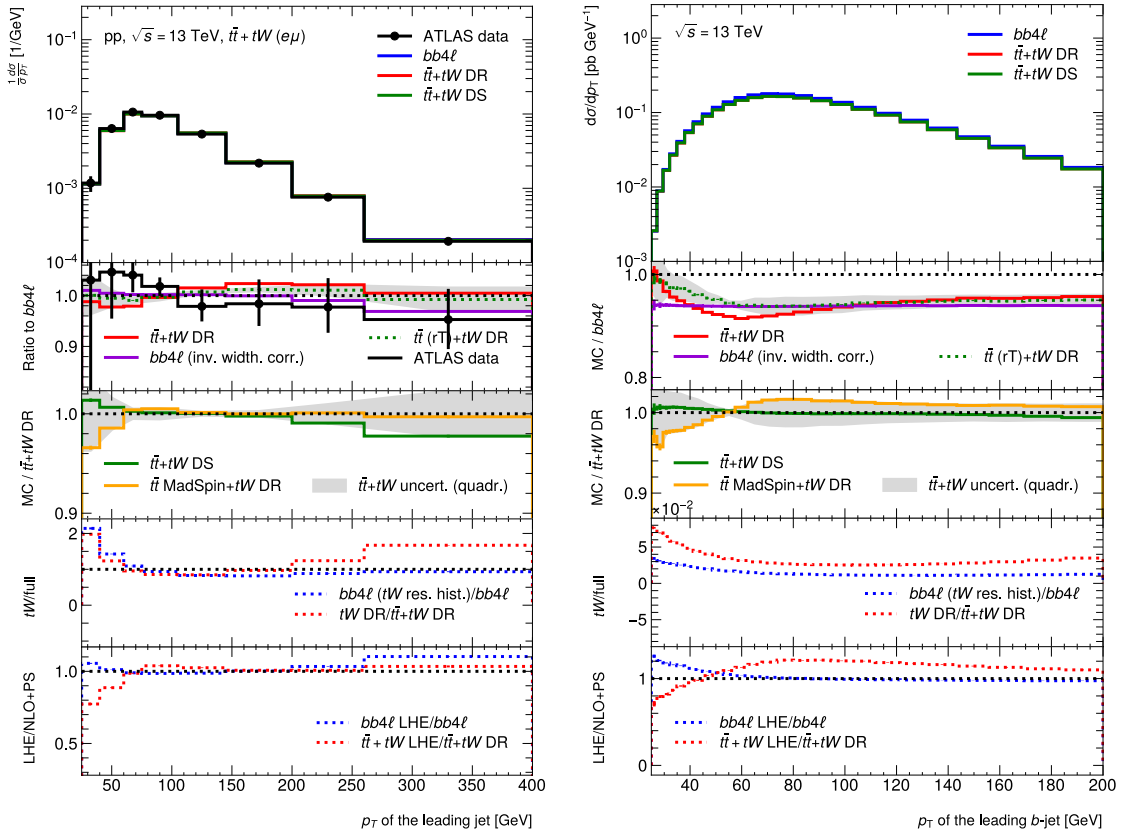


(a) (b)

Figure 3.3: Same comparison as in Figure 3.1, but for (a) the transverse momentum of the leading lepton and for (b) the $\Delta\phi$ distribution of the two charged leptons — including the comparison of the latter to unfolded CMS $t\bar{t}$ data [107]. To validate the comparison of the $bb4\ell$ and $t\bar{t}+tW$ predictions to unfolded $t\bar{t}$ data in (b), additionally the $t\bar{t}$ prediction is compared to the $t\bar{t}+tW$ distribution in the second ratio panel, showing a negligible influence of the tW contribution in this distribution.

Finally, Figure 3.4 shows the transverse momentum distribution of the leading jet in (a), including a comparison to unfolded $t\bar{t}+tW$ ATLAS data [102] from the same analysis as shown in Figure 3.1(b), as well as the transverse momentum distribution of the leading b-jet in (b). These two observables show similar features, since the leading jet distribution includes contributions from both, light jets and b-jets. While normalisation differences between the

$bb4\ell$ and $t\bar{t}+tW$ predictions observed in (b) are again resolved by applying inverse width corrections to the $bb4\ell$ predictions, the remaining shape differences between the two nominal predictions are largely covered by the $t\bar{t}+tW$ specific uncertainties, represented by the grey uncertainty band. Both nominal predictions are found to agree well with the unfolded data in (a) within the measurement uncertainties. From the distributions in Figure 3.4, it is evident that the tW contribution is enhanced at low transverse momenta. As anticipated from the previous discussions, the impact of the parton shower differs between the $bb4\ell$ and $t\bar{t}+tW$ predictions. In the $t\bar{t}+tW$ MC simulations, parton-shower emissions from the b-jet originating from the top-quark decay reduce the transverse momentum of these jets in the NLO+PS prediction relative to the corresponding LHE-level distribution. In contrast, in the $bb4\ell$ simulation the first emission from the top-quark decay is already included at the LHE distribution.



(a) (b)

Figure 3.4: Same comparison as in Figure 3.1, but for (a) the transverse momentum of the hardest jet, including also the comparison to unfolded $t\bar{t}+tW$ ATLAS data [102], and (b) the transverse momentum of the hardest b-jet.

In the following, two MC generation parameters are briefly discussed: the minimal transverse momentum generated in POWHEG ($p_{\mathrm{T,min}}$) and the PYTHIA 8 $p_{\mathrm{T}}^{\mathrm{def}}$ matching parameter. For the ATLAS $bb4\ell$ MC sample, these parameters are set to values that differ from their respective defaults in POWHEG and PYTHIA 8. The motivation for these choices is outlined in the following.

Minimal generated Powheg p_{T} In the generation of the ATLAS $bb4\ell$ MC samples, the minimum allowed transverse momentum of a POWHEG emission is set to 1.2 GeV, rather than to the default POWHEG value of $\sqrt{0.8}$ GeV, which is also used in the ATLAS $t\bar{t}$ POWHEG+PYTHIA 8 sample. Further, the $p_{\mathrm{T,min}}$ parameter determines the starting scale of the parton shower in events in which no POWHEG emission is generated.

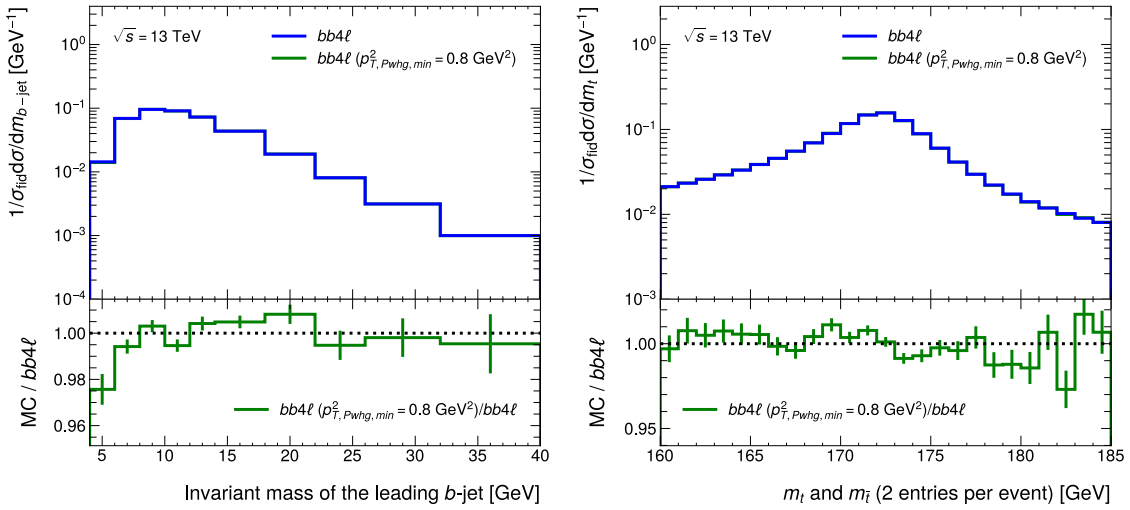
The authors of [109], where the $bb4\ell$ interface to the HERWIG parton shower was studied, recommended to set $p_{\mathrm{T,min}} = 1.2$ GeV, as this value is larger than the shower cut-off scale of 0.958 GeV in the HERWIG 7.2 parton shower [110].[14] Otherwise, POWHEG could generate emissions with transverse momenta below the cut-off scale of the parton shower. Since the $bb4\ell$ sample interfaced with HERWIG is used as part of the modelling uncertainty estimation, the parameter $p_{\mathrm{T,min}}$ is set to 1.2 GeV also in the nominal POWHEG+PYTHIA 8 MC sample to stay consistent in the POWHEG settings of these two samples.

Overall, the impact of varying the $p_{\mathrm{T,min}}$ POWHEG parameter on the differential distributions considered in this analysis is found to be negligible. One of the largest variations was observed in the invariant mass distribution of the leading b-jet, as shown in Figure 3.5(a). Given that soft radiation off the b-quarks is expected, it is plausible that changes in the treatment of the lowest-p_{T} emissions affect b-jet-related observables. However, as demonstrated in Figure 3.5(b), this small variation in the b-jet invariant mass has no impact within statistical uncertainties on the reconstructed top-quark invariant mass. For this reason, variations of the $p_{\mathrm{T,min}}$ parameter are not included as an additional systematic uncertainty.

Pythia 8 $p_{\mathrm{T}}^{\mathrm{def}}$ matching parameter The matching of the POWHEG hard-process LHE to the PYTHIA 8 parton shower is influenced by the $p_{\mathrm{T}}^{\mathrm{def}}$ matching parameter. This parameter controls how the scale of the parton-shower emission is defined, which is subsequently compared to the POWHEG matching scale in the production process.[15] The available settings of the $p_{\mathrm{T}}^{\mathrm{def}}$ parameter and their influence on the transverse momentum definition of initial-state (ISR) and final-state (FSR) parton shower radiation are:

[14] In HERWIG 7.3, this shower cut-off scale is reduced to 0.655 GeV [110], which would also be compatible with the default POWHEG $p_{\mathrm{T,min}}$ value. The choice $p_{\mathrm{T,min}} = 1.2$ GeV is nevertheless retained in the $bb4\ell$ MC samples, which allows for a consistent interface with both the HERWIG 7.2 and HERWIG 7.3 versions.

[15] For parton-shower emissions originating from top- or anti-top-quark decays, a dedicated matching procedure is used, which is independent of the $p_{\mathrm{T}}^{\mathrm{def}}$ parameter, as described in Section 2.5 and Section 3.3.



(a) (b)

Figure 3.5: The invariant mass distribution of the hardest b-jet (a) and of the reconstructed top-quark (b), as predicted by the nominal ATLAS $bb4\ell$ sample and by an alternative $bb4\ell$ simulation, in which the minimum transverse momentum generated within the POWHEG framework is set to its default value of $\sqrt{0.8}\,\text{GeV}$, instead of 1.2 GeV used in the nominal sample.

- $p_{\text{T}}^{\text{def}} = 0$: The POWHEG ISR p_{T} definition (i.e. the p_{T} with respect to the beam axis) is used for both, ISR and FSR.
- $p_{\text{T}}^{\text{def}} = 1$: The POWHEG ISR p_{T} definition is used for ISR, and the POWHEG FSR p_{T} definition is used for FSR. The definition of the POWHEG FSR p_{T}^2 between outgoing particles j and k is given by $(p_j + p_k)^2 \frac{E_j E_k}{(E_j + E_k)^2}$.
- $p_{\text{T}}^{\text{def}} = 2$: The PYTHIA p_{T} definitions are used for both, ISR and FSR.

For a consistent matching of POWHEG-generated hard process events to the parton shower, the p_{T} scale of parton-shower emissions should also be calculated using the POWHEG p_{T} definition ($p_{\text{T}}^{\text{def}} = 1$), such that vetoing emissions with $p_{\text{T}} > p_{\text{T}}^{\text{veto}}$ is most meaningful. However, when applying these consistent matching settings, sizeable discrepancies are observed in unfolded jet multiplicity N_{jet} distributions, which are sensitive to variations in the PYTHIA 8 matching parameter settings, as shown in Figure 3.6(a). While both the $bb4\ell$ and $t\bar{t}$ predictions obtained with the $p_{\text{T}}^{\text{def}} = 1$ setting overestimate the jet multiplicity, the corresponding predictions using $p_{\text{T}}^{\text{def}} = 2$ show a better agreement with the unfolded CMS $t\bar{t}$ data from [107], the same measurement which was discussed in the context of Figure 3.3(b). Thereby, the jet multiplicity distributions predicted by the $bb4\ell$ and $t\bar{t} + tW$ simulation are found agree very well when using the same $p_{\text{T}}^{\text{def}}$ parameter setting. Consequently, as illustrated in the lowest ratio panel of Figure 3.6(a), variations of the $p_{\text{T}}^{\text{def}}$ parameter have a similar impact on both the $t\bar{t}$ and $bb4\ell$ predictions. The transverse momentum distribution of the reconstructed top-quark

pair, unfolded in the same CMS data analysis [107], is shown in Figure 3.6(b). While this observable is also sensitive to the matching of hard-process events to the parton shower and the parton-shower emission generation, the influence of the $p_{\mathrm{T}}^{\mathrm{def}}$ variations is found to be small.

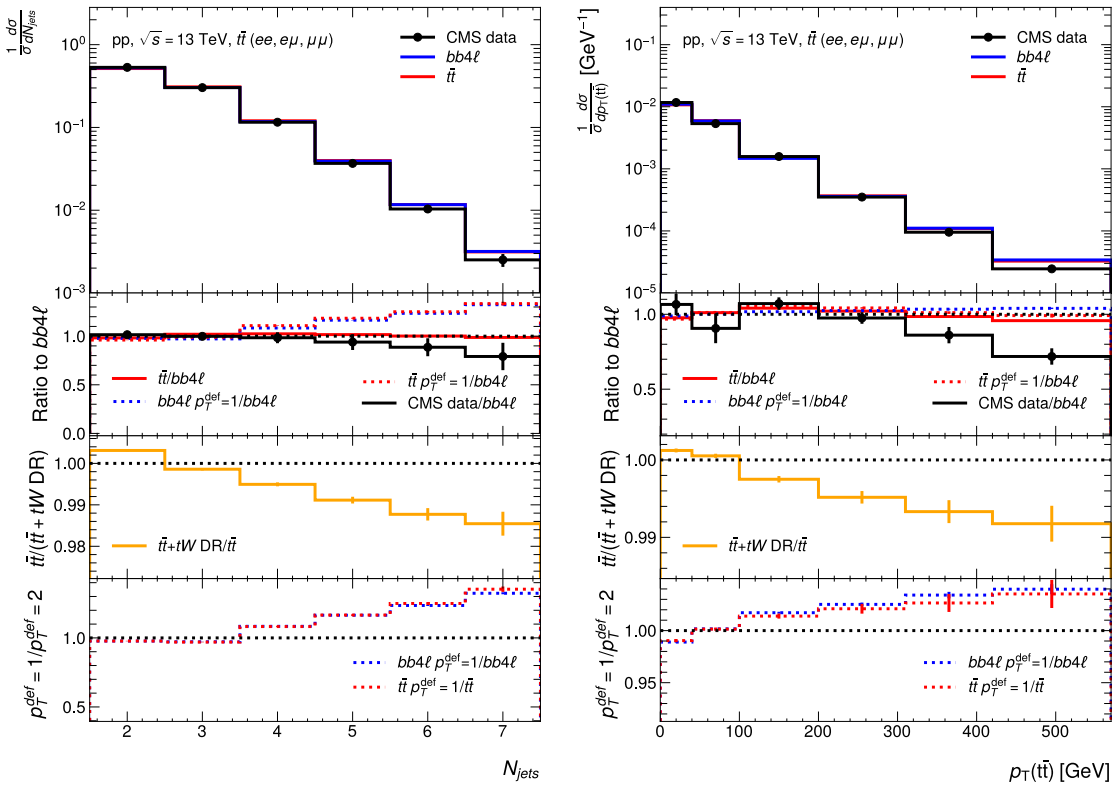


Figure 3.6: The jet multiplicity distribution N_{jet} (a) and the transverse momentum distribution of the reconstructed $t\bar{t}$ pair (b) as predicted with nominal ATLAS Powheg+Pythia 8 $bb4\ell$ and $t\bar{t}$ `hvq` MC simulation, using the Pythia 8 matching parameter $p_{\mathrm{T}}^{\mathrm{def}} = 1$ or $p_{\mathrm{T}}^{\mathrm{def}} = 2$, compared to unfolded CMS data $t\bar{t}$ dilepton data [107].

Given the improved agreement between data and simulation obtained with the $p_{\mathrm{T}}^{\mathrm{def}} = 2$ setting, and the fact that this setting is used in the ATLAS $t\bar{t}$ `hvq` and $t\bar{t}$ Minnlo MC samples, where the former is used also in the derivation of physics object calibrations in ATLAS, it was decided to adopt the $p_{\mathrm{T}}^{\mathrm{def}} = 2$ setting likewise for the $bb4\ell$ MC sample.

3.2 Uncertainties in the hard process generation

In this section, $bb4\ell$ modelling systematics estimated by varying settings in the POWHEG hard process generation are studied: matrix-element scale variations and their comparison to the size of inverse width corrections, PDF uncertainties and the variation of the POWHEG h_{damp} parameter.

Matrix-element scale variations The uncertainty due to unknown higher-order corrections in the NLO matrix-element calculation is estimated by varying the renormalisation (μ_R) and factorisation (μ_F) scales by the factors K_R and K_F, respectively, such that $\mu_R = K_R\mu_0$ and $\mu_F = K_F\mu_0$. Both, the envelopes of the seven-point variations,

$$(K_R, K_F) \in \{(0.5, 0.5), (0.5, 1), (1, 0.5), (1, 1), (1, 2), (2, 1), (2, 2)\} \tag{3.2}$$

and the three-point variations, using $K_R = K_F = 0.5$ and $K_R = K_F = 2$, are compared in the following. In the $bb4\ell$ implementation, the definition of the central scale μ_0 depends on the resonance history assigned to a given LHE. For a $t\bar{t}$ resonance history, the central scale[16] is set to [2]

$$\mu_0 = [(m_t^2 + p_{\text{T},t}^2)(m_{\bar{t}}^2 + p_{\text{T},\bar{t}}^2)]^{1/4}. \tag{3.3}$$

For a tWb resonance history, μ_0 is defined as

$$\mu_0 \quad = \quad [(m_t^2 + p_{\text{T},t}^2)(m_b^2 + p_{\text{T},b}^2)]^{1/4}. \tag{3.4}$$

The transverse momenta p_{T} and the invariant masses m of the (anti-)top and (anti-)bottom quark in the μ_0 definition are thereby evaluated at the underlying Born level in POWHEG, i.e. the Born configuration without any additional real emission. In the $t\bar{t}$ `hvq` simulation, the central scale is set to

$$\mu_0 = \sqrt{p_{\text{T},t}^2 + m_t^2}. \tag{3.5}$$

Also in the `hvq` process, $p_{\text{T},t}$ is defined as the transverse momentum of the top quark before any real radiation, such that $p_{\text{T},t} = p_{\text{T},\bar{t}}$. The $t\bar{t}$ central scale definition is thus very close to the scale definition used for double-resonant topologies in $bb4\ell$, except that in the `hvq` $t\bar{t}$ simulation always a fixed top mass of $m_t = 172.5\,\text{GeV}$ is used in the μ_0 definition, since the calculation assumes on-shell top-quarks, whereas in $bb4\ell$ off-shell effects are fully accounted for at NLO. For the tW process calculation [75], a dynamic scale definition [111] is used also, given by

$$\mu_0 = H_{\text{T}}/2 \text{ with } H_{\text{T}} = \sqrt{m_t^2 + p_{\text{T},t}^2} + \sqrt{m_W^2 + p_{\text{T},W}^2} + p_{\text{T}}^j,\ p_{\text{T}}^j = 0 \text{ if no real emission.} \tag{3.6}$$

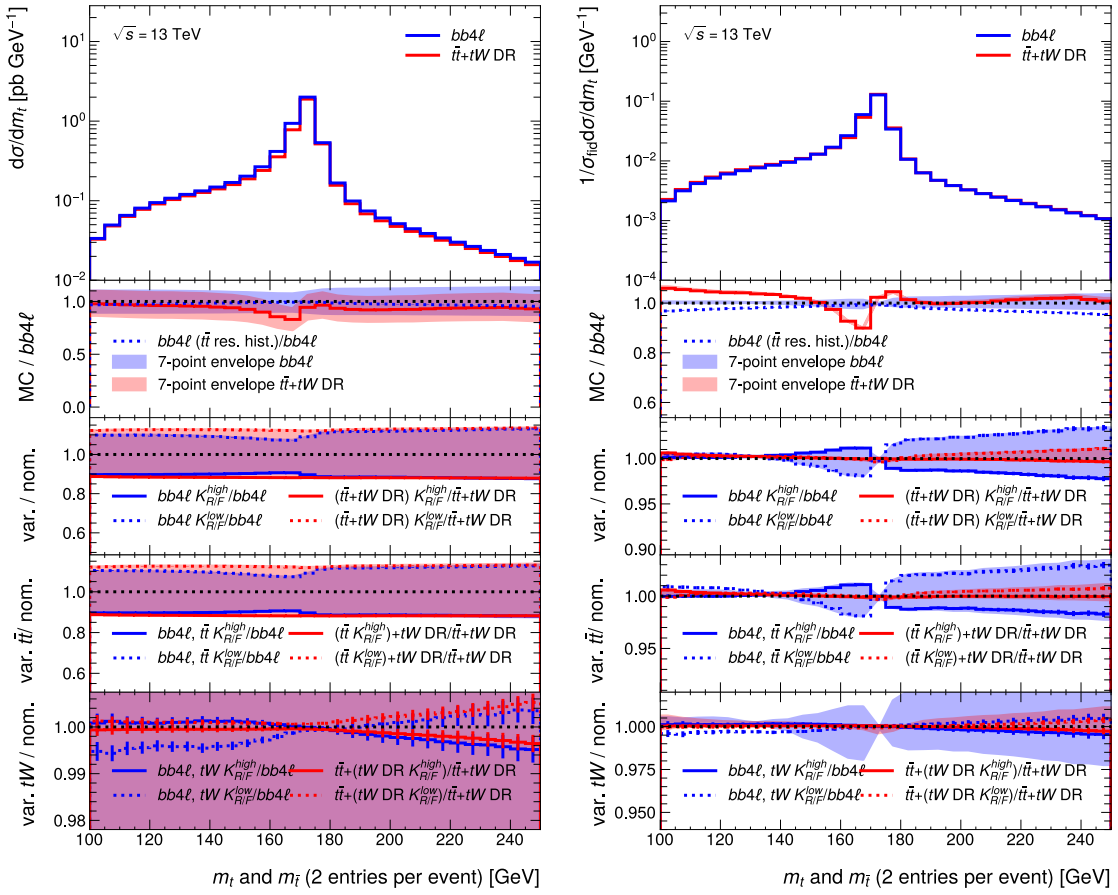
[16] In a private communication, the authors of [2] recommended to use a reduced central scale $\mu_0' = \frac{1}{2}\mu_0$ for future productions of the $bb4\ell$ sample, as this alternative scale should improve the agreement of the nominal prediction with higher-order QCD calculations. It was confirmed that the scale variation uncertainties obtained with variations around the central scales μ_0 and μ_0' are in reasonable agreement.

Thus, the transverse momentum of an additional jet p_{T}^{j} is included, if a real emission was generated in the POWHEG framework.

In Figure 3.7, the ME scale-variation uncertainty in the (a) absolute and the (b) normalised invariant mass distribution of the reconstructed top-quark is shown. The envelopes of the seven-point scale variation obtained with the full $bb4\ell$ and the $t\bar{t} + tW$ predictions are represented in the four ratio panels as blue and red bands, respectively. The comparison in the second ratio panel shows that the seven-point uncertainties are well described by their corresponding three-point uncertainties. The scale-variation uncertainty in $t\bar{t} + tW$ remains flat over the entire m_t range, while the $bb4\ell$ scale-variation uncertainty has a shape dependence around the top-quark invariant-mass peak. For the normalised distribution (b), this leads to a significantly larger ME scale-variation uncertainty in the $bb4\ell$ compared to the $t\bar{t} + tW$ predictions. In a recent ATLAS top-quark mass measurement [112] using unfolded normalised distributions, an increased ME scale-variation uncertainty on the extracted top-quark mass parameter was found, when using off-shell $(2 \to 7)$ $t\bar{t}j$ predictions, instead of the corresponding on-shell $(2 \to 3)$ predictions, consistent with the studies presented here. Further, the ME scale-variation uncertainty is dominated by the corresponding variation in the $t\bar{t}$ process, as can be seen by comparing the third and fourth ratio panel. These panels compare the uncertainty from three-point ME scale variations, applied exclusively to either the $t\bar{t}$ or the tW process (or, in the case of the $bb4\ell$ simulation, to events corresponding to the respective resonance histories), with the seven-point scale variation envelopes obtained for the full $bb4\ell$ and $t\bar{t} + tW$ predictions.

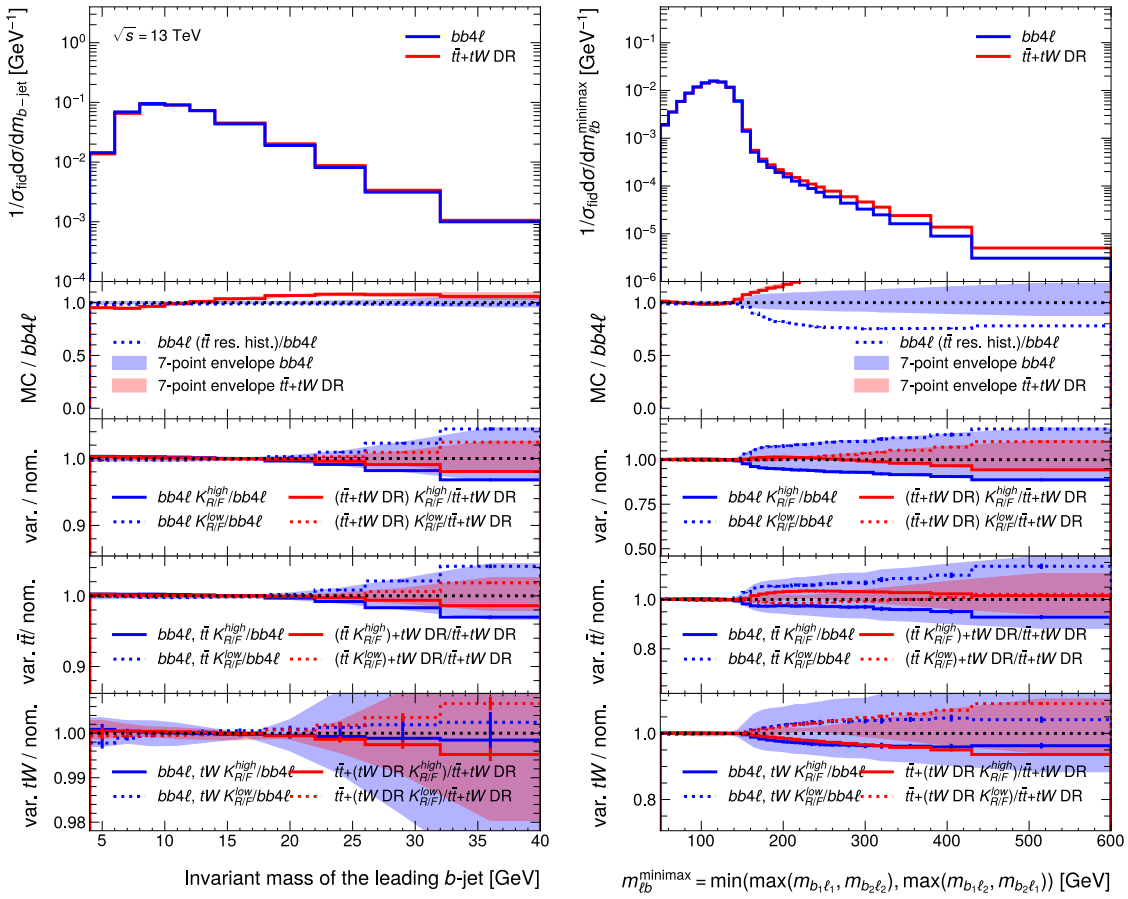
Figure 3.8 shows the same comparison as Figure 3.7, but for the invariant mass of the leading b-jet (a) and for the $m_{\ell b}^{\mathrm{minimax}}$ distribution (b), which is constructed from the two charged leptons and the two leading b-jets. Again, larger ME scale-variation uncertainties are observed for the $bb4\ell$ predictions, in particular when comparing the $t\bar{t}$ components of the $bb4\ell$ and $t\bar{t} + tW$ predictions. The larger size of the ME scale variations for $bb4\ell$ relative to the $t\bar{t}$ simulation in these distributions is not unexpected, since the top-quark decay and the $t\bar{t}/tW$ interference are modelled consistently at NLO QCD accuracy only in the $bb4\ell$ matrix-element calculation and are therefore sensitive to variations of the μ_R and μ_F scales.

In Figure 3.9, the $bb4\ell$ and $t\bar{t}+tW$ predictions are compared to the unfolded CMS $t\bar{t}$ differential cross-section of the transverse momentum of the reconstructed top quark [107], which was reconstructed using the positively charged lepton. Again both the absolute (a) and the normalised (b) distributions are shown. The corresponding three-point ME scale variation uncertainty is found to be significantly reduced in the normalised distribution, in this case both for the $bb4\ell$ and the $t\bar{t} + tW$ prediction. This emphasizes that the ME scale variations have only a limited impact on the shape of the distribution. In case of observables which are not very sensitive to the top-quark line shape, to $t\bar{t}/tW$ interference modelling or to b-jet properties, a similar impact of the ME scale-variation uncertainties on the $bb4\ell$ and $t\bar{t}$ predictions is found. While the data–MC disagreement is mostly covered by the normalisation impact of the ME scale-variation uncertainties in the absolute distribution in (a), this is clearly not the case for the normalised distribution in (b), emphasizing the small ME scale-variation impact on the shape of this distribution. In ATLAS top-quark data analyses using $t\bar{t}$ predictions, it is therefore recommended to apply a NNLO QCD reweighting uncertainty [90],



(a) (b)

Figure 3.7: (a) Absolute and (b) normalised distribution of the invariant mass of the reconstructed top quark as predicted by the $bb4\ell$ and $t\bar{t}+tW$ simulations. The envelope of the seven-point ME scale variations for the full $bb4\ell$ and $t\bar{t}+tW$ predictions is shown in all ratio panels as blue and red bands, respectively. These are compared to the three-point ME scale variations ($K_{R/F}^{\text{high}}=2$, $K_{R/F}^{\text{low}}=0.5$), applied in different configurations: in the second ratio panel to the full $bb4\ell$ and $t\bar{t}+tW$ predictions, in the third ratio panel to the $bb4\ell$ prediction projected onto $t\bar{t}$ resonance histories and to the $t\bar{t}$ prediction, and in the fourth ratio panel to the $bb4\ell$ prediction projected onto tW resonance histories and to the tW prediction.

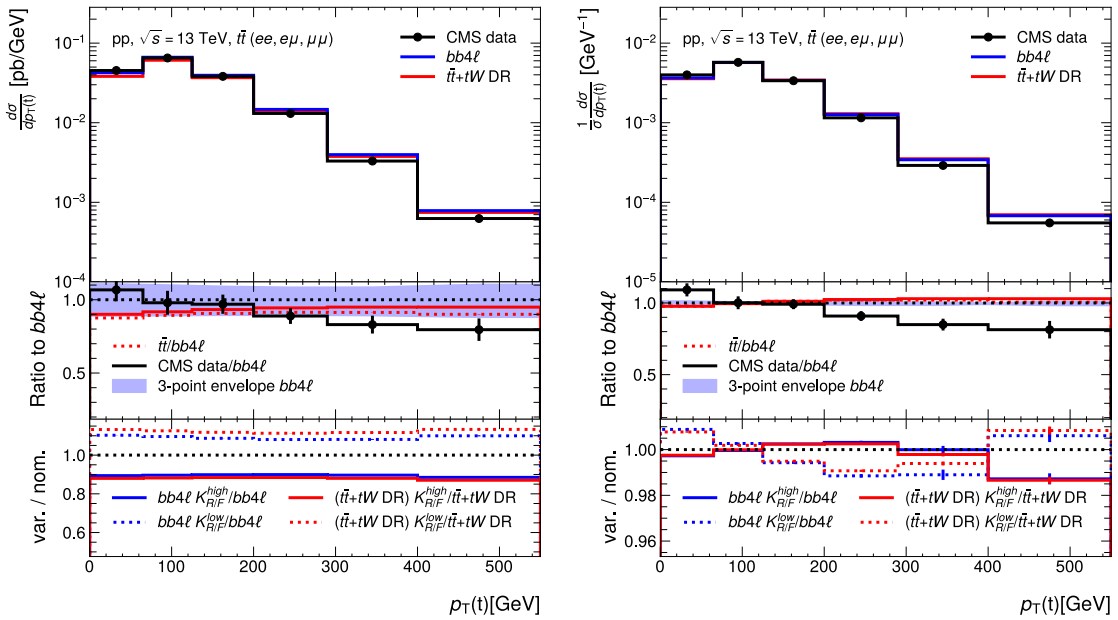


(a) (b)

Figure 3.8: Same comparison as in Figure 3.7, but for (a) the invariant mass of the leading b-jet and (b) the $m_{\ell b}^{\text{minimax}}$ distribution.

especially for variables like the transverse momentum of the reconstructed top-quark or of charged leptons, for which a mismodelling in MC was observed. This reweighting corrects the parton-level kinematics of the top-quark transverse momenta and the $t\bar{t}$ mass of the NLO+PS $t\bar{t}$ MC sample to the fixed-order NNLO QCD predictions. In contrast, the reweighting of the parton-level kinematics of $bb4\ell$ MC events, which are assigned a $t\bar{t}$ resonance history, is currently not recommended by the authors of [2].[17]

[17] The validity of reweighting $bb4\ell$ MC events simulated according to a NLO QCD off-shell $WWbb$ calculation, albeit only those projected onto a $t\bar{t}$ resonance history, to NNLO QCD on-shell predictions needs further study. A necessary requirement is the application of the inverse width corrections in the $bb4\ell$ simulation, but also the LO QCD accurate resonance history projectors need to be further studied. In a first step, the $bb4\ell$ predictions of differential cross-sections for different resonance history projectors have been compared and very good agreement has been found, as already noted in [2].



(a) (b)

Figure 3.9: (a) Absolute and (b) normalised transverse momentum distribution of the reconstructed top-quark as predicted by the $bb4\ell$, $t\bar{t}+tW$ and $t\bar{t}$ simulations and compared to unfolded CMS $t\bar{t}$ data [107]. Additionally, in the second ratio panel, the three-point ME scale variations are shown with $K^{\text{high}}_{R/F} = 2$ and $K^{\text{low}}_{R/F} = 0.5$. In the case of the $bb4\ell$ ME scale uncertainty, this is also indicated by the uncertainty band in the first ratio plot.

Inverse width corrections A feature introduced in the new version of the $bb4\ell$ POWHEG process implementation [2] is the correction for spurious top-quark width effects when comparing, at a fixed order in perturbation theory, the full off-shell calculation in the zero-width limit to the narrow-width-approximation (NWA) of this process.

The discussion in [2] is briefly summarised in the following. In the NWA of the production and decay of a resonance, the cross-section factorises in an independent production and decay contribution as

$$\mathrm{d}\sigma_{\text{prod}\times\text{dec}} = \mathrm{d}\sigma\frac{\mathrm{d}\Gamma}{\Gamma}, \text{ with } \Gamma = \int_{\text{dec}} \mathrm{d}\Gamma \text{ and } \mathrm{d}\sigma = \int_{\text{dec}} \mathrm{d}\sigma_{\text{prod}\times\text{dec}}, \tag{3.7}$$

where $\mathrm{d}\sigma$ is the differential production cross-section. It can be shown, that when using $1/\Gamma_{\text{NLO}}$ in the formula above for an NLO QCD calculation, a spurious term of the order $\mathcal{O}(\alpha_S^2)$ remains after integration over the full decay phase space:

$$\int_{\text{dec}} \mathrm{d}\sigma^{\text{NLO}}_{\text{prod}\times\text{dec}} = \mathrm{d}\sigma^{\text{NLO}}_{\text{prod}}\frac{\mathrm{d}\Gamma^{\text{NLO}}}{\Gamma^{\text{NLO}}} = \mathrm{d}\sigma_0 + \mathrm{d}\sigma_1 - \mathrm{d}\sigma_1\frac{\Gamma_1}{\Gamma_{\text{NLO}}}. \tag{3.8}$$

Thereby σ_0, Γ_0 are the LO contributions to the production cross-section and the top-quark decay width respectively and σ_1, Γ_1 the NLO QCD corrections to them. This spurious width term

$$\mathrm{d}\sigma_{\text{spurious}} = -\mathrm{d}\sigma_1 \frac{\Gamma_1}{\Gamma_{\text{NLO}}} \tag{3.9}$$

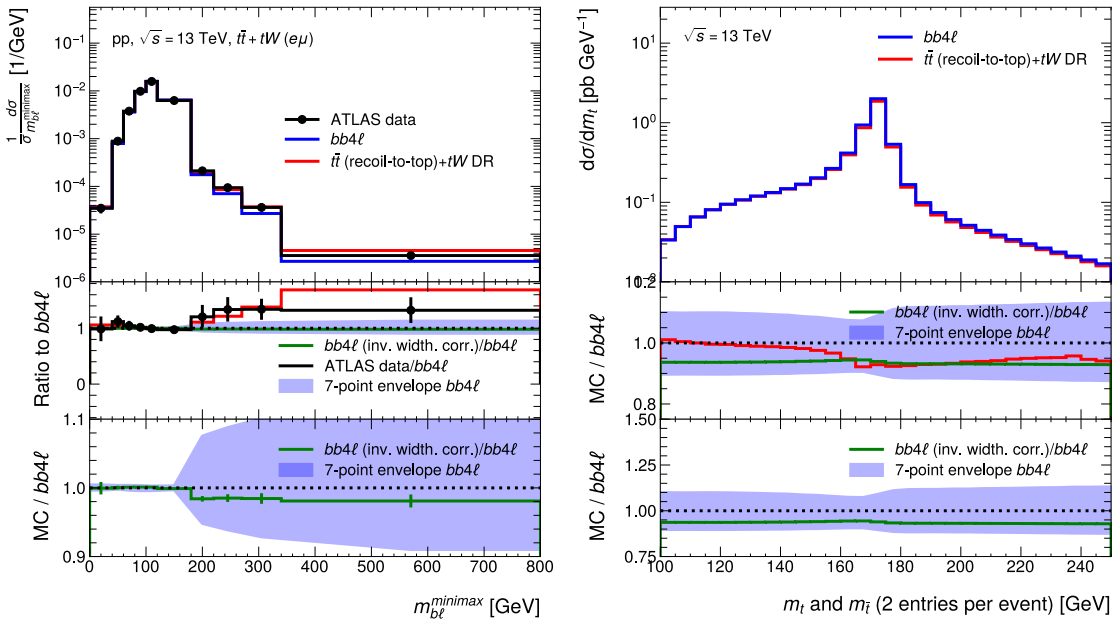
is removed when using instead the top-quark width expansion

$$\frac{1}{\Gamma_{\text{NLO}}} \to \frac{1}{\Gamma_0}\left(1 - \frac{\Gamma_1}{\Gamma_0}\right) \tag{3.10}$$

in Equation 3.7 and thus satisfying Equation 3.7 order-by-order in the perturbative series. Following this procedure of the NWA, the option to apply similar spurious width corrections in the off-shell $bb4\ell$ calculation was implemented in [2]. This ensures consistency of the off-shell calculation in the zero-width limit with the calculations performed in the NWA, which generally use these spurious width corrections.

As discussed in Section 3.1, finding good agreement in the normalisation of the absolute $bb4\ell$ and $t\bar{t}+tW$ predictions requires the application of the inverse width correction to the $bb4\ell$ prediction. Since the inverse width corrections are a higher-order effect, they are expected to be covered by the matrix-element scale variations, and therefore no additional uncertainty needs to be defined. That this expectation is fulfilled is demonstrated by comparing the seven-point ME scale variation uncertainty band with the effect of the inverse width corrections in Figure 3.10. The comparison is shown for (a) the $m_{\ell b}^{\text{minimax}}$ distribution, including unfolded ATLAS $t\bar{t}+tW$ data [102], and for (b) the invariant mass distribution of the reconstructed top quark. The $m_{\ell b}^{\text{minimax}}$ distribution was already discussed in Figure 3.1(b). As seen in Figure 3.10(a), the matrix-element scale variations are sizeable in the $t\bar{t}/tW$ interference region in the $m_{\ell b}^{\text{minimax}}$ tail. Both the normalisation and the shape differences induced by the application of the inverse width corrections are fully covered by the ME scale-variation uncertainty. Although the authors of [2] recommend applying the inverse width corrections to the nominal $bb4\ell$ and all systematically varied $bb4\ell$ MC samples, as is planned for future ATLAS MC productions, this procedure is not adopted for the $bb4\ell$ MC sample presented in this thesis. As demonstrated above, the resulting difference is nonetheless adequately covered by the applied modelling uncertainties.

PDF uncertainties A precise assessment of theoretical uncertainties associated with PDFs is essential for reliable predictions at hadron colliders. For the $bb4\ell$ prediction, the PDF uncertainty is evaluated using the PDF4LHC21 eigenvector variations [113]. To validate this modelling uncertainty prescription, the size of the PDF4LHC15 eigenvector variations [114] is compared between the $bb4\ell$ and $t\bar{t}$ predictions, and they are observed to agree well. In addition, the variation of the nominal prediction based on NNPDF3.0 is studied when alternative PDF sets — PDF4LHC15, MMHT14 [115], and CT14 [116] — are used. These variations are compared between the $bb4\ell$ and $t\bar{t}$ simulations and are again found to be in excellent agreement. It is therefore concluded that the considered PDF uncertainty estimation is adequate. Consequently, the more recent PDF4LHC21 prescription is adopted, in contrast to the PDF4LHC15 prescription used currently for the ATLAS $t\bar{t}$ prediction.



(a) (b)

Figure 3.10: Comparison of the nominal $bb4\ell$ and $t\bar{t}+tW$ prediction, as well as the seven-point matrix-element scale-variation uncertainty band evaluated with the $bb4\ell$ MC sample and the $bb4\ell$ prediction, in which the inverse width correction weights are applied for the distribution of (a) $m_{\ell b}^{\text{minimax}}$— including the comparison to unfolded ATLAS $t\bar{t}+tW$ data [102] and (b) the invariant mass of the reconstructed top quark.

h_{damp} variations Within the POWHEG BOX framework, real emissions can be generated in two distinct ways: soft emissions through the POWHEG formula given in Equation 2.60, and hard emissions generated according to the real-emission matrix element. For soft emissions, the ratio R/B appearing in the Sudakov exponent of the POWHEG formula approaches the Altarelli–Parisi splitting functions in the infrared limit. In this regime, the generation of emissions via the POWHEG Sudakov form factor is required in order to preserve the logarithmic accuracy of the parton shower of the full NLO+PS prediction. This interpretation does not apply to hard emissions generated within POWHEG. In the case of hard radiation, the exponentiation of the restricted integral of the R/B ratio over the real-emission phase space cannot be interpreted as an approximation of a parton-shower Sudakov form factor. It is nevertheless important to emphasize that the resulting prediction remains NLO+PS accurate, independently of how the real-emission phase space is partitioned into soft and hard contributions during the POWHEG event generation. Even if the R/B factor were exponentiated for hard-emission events, the induced differences would only appear at higher orders in the strong coupling constant, namely at NNLO in QCD. Accordingly, soft emissions are generated using the POWHEG approach, while events with hard emissions have the corresponding four momenta distributed according to the real matrix-element squared in the hard-emission phase space. This separation is

implemented through a split of the real-emission contribution into a singular part, $\mathcal{R}^s$, and a finite remainder, $\mathcal{R}^f$, which is controlled by the h_{damp} parameter via

$$\mathcal{R} = \mathcal{R}^s + \mathcal{R}^f \text{ with } \mathcal{R}^s = \mathcal{R}F(k_{\text{T}}^2, h_{\text{damp}}) \text{ and } \mathcal{R}^f = \mathcal{R}(1 - F(k_{\text{T}}^2, h_{\text{damp}})) \tag{3.11}$$

and

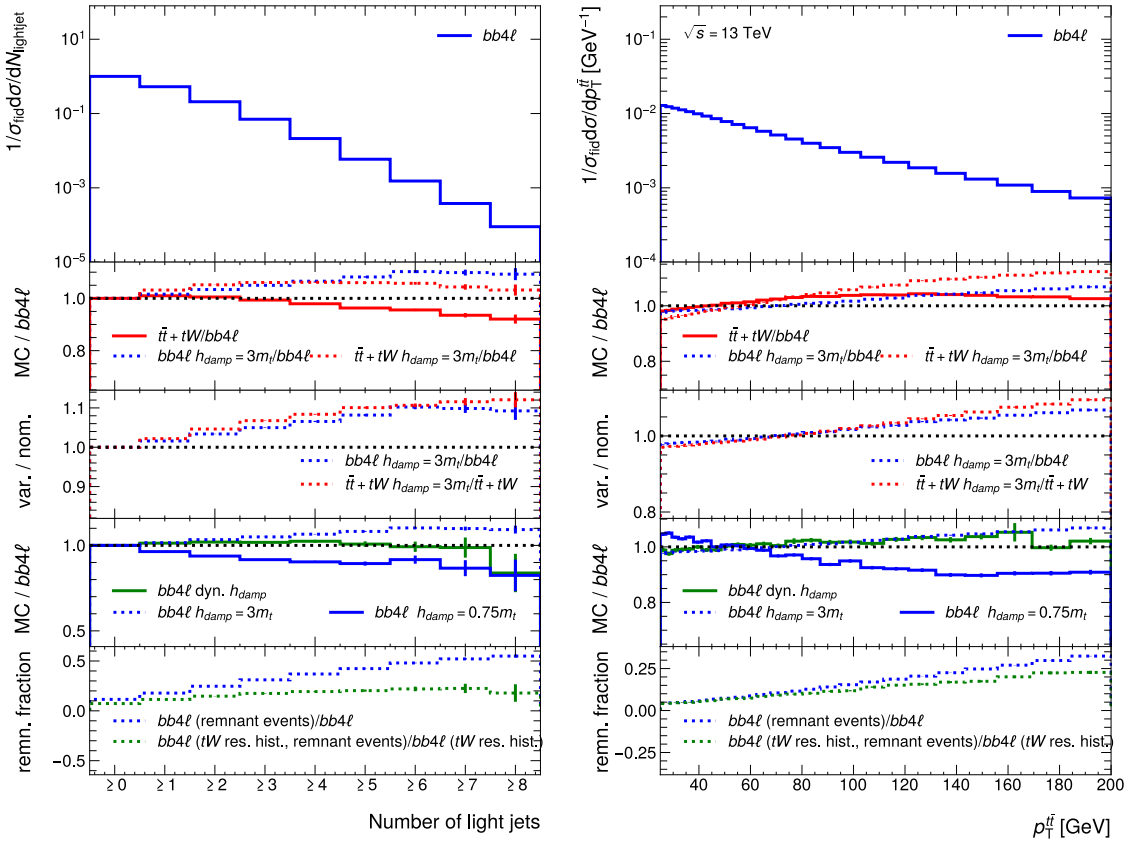
$$F(k_{\text{T}}^2, h_{\text{damp}}) = \frac{h_{\text{damp}}^2}{k_{\text{T}}^2 + h_{\text{damp}}^2}. \tag{3.12}$$

From the functional form of $F(k_{\text{T}}^2, h_{\text{damp}})$ it follows that $\mathcal{R}^s$ approaches the full real-emission contribution in the limit $h_{\text{damp}} \to \infty$. Consequently, larger values of h_{damp} result in a larger fraction of emissions being generated via the POWHEG Sudakov form factor, as the singular component $\mathcal{R}^s$ increases. The value of the h_{damp} parameter cannot be derived from first principles, but has to be tuned in order to improve the agreement between data and MC. This possibility was already discussed in the original description of the POWHEG method [46] and was first implemented in POWHEG in the context of Higgs boson production via gluon fusion [117].

The nominal h_{damp} value chosen for the ATLAS $bb4\ell$ MC sample is $h_{\text{damp}} = 1.5\,m_{\text{top}}$,[18] matching the h_{damp} parameter used in the $t\bar{t}$ and tW POWHEG generation in ATLAS. The h_{damp} modelling uncertainty prescription applied to the $t\bar{t}$ and tW MC samples, obtained from a comparison between the nominal prediction and a prediction with $h_{\text{damp}} = 3\,m_{\text{top}}$, is also tested for the $bb4\ell$ simulation. The results are shown in Figure 3.11 for (a) the light-jet multiplicity and (b) the transverse momentum of the reconstructed top-quark pair. The size of the variation obtained by setting $h_{\text{damp}} = 3\,m_{\text{top}}$ in the $bb4\ell$, $t\bar{t}$, and tW simulations, relative to the nominal choice, is evaluated in the second ratio panel. Good agreement is observed in the resulting uncertainty estimates for the $bb4\ell$ and $t\bar{t} + tW$ samples. The third ratio panel compares different h_{damp} variations within the $bb4\ell$ MC sample: the default variation to $h_{\text{damp}} = 3\,m_{\text{top}}$, a downward variation to $h_{\text{damp}} = 0.75\,m_{\text{top}}$, and a dynamical h_{damp} choice. In the latter case, h_{damp} is set equal to the renormalisation and factorisation scale, whose definitions depend on the resonance history of the event, as discussed previously. The predictions obtained with the dynamical h_{damp} setting show reasonable agreement with those obtained using the nominal h_{damp} parameter value. In contrast, the downward variation to $h_{\text{damp}} = 0.75\,m_{\text{top}}$ leads to slightly larger deviations, and the validity of applying an $h_{\text{damp}} < m_{\text{top}}$ value remains under discussion. Consequently, the uncertainty prescription adopted for the $bb4\ell$ MC sample follows that of the $t\bar{t} + tW$ simulation, namely a variation to $h_{\text{damp}} = 3\,m_{\text{top}}$. It should be emphasized that in top-quark analyses in ATLAS these one-sided modelling uncertainties are typically symmetrised by applying the same variation in the opposite direction. The fourth ratio panel shows the fraction of remnant events, defined as events generated according to $\mathcal{R}^f$ in Equation 3.11, either for all $bb4\ell$ events or restricted to $bb4\ell$ events assigned a tW resonance history. As expected, the fraction of remnant events increases in regions of high jet multiplicity and large recoil of the $t\bar{t}$ system (high $p_{\text{T}}^{t\bar{t}}$), which corresponds to phase spaces that are dominated by hard emissions. Similar trends are observed

[18] In the following, m_{top} refers to the fixed input top-quark mass of 172.5 GeV in the MC generation and not to the invariant mass of the top-quark four momentum m_t.

for the full $bb4\ell$ sample and for the subset of tW-like events, although the latter exhibits a systematically smaller remnant fraction.



(a) (b)

Figure 3.11: Distribution of (a) the number of light jets and (b) the transverse momentum distribution of the reconstructed $t\bar{t}$ system, as predicted by the $bb4\ell$ and $t\bar{t}+tW$ simulations. The first and second ratio panel compare the size of the variation obtained by changing the h_{damp} parameter from its nominal value to $h_{\text{damp}} = 3\,m_{\text{top}}$. The third ratio panel shows for the $bb4\ell$ prediction additionally a downwards variation to $h_{\text{damp}} = 0.75\,m_{\text{top}}$ and a dynamical h_{damp} choice, where $h_{\text{damp}} = \mu_0$ as defined in Equations 3.3 and 3.4. The fourth ratio panel shows the fraction of remnant events for the full $bb4\ell$ prediction and for the subset of $bb4\ell$ events projected onto a tW resonance history.

3.3 Uncertainties in the parton-shower matching

LHEs generated with the original POWHEG method [46, 47] are matched to a parton shower by requiring that all parton-shower emissions are generated at scales below the POWHEG matching scale, as described in Section 2.4. For LHEs containing real radiation generated by POWHEG, this matching scale is set to the scale of the real emission, as calculated with the POWHEG p_{T} definition, and stored in the `scalup` variable of the LHE format [86, 87]. In the $bb4\ell$ POWHEG process implementation, up to three additional real emissions with respect to the $b\bar{b}\ell^{+}\ell^{-}\bar{\nu}_{\ell}\nu_{\ell}$ final state can be generated: one from the production process, one from each of the top- and one from the anti-top-quark decay.[19] Consequently, up to three distinct matching scales must be respected during the PS matching, which requires a dedicated $bb4\ell$-PS matching interface. Parton-shower emissions that can be associated with the production process are treated analogously to the default POWHEG approach, with the corresponding matching scale stored in the `scalup` variable of the LHE. In contrast, the matching scales for PS emissions originating from the top- or anti-top-quark decays must be reconstructed from the kinematics of the final-state particles in the LHE.[20] Since QCD radiation off heavy quarks is suppressed relative to radiation from light quarks and gluons, real emissions generated in the top- or anti-top-quark decays within the POWHEG BOX RES framework typically occur at lower scales than emissions from the production process. If no dedicated $bb4\ell$ PS matching interface is applied, the PS matching scale for decay emissions is instead set by the generally higher matching scale of the POWHEG emission from the production process. This leads to an overestimate of radiation off the b quark in the parton shower.[21]

This effect is illustrated in Figure 3.12, which compares $bb4\ell$ LHEs matched to the PYTHIA 8 and HERWIG 7 parton showers with and without the required dedicated $bb4\ell$ matching interface. The impact is shown for the b-jet fragmentation function (a) and the b-jet width (b), using the publicly available RIVET implementation of these observables [118]. Loose event selection criteria are applied by requiring at least one b-jet with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.5$. The b-jet fragmentation function is defined as the fraction of the jet three-momentum carried by the B-hadron, $\vec{p}_{\mathrm{had}} \cdot \vec{p}_{\mathrm{jet}}/\vec{p}^{\,2}_{\mathrm{jet}}$ and is evaluated for each B-hadron in each b-jet in the event. The jet width, also calculated for each b-jet in the event, is defined as the sum of the p_{T}-weighted angular distances of the particles in a jet to the jet-axis, normalised to the p_{T} sum of all particles in the jet

$$\frac{\sum_{i\in\mathrm{jet}} p_{\mathrm{T}}^{i}\Delta R(i,\text{jet-axis})}{\sum_{i\in\mathrm{jet}} p_{\mathrm{T}}^{i}}. \tag{3.13}$$

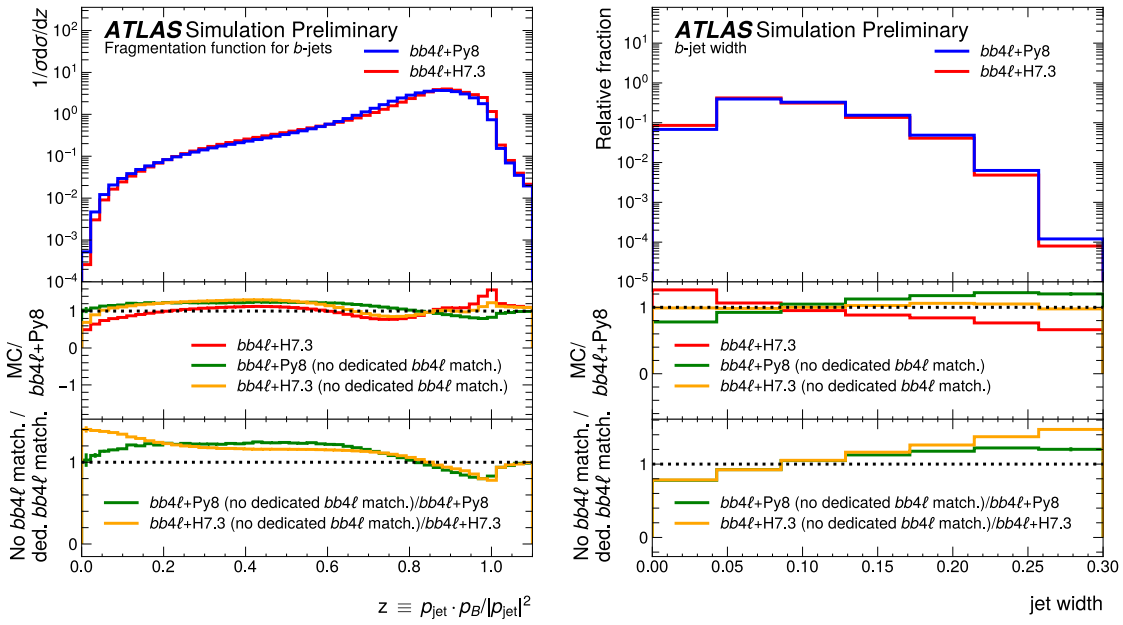
When the dedicated $bb4\ell$-PS matching interface is not applied in Figure 3.12, fewer B-hadrons carrying a large fraction of the jet momentum (large z) are observed, and the resulting

[19] This option is enabled in the ATLAS $bb4\ell$ MC samples by using the `allrad` mode [69]. The maximum number of real emissions depends on the projected resonance history. For a $t\bar{t}$ resonance history, up to three emissions are possible (from $t\bar{t}$ production and from the top- and anti-top-quark decays), while for a $tW^{-}\bar{b}$ topology the maximum number is two (from $tW^{-}\bar{b}$ production and top-quark decay).

[20] Technically, this is implemented for the PYTHIA 8 parton shower by a dedicated $bb4\ell$ UserHook [109], and in HERWIG 7 through a shared library implementing a `FullShowerVeto` [109].

[21] This is the case only when the `allrad` scheme is applied in the $bb4\ell$ POWHEG generation.

b-jets are broader due to the increased amount of PS radiation. The effect of the dedicated matching is found to be very similar in Pythia 8 and Herwig 7 for the b-jet observables shown in Figure 3.12. These comparisons validate the two independent $bb4\ell$-PS matching implementations and their integration into the ATLAS Athena software framework. However, they do not define a systematic uncertainty, as the dedicated $bb4\ell$ matching is required to avoid double counting of emissions.



(a) (b)

Figure 3.12: Comparison of the effect of applying or not applying the dedicated $bb4\ell$ matching interface, when matching $bb4\ell$ LHE with the Pythia 8 and Herwig 7 parton shower on the b-jet fragmentation function (a) and the width of the b-jet (b). The publicly available Rivet implementation [118] of these observables was used. Published in [119].

Due to the use of different Powheg+Pythia 8 matching interfaces for parton-shower emissions originating from the production and top-quark decay subprocesses in $bb4\ell$ MC events, the evaluation of matching uncertainties is correspondingly separated into contributions associated with the production and decay subprocesses.

Definition of matching uncertainty for emissions from the production process As introduced in [90], the ATLAS Powheg+Pythia 8 matching uncertainty in top-quark pair and single top-quark production processes is evaluated by varying the $p_{\mathrm{T}}^{\mathrm{hard}}$ parameter from its nominal value of $p_{\mathrm{T}}^{\mathrm{hard}} = 0$ to $p_{\mathrm{T}}^{\mathrm{hard}} = 1$. This Pythia 8 parameter controls the definition of the matching scale. For the default setting $p_{\mathrm{T}}^{\mathrm{hard}} = 0$, the matching scale is taken directly from the `scalup` value stored in the LHE file, which corresponds to the transverse momentum

of the POWHEG emission relative to its emitter, calculated using the POWHEG p_{T} definition. In contrast, for $p_{\mathrm{T}}^{\mathrm{hard}} = 1$, the transverse momentum of the POWHEG emission is computed with respect to all final-state particles and the beam axis, and the smallest resulting scale is used as the matching scale. As described in the beginning of this section, when matching $bb4\ell$ events to the PYTHIA 8 parton shower, emissions associated to the production process are treated identically to those in the $t\bar{t}$ `hvq` process, which follows the default POWHEG emission generation scheme. Consequently, parton-shower emissions from the production process in $bb4\ell$ events are expected to be affected in the same way as those in $t\bar{t}$ events when varying the $p_{\mathrm{T}}^{\mathrm{hard}}$ parameter. In addition to the $p_{\mathrm{T}}^{\mathrm{hard}} = 1$ variation used to define the ATLAS matching uncertainty, predictions obtained with $p_{\mathrm{T}}^{\mathrm{hard}} = 2$ are also studied. This corresponds to a more extreme variation, in which the matching scale is defined as the minimum transverse momentum found by computing the p_{T} of all final-state particles with respect to each other and the beam axis.

Definition of matching uncertainty for emissions from the decay process In the nominal $bb4\ell$ MC sample, parton-shower emissions from the top- and anti-top-quark decays are vetoed using the dedicated $bb4\ell$ POWHEG-PYTHIA 8 vetoing procedure described in the beginning of this section. To evaluate the matching uncertainty associated with emissions from the decay subprocess, an alternative approach is employed, following the prescription proposed in [109]. Thereby, the PYTHIA 8 resonance scale is reset to the value calculated as the matching scale for emissions from the top-quark decay subprocess. By default, the PYTHIA 8 resonance scale limits the maximum scale of parton-shower emissions originating from a given resonance decay. The key difference compared to the nominal $bb4\ell$ vetoing procedure is that, in the varied setup, the scale of the parton-shower emission, against which the matching scale is compared, is calculated using the PYTHIA 8 p_{T} definition rather than the POWHEG p_{T} definition in the top-quark decay, which is used in the nominal matching procedure.[22]

Comparison of matching uncertainties in $bb4\ell$ and $t\bar{t}$ Figure 3.13 shows the impact of the dedicated $bb4\ell$ matching procedure, as well as the size of the matching uncertainties evaluated using either the $bb4\ell$ or the $t\bar{t}$ MC simulation, for the transverse momentum distribution of the second leading light jet (a) and the reconstructed top-quark invariant mass distribution (b). The former observable is primarily sensitive to initial-state radiation, while the latter is sensitive to radiation from the top-quark decay products, which is directly reflected in the size of the corresponding matching uncertainties.

While the leading light jet generally corresponds to the POWHEG real emission, the second-leading light jet is typically generated by the parton shower. Since radiation from the b-quark is suppressed by its large mass, the second-leading light jet predominantly originates from

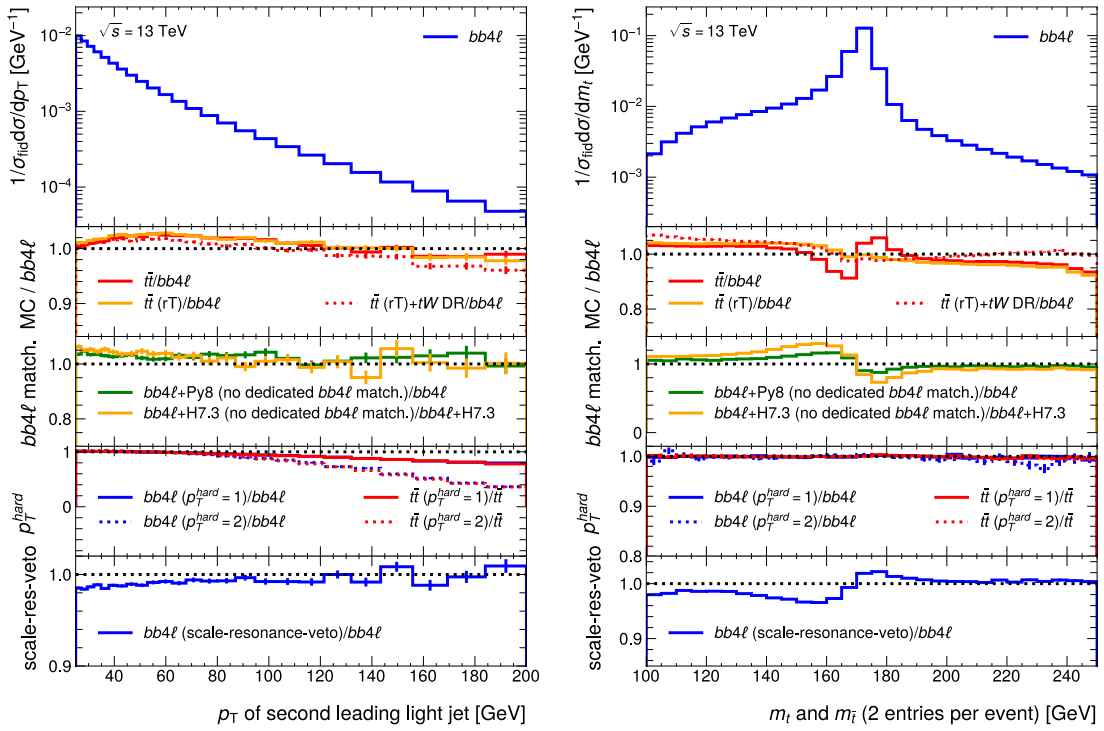
[22] As described in Section 3.1, the nominal $bb4\ell$ MC sample uses the matching setting $p_{\mathrm{T}}^{\mathrm{def}} = 2$ instead of the default $p_{\mathrm{T}}^{\mathrm{def}} = 1$. Although this choice could in principle affect the uncertainty evaluated using the resonance scale veto, since the default $bb4\ell$ decay matching is independent of the $p_{\mathrm{T}}^{\mathrm{def}}$ setting, good agreement is observed between the decay matching uncertainties obtained with $p_{\mathrm{T}}^{\mathrm{def}} = 1$ and $p_{\mathrm{T}}^{\mathrm{def}} = 2$. For this comparison, the $p_{\mathrm{T}}^{\mathrm{def}}$ setting was varied consistently in both the $bb4\ell$ predictions using the default $bb4\ell$ matching and those using the scale-resonance-veto setting.

initial-state radiation. As a result, its transverse momentum distribution is largely insensitive to the impact of the dedicated $bb4\ell$ parton-shower matching interface and to the matching uncertainty associated with the decay subprocess. Instead, this observable is strongly affected by the matching uncertainty in the production subprocess, in particular by the more extreme $p_{\mathrm{T}}^{\mathrm{hard}} = 2$ variation. The size of the matching uncertainty in the production process is found to be in excellent agreement between the $bb4\ell$ and $t\bar{t}$ predictions.

In contrast, the invariant mass distribution m_t of the reconstructed top-quark is largely insensitive to the matching uncertainty in the production subprocess,[23] but is sensitive to variations of the matching in the top-quark decay subprocess. As anticipated from the observations in Figure 3.12, not applying the required dedicated $bb4\ell$ matching, results in a shift of the reconstructed m_t distribution towards lower invariant masses, driven by increased radiation off the b quark and enhanced out-of-cone radiation. The difference between the dedicated $bb4\ell$ matching and the default POWHEG parton-shower matching is more pronounced when matching to the HERWIG 7 parton shower than to PYTHIA 8. This can be attributed to a larger impact on the transverse momentum distribution of the leading b-jet in HERWIG 7 when applying or omitting the dedicated $bb4\ell$ parton-shower interface. Since this distribution is sensitive to radiation from top-quark decay products, it is also expected to be sensitive to the corresponding decay matching uncertainty, which is indeed observed.

Both matching uncertainties definitions, those associated with the production and top-quark decay subprocesses, produce holes or overlaps in the emission phase space covered by POWHEG and the PYTHIA 8 parton shower. This is interpreted as a conservative uncertainty estimate. Alternative uncertainty definitions for $bb4\ell$ hard-process events are currently under discussion.

[23] In the $t\bar{t}$ MC sample, all parton-shower emissions are limited by the alternative matching scale defined via the $p_{\mathrm{T}}^{\mathrm{hard}}$ prescription, which could in principle affect b-jet-related observables. However, this matching scale is typically significantly larger than the characteristic scales of emissions from b-quarks. As a consequence, variations of this high matching scale do not significantly impact parton-shower emissions from b-quarks, whose relevant emission scales lie well below both the default and alternative matching scales.



(a) (b)

Figure 3.13: (a) Transverse momentum distribution of the second hardest light jet and (b) invariant mass distribution of the reconstructed top-quark as predicted with POWHEG+PYTHIA 8 $bb4\ell$, $t\bar{t}$, $t\bar{t}$ (rT) (recoil-to-top PYTHIA 8 setting) and $t\bar{t}$ (rT)+tW DR. The comparison of these predictions is shown in the first ratio panel. The second ratio panel shows in the influence of applying or not applying the dedicated $bb4\ell$ parton shower interface in PYTHIA 8 and HERWIG 7, the third ratio panel shows the matching uncertainty in the production process and the fourth ratio panel shows the matching uncertainty in the top-quark decay sub-process, which can only be evaluated in case of the $bb4\ell$ MC sample.

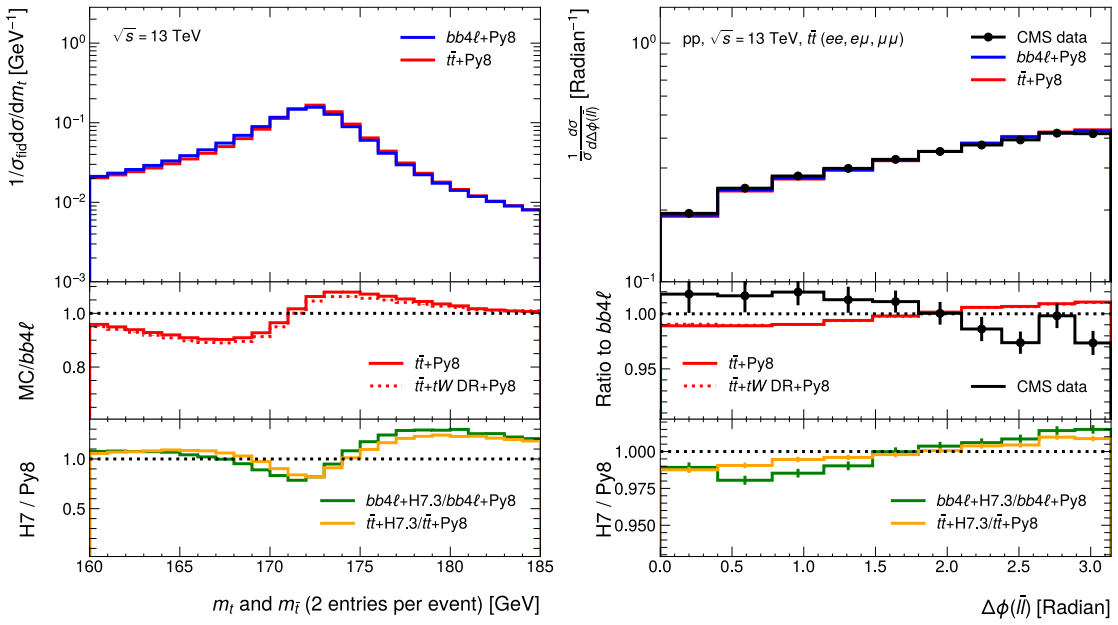
3.4 Uncertainties in the parton showering

This section discusses modelling uncertainties associated with the parton shower in the $bb4\ell$ MC simulation. For many of these uncertainties, a similar impact on the $bb4\ell$ and $t\bar{t}$ predictions is expected, since they primarily probe aspects of the parton shower that are largely independent of the details of the hard-process generation. This expectation is explicitly tested in the following, thereby validating the applicability of the corresponding uncertainty prescriptions to the $bb4\ell$ simulation. For parton-shower modelling uncertainties that are sensitive to radiation off b-quarks originating from top-quark decays, differences between the $bb4\ell$ and $t\bar{t}$ predictions can arise. This is studied in detail and the uncertainty prescription is adapted to ensure that all relevant modelling effects are adequately covered.

Pythia 8 vs. Herwig 7 A systematic uncertainty including a variety of different effects is obtained by comparing Monte Carlo samples in which the same hard-process events are interfaced either with the nominal Pythia 8 or with the alternative Herwig 7.3 parton shower. These two parton-shower implementations differ in several aspects, in particular in their shower ordering (dipole-ordered versus angular-ordered shower) and in their hadronisation models (Lund string versus cluster model). In addition, the underlying modelling parameters are tuned to data independently in the two generators. The parton-shower uncertainty estimate derived from comparing $bb4\ell$ and $t\bar{t}$ predictions interfaced with either the Pythia 8 or Herwig 7 parton shower, is found to agree well for b-jet-related observables, provided that the dedicated $bb4\ell$ parton-shower matching is applied consistently for both shower models. In [109], where the $bb4\ell$ Powheg+Herwig 7 matching was first introduced, a significantly enhanced parton-shower uncertainty was observed for top-quark mass sensitive observables in the $bb4\ell$ simulation compared to the $t\bar{t}$ `hvq` simulation. Using an improved version of the $bb4\ell$ dedicated Herwig 7 matching,[24] the parton-shower uncertainty obtained in the $bb4\ell$ and $t\bar{t}$ `hvq` simulations presented here is found to be comparable. This is illustrated in Figure 3.14(a) for the reconstructed top-quark invariant-mass distribution. The slight difference in the effect of the dedicated $bb4\ell$ matching in Pythia 8 and Herwig 7 in the top-quark invariant mass distribution in Figure 3.13(b) leads to very good agreement in the parton-shower uncertainty evaluated in the $bb4\ell$ and $t\bar{t}$ MC samples. Figure 3.14(b) shows the distribution of the azimuthal angle difference $\Delta\phi$ between the two charged leptons, including a comparison to unfolded CMS $t\bar{t}$ data from [107], which was already discussed in Figure 3.3(b). The parton-shower uncertainty for a related observable was studied in detail in the ATLAS $t\bar{t}$ entanglement observation analysis [120]. As seen in Figure 3.14(b), the impact of the parton-shower uncertainty on the $\Delta\phi$-observable is similar in the $bb4\ell$ and $t\bar{t}$ predictions.

Recoil uncertainty The Pythia 8 recoil setting affects the emission pattern of the second and subsequent emissions from the top-quark decay products. In the Pythia 8 parton shower, the colour connected parton r of e.g. an emitting quark q is used as a recoiler in a $q \to qg$ branching to preserve energy–momentum. Further, colour coherence is respected in the Pythia 8 parton

[24] Kindly provided to the ATLAS Collaboration by Silvia Ferrario Ravasio.

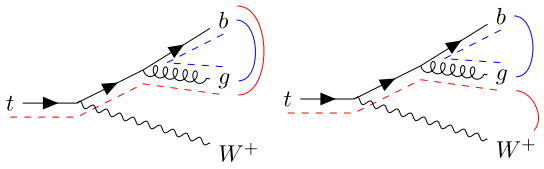


(a) (b)

Figure 3.14: Comparison of $bb4\ell$ and $t\bar{t}$ predictions, interfaced with the Pythia 8 (nominal) and the Herwig 7 parton shower, of (a) the invariant mass of the reconstructed top-quark mass and (b) the azimuthal angle between the two charged leptons, including the comparison to unfolded CMS data [107].

shower, by suppressing radiation outside an angle θ around the emitting particle direction, where θ corresponds to the angle between the two colour-connected partons. In the Pythia 8 parton shower, intermediate particles cannot act as a recoiler. When considering now a $t \to Wb$ branching, the b quark is colour-connected to the top-quark, which is an intermediate particle and therefore cannot serve as recoiler. For the first $b \to bg$ branching, the only possibility is to assign the W-boson as a recoiler. For an emitted gluon, which would be colour-connected to the top quark, a choice can be made: The colour-connection to the top-quark can be moved again to the W-boson or instead to the b quark (amounting to in total two colour connections to the b quark), as illustrated in Figure 3.15.

The phase space and the recoil is set by this choice, where moving the colour connection to the b quark leads in general to b-jets with less out-of-cone radiation [121], since the angle between the emitted gluon and the b quark is typically smaller than the angle between the emitted gluon and the W-boson. The option to use the b quark as a recoiler is known as recoil-to-colour and is the default option in the on-shell $t\bar{t}$ ATLAS MC sample. Another recoil scheme, called recoil-to-top, can also be applied, where the W-boson is assigned as a recoiler, but the gluon emission pattern is corrected to reflect more closely the emission pattern expected for an emission from a colour-connected parton to the top-quark. This



(a) (b)

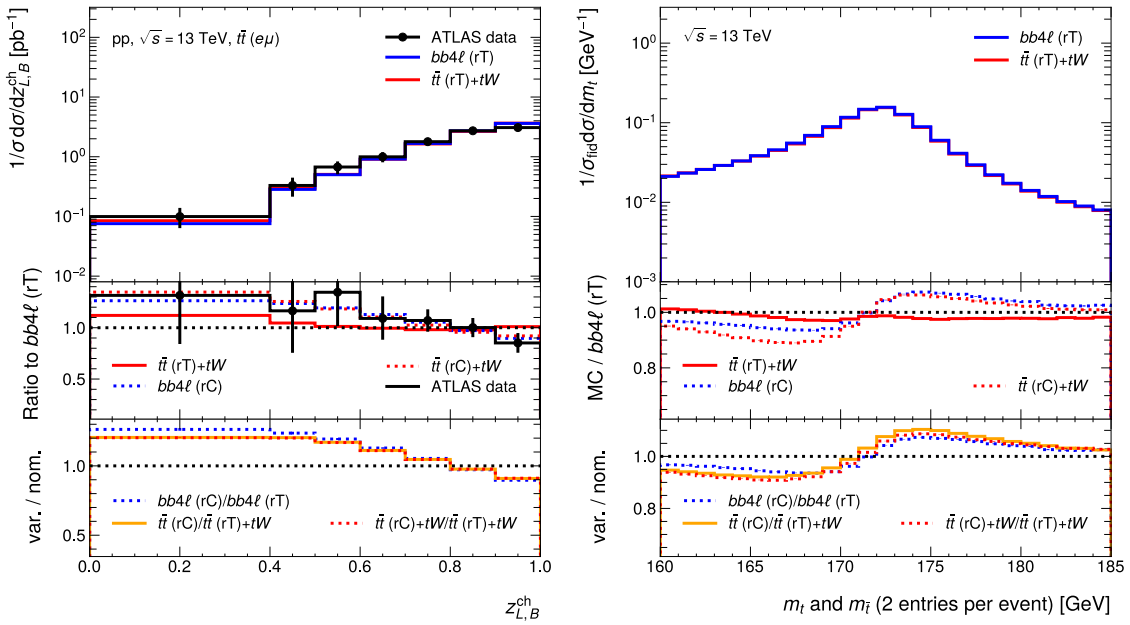
Figure 3.15: Sketch of the (a) recoil-to-colour and (b) recoil-to-W schemes in PYTHIA 8, where the colour connection to the top-quark is moved to the b quark or the W-boson, respectively.

should in principle lead to a better description of the second and subsequent gluon-emission pattern from the top-quark decay. The recoil-to-top setting is applied as the default setting in the ATLAS $bb4\ell$ MC sample, while it is used to define the systematic recoil uncertainty in the $t\bar{t}$ ATLAS MC prediction. The corresponding recoil uncertainty in case of the $bb4\ell$ MC sample is defined by comparing the nominal $bb4\ell$ predictions to the $bb4\ell$ predictions obtained with the recoil-to-colour setting.

Distributions sensitive to the recoil uncertainty include b-jet property observables, such as the charged b-jet fragmentation function $z^{\text{ch}}_{L,b}$ shown in Figure 3.16(a). This observable is defined as

$$z^{\text{ch}}_{L,b} = \frac{\vec{p}^{\,\text{ch}}_{b} \cdot \vec{p}^{\,\text{ch}}_{\text{jet}}}{|\vec{p}^{\,\text{ch}}_{\text{jet}}|^2}, \tag{3.14}$$

where $\vec{p}^{\,\text{ch}}_{\text{jet}}$ denotes the vectorial sum of the charged constituents momenta of the b-jet, and $\vec{p}^{\,\text{ch}}_{\text{b}}$ that of the charged decay products of the B-hadron. The particle-level event selection requires exactly one muon and one electron of opposite charge with $p_{\text{T}} > 25\,\text{GeV}$ and $|\eta| < 2.5$. Additionally, exactly two b-jets with $p_{\text{T}} > 30\,\text{GeV}$, $|\eta| < 2.5$ and $\Delta R(\text{jet}, \text{jet}) > 0.5$ are required. In the observable definition, only charged particles with $p_{\text{T}} > 500\,\text{MeV}$ are considered. Each selected b-jet must contain exactly one weakly decaying B-hadron with $p_{\text{T}} > 5\,\text{GeV}$ and at least three charged decay products of the decaying B-hadron. A comparison of the unfolded ATLAS $t\bar{t}$ data [122] with MC predictions obtained with the $bb4\ell$ and $t\bar{t} + tW$ simulations is shown for both the PYTHIA 8 recoil-to-colour and recoil-to-top settings. For the $t\bar{t} + tW$ prediction, the recoil model is varied only in the $t\bar{t}$ component, while the tW contribution is always showered using the recoil-to-colour setting. This choice is justified by the negligible impact of the tW contribution on this observable, as demonstrated by the close agreement between the $t\bar{t}$ and $t\bar{t} + tW$ predictions in Figure 3.16(a). Very good agreement is observed between the recoil uncertainty in the $bb4\ell$ and $t\bar{t}$ ATLAS MC samples. The recoil uncertainty is obtained by comparing the recoil-to-top and recoil-to-colour predictions and shown in the second ratio panels of Figure 3.16. The good agreement seems reasonable, since the recoil-to-top and recoil-to-colour configurations differ only in the treatment of the second and subsequent emissions from the top-quark decay, which are performed by the parton shower in both MC simulations. Both recoil models are compatible with the unfolded ATLAS data [122] within uncertainties. However, a slight trend towards better agreement with data is observed for distributions generated using the recoil-to-colour setting.



(a) (b)

Figure 3.16: (a) Fragmentation function of the b-jet, evaluated with the charged constituents $z^{\mathrm{ch}}_{L,b}$ and (b) the invariant mass of the reconstructed top-quark as predicted by the $t\bar{t}+tW$, $bb4\ell$ and $t\bar{t}$ simulation, using either the recoil-to-top (rT) or recoil-to-colour (rC) Pythia 8 setting in the $t\bar{t}$ and $bb4\ell$ simulation. In (a) also the comparison to unfolded ATLAS data [122] is shown.

Due to the impact of the recoil uncertainty in b-jet property distributions, observables sensitive to the top-quark mass, such as the invariant mass of the reconstructed top-quark shown in Figure 3.16(b), are also affected by this uncertainty. For this reason, the recoil uncertainty can constitute one of the dominant systematic uncertainties in top-quark mass measurements [123, 124]. As seen in Figure 5.3, the LHCTopWG summary of top-quark mass measurement lists this uncertainty as a separate contribution in the total uncertainty breakdown. A recent ATLAS top-quark mass measurement [125] introduced a dedicated control region to determine which recoil model is preferred within the specific analysis phase space. In that measurement, the recoil-to-colour configuration was favoured over the recoil-to-top setting. While the most theoretically precise description of the radiation pattern from the top-quark decays would be provided by the Vincia parton shower [126], its employment and the corresponding tuning in ATLAS is a topic of current studies. Within Pythia 8, the recoil-to-top configuration is expected to model emissions from top-quark decay products more accurately than the recoil-to-colour setting. However, some comparisons to data indicate that the recoil-to-colour option might yield better agreement. The extent to which this trend is influenced by ATLAS physics-object calibrations and parton-shower tuning, both of which were derived using $t\bar{t}$ MC samples with the recoil-to-colour configuration, is difficult to quantify.

Pythia 8 splitting kernel variations To estimate uncertainties associated to variations of the parton-shower splitting kernel, the terms introduced in Equation 2.58 and Equation 2.59 can be reparameterised for the PYTHIA 8 transverse momentum-ordered shower ($t = p_\perp$) as

$$\frac{\alpha_S(t)}{2\pi} K_{\tilde{i}\to ij} \to \frac{\alpha_s(kt)}{2\pi} \cdot \mathcal{C} \cdot \left(\frac{P(z) + c_{\text{NS}} Q^2/m_{\text{dip}}^2}{t} \right), \tag{3.15}$$

where $P(z)$ denotes the DGLAP splitting kernels, $\mathcal{C}$ is a colour factor, m_{dip} is the invariant mass of the dipole emitter–recoiler system and Q^2 an energy scale, which is proportional to the shower ordering variable t. Uncertainties due to missing higher-order contributions in the perturbative expansion are estimated by varying the scale at which the strong coupling α_s is evaluated, through variations of the factor k. In the following, these are referred to as μ_R variations. By construction, they are proportional to the original splitting kernels [127]. In contrast, variations of the parameter c_{NS}, which is set to zero in the nominal configuration, probe the uncertainty associated with the choice of splitting-kernel parameterisation. While all parameterisations coincide in the soft-collinear limit, they differ in the non-singular region. The uncertainty estimate obtained by varying c_{NS} to ± 2 is therefore particularly relevant in regions of phase space dominated by hard parton shower emissions [127].

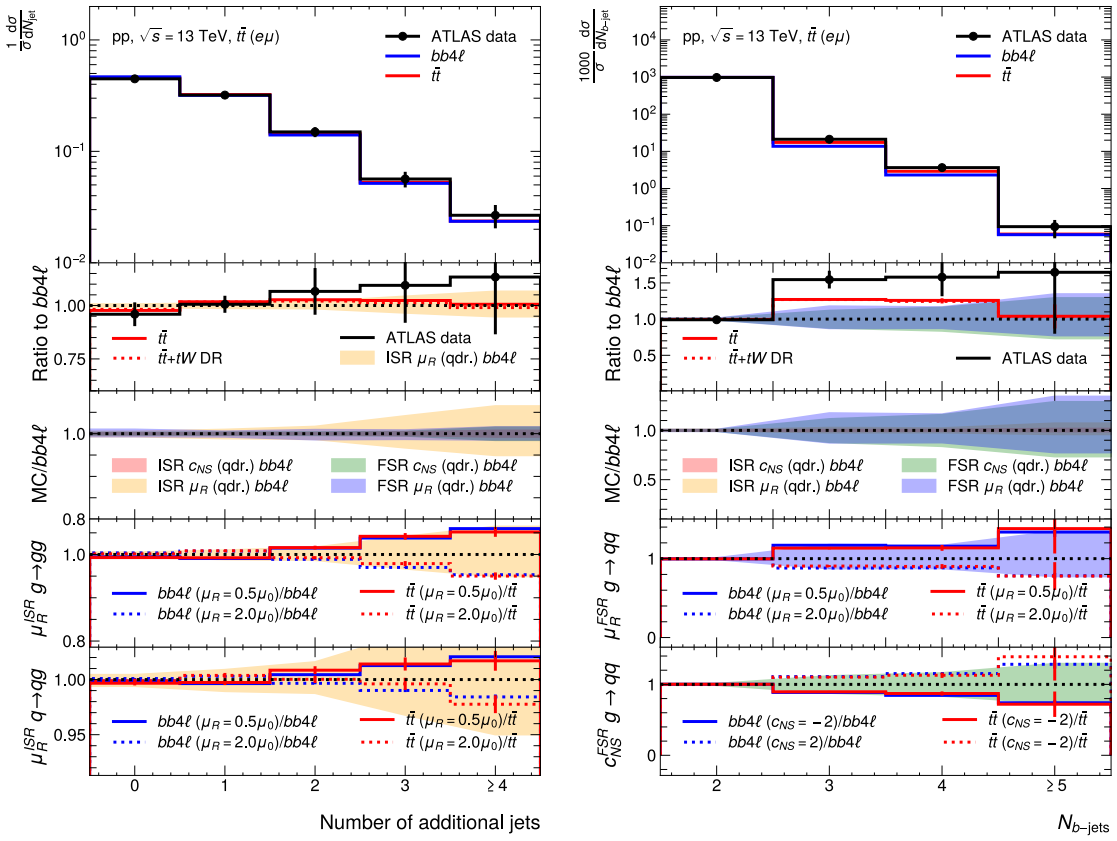
Figure 3.17(a) shows the distribution of the number of additional jets with $p_{\text{T}} > 25\,\text{GeV}$, as predicted by the $bb4\ell$, $t\bar{t}$ and the $t\bar{t} + tW$ DR simulations. The particle-level event selection requires exactly one electron and one muon of opposite electric charge with $p_{\text{T}} > 25\,\text{GeV}$ and $|\eta| < 2.5$, as well as at least two b-jets with $p_{\text{T}} > 25\,\text{GeV}$ and $|\eta| < 2.5$. These MC predictions are compared to unfolded ATLAS $t\bar{t}$ data [128, 129]. The PYTHIA 8 splitting kernel variation uncertainty is evaluated by independently varying μ_R and c_{NS} in each of the ISR and FSR splitting kernels ($g \to gg$, $g \to q\bar{q}$, $q \to qg$, and $Q \to Qg$ with $Q \in t, b$). For the $bb4\ell$ prediction, uncertainty bands are displayed in the second ratio panel, which are build by adding in quadrature the variations obtained from the individual splitting kernels for ISR μ_R, ISR c_{NS}, FSR μ_R and FSR c_{NS}. As expected, the uncertainty associated to the ISR splitting kernel variations is dominant in the N_{jet} distribution and its impact increases with increasing jet multiplicity, while the zero- and one-additional-jet bins, corresponding to events with only two b-jets or one additional light jet, should be largely described by the NLO QCD matrix-element calculation. For the dominant ISR μ_R uncertainty, the variation of the $g \to gg$ splitting kernel provides the largest contribution, followed by the $q \to qg$ variation. These two dominant contributions are shown in the third and fourth ratio panels, respectively. Those splitting kernel variations are also shown for the $t\bar{t}$ prediction in the third and fourth ratio panel, demonstrating excellent agreement in the size of the corresponding uncertainties between the $bb4\ell$ and $t\bar{t}$ MC simulation.

A quite different picture is observed for the b-jet multiplicity distribution. Figure 3.17(b) shows the $bb4\ell$, $t\bar{t} + tW$ DR and $t\bar{t}$ predictions compared to unfolded $t\bar{t}$ ATLAS data [130, 131], as well as the size of the separate $bb4\ell$ splitting kernel variations added in quadrature for ISR and FSR (c_{NS} and μ_R). The particle-level event selection requires exactly one muon and one electron of opposite sign with $p_{\text{T}} > 28\,\text{GeV}$ and $|\eta| < 2.5$, as well as at least two b-jets with $p_{\text{T}} > 25\,\text{GeV}$ and $|\eta| < 2.5$. In contrast to the conclusions drawn from the number

of additional jets in Figure 3.17, the splitting kernel variation uncertainty in the number of additional b-jets i.e. beyond the two b-jets which originate from the $b\bar{b}\ell^+\ell^-\bar{\nu}_\ell\nu_\ell$ final state, is entirely dominated by the FSR splitting kernel variations. Moreover, the size of these variations is considerably larger than for the additional N_{jet} distribution, reaching roughly 20 % for both the FSR c_{NS} and FSR μ_R uncertainty in the $N_{b\text{-jet}} = 4$ bin. By comparison, the ISR μ_R variation amounts to only about 6 % in the highest additional-jet multiplicity bin of $N_{\text{jet}} \geq 4$ with $p_{\text{T}} > 25\,\text{GeV}$. While (ISR) μ_R variations dominate the N_{jet} distribution, both FSR μ_R and FSR c_{NS} variations are equally important for the $N_{b\text{-jet}}$ distribution. The most sizeable contribution to the FSR μ_R and c_{NS} uncertainties arise from the $g \to q\bar{q}$ splitting kernel variations. Indeed, additional b-jets in the event are expected to originate from $g \to b\bar{b}$ parton-shower splittings. The size of these most dominant splitting kernel variations in the b-jet multiplicity distribution are in very good agreement between the $bb4\ell$ and $t\bar{t}$ predictions, as shown in the third and fourth ratio panel. However, the nominal b-jet multiplicity predictions differ by up to 30 %, which is partly covered by the large FSR uncertainty bands. Still, both MC simulations underestimate the production of additional b-jets, even more severely so in the case of the $bb4\ell$ MC simulation. Further studies to investigate this behaviour are currently being conducted.

To study the impact of splitting kernel variations on b-jet properties, in Figure 3.18(a), the b-jet fragmentation function with FSR μ_R splitting kernel variations are displayed for the $bb4\ell$ and $t\bar{t}$ simulation. The total splitting kernel variation uncertainty is dominated by the FSR μ_R contribution, for which significant differences are observed between the $bb4\ell$ and $t\bar{t}$ predictions. For the $g \to gg$, $g \to q\bar{q}$ and $q \to qg$ splitting kernels, the FSR μ_R variation leads to comparable uncertainty estimates in both $bb4\ell$ and $t\bar{t}$. In contrast, for the heavy-quark splitting kernels $Q \to Qg$ with $Q \in \{t, b\}$, the FSR μ_R variation is substantially larger in $t\bar{t}$ than in $bb4\ell$, as shown in Figure 3.18(a). The same pattern is observed for the combined FSR μ_R uncertainty, obtained by adding the FSR μ_R splitting kernel variations in quadrature, since the $Q \to Qg$ contribution dominates the total. This behaviour is physically intuitive: the $Q \to Qg$ splitting directly affects parton emissions off the b quark and therefore the remaining momentum of the B-hadron after FSR. The difference between $t\bar{t}$ and $bb4\ell$ can be explained by the different treatment of top-quark decays: In the $bb4\ell$ MC simulation, the top-quark decay is described with NLO QCD accuracy, and the hardest emission from the decay is generated with the POWHEG method. As a result, the parton shower only affects subsequent $b \to bg$ splittings at a lower scale. In contrast, in the $t\bar{t}$ `hvq` simulation, all emissions from the top-quark decay are generated by the parton shower, such that variations of the $b \to bg$ splitting kernels have a larger impact. Within the POWHEG framework, the real emission is generated via the second term in Equation 2.60, which depends on the POWHEG Sudakov factor, similar to the dependence of the parton showering algorithms on the Sudakov factor.[25] A direct variation of the scale at which α_S is evaluated in the POWHEG Sudakov is currently not implemented. However, a workaround can be employed by using PDF sets with different $\alpha_S(M_Z)$ values as suggested in [109]. This is performed by using the NNPDF3.0 NLO PDF sets with alternative $\alpha_S(M_Z)$ values of $\alpha_S(M_Z) = 0.115$ and $\alpha_S(M_Z) = 0.121$ instead of the nominal $\alpha_S(M_Z) = 0.118$ value. In the third ratio panel of Figure 3.18(a), the POWHEG

[25] This applies only to events generated via the POWHEG method, not to remnant events, as explained in the h_{damp} uncertainty discussion in Section 3.2.



(a) (b)

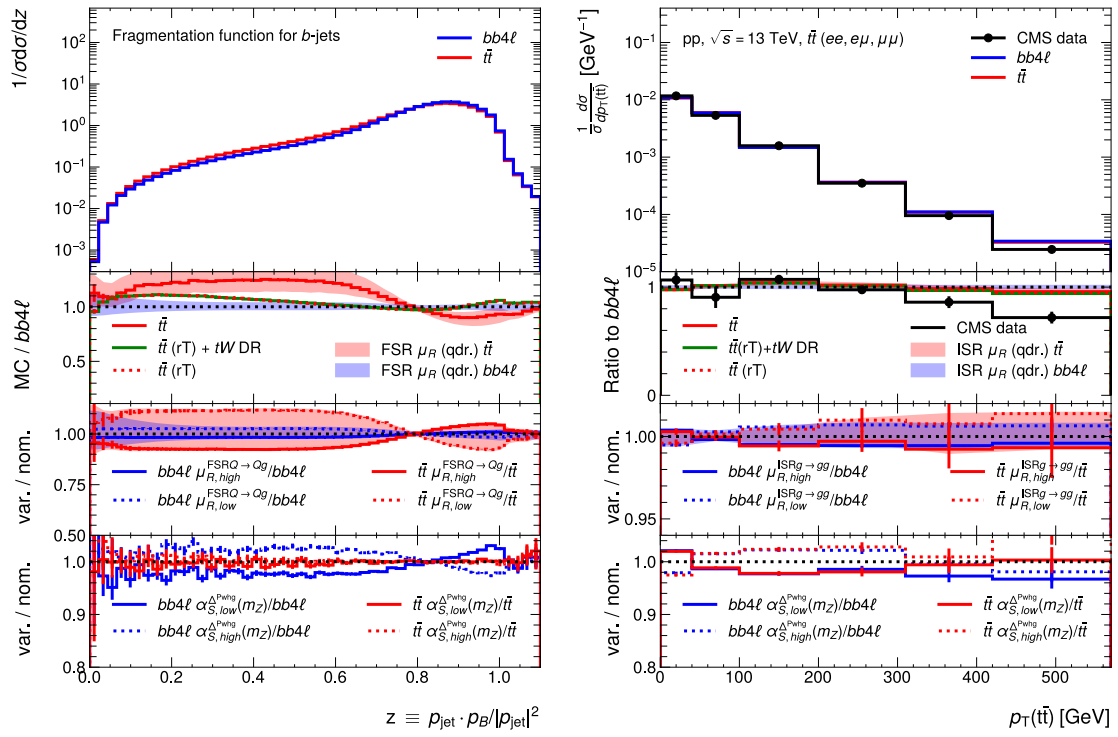
Figure 3.17: (a) Number of additional jets, and (b) number of b-jets, as predicted with the $bb4\ell$, $t\bar{t}+tW$ DR, and $t\bar{t}$ MC simulations, compared to unfolded ATLAS data for the additional-jet multiplicity [128, 129] and the b-jet multiplicity [132, 133]. The ISR and FSR uncertainty bands shown in the ratio panels are obtained by varying the parameters μ_R and c_{NS} independently in the individual splitting kernels ($g \to gg$, $g \to q\bar{q}$, $q \to qg$, and $Q \to Qg$ with $Q \in t, b$) and adding the resulting variations in quadrature. The third and fourth ratio panels show the two dominant splitting-kernel variations for the $bb4\ell$ and $t\bar{t}$ predictions.

Sudakov α_S variation uncertainty is shown for both the $bb4\ell$ and $t\bar{t}$ predictions. To evaluate this uncertainty, the difference between the nominal sample reweighted to an alternative PDF set (with a different $\alpha_S(M_Z)$) and a sample explicitly generated with the same alternative PDF is used. Since PDF reweighting does not affect the POWHEG Sudakov factor, the observed difference reflects the variation in the $\alpha_S(M_Z)$ value used in the Sudakov, while the $\alpha_S(M_Z)$ value in the underlying Born matrix element is the same for both predictions. As expected, the POWHEG Sudakov variation has no effect in the $t\bar{t}$ sample, because the top-quark decay is not part of the matrix-element calculation. However, a clear shape dependence is observed in the $bb4\ell$ prediction. Moreover, the shape of this variation closely follows that induced by the FSR μ_R $Q \to Qg$ splitting-kernel variations, consistent with the discussion above.

Figure 3.18(b) shows the impact of varying the $\alpha_S(M_Z)$ value in the POWHEG Sudakov on the transverse momentum of the reconstructed top-quark pair, for $bb4\ell$, $t\bar{t}+tW$ and $t\bar{t}$ MC predictions. Further, the comparison to unfolded CMS $t\bar{t}$ data [107] is shown, the same measurement as described in Figure 3.3(b). In the low-p_{T} region, the splitting-kernel variation uncertainties are dominated by ISR μ_R variations, although they are generally small. Among these, the $g \to gg$ ISR μ_R variation provides the largest single contribution, which is shown in the third ratio panel, and exhibits very good agreement between the $bb4\ell$ and $t\bar{t}$ predictions. Both the $bb4\ell$ and $t\bar{t}$ distributions are affected by the POWHEG Sudakov $\alpha_S(M_Z)$ variation, which is expected behaviour because the transverse momentum of the $t\bar{t}$ system is sensitive to the hardest POWHEG emission, typically originating from the production process. The emission from the production process is generated both in the $bb4\ell$ and the $t\bar{t}$ `hvq` simulation via the POWHEG method, explaining the good agreement in the POWHEG Sudakov $\alpha_S(M_Z)$ variation of the two. Furthermore, the shape induced by the POWHEG Sudakov $\alpha_S(M_Z)$ variation closely follows that of the ISR μ_R $g \to gg$ splitting-kernel variation, suggesting a similar effect as in the case of the FSR $Q \to Qg$ splitting-kernel variations in $bb4\ell$.

The distributions in Figure 3.18 illustrate the sensitivity of the POWHEG emission generation to variations in the POWHEG Sudakov $\alpha_S(M_Z)$. Currently, these variations are implemented indirectly using PDF sets with alternative $\alpha_S(M_Z)$ values, as described above. A direct implementation of $\alpha_S(M_Z)$ variations in the POWHEG Sudakov form factor would be desirable, as it would allow for a consistent treatment of α_S variations in the POWHEG Sudakov factor and the parton shower. At present, the range of Sudakov α_S variations is limited by the available PDF sets.

Underlying event Several aspects of a parton shower implementation rely on approximate models of non-perturbative phenomena or perturbative QCD phenomena, which in practice cannot be calculated due to the large number of final state particles. The implemented models contain several free parameters, whose values cannot be predicted from first principles, but need to be adjusted, i.e. tuned, to obtain simulated distributions, which reasonably agree with data. In the PYTHIA 8 A14 tune [95], which is widely used in the ATLAS collaboration, ten PYTHIA parameters were tuned, which influence the modelling of multi-parton interactions (MPI) and the modelling of the initial-state and final-state radiation. Thereby, MC samples of dijet, $t\bar{t}$ and Z-boson events simulated at LO with PYTHIA 8, were compared to unfolded data from ATLAS measurements, which were targeting phase space regions sensitive to the



(a) (b)

Figure 3.18: (a) The b-jet fragmentation function and (b) the transverse momentum of the reconstructed top-quark pair, as predicted by the $bb4\ell$ and $t\bar{t}$ simulations, as well as the $t\bar{t}$ (rT) and $t\bar{t}$ (rT)+tW predictions using the recoil-to-top setting in the PYTHIA 8 parton shower. For (b), a comparison to unfolded CMS data [107] is also shown. The FSR μ_R (a) and ISR μ_R (b) uncertainty bands shown in the ratio panels are obtained by varying the individual splitting kernels ($g \to gg$, $g \to q\bar{q}$, $q \to qg$, and $Q \to Qg$ with $Q \in t, b$) and adding the resulting variations in quadrature. In the second ratio panel, the total FSR or ISR μ_R uncertainty is compared to the dominant single splitting-kernel variation, $Q \to Qg$ for (a) and $g \to gg$ for (b), for both, $bb4\ell$ and $t\bar{t}$. The third ratio panel shows the POWHEG Sudakov $\alpha_S(M_Z)$ variation uncertainty for the $bb4\ell$ and $t\bar{t}$ predictions, evaluated by comparing the nominal sample reweighted to an alternative PDF set with a different $\alpha_S(M_Z)$ value to a sample generated explicitly with that PDF. The values $\alpha_S(M_Z) = 0.115$ (low) and 0.121 (high) are used.

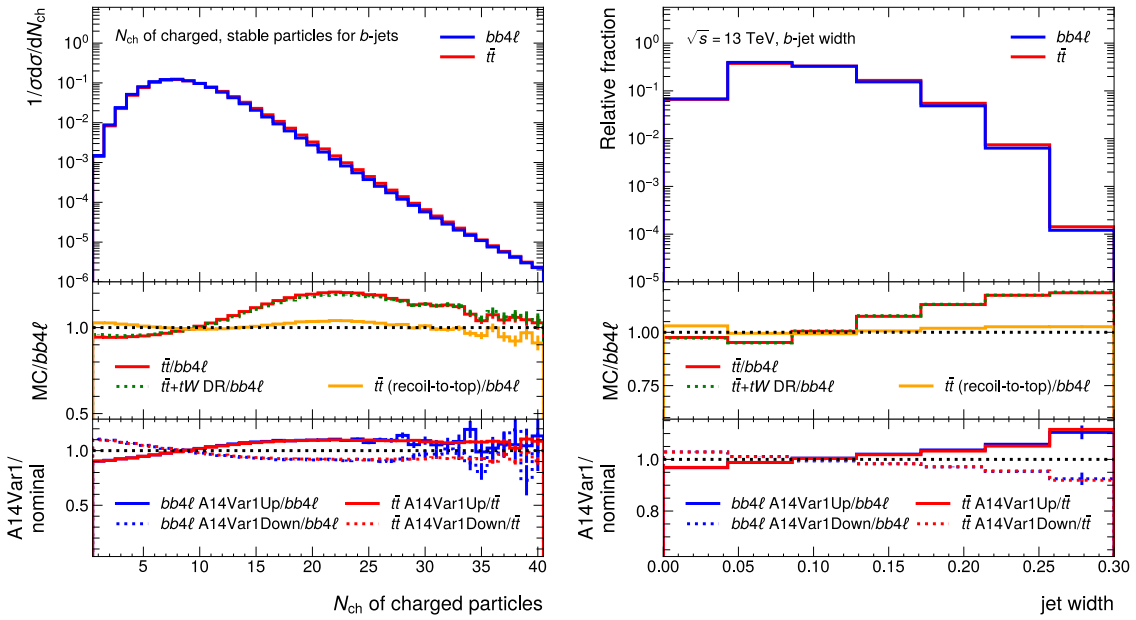
underlying event, the jet structure and the emission of additional jets. The uncertainty in the underlying event modelling is estimated by varying two out of these ten parameters, which were tuned to data with the A14 tune: the $\alpha_S(M_Z)$ value used in the multi-parton interactions and the colour-reconnection parameter `BeamRemnants:reconnectRange`. The nominal values of the considered parameters and the variation ranges as determined in the A14 tuning [95] are listed in Table 3.2.

PYTHIA 8 parameter name	nominal	Var1Up	Var1Down
`BeamRemnants:reconnectRange`	1.71	1.73	1.69
`MultipartonInteractions:alphaSvalue`	0.126	0.131	0.121

Table 3.2: PYTHIA 8 parameter settings for the A14 nominal and Var1 variation [95].

Since these modelling parameters are neither directly connected to the hard process, nor to the matching of the matrix element calculation and the parton shower, very similar results are expected when comparing the A14Var1 variation in the $bb4\ell$ and $t\bar{t}$ predictions. This is indeed shown in Figure 3.19, in which the underlying event uncertainty estimate is shown for both, the $bb4\ell$ and $t\bar{t}$ predictions in the charge multiplicity distribution in b-jets (a) and in the b-jet width distribution (b). The RIVET implementation used for these observables were described in the context of Figure 3.12. The nominal predictions obtained with the $bb4\ell$ and $t\bar{t}$ simulation are in reasonable agreement, when using the same recoil-to-top PYTHIA 8 setting. General jet properties can be influenced by variation of the underlying event modelling due to clustering of particles produced in the underlying event in the reconstructed jet. Increasing the underlying event activity (A14Var1Up) leads to an increase in the number of charged particles per b-jet and an increasing width of the b-jet, since the additional UE particles do not originate from radiation off of the initiating b-parton and thus do not have a preferential direction with respect to the jet cone.

Colour reconnection The parton shower implementations of PYTHIA and HERWIG rely on the leading-colour approximation ($N_c \to \infty$) during the showering of the hard process event, in which each quark is colour-connected to exactly one other parton in the event (gluons have two colour connections, one colour and one anti-colour connection). The string hadronisation used in PYTHIA relies on the notion of a QCD potential between the ends of such a colour dipole, which increases with the distance of the dipole ends and fragments into hadrons. As described in [134], the question underlying the different colour-reconnection models is: which partons lead to confining QCD potentials or, formulated differently, which partons are colour connected? As was realised for example through the comparison of predictions without colour reconnections to unfolded data of the average (charged) particle p_{T} w.r.t. the particle multiplicity [135–140], a description, e.g. of multiple-parton interactions (MPIs), as completely separated processes in colour space does not model experimental data well. Thereby, the colour reconnection of the partons generated during the parton shower occurs after the showering and before the hadronisation. In the first colour-reconnection (CR) model implemented in PYTHIA 8 (CR0 or MPI-based model), gluons from a lower p_{T} MPI interactions can be inserted into colour dipoles of a higher p_{T} MPI interaction, if this colour reconnection



(a) (b)

Figure 3.19: Comparison of the multiplicity of charged particles in a b-jet (a) and of the b-jet width (b) as predicted with the $bb4\ell$, $t\bar{t}$, $t\bar{t}+tW$ simulation and the $t\bar{t}$ simulation using the recoil-to-top Pythia 8 setting. Additionally, the uncertainty estimate in the underlying event modelling is shown in the $bb4\ell$ and $t\bar{t}$ MC sample according to the A14Var1 tune variation.

decreases the total string length λ, which is defined in terms of the invariant masses of the colour dipoles. The model CR2 [141] is a variant of the MPI-based model, in which a gluon originating from any process (not only from the MPI-process) can be completely moved to the colour dipole of any other dipole system, if it decreases the total string length (also called gluon move model). The QCD CR model (CR1) [134] additionally allows the formation of so-called junctions of three quark lines, which increases the probability of baryon formation. To estimate the uncertainty originating from the freedom to choose a particular model for the colour reconnection in Pythia, the three different CR tunes CR0, CR1 and CR2 described in [142] are studied. As detailed in the original publication, additionally to Pythia parameters controlling the colour reconnection itself, also MPI parameters are considered in the tuning due to their strong correlation. The CR tunes were determined by starting from the A14 tune [95], which was obtained with the CR0 model. For the CR1 and CR2 tunes, the CR model was changed accordingly and recommended parameter settings for these CR models applied. In a second step, for each CR model, including CR0, the Pythia parameters considered in [142] were tuned. Even for the CR0 tune, some differences w.r.t. the A14 tuned parameter were found, due to a different tuning parameter choice and due to unfolded data included in the CR0 tune, which was not yet available for the A14 tune. The ATLAS recommendation to evaluate this uncertainty for the $t\bar{t}$ Powheg+Pythia 8 process, which prescribes to compare

the sample with the CR0 tune with the predictions obtained with the CR1 and CR2 tune, is evaluated also for the $bb4\ell$ process.

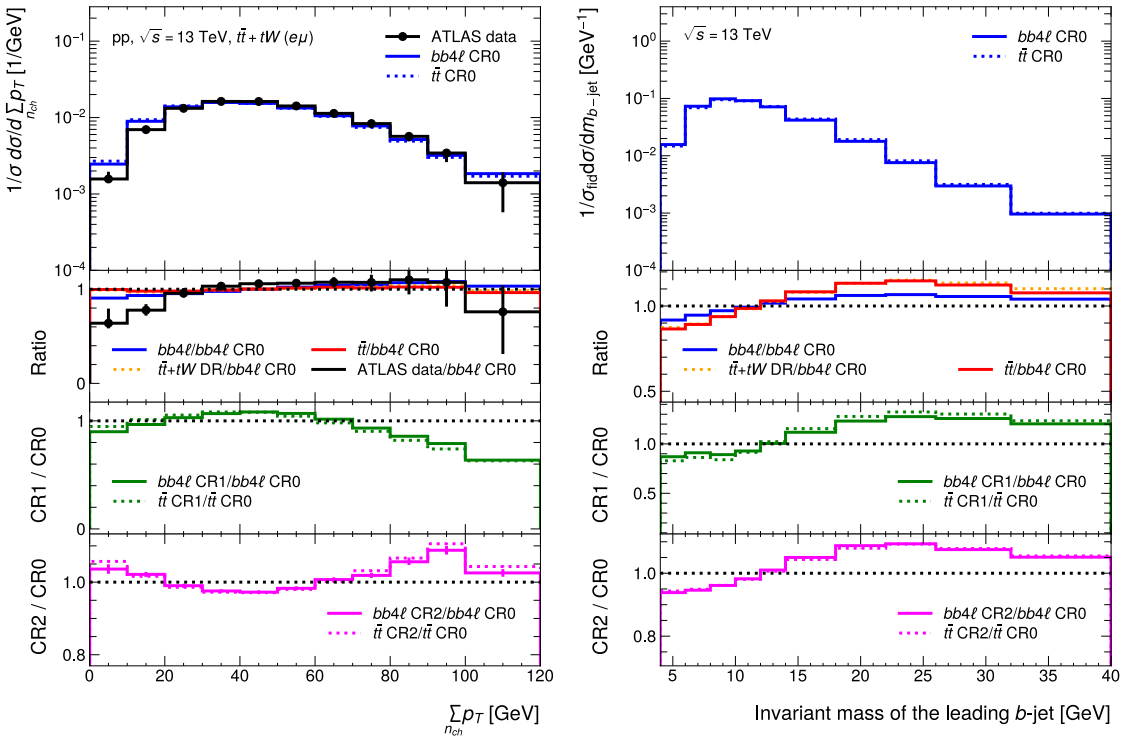
In Figure 3.20(a), predictions from the $t\bar{t}$ and $bb4\ell$ simulations obtained with different CR tunes are compared to unfolded ATLAS $t\bar{t}$ data from [143, 144]. This measurement focuses on observables, which are sensitive to colour reconnection effects. The particle-level event selection requires exactly one muon and one electron of opposite electric charge with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.5$. Selected events must contain two or three jets with $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.5$, of which exactly two are required to be b-jets. Additionally, events with a charged particle multiplicity exceeding 100 are rejected, where charged stable particle objects need to satisfy $p_{\mathrm{T}} > 500\,\mathrm{MeV}$ and $|\eta| < 2.5$. The measurement targets charged-particle activity outside of jets. Consequently, charged particles within a cone of $\Delta R = 0.4$ around a jet or within a cone of $\Delta R = 0.1$ around the leptons are excluded. Using these definitions, the observable shown in Figure 3.20(a) is defined as the scalar p_{T} sum of the charged particles outside of jets. Good agreement[26] is found in the colour-reconnection uncertainty evaluated with the $t\bar{t}$ and $bb4\ell$ simulation, as expected since these variations are largely independent of the hard process and the parton-shower matching procedure. Colour-reconnection modelling also affects jet properties, as illustrated in Figure 3.20(b) for the invariant mass of the leading b-jet. Consequently, this has a direct impact on observables of reconstructed top-quarks that are built from b-jets. In both distributions in Figure 3.20, the CR1 uncertainty results in a larger variation than the CR2 uncertainty, which is plausible given the closer similarity between the CR0 and CR2 models.

3.5 Conclusions

The use of a nominal $bb4\ell$ MC sample in data analyses provides improved theoretical modelling and allows the removal of two systematic uncertainties previously required in analyses based on on-shell $t\bar{t} + tW$ predictions: the top-quark lineshape uncertainty and the $t\bar{t}/tW$ interference uncertainty. These uncertainties were designed to account for limitations of the on-shell predictions and, as shown in Section 3.1, are significant in phase space regions where the on-shell and off-shell $WWbb$ predictions differ, validating their use in current analyses. Replacing the on-shell prediction with the more accurate $bb4\ell$ simulation eliminates the need for these modelling uncertainties entirely.

In Section 3.2, systematic uncertainties related to the hard process were compared. It was found that ME scale variations in $bb4\ell$ lead to larger shape variations in distributions sensitive to b-jet properties, the top-quark lineshape, and interference modelling, resulting in increased ME scale uncertainties in normalised distributions. The effect of the inverse width correction in $bb4\ell$ is well covered by these scale variations and no additional systematic uncertainty

[26] Small differences were observed in the uncertainties evaluated with the CR1 tune in the $bb4\ell$ and $t\bar{t}$ samples. These differences arise due to technical corrections in the implementation of the Pythia QCD CR1 model starting from the Pythia 8.307 version. While these corrections are therefore included in the Pythia 8.312 version used for the $bb4\ell$ CR1 sample, this is not the case for the $t\bar{t}$ CR1 sample, which was generated using Pythia 8.230. It was verified that using a Pythia 8 version before the updates in Pythia 8.307, the $bb4\ell$ and $t\bar{t}$ CR1 uncertainty agree well.



(a) (b)

Figure 3.20: (a) The scalar sum of transverse momenta of charged particles outside of jets including the comparison to unfolded ATLAS data [143] and (b) the invariant mass of the leading b-jet, as predicted by the $bb4\ell$ and $t\bar{t}$ MC simulations, thereby showing the CR uncertainty defined by comparing the CR0 tune predictions to the CR1 and CR2 tune predictions in the second and third ratio panel.

needs to be defined. The PDF uncertainty and the h_{damp} variation uncertainty are consistent between $bb4\ell$ and $t\bar{t}$.

The matching of $bb4\ell$ LHE to the parton shower was discussed in Section 3.3, and it was verified that the dedicated $bb4\ell$ parton-shower matching procedure to either the Pythia 8 or Herwig 7 parton shower has a similar effect in b-jet property distributions, which are most affected by this dedicated $bb4\ell$ matching interface. Furthermore, the matching uncertainty in the top-quark production process and the top-quark decay process was studied. While for the former very good agreement between the uncertainty estimate in $bb4\ell$ and $t\bar{t}$ was found, the latter is only applicable in the case of the $bb4\ell$ simulation and leads to non-negligible shape variations in b-jet observables.

The broad parton-shower uncertainty evaluated by comparing predictions interfaced with either Pythia 8 or Herwig 7 is found to be consistent between $t\bar{t}$ and $bb4\ell$ in Section 3.4. Parton-shower uncertainties, evaluated in Pythia 8 and using the same prescriptions in both, $t\bar{t}$ and $bb4\ell$ simulations, are found to be in excellent agreement overall. This includes the uncertainty related to the underlying event modelling, the colour reconnection uncertainty, the Pythia 8 recoil uncertainty, and the μ_R/c_{NS} variations in the Pythia 8 splitting kernels.

The notable exception is the FSR μ_R heavy quark $Q \to Qg$ splitting kernel variation, for which a significantly reduced uncertainty is observed in $bb4\ell$ compared to $t\bar{t}$. This reduction is attributed to the NLO QCD treatment of the top-quark decay in $bb4\ell$. Varying $\alpha_S(M_Z)$ in the POWHEG Sudakov form factor yields a shape uncertainty in $bb4\ell$ predictions that closely matches the $Q \to Qg$ FSR μ_R variation. In contrast, this variation has a negligible impact on $t\bar{t}$ predictions, where the top-quark decay is not included in the ME calculation.

In conclusion, ATLAS data analyses targeting the $WWbb$ process in the dileptonic final state stand to benefit significantly from the improved modelling of the ATLAS $bb4\ell$ sample presented in this chapter, particularly due to the elimination of the lineshape and $t\bar{t}/tW$ interference uncertainties, which dominate in certain phase-space regions. Although two new uncertainties, related to the top-quark decay matching and $\alpha_S(M_Z)$ variation in the POWHEG Sudakov, are introduced in the systematic uncertainty model for the $bb4\ell$ prediction, they do not lead to a substantial increase in the total modelling uncertainty in the studied observables. Other modelling uncertainties evaluated consistently between $t\bar{t}$ and $bb4\ell$ are generally comparable in size, with the notable exception of the reduced FSR μ_R uncertainty in the $Q \to Qg$ splitting and an increased ME scale-variation uncertainty in some normalised distributions.

In collaboration with the authors of [2], further studies on the modelling uncertainty prescription are being conducted, e.g. the matching uncertainty prescription is planned to be refined, the influence of alternative resonance history projectors is studied and the modelling of $bb4\ell$ events with additional b-jets is scrutinized.

CHAPTER 4

From proton–proton collisions to reconstructed physics objects

4.1 The CERN accelerator complex

Proton–proton collisions at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ were provided by the CERN accelerator complex to the ATLAS experiment in the data taking period of Run 2 (2015–2018), which is relevant for the analysis presented in this thesis. The Large Hadron Collider (LHC), which is part of the CERN accelerator complex, is a circular proton accelerator ring, located 50–175 m underground [145] and crosses the border between France and Switzerland, as depicted in Figure 4.1.

The acceleration history of protons before arriving at the collision points at the four experimental caverns of ALICE, ATLAS, CMS and LHCb starts as hydrogen molecules in a bottle. The hydrogen atoms capture an additional electron to form H^- ions, which are accelerated from 45 keV when exiting the H^- source to 160 MeV in the linear accelerator LINAC4. After the LINAC4, the accelerated H^- ions pass a stripper foil, which removes the two bound electrons of each H^- ion with an efficiency of 99 % [147] and as such produces the desired 160 MeV protons. The proton energy after the LINAC4 is thus still four orders of magnitude smaller than at the collision point in e.g. the ATLAS detector. As shown in Figure 4.1, the protons are accelerated to 2 GeV in the Proton Synchrotron (PS) Booster, then to 26 GeV in the PS and finally to 450 GeV in the Super Proton Synchrotron (SPS) [145]. At this point, the proton beams are injected in the LHC ring and accelerated to the final energy of 6.5 TeV.

The two proton beams are brought to collision at the four experimental caverns, while in the LHC accelerator ring the proton beams run in two separate beam pipes,[1] one beam running clockwise and the other beam running anti-clockwise. The protons are accelerated to energies

[1] The LHC uses *two-in-one* magnets, where both beam pipes are placed within the same magnets and cryostats.

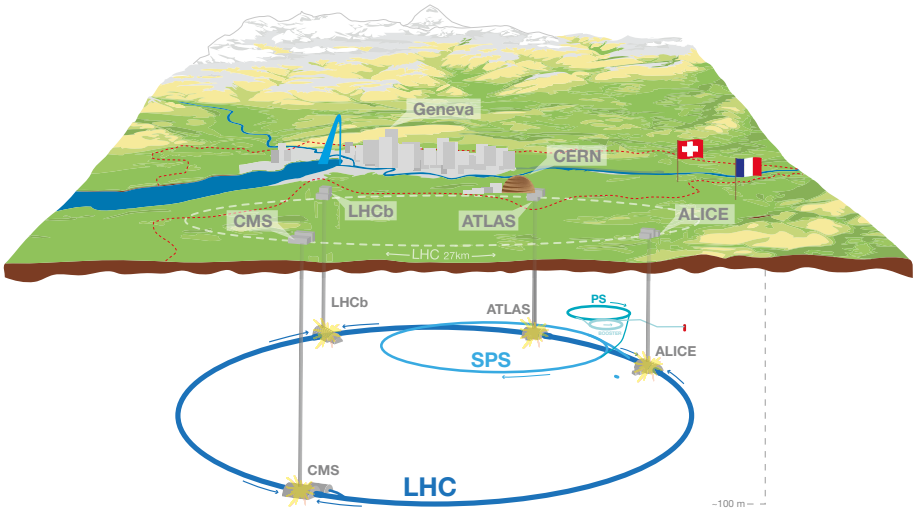


Figure 4.1: Sketch of the CERN accelerator complex, including the position of the four LHC experiments ALICE, ATLAS, LHCb and CMS [146].

of 6.5 TeV by eight radio-frequency cavities per beam, which are installed in straight sections of the LHC accelerator ring. Therefore, the proton beams circulating in the LHC are not continuous, but instead divided into proton bunches with a longitudinal distance between two proton bunches corresponding to a 25 ns time difference.[2] Up to 2808 proton bunches circulate in the LHC ring with 10^{11} protons each [148]. In the LHC collider ring roughly 1200 superconducting dipole magnets keep the protons on a circular trajectory by means of the Lorentz force, while in total around 10 000 magnets are installed, including higher order magnets for focussing of the proton beams and ensuring beam stability. Since the LHC has a fixed circumference of approximately 27 km, increasing the proton energy requires an increase in the magnetic field strength. This ensures that the accelerated protons remain on the trajectory dictated by the geometry of the LHC ring. The required magnetic field strength is given by the Lorentz force, such that

$$q\vec{v} \times \vec{B} = -\frac{m|\vec{v}|^2}{R}\vec{e}_r \Rightarrow |\vec{B}| = \frac{\gamma m_p |\vec{v}|}{qR}, \tag{4.1}$$

where q is the proton charge, $\vec{v}$ its velocity, $\vec{B}$ the magnetic field, R the average bending radius of the LHC ring and m the relativistic mass of the proton $m = \gamma m_p$, with m_p being the proton rest mass and γ the Lorentz factor. The superconducting bending dipole magnets at the LHC have a magnetic field strength B of up to 8.3 T and are operated at a temperature of 1.9 K.

[2] Or to a multiple of a 25 ns time difference, depending on the LHC filling scheme.

4.2 Luminosity and pile-up

After the brief overview of how protons with an energy of 6.5 TeV are produced, in this section the quantities related to the total number of interactions recorded by the ATLAS detector are described: luminosity and pile-up.

Luminosity The ultimate objective of the data analysis presented in this thesis is to minimize the uncertainty in the measured parameters of the top-quark mass and decay width. One contribution to the total uncertainty in these parameters is the statistical component, which can be reduced by increasing the size of the dataset. The amount of data events from a given physics process scales with its total cross-section σ, which is given in units of barn with 1 b corresponding to $10^{-28}\,\mathrm{m}^2$. The number of expected data events N_{WWbb} producing a $WWbb$ final state is given by

$$N_{WWbb} = \mathcal{L}_{\mathrm{int}}\sigma_{WWbb}, \tag{4.2}$$

where $\mathcal{L}_{\mathrm{int}}$ is the integrated luminosity, which has the unit of inverse barn. This integrated luminosity describes the total amount of data collected over time and is related to the instantaneous luminosity $\mathcal{L}$ by

$$\mathcal{L}_{\mathrm{int}} = \int_0^T \mathrm{d}t\mathcal{L}. \tag{4.3}$$

The instantaneous luminosity is calculated from the frequency at which a proton bunch is circulating in the LHC ring f, the number of proton bunches N_b, the number of protons in the two colliding bunches n_1 and n_2, as well as the effective area A which measures the overlap of the two colliding proton bunches:

$$\mathcal{L} = \frac{fN_b n_1 n_2}{A}. \tag{4.4}$$

In Run 2, instantaneous luminosities of up to two times the design luminosity of the LHC ($10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$) were reached. In Figure 4.2, the integrated luminosity of the years 2015–2018, corresponding to the full Run 2 data taking period is shown, amounting in total to a dataset of $140\,\mathrm{fb}^{-1}$. In terms of our process of interest, 115 million $WWbb$ events were produced in Run 2.

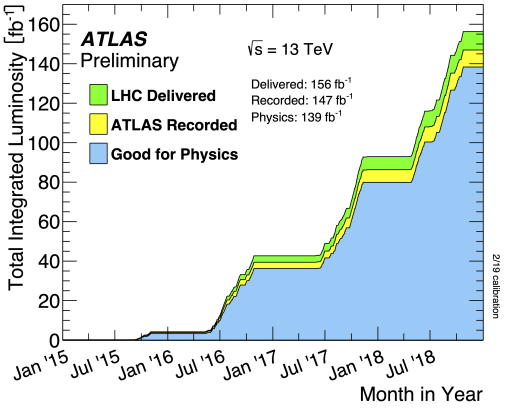


Figure 4.2: Integrated luminosity versus data taking period in Run 2 [149]. For data analysis, only the data labelled as good for physics is used, meaning that all reconstructed physics objects in the data events are of good quality.

Pile-up Since the goal is to collect as many interesting data events[3] as possible, increasing the instantaneous luminosity $\mathcal{L}$ seems to be the way forward. Looking back at Equation 4.4, one possibility would be to increase the number of protons per bunch n_i or decrease the effective area of the bunch crossing A. One of the drawbacks,[4] is the increase of the number of multiple proton-proton interactions during the same bunch crossing with increasing collision intensity, also called *in-time pile-up*.

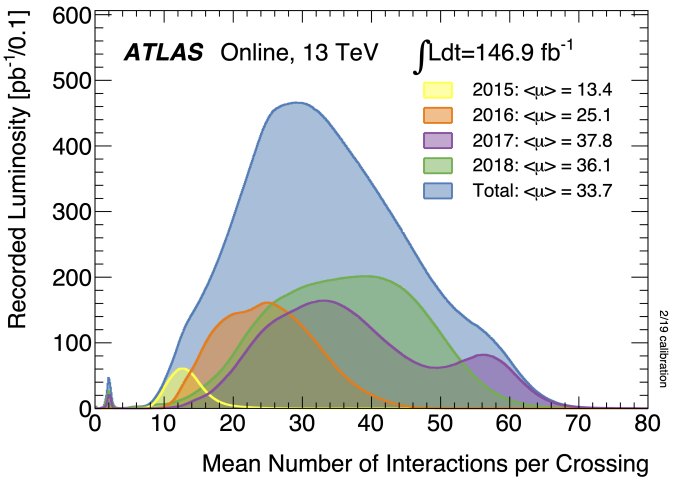


Figure 4.3: Luminosity-weighted distribution of the number of interactions per bunch crossing in Run 2 (2015–2018) per data-taking year and the overall total. The total recorded data by ATLAS during stable beams is shown, thus the integrated luminosity is higher than the good-for-physics dataset of 140 fb^{-1} (compare to Figure 4.2) [149].

The luminosity-weighted distribution of the number of interactions per bunch crossing for Run 2 is shown in Figure 4.3, amounting to an average of 33.7 *pp* collisions per bunch crossing in Run 2. To illustrate the challenges of reconstructing tracks and physics objects under high pile-up conditions, Figure 4.4 displays a data event featuring 25 collision vertices with outgoing tracks. A possibility to increase the integrated luminosity, but not the number of *pp* collisions per bunch crossing, is to increase the number of bunches N_b in Equation 4.4. There are inherent limitations on N_b related to, for example, the pre-acceleration stages and the bunch injections, but also beyond constraints of the machine, increasing the number of bunches, increases the *out-of-time pile-up*. This refers to overlaps of detector signals caused by charged particles produced in two different bunch crossings, due to the finite read-out time of the ATLAS detector, which is slower than the minimum bunch spacing time of 25 ns, combined with long signal shapes in the detector.

[3] What constitutes an interesting event is of course analysis dependent, but often, hard (high-energy) processes are studied in LHC analyses, which have highly energetic particles in their final-states. Typically, the soft (low-energy) QCD processes have the highest production cross-sections, and thus dedicated triggers (see Section 4.3.8) select mostly data events with reconstructed physics objects with high transverse momentum.

[4] Another drawback being the increase of radiation damage to the detector material with increasing integrated luminosity.

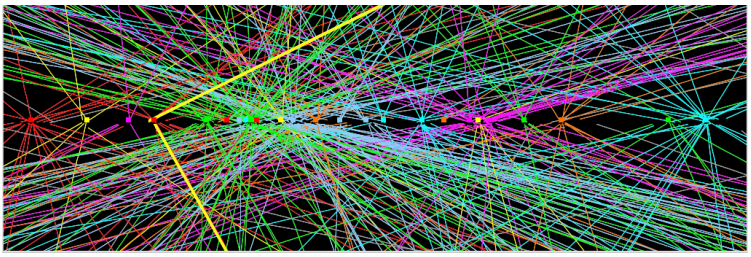


Figure 4.4: Part of an ATLAS event display, showing a $Z \to \mu\mu$ candidate event in 2012 (Run 1) with 25 reconstructed pp collision vertices [150], which is similar to typical pile-up conditions in Run 2 with a mean number of pile-up interactions of 33.7, as shown in Figure 4.3. The horizontal direction corresponds to the direction of the beam-axis.

4.3 The ATLAS detector and low-level object reconstruction

The ATLAS detector is one of the two general-purpose experiments at the LHC. It is therefore designed to detect and accurately measure a broad range of physics objects: light and b-jets, electrons, muons, taus, and photons. A schematic view of the ATLAS detector[5] geometry is shown in Figure 4.5, which follows standard design concepts in high-energy particle physics:

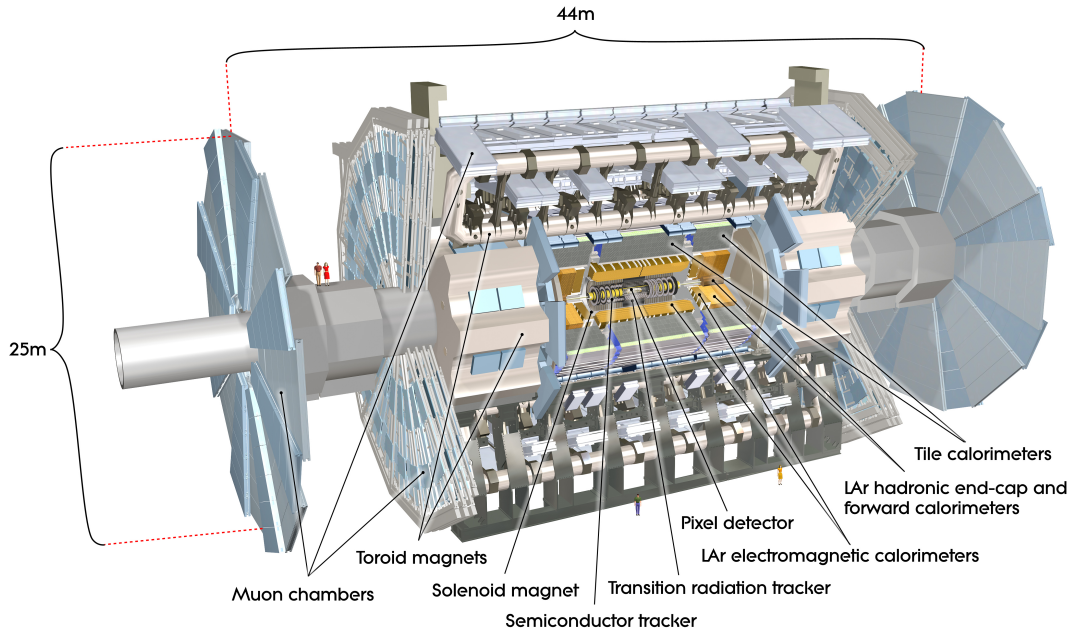


Figure 4.5: Sketch of the ATLAS detector [152].

[5] Since this analysis is based on ATLAS data recorded during the Run 2 data-taking period, the Run 3 detector configuration [151] is not discussed in this section.

The innermost component, closest to the interaction point, is the tracking detector. It records the positions where charged particles intersect the detector material, providing input for the particles' trajectory reconstruction. To perform a momentum measurement of these charged particles, the tracking detector is immersed in a magnetic field. Surrounding the tracker are the electromagnetic and hadronic calorimeters, which measure the energy of electrons and photons, and of particles forming jets, respectively. These physics objects are typically stopped in the calorimeter, i.e. they lose their full kinetic energy through interaction with the detector material and part of this deposited kinetic energy is recorded by the calorimeters. The outermost detector component is the muon spectrometer, which measures the trajectories of muons — particles that typically penetrate beyond the calorimeters. As the innermost tracking detector, also the muon spectrometer is inside a magnetic field. As indicated in Figure 4.5 the ATLAS detector is large: it has a cylindrical geometry, with a length of 44 m and a diameter of 25 m. These dimensions are necessary to achieve high momentum resolution in the tracking detectors (which improves with the track length, but also with the granularity of the tracking detector components) and to fully absorb the kinetic energy of electrons, photons and jets for precise energy measurements. The detector subsystems are arranged cylindrically in the barrel region and concentrically around the beam axis in the end-caps, providing close to 4π coverage. The Cartesian coordinate system in ATLAS [153] has its origin in the nominal pp interaction point, which is defined as the geometric centre of the detector. The z-axis is aligned with the beam direction, the positive x-direction points from the nominal interaction point to the centre of the LHC ring, and the positive y-direction upward. For describing particle kinematics, cylindrical coordinates are often more convenient: r is the radial distance in the x-y plane, ϕ the azimuthal angle from the x-axis around the beam axis and θ the polar angle from the z-axis. Rather than using the polar angle θ, it is common to use the pseudorapidity: $\eta = -\ln(\tan\theta/2)$. For massless particles, this is the same as using the rapidity

$$y = \frac{1}{2}\ln\frac{E + p_z}{E - p_z}, \tag{4.5}$$

with differences in rapidity Δy being Lorentz invariant with respect to boosts along the beam axis (z-axis). For massive, highly relativistic particles with $p \gg m$, approximately $y \approx \eta$. While the LHC collides protons, the actual interactions involve partons (quarks and gluons), which are the proton's constituents. Even though the total proton momentum is known, the momentum fractions carried by the interacting particles are not. However, the transverse momentum p_{T} — the momentum in the x-y-plane — of the colliding partons is approximately zero.[6] Thus, quantities, like Δy, which are invariant under Lorentz-boosts along the beam-axis are useful in physics analyses. Moreover, conservation of transverse momentum allows the inference of undetected particles via the missing transverse momentum, defined as the vector to balance the total $\vec{p}_{\mathrm{T}}$ in the event. As mentioned above, to measure particle momenta and determine particle charges from the curvature of their trajectories, the inner tracking detector and the muon spectrometer are placed within magnetic fields. The ATLAS magnetic field system consists of two components [153]:

[6] Neglecting initial-state radiation, which does provide a source of initial transverse momentum prior to the hard scattering.

- **Central Solenoid Magnet** Magnetic field lines are aligned with the beam axis, this superconducting solenoid provides a 2 T axial magnetic field [153] in the inner detector. Charged particles in this region bend in the r-ϕ plane perpendicular to the beam axis.
- **Toroidal Magnet System** Consisting of one barrel and two end-cap air-core toroids, each with eight superconducting coils, this system produces magnetic fields in the range of 0.5–1 T in the tracking volume of the muon spectrometer [153]. The field is primarily in the azimuthal ϕ direction, causing muons to bend in the r-z plane.

4.3.1 Inner detector

The main task of the Inner Detector (ID) is to measure spatial hit positions (*space points*) along the trajectory of charged particles in the ID material within a pseudorapidity range of $|\eta| < 2.5$. These measurements are the input for track and vertex reconstruction. The ID itself consists of a barrel and two endcap regions. In the barrel region, detector elements surround the beam pipe in form of concentric cylindrical layers, while in the endcap regions the detectors modules are arranged perpendicularly to the beam axis. This geometry is illustrated in Figure 4.6. These detector components are further subdivided into three subsystems, the Pixel Detector [154, 155], the Strip Semiconductor Tracker (SCT) [156] and the Transition Radiation Tracker (TRT) [157], as shown in more detail for the barrel region in Figure 4.7.

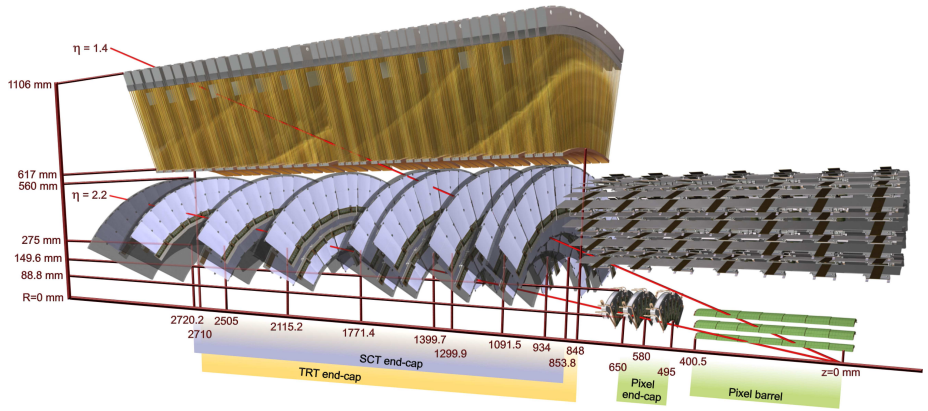
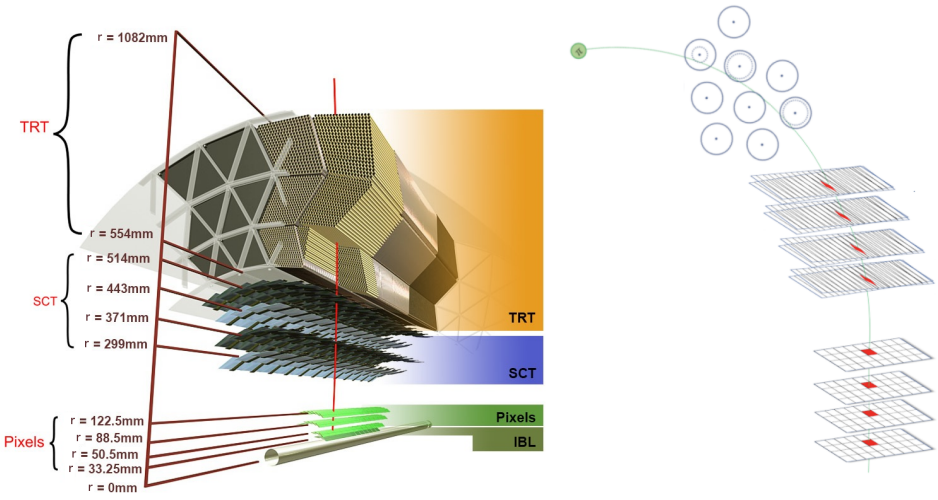


Figure 4.6: Sketch of the ATLAS Inner Detector barrel region and endcap region as operated in Run 1 [153].

Pixel Detector and Strip Semiconductor Tracker (SCT) The two subsystems closest to the beam pipe, the Pixel Detector and the SCT, are based on semiconductor silicon detectors. This detection principle relies on combining pure silicon lattices infused with non-silicon atoms, which either add an additional valence electron (n-type doping) or an additional hole (p-type doping) in the silicon band structure. When combining two such differently doped materials, at the p-n-junction the free electrons and holes recombine and create a zone depleted of free charge carriers. Due to the electron–hole recombination, an electric field is

created by the space charge of the remaining positively and negatively charged ions, which can be enlarged by applying a depletion voltage. This *depletion zone* is the active material of the silicon detector, in which electron–hole pairs are created when a charged particle passes the sensor and drifting charges in the electric field result in a current, which can be measured as signal in the pixel and strip layers. Along a charged-particle trajectory, in total up to eight three-dimensional space points per charged particle are recorded by the pixel detector (four hits, four space points) and the SCT (four times two hits in adjacent layers, four space points). The high granularity[7] of the Pixel Detector provides precise measurements close to the interaction point and therefore strongly influences the impact parameter resolution as well as the reconstruction of primary and secondary vertices. In contrast, since the SCT measurements extend to larger radii, they are particularly important for determining the charged-particle momentum and further constraining the impact parameter. In the pixel detector, three-dimensional hit information is obtained directly from the pixel cell geometry and the known sensor position within the detector. Spatial resolutions of 10 μm in the $R-\phi$ direction and 66.5 μm and 115 μm in the z-direction are achieved for the innermost and three outer pixel layers, respectively [158, 159]. In contrast, a single layer of the SCT strip detector constrains only the $R-\phi$ position to 17 μm [158], as the silicon strips extend parallel to the beam axis and are segmented only once at $z = 0$. This limitation is addressed by using two layers of silicon strips stacked within each strip module, with the layers rotated by a 40 mrad stereo angle relative to the z-axis. The z-position of a charged particle can then be inferred with a resolution of 580 μm [158] by identifying the two strips that registered a signal and locating their intersection point within the module, as illustrated in Figure 4.7(b).



(a) (b)

Figure 4.7: Illustration of the ATLAS Inner Detector barrel region, including the IBL Pixel layer (a), which was inserted during the long shutdown between Run 1 and Run 2 [158] and a sketch (b) of recorded space points in the different detector layers from a charged particle trajectory [160].

[7] Typical pitch-sizes for the pixel sensors in the Insertable B-Layer (IBL) are $50 \times 250\,\mu\text{m}^2$ and for the sensors in the outer layers of the pixel detector $50 \times 400\,\mu\text{m}^2$.

Transition Radiation Tracker (TRT) The TRT [157] forms the outermost subsystem of the ID. Despite its comparatively modest spatial resolution (for example, about 130 μm in the $R-\phi$ direction in the barrel [158]) with respect to the silicon detectors, the TRT plays an important role in the precise determination of charged-particle momenta. This is due to the large number of measurement points it provides — on average about 36 hits per track — combined with its substantial radial extent of approximately 0.5 m, which yields a long lever arm for the curvature measurement (see Figure 4.7(a)). The TRT is composed of drift (straw) tubes of 4 mm diameter, which are filled with a xenon-argon gas mixture [161]. An anode wire is placed in the centre of the tube, while the walls of the straw tube are utilized as cathode. When a charged particle traverses a straw tube, it ionizes atoms in the gas mixture (5–6 ionizations per millimetre [161]) and the free electrons are accelerated towards the central anode. Thereby, an avalanche of secondary ionizations is created, such that a sufficiently large charge for detecting a signal is collected. The TRT detectors in the barrel region only provide an $R-\phi$ measurement, since the straw tubes run parallel to the beam pipe with their wires divided into two halves at approximately $z=0$ [158], while in the end-cap a $z-\phi$ measurement is made with the straw tubes being arranged perpendicularly to the beam axis. Besides the measurement of space points, the TRT contributes to the particle identification. This is achieved by making use of transition radiation, which occurs when a charged particle passes the boundary of two media with different dielectric constants. In contrast to using the ionization process for the detection of a charged particle, which follows the Bethe–Bloch equation and depends on the velocity of the particle, the emission of transition radiation increases with the relativistic γ factor $E/(m_0c^2)$ and can thus provide input for particle identification at high energies [162]. The transition radiation is emitted in dedicated radiator foils, which are interleaved with the straw tubes, in which the transition radiation is detected [163]. This measurement technique is particularly effective for distinguishing electrons from charged pions, as electrons are more likely to emit transition radiation, resulting in a stronger TRT response.

4.3.2 Track reconstruction

The ATLAS ID tracking algorithms [164] reconstruct the trajectories of charged particles with transverse momenta $p_{\mathrm{T}} > 500\,\mathrm{MeV}$ within a pseudorapidity range of $|\eta| < 2.5$, using mainly the measurements provided by the ID subsystem as input. The amount of input data to the track reconstruction can be substantial, reaching up to approximately 15 000 silicon hits per bunch crossing. Pattern recognition algorithms are used to identify track candidates from the hit pattern recorded in the ID. Starting from three track seeds, which are three hits in the pixel detector, hits in the outer layers of the ID are included, which are compatible with these seeds. A range of quality requirements, which are either based on kinematic selection criteria (cut based) or based on a machine-learning algorithm, are applied to reject fake tracks. This reconstruction method is called inside–out track reconstruction. Also, an outside–in procedure is applied in ATLAS, where track seeds are formed with hits in the TRT and compatible hits in the silicon detectors are collected into a track candidate. This procedure should recover tracks, which are not directly produced at the pp collision point, such as electrons produced through photon conversion in the detector material. The TRT track seeds

should be matched to deposits in the electromagnetic calorimeter, which is described later in this section. Once a charged-particle trajectory has been reconstructed, its momentum is determined from the curvature of the track in the transverse plane ($r - \phi$) in the magnetic field in the ID. This curvature can be quantified using the sagitta parameter s, defined as the maximum distance between an arch of a circular trajectory — which corresponds to the reconstructed track trajectory — between two points and a straight line connecting these points. The general expression for the uncertainty in the sagitta measurement σ_s depends on the number of measurement points N and the spatial uncertainty of the hits σ_x [165], which is dictated by the detector geometry and strip or sensor pitch. The uncertainty in the momentum measurement σ_p from the track resolution can then be parameterised as

$$\frac{\sigma_p}{p} = \frac{\sigma_s}{s} \propto \frac{\sigma_x}{BL^2\sqrt{N}} \cdot p, \tag{4.6}$$

where B is the magnetic field and L the path length of the particle in the ID (see e.g. Chapter 8 of [162]). The transverse momentum resolution of the ATLAS tracker is given by [153]

$$\frac{\sigma_{p_\mathrm{T}}}{p_\mathrm{T}} = 0.05\% p_\mathrm{T} \oplus 1\%, \tag{4.7}$$

where the symbol $\oplus$ denotes quadratic summation and the transverse momentum is given in GeV. The constant term arises from multiple scattering in the detector material and dominates the resolution at low transverse momenta.

4.3.3 Primary vertex reconstruction

A vertex is defined as a space point in which a set of reconstructed tracks intersect.[8] This space point, the result of the position estimation given a set of trajectories, is assumed to be the point of origin of these tracks. In general, two types of vertices are distinguished: primary and secondary vertices. The primary vertices are space points at which the pp collisions occur, while secondary vertices are space points at which a particle produced at a primary vertex decays or interacts with detector material at a sufficient distance from the primary vertex to be identified as a separate vertex. While secondary vertex reconstruction was a crucial input to legacy algorithms to identify heavy-flavour jets,[9] this is not the case for the corresponding algorithm used in this thesis. Therefore, secondary vertexing will not be discussed further. The primary vertex reconstruction is an iterative procedure including vertex finding and vertex fitting. Two algorithms for the vertex reconstruction are briefly outlined here, the iterative vertex reconstruction (IVF) [166], used in ATLAS Run 1 and Run 2 analyses, and the adaptive multi-vertex finder and fitter (AMVF) [167], which is the default vertexing algorithm for reprocessed Run 2 and Run 3 data. Still, both algorithms are discussed, since the jet

[8] In three dimensions, the reconstructed tracks typically do not intersect, but instead are close enough (in three dimensions) to form a vertex taking into account uncertainties in the trajectory measurement.

[9] Heavy-flavour jets are defined as jets, which are initiated from heavy-flavour partons, i.e. either a b- or a c-quark. More information is given in Section 4.4.1 and Section 4.4.2.

calibrations for the reprocessed Run 2 data used in this thesis still have some dependence on the legacy IVF vertex finding algorithm.

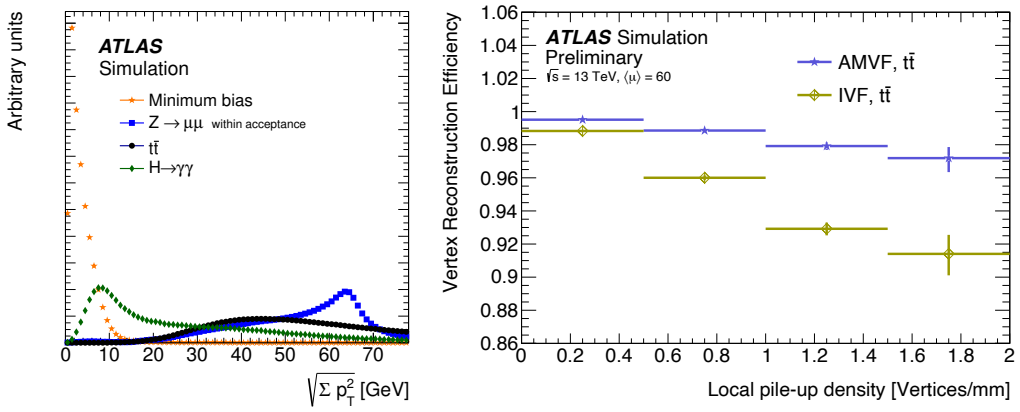
Both algorithms follow a common reconstruction strategy. Primary vertex seeds are identified from regions of high track density using tracks that satisfy quality requirements for vertex reconstruction. Tracks compatible with a given seed, based on their distance of closest approach to the vertex seed, are used to fit a primary vertex candidate, which is improved by using an iterative vertex fitting procedure. This iterative procedure is based on assigning weights to the selected tracks according to their compatibility with the primary vertex candidate and refitting the vertex candidate. The key conceptual difference between the two algorithms is that the IVF assigns tracks exclusively to a single vertex and reconstructs vertices sequentially, whereas the AMVF allows tracks to contribute to multiple vertex candidates and refits all vertices with shared tracks simultaneously. During the vertex fitting procedure in the AMVF, the track weights generally develop such that each track has a non-negligible weight only in a single vertex fit.

A reconstructed primary vertex is required to have at least two associated tracks with significant weights. In the IVF, tracks associated with an accepted vertex are removed from further consideration, while in the AMVF they are excluded only from the seeding stage and may be reused in subsequent fits. The primary vertex of the hard-scatter interaction is typically defined as the vertex with the largest sum of the squared transverse momentum of the associated tracks, $\sum p_{\mathrm{T}}^2$. All other reconstructed primary vertices are assumed to be from minimum-bias events. This criterion is well suited for $t\bar{t}$ events, which dominate the $WWbb$ final state studied in this thesis, as illustrated in Figure 4.8(a). Reconstructed tracks from pile-up primary vertices can also degrade the precision in the longitudinal spatial resolution of hard-scatter primary vertices and reduce their reconstruction efficiency. In Figure 4.8(b), the excellent reconstruction efficiency of hard scatter primary vertices of $t\bar{t}$ events even in the presence of a high local pile-up density is shown, which can be significantly improved by using the AMVF instead of the legacy IVF algorithm.

4.3.4 Calorimeters

Calorimeters measure both the deposited energy and the position of the energy deposits of charged *and* neutral particles. While their spatial resolution is limited, calorimeters provide a high accuracy measurement of the particle energy. The energy measurement is most meaningful, if the total kinetic energy of the incident particle is deposited in the calorimeter, which requires a substantial radial extent. In ATLAS, the calorimeter subsystems extend radially to approximately 3 m [168, 169], roughly three times the size of the ID (compare to Figure 4.7(a)).

The general principle of a calorimeter detector is based on the development of a particle cascade, or shower, resulting from the interactions of primary and secondary particles with the detector material. Charged particles created in this shower development are detected in the active area of the calorimeter via ionization processes. The shower development depends strongly on the type of incident particle. Electron and photons initiate electromagnetic



(a) (b)

Figure 4.8: (a) Simulated distributions of the sum of the transverse momentum squared of reconstructed tracks assigned to the primary vertex of minimum bias and several hard-scatter processes including $t\bar{t}$ production. Minimum bias events represent an unbiased event sample of inelastic collisions dominated by soft QCD processes. The primary vertices were reconstructed using the IVF algorithm [166]. (b) The efficiency to reconstruct the hard-scatter primary vertex in $t\bar{t}$ events as a function of the local pile-up density when using the IVF or the AMVF vertex reconstruction [167].

showers, driven mainly by bremsstrahlung ($e^{\pm} \to e^{\pm}\gamma$) and pair production ($\gamma \to e^{+}e^{-}$). In contrast, hadronic showers, initiated by hadrons, involve a broader range of interactions. These include electromagnetic processes such as ionisation and atomic excitation, but are dominated by hadronic interactions with the nuclei in the detector material — governed by the strong force. Inelastic hadronic interactions produce a variety of secondary particles, including pions and excited nuclei. These nuclei de-excite through various channels, and part of the initial kinetic energy is lost in the form of nuclear binding energy, recoil, and neutrinos, which can not be directly detected. In such hadronic showers, inelastic interactions often produce neutral pions, which decay rapidly ($\sim 10^{-16}$ s) into two photons, each initiating a secondary electromagnetic sub-shower. A defining characteristic of electromagnetic (hadronic) calorimeters is their thickness in units of radiation lengths X_0 (interaction lengths λ). The radiation length defines the distance after which the energy of a charged particle is reduced by $1/e$ through electromagnetic interactions, while the interaction length is the average distance after which a hadronic interaction takes place. The length of a particle cascade increases thereby logarithmically with the particle's energy. To contain a large fraction of the full electromagnetic (hadronic) shower, the thickness of the ATLAS electromagnetic calorimeter (ECAL) corresponds to $> 22X_0$ in the barrel ($> 24X_0$ in the end-caps) and the hadronic calorimeter (HCAL) thickness to 9.7λ in the barrel (10λ in the end-caps) [153]. The values of the radiation X_0 and interaction length λ depend on the detector material and scale with the atomic number Z and the number of nucleons A [162, 170]:

$$X_0 \sim \frac{716.4\,\mathrm{g\,cm^{-2}A}}{\mathrm{Z(Z+1)\ln(287/\sqrt{Z})}} \qquad \text{and} \qquad \lambda \sim 35\,\mathrm{g/cm^2 A^{1/3}}. \tag{4.8}$$

Often, for the same material the hadronic interaction length λ is about an order of magnitude larger than the electromagnetic radiation length X_0, leading to a larger longitudinal and also larger radial extent of the hadronic shower. Consequently, the inner calorimeter layer consists of the ECAL, which is designed to contain the bulk of the electromagnetic shower initiated by electrons or photons, while the HCAL is placed in the outer layer. Hadrons typically leave a signal in the ECAL and are stopped, i.e. deposit their full kinetic energy, only in the HCAL. For this reason, the HCAL will in general have a larger depth compared to the ECAL. In ATLAS the ECAL has a radial extent of 1.1 m [168], while the HCAL extends to nearly two meters [169].[10]

This layout is shown in Figure 4.9 for the barrel, the end-cap and the forward region. In the following, only the barrel region will be described in more detail. Both, the ECAL and the HCAL in ATLAS are *sampling* calorimeters. In this design, the particle shower develops in the passive absorber material (high-Z), with active layers being interspaced between layers of this passive material. In the active layers, a fraction of the ionization signal of the charged particles produced in the particle shower is detected.

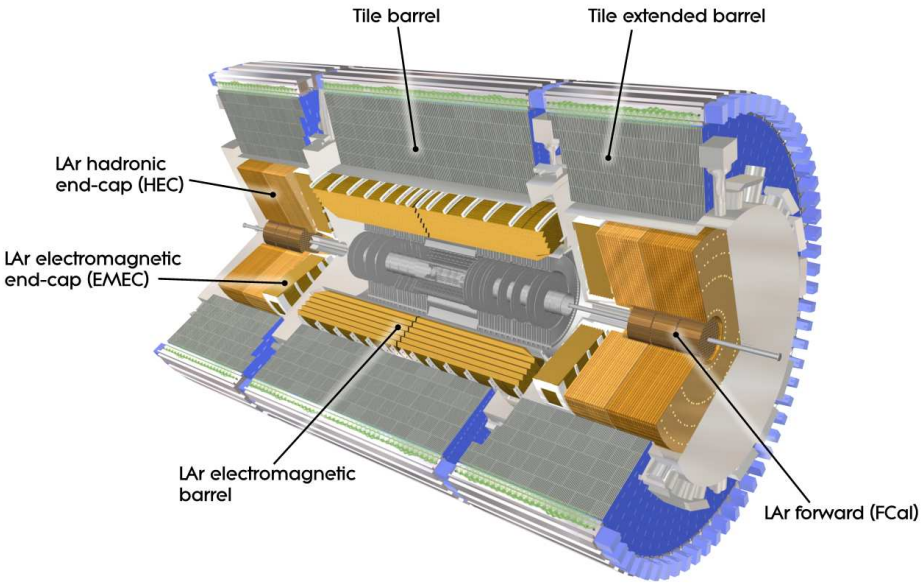


Figure 4.9: Illustration of the ATLAS calorimeter system [153].

Liquid Argon Electromagnetic Calorimeter (LAr ECAL) The ATLAS ECAL uses liquid argon (LAr) in its active layers and lead as absorber material, which is assembled in an accordion shape, as shown for one wedge of the ECAL of the barrel region in Figure 4.10(a). Between two lead absorber plates, accordion shaped electrodes are placed to collect the ionization signal from charged particles passing through the liquid argon. In the central region ($|\eta| < 1.475$), the ECAL consists of two half barrels, which are separated by a small gap

[10] Since not the same detector material is used in the ECAL and in the HCAL as described below, the scaling of the radial extent of the ECAL to HCAL is not as dramatic as an order of magnitude.

at $z = 0$. The calorimeter is segmented into three layers in radial direction, such that also the shower development along the shower axis can be monitored. The first layer has a fine granularity in pseudorapidity and is designed to discriminate between single electromagnetic showers and two close showers from $\pi^0 \to \gamma\gamma$ decays. The second layer collects the bulk of the energy deposited in the electromagnetic shower, while the third layer provides information about the shower tails. To correct for energy losses before the incident particles reach the ECAL, a presampling layer is added for $|\eta| < 1.8$, which is a thin (1 cm) active layer of liquid argon. For larger pseudorapidities ($|\eta| > 1.8$), this presampling layer is not necessary due to the reduced amount of inactive material in front of the ECAL, as the solenoid magnet is placed only in the barrel region [168]. Particles in the transition region of $1.37 < |\eta| < 1.52$ between the barrel ($|\eta| < 1.475$) and the end-cap of the ECAL $1.375 < |\eta| < 3.2$ traverse a large amount of inactive material from 5 to almost 10 radiation lengths X_0 [171]. As a result, the energy reconstruction performance in this region is significantly degraded and electron object candidates in this pseudorapidity range are thus often excluded from data analyses, which is also the case in this thesis.

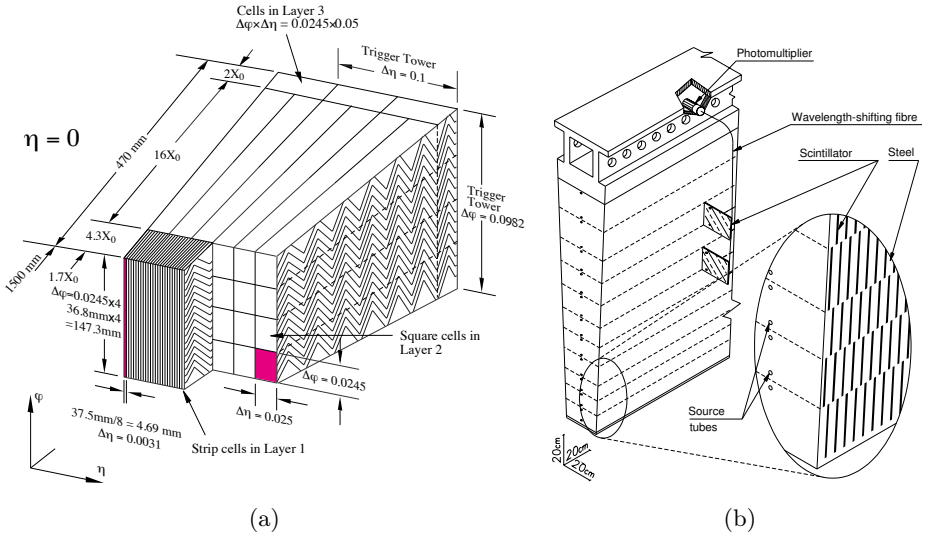


Figure 4.10: Sketch of a wedge of the ATLAS LAr ECAL (a) and of one of the tile HCAL modules (b) [153].

Tile Hadronic Calorimeter (Tile HCAL) The barrel HCAL in ATLAS is constructed of alternating steel absorber tiles and scintillating tiles, which function as active material. In each tile, the scintillation light is collected by wavelength-shifting optical fibres that convert the ultraviolet light into visible light, which is measured by a photomultiplier. This is illustrated in Figure 4.10(b) for one of the 64 HCAL modules in the barrel. Multiple optical fibres are grouped together to form a three-dimensional cell structure, with three divisions in the radial direction as in the ECAL. Compared to the ECAL, the readout granularity of the HCAL is coarser, reflecting the intrinsically broader spatial extent and larger fluctuations in the shower development in hadronic showers.

The ATLAS calorimeters have been designed to achieve an energy resolution of[11]

$$\frac{\sigma_{\mathrm{ECAL}}(E)}{E} = \frac{10\%}{\sqrt{E}} \oplus \frac{0.1\%}{E} \oplus 0.7\% \quad \text{and} \tag{4.9}$$

$$\frac{\sigma_{\mathrm{HCAL}}(E)}{E} = \frac{50\%}{\sqrt{E}} \oplus \frac{1\%}{E} \oplus 3\%, \tag{4.10}$$

with the energy E in GeV [153]. The stochastic term ($\propto 1/\sqrt{E}$) accounts for statistical fluctuations in the shower development. More specifically for sampling calorimeters, this is dominated by the variation of the number of charged particles, which pass the active layers. The noise term ($\propto 1/E$) is due to electronic noise and pile-up contributions. The constant term originates from response non-uniformities of the calorimeter in different detector regions, which can give rise to non-uniformities in the energy calibration and non-linearities in the read-out [172]. As expected, the energy resolution is significantly better in the ECAL compared to the HCAL.

4.3.5 Topological energy cell cluster reconstruction

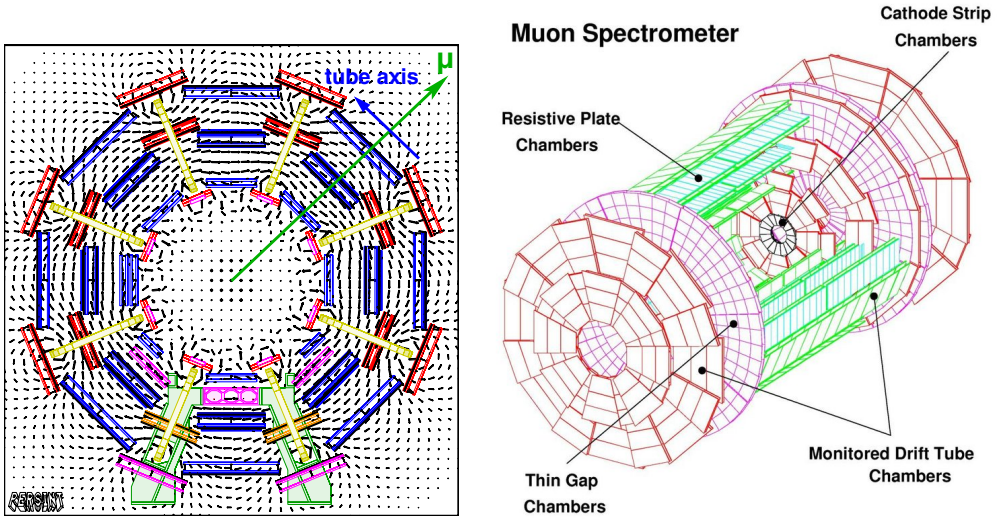
The measured signals from individual, topologically connected calorimeter cells are grouped into three-dimensional energy cell clusters to distinguish significant energy deposits from background contributions such as electronic noise or pile-up. The reconstructed *topo-clusters* are input for physics object reconstruction, such as the reconstruction of jets or electrons. The clustering algorithm begins by identifying seed cells — calorimeter cells with high signal significance, $\xi^{\mathrm{EM}}_{\mathrm{cell}}$, defined as the ratio of the measured cell energy $E^{\mathrm{EM}}_{\mathrm{cell}}$ to the average expected noise per cell from electronics and pile-up, $\sigma^{\mathrm{EM}}_{\mathrm{noise,cell}}$ [173]. Both, the cell energy and noise are evaluated at the electromagnetic (EM) energy scale, which accurately represents the deposited energy of electrons and photons, but does not include compensating terms for the expected fraction of the energy loss of hadrons, which cannot be measured by the calorimeter. Consequently, for the same energy of an incident electron and hadron, $E^{\mathrm{EM},e}_{\mathrm{cell}} > E^{\mathrm{EM,hadron}}_{\mathrm{cell}}$. Seed cells are required to satisfy $|\xi^{\mathrm{EM}}_{\mathrm{cell}}| > 4$. Neighbouring cells are iteratively added if they are geometrically adjacent and exceed a lower significance threshold of $|\xi^{\mathrm{EM}}_{\mathrm{cell}}| > 2$, allowing clusters to span calorimeter layers and subsystems. In a final step, all directly adjacent cells are included in the topo-clusters regardless of their energy. To account for overlapping energy deposits from nearby particles, a cluster-splitting algorithm based on local energy maxima is applied, with energy shared between clusters according to their geometry and energy.

4.3.6 Muon spectrometer

The ATLAS muon spectrometer is designed to achieve two primary objectives: providing precise measurements of space points along the muon trajectory to determine its momentum within the pseudorapidity range of $|\eta| < 2.7$, and enabling muon candidate triggering for $|\eta| < 2.4$. Achieving these goals requires the use of different detector technologies, which will

[11] The notation $\oplus$ implies addition in quadrature.

be briefly outlined. The muon spectrometer is located in the outermost region of the ATLAS detector, after the calorimeter system as depicted in Figure 4.5. The muon spectrometer also contributes to particle identification, because only muons and hadronic showers, which extended beyond the HCAL (*punch-through*), are expected to produce signals in this detector subsystem. The muon momenta are determined by measuring the bending radius in the $r - z$ plane of their tracks in the 0.5–1 T toroidal magnetic field inside the tracking region, which is illustrated for the muon spectrometer in the barrel region in Figure 4.11(a). Two detector technologies are employed for the precision muon p_{T} measurement: the Monitored Drift Tubes (MDT) for $|\eta| < 2.7$ and the Cathode Strip Chambers (CSC) in the innermost layer of the forward region ($2.0 < |\eta| < 2.7$), where high particle fluxes are expected. For triggering on muon candidates, Resistive Plate Chambers (RPC) are used in the barrel region, while Thin Gap Chambers (TGC) are installed in the end-cap region, as depicted in Figure 4.11(b). These different subdetector tasks — precision p_{T} measurement and triggering — are reflected in the parameters of the muon sub-systems listed in Table 4.1.



(a) (b)

Figure 4.11: (a) Magnetic field map in the ATLAS barrel muon spectrometer for $\eta = 0$ in the transverse plane [174]. (b) Schematic view of the Run 2 configuration — relevant for this thesis — of the ATLAS muon spectrometer [175]: the Monitored Drift Tubes (MDT) and Cathode Strip Chambers (CSC) for precision measurements of the muon momentum and the Resistive Plate Chambers (RPC) and Thin Gap Chambers (TGC) for triggering on muon candidates.

All the detector technologies used in the muon spectrometer rely on charged particles ionizing gas molecules and the measurement of the produced ionization signal. The MDTs are drift tubes, similar to those employed in the TRT of the ID, with excellent spatial resolution, but with slow charge collection times. The maximum drift time, from tube wall to anode wire, is 700 ns [153], thus the MDTs do not provide timing resolution. The direction of the tubes is along the ϕ-direction in the barrel and end-caps, meaning tangential to circles around the beam axis and thus only provide a one-dimensional space-point measurement in z-direction, without any spatial resolution in the ϕ-direction. The other muon detector systems — the

Type	Function	Chamber resolution z/R	ϕ	time	Measurements per track barrel	end-cap
MDT	tracking	35 μm (z)	–	–	20	20
CSC	tracking	40 μm (R)	5 mm	7 ns	–	4
RPC	trigger	10 mm (z)	10 mm	1.5 ns	6	–
TGC	trigger	2–6 mm (R)	3–7 mm	4 ns	–	9

Table 4.1: Spatial and time resolution and number of measurements per muon track of the MDT and CSC sub-systems of the muon detector for precision muon p_{T} measurements and the RPC and TGC for triggering on muon candidates [153].

CSC, RPC and TGC — are multiwire proportional chambers or derivates of this technology. In CSCs, the cathodes are segmented into strips orthogonal to the anode wire direction, enabling a two-dimensional space point measurement. This is achieved by detecting the induced charge created by the avalanche from the free electron from ionization, when drifting to the anode wire. The TGCs are also multiwire proportional chambers, but with a very small gap between the wires and the graphite-cathode of 1.4 mm, which is smaller than the wire-to-wire distance of 1.8 mm [153]. A two-dimensional measurement is possible due to pickup strips, which are segmented orthogonally to the anode wires as in the CSCs. A high electric field around the anode wires and the small wire distances allow for good time resolution, with a 99 % probability of collecting the signal in the 25 ns time window between two bunch crossings [153]. In contrast to the other three muon detector technologies discussed so far, the RPCs operate without wires and instead are gaseous parallel electrode-plate detectors. In the gas volume between two resistive plates, which are kept at a distance of 2 mm to each other, charged particles interact and ionize the gas molecules. Due to the electric field of 4.9 kV/mm between the plates, avalanches form from charged particles drifting to the anode. The signal read-out is done via capacitive coupling to metallic strips, located behind the resistive plates. An RPC unit consists of two detector layers with sets of pick-up strips orthogonal to each other, one layer measuring the η and the other layer the ϕ coordinate [153]. In addition to providing data for the trigger system, the TGC and RPC muon trigger chambers also supply the second space-point coordinate to complement the one-dimensional precision space point measurement performed by the MDT.

4.3.7 Muon track reconstruction

In the first step of the muon track reconstruction[12] [164], straight line local track segments are identified by combining measurements from the precision tracking detectors, the MDT and CSC detectors, and nearby muon trigger detectors. There are three MDT stations, providing 6–8 measurements along the muon trajectory in η direction per station, and one CSC station,

[12] In this section, the track reconstruction using only information from the muon spectrometer is described. For the reconstruction of muon objects both, the reconstructed tracks in the muon spectrometer and in the ID, are combined, as will be discussed in Section 4.4.4.

with four space point measurements in both η and ϕ-direction. The track segments, which are reconstructed in the different stations, are combined to form a track candidate, using a parabolic approximation of the track curvature due to the magnetic field and applying a loose constraint based on the interaction point. The measurements of the muon trigger detectors are included to obtain an estimate of the three-dimensional trajectory. A global χ^2 fit is performed to assess the track quality, taking into account interactions with the detector material in the muon spectrometer and misalignment of the muon spectrometer detectors. The final track candidate measurement is improved by removing outlier hits and incorporating hits along the muon trajectory, which are compatible with the track, but which were not included in the track fit previously. The track candidate is then extrapolated to the beam line and refit while accounting for the energy loss in the calorimeter system. The track parameters of the track reconstructed in the muon spectrometer are thereby expressed with respect to the beam line.

Similarly to the case of reconstructed tracks in the ID, a momentum measurement of muons with a reconstructed track in the muon spectrometer is possible due to the bending curve of the charged particle trajectory in the toroidal magnetic field. The contributions to the resolution of the muon transverse momentum in dependence of the transverse momentum for pseudorapidities of $|\eta| < 1.5$ is shown in Figure 4.12. At low transverse momenta, the fluctuations of the energy loss in the ID and the calorimeter system dominate [177]. In the intermediate p_{T} range, the resolution is primarily limited by multiple scattering in the muon spectrometer. This contribution is roughly constant over the whole p_{T} spectrum and is minimized in ATLAS by using an air-core toroidal magnet system, i.e. reducing the traversed material in the muon spectrometer subsystem. At high p_{T}, above approximately 300 GeV, the dominant factor is the single-hit resolution, which is determined by the space point resolution and alignment.

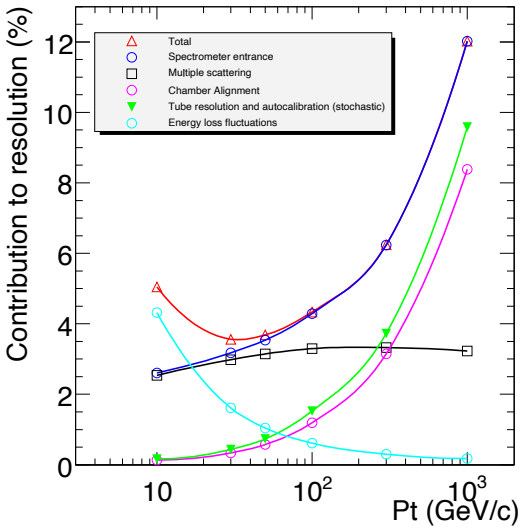


Figure 4.12: Contribution to the transverse momentum resolution for muons reconstructed in the muon spectrometer in % as a function of the muon transverse momentum as predicted by Monte-Carlo simulation [176].

4.3.8 Trigger and data acquisition

In the data taking conditions of Run 2, proton bunch spacings of at least 25 ns led to a peak event rate of 40 MHz, which in turn corresponds to a peak data rate of around 60 TB/s [178]. Besides representing an unmanageable data stream, the storage of all events is not necessary for the physics program of ATLAS. The bulk of the data events are from minimum-bias processes with the corresponding production cross-section being close to the total inelastic pp production cross-section of almost 80 mb at $\sqrt{s} = 13$ TeV [179], which is roughly eight orders of magnitude larger than the production cross-section of $WWbb$ production, which is of interest in this thesis, of roughly 800 pb. Interesting data events from hard processes,

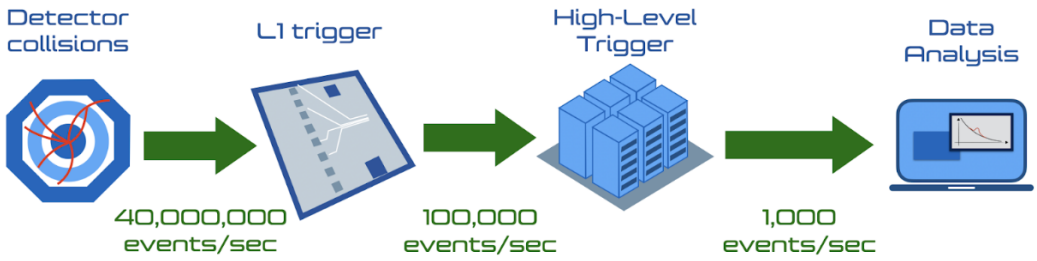


Figure 4.13: Sketch of the event rates in ATLAS before and after the trigger system [181].

which are typically defined by high p_{T} objects, are selected by the ATLAS trigger system [180] and saved by the data acquisition system. The trigger decision, i.e. whether a data event should be stored or not, is taken in a two-step procedure, as illustrated in Figure 4.13. First, the fast Level 1 (L1) trigger preselects interesting events based on coarse information from mainly the muon and calorimeter system and reduces the event rate to approximately 100 kHz [180]. During the determination of the L1 trigger decision, the event information is stored on front-end buffers and then passed on to read-out-drivers outside the detector if the L1 trigger accepts the event. Due to limited buffer resources and the high incoming data stream, the L1 trigger decision needs to be fast, which is achieved by the L1 trigger being designed as hardware trigger, making a trigger decision within 2.5 μs possible [153]. The final trigger selection is performed by the High Level Trigger (HLT), which is a software based trigger using the full detector information and focusing on regions-of-interest, which are identified by the L1 trigger and are regions in $\eta - \phi$-space with e.g. high transverse energy deposits in the calorimeter or contain a high-p_{T} muon candidate. After the HLT trigger decision, which takes roughly 200 ms [180], an event rate of around 1–2 kHz is saved, corresponding to 1 GB/s of data, which is moved from the read-out buffer system to permanent storage [178].

4.4 Physics object reconstruction

In the data analysis presented in this thesis, the final state of interest is $WWbb$, with both W-bosons decaying leptonically in the $e\mu$ channel. Consequently, the relevant reconstructed physics objects are b-tagged jets, electrons and muons. The reconstruction and identification of these objects are described in the following sections.

4.4.1 Jets

A parton — either a quark q or a gluon g — originating from the hard process initiates a sequence of parton emissions, such as $g \to q\bar{q}$ or $q \to qg$ splittings. The emission probabilities are enhanced in the soft and collinear region of phase space. Emissions are referred to as soft, when the energy fraction of the emitted parton approaches zero, and as collinear when the angle between the emitted and the emitting partons becomes vanishingly small. As a result, partons are mostly produced within a narrow cone around the direction of the initial parton.

After hadronisation, this gives rise to a collimated bundle of hadrons, which is referred to as a jet.

Anti-k_t clustering algorithm Different algorithms are available for clustering jets from their constituents. These jet constituents can be either truth-level hadrons in pure MC simulation studies or reconstructed quantities at detector level. In this analysis, reconstructed jet constituents are topological energy cell clusters and reconstructed tracks. Due to the property of soft and collinear enhancement of parton showers, jet reconstruction algorithms should be insensitive to individual soft and collinear splittings. Sequential jet clustering algorithms satisfy this requirement of infrared and collinear (IRC) safety. In these algorithms, a distance metric is defined between jet constituents w.r.t. each other (d_{ij}) and w.r.t. the beam axis (d_i). These distances are expressed in terms of the transverse momentum and the angular distance $\Delta R_{ij}^2 = (y_i - y_j)^2 + (\phi_i - \phi_j)^2$, where i and j refer to two constituents:

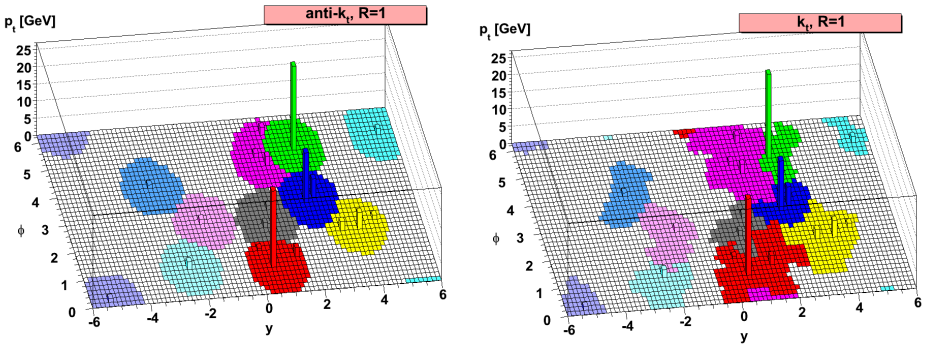
$$d_{ij} = \min\left(p_{\mathrm{T},i}^{2p}, p_{\mathrm{T},j}^{2p}\right) \frac{\Delta R_{ij}^2}{R^2} \quad \text{and} \quad d_i = p_{\mathrm{T},i}^{2p}. \tag{4.11}$$

Here, R denotes the jet radius, which is set to $R = 0.4$ in this analysis. A larger jet radius leads to wider jet cones and increases the number of jet constituents originating from pile up and the underlying event. For small values of R, hadrons originating from radiation off the initiating parton may not be clustered into the reconstructed jet, leading to a reduced reconstructed jet energy compared to that of the initial parton. The parameter p in Equation 4.11 determines the type of sequential jet clustering algorithm. In the following the case $p = -1$, corresponding to the anti-k_t algorithm [99] used in this analysis, is briefly compared to the case $p = 1$, which defines the k_t algorithm. The jet clustering proceeds as follows:

- For all jet constituents, the distance metrics d_i and d_{ij} are calculated and the minimum distance metric $d = \min(d_i, d_{ij})$ is identified.
  - If $d = d_{ij}$, the constituents i and j are removed from the list of constituents and recombined into a new constituent k with $p_k = p_i + p_j$.
  - If $d = d_i$, the constituent i is declared a jet and removed from the list of constituents.[13]
- These steps are repeated until no constituents remain in the list.

The choice of the p-parameter has a significant impact on the shapes of the resulting jets. For the anti-k_t algorithm ($p = -1$), the jet clustering starts around the highest-p_{T} jet constituent. When adding the four-momenta of softer constituents, the jet direction will in general not significantly change during the clustering procedure. In contrast, the k_t algorithm ($p = 1$) is conceptually closer to an approximation of an inverse parton shower clustering, as the softest and most collinear emissions are clustered first. However, it leads to very complex jet area shapes as shown in Figure 4.14, which are impractical for experimental analyses.

[13] Looking at Equation 4.11, this implies that for the constituent i the angular distance to any other constituent j from the list satisfies $\Delta R_{ij} > R$, since otherwise $d_{ij} < d_i$.



(a) (b)

Figure 4.14: Shapes of partonic jets clustered by the anti-k_t (a) and by the k_t algorithm (b) using in both cases the same input of hard parton level objects and infinitesimally soft *ghosts*, which have vanishing four momenta and are distributed randomly in ϕ and y. The *ghosts* are used to visualise the $\phi - y$ area of the clustered jet [99], since they are assigned to a particular jet during the jet clustering algorithm, but do not change the kinematics of this jet.

Particle Flow Jets The particle flow (PFlow) algorithm [182] combines topological energy cell clusters (topo-clusters) and ID tracks to form jet constituents. It takes advantage of the complementarity of the ID and the calorimeter, which provide precise measurements of low-p_{T} tracks of charged particles and of high-energy deposits of charged and neutral particles, respectively. To avoid double-counting of energy deposits, topo-clusters associated with charged tracks are removed. In addition, charged jet constituents originating from pile-up interactions can be identified by their trajectory's distance of closest approach to the hard-scatter primary vertex.

The PFlow algorithm is based on the following steps, which are also illustrated in Figure 4.15:

- *Track selection:* All ID tracks passing stringent requirements on quality criteria, such as for example a minimum number of hits in the ID, are selected if they satisfy $0.5\,\mathrm{GeV} < p_{\mathrm{T}} < 40\,\mathrm{GeV}$ and $|\eta| < 2.5$. At this point, no requirement is imposed on the association of the tracks to the primary hard-scatter vertex. This implies that also pile-up tracks are included.
- *Matching ID tracks and topo-clusters:*
  - *Single topo-cluster hypothesis:* First, reconstructed tracks are extrapolated to the calorimeter layer and matched with an angular-distance criterion to the closest single topo-cluster. As initial requirement, the energy deposited in the topo-cluster needs to correspond to at least 10 % of the track momentum.
  - *Multiple topo-cluster hypothesis:* For hadronic jets it is very common that the energy is deposited in two separate topo-cluster.[14] In case, the energy in the

[14] As evaluated with a dijet MC sample, only for 60 % (25 %) of particles with a true transverse momentum in the range of $1\,\mathrm{GeV} < p_{\mathrm{T}} < 2\,\mathrm{GeV}$ ($5\,\mathrm{GeV} < p_{\mathrm{T}} < 10\,\mathrm{GeV}$) the shower is contained in a single topo-cluster. A very common case is that the energy is deposited in two topo-clusters, e.g. for particles with $p_{\mathrm{T}} < 2\,\mathrm{GeV}$

initially matched topo-cluster is significantly lower than expected from the track momentum, additional topo-clusters within a cone of $\Delta R = 0.2$ are considered in the track-to-topo-cluster matching.

- *Subtracting matched cells from topo-clusters:* To remove from the clustering algorithm input the energy deposit in the calorimeter, which is expected from a selected charged track, the subtraction algorithm starts from the cell in the matched topo-cluster that contains the shower maximum. Following a parameterized shower shape, energy is sequentially removed from the surrounding cells in the topo-cluster, until the expected energy deposition of the charged track in the calorimeter is fully subtracted.
- *Remnant removal:* After the subtraction procedure, any remaining energy cells in the affected topo-clusters are removed from the PFlow clustering input if they are consistent with expected fluctuations of the deposited shower energy.
- *PFlow jet clustering:* The PFlow jets are clustered with the anti-k_t algorithm using as jet constituent input the remaining topo-clusters and the ID tracks, which are matched to the primary hard scatter vertex.[15] Since topo-clusters associated with pile-up tracks have been removed in the previous steps, charged pile-up contributions are largely suppressed in the reconstructed PFlow jets.

4.4.2 Heavy-flavour jets

The long lifetimes of heavy-flavour hadrons and particularly of hadrons containing a valence b-quark, so called B-hadrons, ($\tau \approx 1.5\,\text{ps}$ [185–187]) lead to characteristic secondary decay vertices, displaced from the primary vertex by the order of a few millimetres to even centimetres.[16] Therefore, it is possible to distinguish heavy-flavour jets from jets which are initiated by light partons.[17] For the heavy-flavour jet tagging task a machine learning algorithm, a graph neural network (GNN) is used. The main objective of the GNN is the jet-flavour identification. For each jet, three outputs are given, which describe how likely a given jet is either a light jet (p_l), a c-jet (p_c) or a b-jet (p_b). A fourth output in the jet-flavour identification task predicts if the reconstructed jet object is coming from a hadronically decaying tau lepton (p_τ). Additionally, two auxiliary tasks are implemented, which predict the origin of a given track[18] and whether two tracks originate from a common vertex. The inclusion of these

this is true for 95 % of the hadronic showers [182].

[15] More specifically, the selection is performed by requiring $|z_0 \sin\theta| < 2\,\text{mm}$. The track selection for modern b-tagging algorithms in ATLAS is looser with a cut on the longitudinal impact parameter of $|z_0 \sin\theta| < 5\,\text{mm}$ [183]. A large subtraction of energy deposited by B-hadrons in the jet by the PFlow algorithm is not necessarily expected though, since the z_0 displacement is highly peaked at zero also for tracks from B-hadrons, see e.g. [184].

[16] This corresponds to the mean flight length $\langle l \rangle = \beta\gamma c\tau$ for B-hadrons with transverse momenta in the range of tens to hundreds of GeV [188].

[17] In Appendix C, a study performed targeting explicitly b-tagging at high p_T is presented.

[18] The track truth origin categories are tracks from pile-up interactions and from the primary hard-scatter vertex, as well as tracks from a hadronically decaying tau, from a B-hadron, from D-hadron (hadron containing a valence c-quark) decay and from D-hadron decay, which itself originates from a B-hadron decay. The remaining tracks are grouped into the *other secondary* class [188].

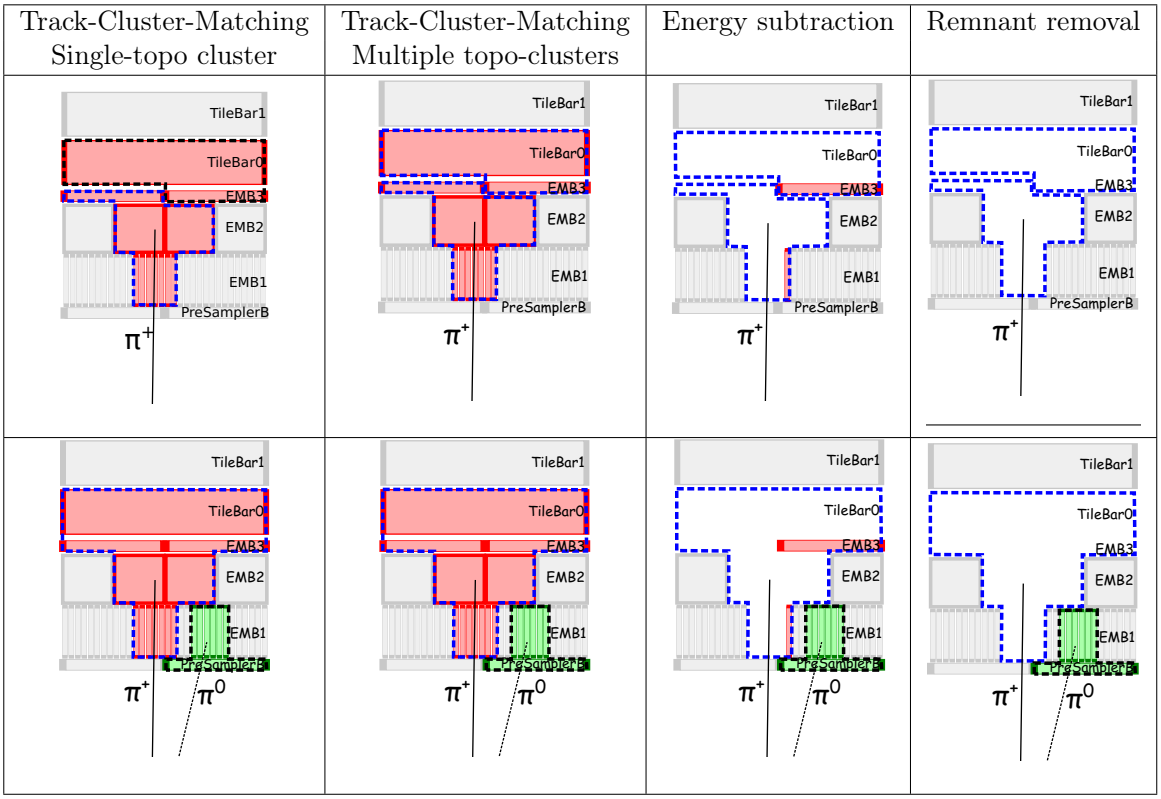


Figure 4.15: Visualisation of the particle flow algorithm for two examples: (top) a single incident charged particle (π^+), which produces one track in the ID and two topo-clusters in the calorimeter; (bottom) two nearby incident particles — a charged particle (π^+) and a neutral particle (π^0) — resulting in one ID track and two separate topo-clusters [182]. In this illustration, topo-clusters with blue coloured edges are matched to an ID track. After the particle flow algorithm the jet constituent inputs to the jet clustering algorithm are reconstructed tracks for charged particles and topo-clusters for neutral particles.

auxiliary training objectives was found to improve the flavour-tagging performance [188]. The input to the GNN are parameters of the tracks associated to a given jet, like the tracks' perigee representation and the associated uncertainties from the track fitting procedure and the number of hits in the ID silicon layers, as well as the jet p_{T} and jet η. After preprocessing the track input parameters, each track is represented by a node in a fully connected GNN. The track representation is updated through a sequence of fully connected neural network layers, which use as input the node features (track representation) and the combination of the features of two nodes. The final jet representation used in the jet flavour classification is obtained by calculating a weighted sum of the transformed track representations, where also the corresponding weights are learned during the training. A final graph classification neural network provides the (p_τ, p_l, p_c, p_b)-quantities as output. These outputs are combined into a b-tagging discriminant D_b by [188]

$$D_b = \log \frac{p_b}{(1 - f_c - f_\tau)p_l + f_c p_c + f_\tau p_\tau} \quad \text{with} \quad f_c = 0.2, f_\tau = 0.01. \tag{4.12}$$

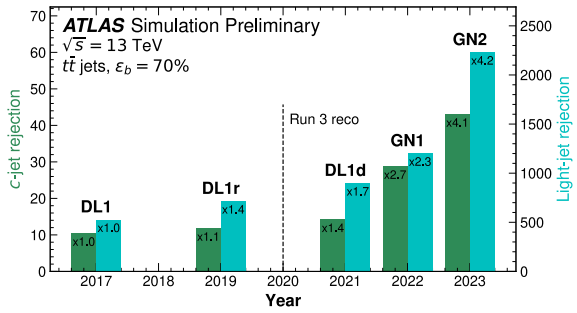


Figure 4.16: Improvement of the light and c-jet rejection from the DL1 flavour tagging algorithms to the GN2 tagger evaluated on $t\bar{t}$ events for the 70 % b-jet efficiency working point [189].

Thereby, the f_c value is optimised to obtain a balance between light and c-jet rejection, while the f_τ value is determined by maximising the tau-rejection, but having negligible impact on the light and c-jet rejection. Depending on the b-jet efficiency used in the analysis, different cuts on D_b are applied. In the analysis presented in this thesis, b-jets are required to pass the tight 65 % b-jet efficiency working point of the GN2 b-tagging algorithm. A higher b-jet efficiency comes at the price of a lower background rejection. Using the GN2 flavour-tagging algorithm, the background rejection for the same b-jet efficiency is greatly reduced with respect to the previous algorithms in ATLAS, as shown in Figure 4.16. Thus, compared to ATLAS analyses using the previous algorithms, the background events entering the event selection due to light jets mistagged as b-jets is strongly reduced.

4.4.3 Electrons

Electron reconstruction [190] is based on matching ID tracks to corresponding energy deposits in the calorimeter, forming the characteristic signature of an electron or positron in the ATLAS detector. Tracks in the ID are first reconstructed using the standard ID track reconstruction (see Section 4.3.2). If ID tracks can be loosely associated to a topo-cluster (see Section 4.3.5) by angular matching, these tracks are refit to account for energy loss through bremsstrahlung. Further, track seeds which are compatible with an energy deposit in the calorimeter characteristic for an electromagnetic shower, but did not lead to a reconstructed track, are refit by allowing for 30 % energy loss when passing material layers. Energy deposits from electrons in the calorimeter are reconstructed by first selecting topo-clusters, which deposited at least 50 % of their energy in the ECAL, since electromagnetic showers are expected to be almost fully contained in the ECAL.[19] Out of the selected topo-clusters, superclusters are formed out of multiple reconstructed topo-clusters using only energy contributions from the ECAL. Starting from the topo-cluster with the highest energy deposit in the ECAL, additional topo-clusters, called satellite clusters in this context, are added to the original seed cluster, if they fall into a window of $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$ of the barycentre of the seed cluster, as shown in Figure 4.17. This is done to recover the energy of the primary

[19] Further, the energy deposit in the ECAL needs to be above 400 MeV.

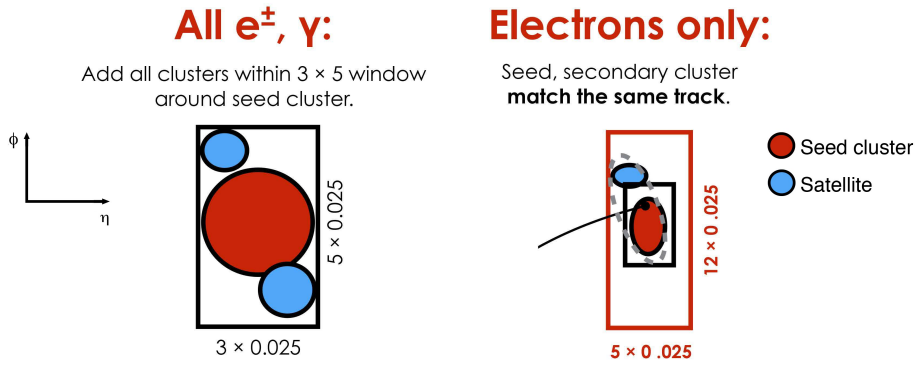


Figure 4.17: Sketch of the superclustering algorithm [190].

electron in case secondary particles produced in the electromagnetic shower formed a separate topo-cluster. Additional satellite clusters are included in the supercluster using a larger window of $\Delta\eta \times \Delta\phi = 0.125 \times 0.300$, if the best matched ID track of the satellite cluster corresponds to the best matched ID track of the seed cluster. The size of the allowed angular window around the seed cluster is a compromise between collecting the full energy of the primary electron and reducing sensitivity to pile-up. Since the magnetic field in the ID is parallel to the beam pipe, the entry point in the ECAL of the primary electron, photons originating from bremsstrahlung from this electron, and electron-positron pairs from photon conversions, are spread predominantly in ϕ. Consequently, the resulting electromagnetic shower tends to be also extended in ϕ- rather than in η-direction. Thus, a wider angular window in ϕ-direction for the super-cluster formation is chosen. All selected topo-clusters are tested as seed-clusters in descending E_{T} order, if they were not assigned as satellite clusters to another super-cluster previously.

Electron identification Prompt electrons are distinguished from hadronic jets, converted photons, as well as from electrons which originate from the decay of a heavy-flavour hadron by the electron identification [190]. For the electron identification task typically 13 observables are utilized, which describe the track of the electron in the ID, the shape of the electromagnetic shower in the ECAL, as well as particle identification information from transition radiation in the TRT. An electron is identified through an $|\eta|$- and E_{T}-dependent cut on the discriminant d_L, which is defined in terms of the likelihoods for a reconstructed electron object to originate from a prompt electron L_S or from other sources L_B

$$d_L = \ln \frac{L_S}{L_B} \quad \text{with} \quad L_{S/B}(\vec{x}) = \prod_{i=1}^{n} P_{S/B,i}(x_i). \tag{4.13}$$

The likelihoods $L_{S/B}$ are constructed with probability density functions (p.d.f.s) $P_{S/B,i}$, which are derived from smoothed histograms of the n discriminating variables x_i of prompt electrons (signal S) or other, background (B) objects. These histograms are obtained from data events. For signal objects, a tag-and-probe method is used in $Z \to ee$ and $J/\psi \to ee$ data events from

2015 and 2016 for $E_{\mathrm{T}} > 15\,\mathrm{GeV}$ and $E_{\mathrm{T}} < 15\,\mathrm{GeV}$, respectively. In this tag-and-probe method, one electron (the tag electron) needs to satisfy a previously used tight electron identification criterion [191], while the properties of the other (probe) electron are used to define the pdfs $P_{S,i}$. Very loose cuts on the previously used electron identification discriminant are used on the probe electron in order to reduce the background contamination of the sample, but not to distort the shape of the resulting signal pdfs. Background pdfs are obtained from data events with at least one reconstructed electron, which are mostly dijet events. Due to a known mismodelling of calorimeter shower shapes in the GEANT4 detector simulation, the determination of the pdfs from data is an improvement w.r.t. the previous approach, in which they were derived from simulation [191]. Three electron identification working points are provided, of which the most stringent *Tight* working point is used in this analysis. Additionally to the cut on the discriminant d_L, the electron objects need to pass quality criteria on the associated ID tracks.

Electron isolation The isolation of an electron object [190] from other activity in the detector is defined by the isolation of the energy deposit of the electron in the calorimeter and by the isolation of the electron track in the ID. The isolation criterion for the electron signal in the calorimeter is based on the sum of the energy deposited in topo-clusters $E_{\mathrm{T,raw}}^{\mathrm{isol20}}$ around the barycentre of the energy deposit associated to the reconstructed electron in a cone of $\Delta R = 0.2$. Since the energy sum $E_{\mathrm{T,raw}}^{\mathrm{isol20}}$ contains also the energy deposit of the reconstructed electron, the energy deposited in the 5×7 calorimeter cells in $\Delta\eta \times \Delta\phi$ around the barycentre of the electron energy cluster ($E_{\mathrm{T,core}}$) is subtracted. This rather simple subtraction technique is robust against differences in signal and background objects for any transverse momenta and pile-up contributions. But, some of the energy of the reconstructed electron object may still be included in the subtracted energy sum $E_{\mathrm{T,raw}}^{\mathrm{isol20}} - E_{\mathrm{T,core}}$ such that an E_{T}- and $|\eta|$-dependent leakage correction is applied, which is derived from simulation. Further, a correction to account for pile-up and the underlying event is used, since the isolation of the electron object from other particles resulting from the hard process and not the isolation from pile-up contributions is of interest. The final calorimeter isolation variable is given by

$$E_{\mathrm{T}}^{\mathrm{cone20}} = E_{\mathrm{T,raw}}^{\mathrm{isol20}} - E_{\mathrm{T,core}} - E_{\mathrm{T,leakage}}(E_{\mathrm{T}}, \eta, \Delta R) - E_{\mathrm{T,pile-up}}(\eta, \Delta R). \tag{4.14}$$

The track isolation $p_{\mathrm{T}}^{\mathrm{varcone20}}$ is defined by summing the transverse momentum of ID tracks within a ΔR cone around the electron track direction, not including the electron track itself. The cone size is shrinking with the p_{T} of the electron as

$$\Delta R = \min\left(\frac{10}{p_{\mathrm{T}}[\mathrm{GeV}]}, 0.2\right) \tag{4.15}$$

to account for electrons from the decay of strongly boosted resonances, which will lead to decay products with small angular separation. Only ID tracks passing several quality criteria, regarding the number of hits in the ID subsystems and the primary vertex association, i.e. no pile-up tracks, are used in the calculation of $p_{\mathrm{T}}^{\mathrm{varcone20}}$. An electron passes the *Tight* electron isolation working point, which is used in this analysis, if it satisfies $E_{\mathrm{T}}^{\mathrm{cone20}}/p_{\mathrm{T}} < 0.06$ and $p_{\mathrm{T}}^{\mathrm{varcone20}}/p_{\mathrm{T}} < 0.06$.

4.4.4 Muons

A muon produces signals in all ATLAS detector subsystems. While muon reconstruction is based primarily on the tracking information from the ID and the muon spectrometer, also calorimeter information is utilized to correct for energy losses in the calorimeter. Different types of reconstructed muon objects are available within ATLAS, which are obtained using stand-alone or combined detector subsystem information. For the measurement in this thesis, the *combined* and the *inside-out combined* muon reconstruction algorithms are most relevant [192]. Both muon reconstruction types use information from all detector subsystems. In the combined muon reconstruction, tracks reconstructed in the muon spectrometer are matched to reconstructed tracks in the ID. In contrast, inside-out combined muons are reconstructed by extrapolating the remaining ID tracks into the muon spectrometer and requiring at least three associated muon spectrometer hits along the extrapolated trajectory. In both cases, the energy loss in the calorimeter is taken into account during the track extrapolation. The hits from the matched ID and muon spectrometer tracks are then used as input to a combined track fit. The inside-out combined muon reconstruction can recover some muon objects, which did not lead to a reconstructed track in the muon spectrometer due to low p_{T} or spatial regions with limited coverage by the muon spectrometer. The muon objects of interest in this analysis, are prompt muons originating from the W decay from the $WWbb$ final state. These must be distinguished from non-prompt muons originating from hadron decays. This separation is achieved by applying appropriate muon identification and isolation working points, which are defined below. Muons from light hadron decays are typically excluded by the muon identification working point, as they result in a poor quality muon track. However, muon tracks from heavy-flavour decays, originating from D- or B-hadron decay, often produce high quality muon tracks, since the muons are generated directly at the secondary vertex and traverse the detector as a single charged particle. Their contribution can be significantly reduced by applying a suitable isolation due to the typically enhanced hadronic activity in the vicinity of the reconstructed muon.

Muon identification In this analysis, the *Medium* identification working point [192] is used, which defines a set of selection criteria. Mainly, combined and combined inside-out muons are considered for this working point. In addition to requirements on the number of hits and missing hits in the ID and the muon spectrometer, the reconstructed tracks in the ID and muon spectrometer are required to be loosely compatible in their q/p measurements.[20] For the forward region in a pseudorapidity range of $2.5 < |\eta| < 2.7$, so-called silicon-associated forward and muon-spectrometer-extrapolated muon objects are also accepted. These muons do not have a reconstructed track in the ID, but may have a short track-segment in the

[20] This compatibility is defined by requiring

$$\frac{|q/p_{\mathrm{ID}} - q/p_{\mathrm{MS}}|}{\sqrt{\sigma^2(q/p_{\mathrm{ID}}) + \sigma^2(q/p_{\mathrm{MS}})}} < 7, \tag{4.16}$$

where q/p_{k} with $k \in \{\mathrm{ID}, \mathrm{MS}\}$ is the charge-to-momentum ratio measured in the ID or in the muon spectrometer and $\sigma(q/p_{\mathrm{k}})$ are the associated uncertainties [192].

ID (silicon-associated forward) or no ID hits at all (muon-spectrometer-extrapolated). This allows to take advantage of the full muon spectrometer coverage of $|\eta| < 2.7$. Still, the majority of reconstructed muons in $t\bar{t}$ simulation events — with 98 % — that pass the *Medium* identification working point are combined muons [192].

Muon track-to-vertex association The misidentification rate for a given muon identification working point is further reduced by applying cuts on the compatibility of the reconstructed muon track to the primary vertex. These selection criteria help to reduce the amount of muons originating from hadron decay or from pile-up interactions, as well as the number of cosmic muons, which are included in the final event selection. The selection cuts are defined in terms of the impact parameter significances, requiring $|d_0|/\sigma(d_0) < 3$ and $|z_0 \sin\theta| < 0.5\,\text{mm}$.

Muon isolation As discussed previously, the rate of light hadrons misidentified as muons or muons from light hadron decays is well controlled by the muon identification and track-to-vertex association working points. The remaining background of non-prompt muons from heavy hadron decay can be reduced by requiring the muon object to be isolated from other hadronic activity in the detector, as non-prompt muons from the heavy hadron decays are typically in the vicinity of the jet initiated by the heavy quark. In this analysis, the *PFlowTight* muon isolation working point [192] is employed, which uses both, track-based and calorimeter information to define hadronic activity around a reconstructed muon. Track-based isolation variables are defined similarly to the isolation of electron tracks: If the muon transverse momentum is larger than 50 GeV, the scalar sum $p_{\text{T}}^{\text{cone20}}$ of all transverse momenta of tracks associated with the primary vertex within $\Delta R = 0.2$ of the muon track is used. For muons with lower p_{T}, instead, a variable ΔR radius of $\Delta R = \min(10\,\text{GeV}/p_{\text{T}}^{\mu}, 0.3)$ is used to define $p_{\text{T}}^{\text{varcone30}}$. While the track-based isolation definition is robust with respect to pile-up due to the hard-scatter-association constraint, neutral energy deposits cannot be accounted for. For this purpose, the sum of the transverse energy $E_{\text{T}}^{\text{neflow20}}$ of the neutral particle-flow objects in a cone of $\Delta R = 0.2$ is used. Pile-up corrections to this neutral energy component are applied on an event-to-event basis with energy-density subtraction in the topo-cluster cells. A reconstructed muon passes the *PflowTight* isolation working point, if $(p_{\text{T}}^{\text{cone20}} + 0.4E_{\text{T}}^{\text{neflow20}})/p_{\text{T}}^{\mu} < 0.045$ for $p_{\text{T}}^{\mu} > 50\,\text{GeV}$ and $(p_{\text{T}}^{\text{varcone30}} + 0.4E_{\text{T}}^{\text{neflow20}})/p_{\text{T}}^{\mu} < 0.045$ for $p_{\text{T}}^{\mu} < 50\,\text{GeV}$. The excellent rejection of non-prompt muons using the isolation criterion is shown in Figure 4.18.

4.4.5 Missing transverse momentum

Since for initial state partons of a hard process, which are aligned with the beam-axis (z-axis) the total momentum in the transverse plane is zero, the same is true for the sum of the momenta of the produced final state particles in this interaction. The magnitude of the transverse momentum required to balance the transverse momenta of all reconstructed objects is defined as the missing transverse momentum $E_{\text{T}}^{\text{miss}}$. Sources of missing transverse momentum are SM neutrinos, misreconstruction of final-state particles, but also possible Beyond the SM particles, which do not interact with or decay within the detector material.

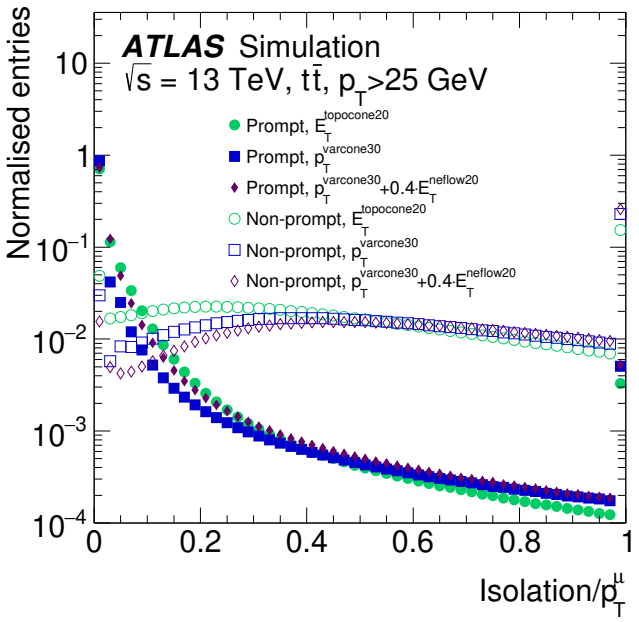


Figure 4.18: Distribution of various isolation variables for prompt and non-prompt muons from simulated $t\bar{t}$ events [192]. The isolation criterion in *PFlowTight* is equal to $(p_{\mathrm{T}}^{\mathrm{varcone30}} + 0.4E_{\mathrm{T}}^{\mathrm{neflow20}})$ for muons with $p_{\mathrm{T}} < 50\,\mathrm{GeV}$.[21] All muon objects with *PFlowTight* isolation$/p_{\mathrm{T}}^{\mu} > 0.045$ are passing this isolation working point. The $E_{\mathrm{T}}^{\mathrm{topocone20}}$ isolation criterion uses only calorimeter information and is defined as the sum of the transverse energy of topological cell clusters in a cone of $\Delta R = 0.2$ around the position of the muon, after subtracting the energy deposits attributed to the muon itself and correcting for pile-up effects. The distributions are obtained with the Run 1 and previous Run 2 software, but the results with the Run 3 and reprocessed Run 2 software are expected to be similar.

In the definition of the $E_{\mathrm{T}}^{\mathrm{miss}}$ [193], reconstructed and calibrated physics objects are used. Since the reconstruction of the different physics objects described above is independent of each other, the same detector signal — e.g. the same energy deposit in a calorimeter cell — can be included in the reconstruction of different physics objects. To avoid double-counting of energy deposits and tracks an overlap removal between these physics objects is applied, and finally only those reconstructed objects, which remain after this procedure, are included in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. Additionally, a soft contribution to the $E_{\mathrm{T}}^{\mathrm{miss}}$ is considered, by including reconstructed tracks, which were not used in the reconstruction of any of the physics objects included in the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. These reconstructed tracks are required to be associated to the hard-scatter primary vertex, to reduce the influence of pile-up tracks. The components of the $E_{\mathrm{T}}^{\mathrm{miss}}$ in x- and y-direction are calculated as

$$\vec{E}_{\mathrm{T}}^{\mathrm{miss}} = -\sum_{\text{selected physics objects}} \vec{p}_{\mathrm{T}} - \sum_{\text{unused tracks}} \vec{p}_{\mathrm{T}}. \tag{4.17}$$

CHAPTER 5

Previous top-quark mass and decay width measurements

In order to study properties of the top quark at the LHC experimentally, first a physics process needs to be identified which leads to a large number of top quarks being produced for a given integrated luminosity. This is achieved by focusing on top-quark pair production, which has by far the largest production cross-section among process involving top quarks, as shown in Figure 5.1. While the LO diagrams of single top-quark production involve electroweak vertices, top-quark pair production involves only QCD interactions. As a consequence, the latter has a significantly larger production cross-sections, reflecting the stronger coupling of QCD relative to the electroweak interaction. For this reason, for example, top-quark mass measurements are often performed by focusing on the top-quark pair production process. The choice of the physics process to study — such as $t\bar{t}$ production — is one of the main decisions defining the overall analysis strategy.

Another key analysis strategy decision is the choice of the decay channel of the W-bosons, which originate from the top-quark decays, as illustrated in Figure 5.2(a). Due to $|V_{tb}| \sim 1$ [4], the top quark in the SM decays in nearly 100 % of cases into a Wb pair. The subsequent decays of the W-bosons in the $WWbb$ final state can be classified into the all-hadronic ($WWbb \to qqqqbb$), the single-lepton ($WWbb \to qq\ell\nu bb$) and the dilepton channels ($WWbb \to \ell\nu\ell\nu bb$), where ℓ denotes an electron, muon or τ-lepton. In most measurements however, only the decay into an electron, muon and leptonically decaying τ-lepton are considered, excluding the hadronic τ-decay channel. While the all-hadronic channel has the largest branching fraction, as shown in Figure 5.2(b), it also suffers from the largest background contributions, since many background processes with larger production cross-sections also produce jets in the final state. The single-lepton channel still leads to good data statistics, but suffers from less background contributions, since fewer background processes produce a lepton in the final state. While in the dilepton channel the kinematic reconstruction of the $t\bar{t}$ system is more challenging due to the presence of two neutrinos in the final state, and it has a smaller branching fraction compared to the other two decay channels, it profits from the more accurate experimental

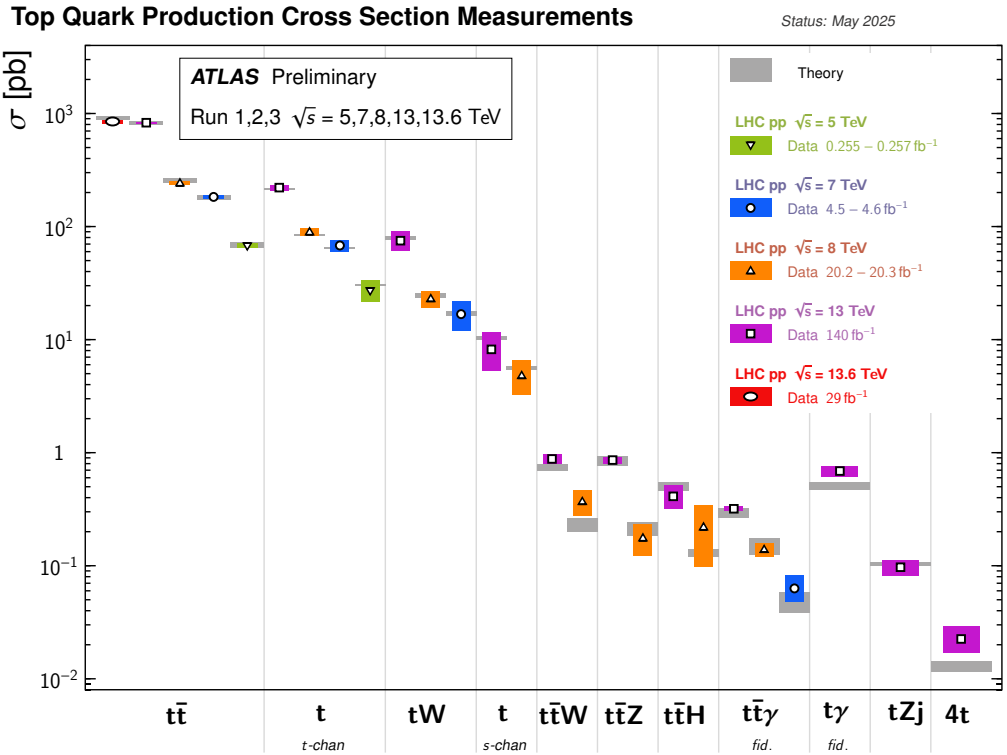
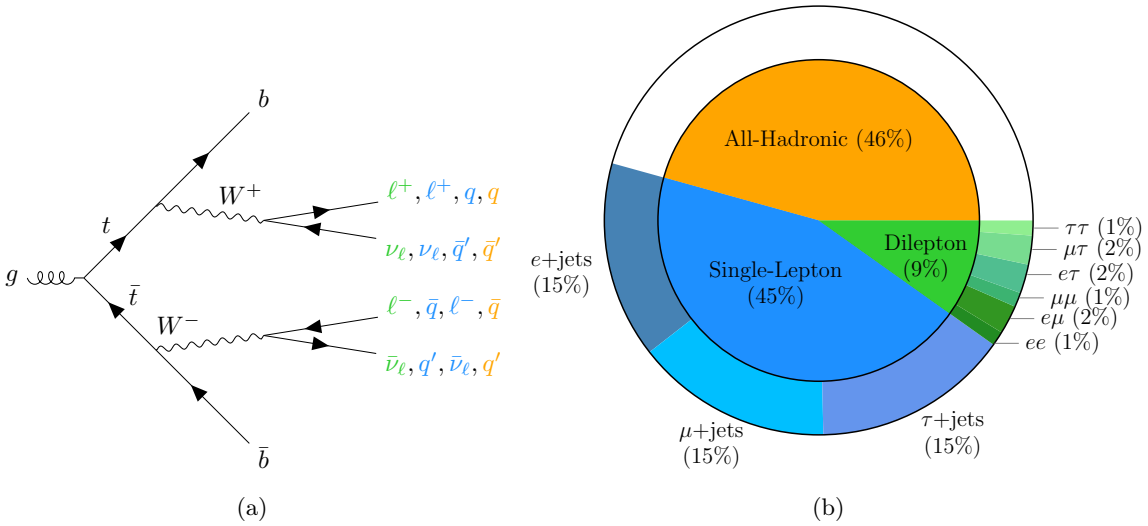


Figure 5.1: Summary of recent ATLAS top-quark production cross-section measurements and comparison to theory predictions [194].



(a) (b)

Figure 5.2: Feynman diagram (a) and branching fractions (b) for different $t\bar{t}$ decay modes: dilepton (green), single-lepton (blue) and all-hadronic (orange). The W-boson decay fractions are taken from [4].

reconstruction of leptons compared to jets and a significantly smaller background contribution. Since precision measurements of the top-quark mass and decay width currently are not limited by the statistical uncertainty, and sensitive observables to these parameters without the full reconstruction of the $t\bar{t}$ system can be constructed, the measurement presented in this thesis is performed in the dilepton decay channel of the $WWbb$ final state.

Nevertheless, the brief overview of previous top-quark mass and decay width measurements, in Section 5.1 and Section 5.2 respectively, also includes results from measurements using other decay channels. Finally, the aspects of the analysis strategy followed in this thesis, which are derived from these previous measurements, are summarised in Section 5.3.

5.1 Top-quark mass measurements

Before giving an overview of recent top-quark mass measurements and the applied analysis methods, a brief discussion of the interpretation of the measured top-quark mass parameter is given.

Interpretation of the measured top-quark mass parameter The top-quark mass measurement presented in this thesis and in the overview of previous measurements given in this section rely on kinematic observables build from the decay products of the top quark at particle-level, meaning after the parton shower and hadronisation. These observables exhibit a high sensitivity to the top-quark mass parameter used in Monte-Carlo simulations and lead currently to the most precise top-quark mass measurements. The interpretation of the measured parameter though, often called m_t^{MC}, is topic of theoretical discussions.[1] While in fixed-order perturbation theory above LO the top-quark mass parameter is defined in a specific renormalisation scheme (see Section 2.2), this is not the case for the NLO+PS MC predictions utilised as templates in these so-called direct top-quark mass measurements. The fixed order NLO computation of these simulations is generally performed in the pole-mass renormalisation scheme [15], but the modelling aspects in the parton shower and hadronisation make the interpretation of m_t^{MC} in a concrete renormalisation scheme difficult. Often, an uncertainty of 0.5-1 GeV in the interpretation of m_t^{MC} as m_t^{pole} is cited [15, 196]. This interpretation problem is of direct practical importance, since for SM or BSM predictions, as well as for EW precision tests of the SM, a top-quark mass parameter in a well-defined renormalisation scheme is needed. Different analysis techniques have been developed to be less sensitive to the modelling and the corresponding interpretation issues, which are called indirect measurements. These measurements focus on the extraction of the top-quark mass by comparing fixed-order calculations (with a definite renormalisation scheme for the top-quark mass parameter) to total or differential cross-sections. Thereby, the extraction of these cross-sections should be largely independent of the top-quark mass parameter used in the MC simulation. These indirect measurements typically lead to larger uncertainties in the measured mass parameter, with the most precise measurement to date being 170.5 ± 0.8 GeV [197]. In contrast, the

[1] See e.g. the review in [15, 195].

most precise direct top-quark mass measurement is given by $171.77 \pm 0.37\,\text{GeV}$ [198]. Also, theoretical work on precision prediction for new observables in the boosted top-quark regime [199, 200] and in parton shower development [201] to mitigate the interpretation problem is currently an active field of research.

Overview of direct top-quark mass measurements An overview of recent direct top-quark mass measurements by ATLAS and CMS, as well as the methods used in these measurements, is shown in Figure 5.3. Three main fit methods are listed in Figure 5.3, which are described below. ATLAS top-quark mass measurements typically use the template method regardless of the decay channel. CMS measurements, on the other hand, tend to use the ideogram method for the single-lepton (lepton+jets) and all-hadronic (all jets) channels, while the Analytic Matrix Weighting Technique (AMWT) is generally used in the dilepton channel. Some of the most recent measurements from both ATLAS and CMS use a profile likelihood template fit to extract the top-quark mass. As can bee seen in Figure 5.3, the precision of the recent top-quark mass measurements is limited by systematic uncertainties, with the most recent measurements and combinations achieving a precision of around 0.2 % compared to a precision of roughly 0.8 % of the first LHC top-quark mass measurements. While uncertainties related to the jet energy scale (JES) and the b-jet energy scale (bJES) are typically among the dominant systematic uncertainties, $t\bar{t}$ modelling uncertainties also play an important role — especially in the most recent measurements, where improvements in the jet energy scale uncertainty treatment have reduced their impact, see e.g. discussion in [125]. In the following, a simplified discussion of different top-quark mass fit strategies is given. While important aspects of data analyses such as the treatment of background processes and the details of the kinematic reconstruction are not considered here, a general overview of previously employed strategies is presented in order to place the current measurement within the wide range of existing top-quark mass measurements. In all fit strategies, the top-quark mass parameter m_t is determined by minimizing the negative log-likelihood of the sample likelihood, $-\ln \mathcal{L}_{\text{sample}}$. This likelihood can also depend on auxiliary parameters $\vec{\beta}$ that describe, for example, the jet energy scale response. The negative log-likelihood minimisation is then performed w.r.t. both the parameter of interest, the top-quark mass m_t, and the additional parameters $\vec{\beta}$.

Matrix element method In the matrix element method, the sample likelihood is defined as the product of per-event likelihoods:

$$\mathcal{L}_{\text{sample}}(x_1, \cdots, x_N, m_t, \vec{\beta}) = \prod_{i=1}^{N_{\text{events}}} \mathcal{L}_{\text{event},i}(x_i, m_t, \vec{\beta}), \tag{5.1}$$

where x_i represents the measured event properties at detector level, including the directions and energies of the reconstructed physics objects and the missing transverse momentum. The general set of parameters $\vec{\beta}$ describe modifications to the detector response.[2] Often, the jet

[2] The parameters $\vec{\beta}$ are thereby not nuisance parameters describing the impact of a systematic uncertainty, as in the profile likelihood fit described later in this section, but additional free parameters, which are included in the fit model.

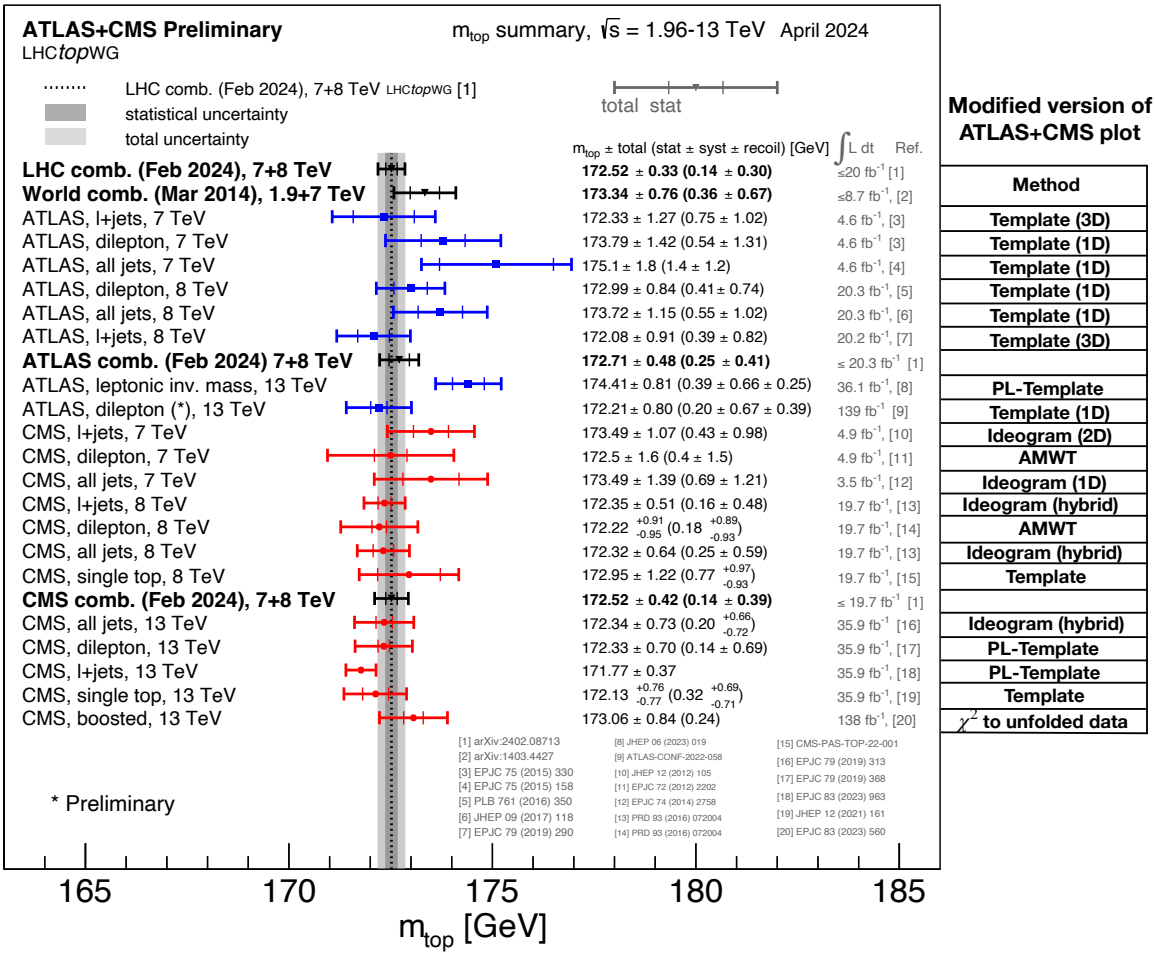


Figure 5.3: Summary of recent LHC top-quark mass measurements [202]. The different analysis techniques of the Template, the Ideogram, the Analytic Matrix Weighting Technique (AMWT) and the Profile Likelihood Template (PL-Template) fit will be discussed in this Section.

energy scale factor (β = JSF) is measured, which is a parameter used to correct the jet energy scale to match the recorded data.[3] The negative log-likelihood minimisation is then performed simultaneously for the top-quark mass parameter m_t and the parameters $\vec{\beta}$. The name matrix element method originates from the expression for the per-event likelihood:

$$\mathcal{L}_{\text{event},i}(x_i, m_t, \vec{\beta}) = \frac{1}{N} \int d\sigma^{\text{LO}}(pp \to t\bar{t}, m_t) W(x_i, \vec{\beta}, \Phi_{t\bar{t}}). \tag{5.2}$$

In Equation 5.2, the differential LO cross-section $d\sigma^{\text{LO}}(pp \to t\bar{t}, m_t)$ for producing a $t\bar{t}$ final state with kinematics $\Phi_{t\bar{t}}$ is convoluted with the detector transfer functions $W(x_i, \vec{\beta}, \Phi_{t\bar{t}})$ and integrated over the full available phase space. The detector transfer function $W(x_i, \vec{\beta}, \Phi_{t\bar{t}})$ describes the probability that a partonic $t\bar{t}$ final state with kinematics $\Phi_{t\bar{t}}$ is reconstructed as a detector-level event x given the parameters $\vec{\beta}$. To obtain a normalised likelihood, a normalisation factor N is introduced.[4] Clearly, the advantage of this method is that it utilises the full kinematic information of the final state. In particular, there is no reduction of dimensionality and hence no loss of information by reconstructing one specific top-quark mass sensitive observable. On the other hand, this method is computationally very expensive and was primarily used for the lower-statistics datasets at the Tevatron collider, e.g. in [203–206], which are input to the Tevatron top-quark mass combination [207]. Furthermore, the top-quark mass measurements based on the matrix element method use typically LO $t\bar{t}$ matrix elements. Extension of this method to incorporate NLO matrix elements have been discussed, e.g. recently in [208]. For a comprehensive description of the matrix element method, see e.g. [209].

Ideogram method and template fit method As in the matrix element method, both the ideogram and unbinned template fit methods define a per-event likelihood. However, instead of using the full kinematic information as in Equation 5.2, these methods compare a single top-quark mass-sensitive observable O_{m_t} to templates derived from MC simulation to allow for faster evaluation. Additionally, other observables O_β can be included in the fit to measure and reduce the impact of parameters related to significant systematic uncertainties, such as the JSF. Schematically, the per-event likelihood can be written as[5]

$$\mathcal{L}_{\text{event},i}(O_{m_t}, O_\beta, m_t, \beta) = P(O_{m_t}, O_\beta | m_t, \beta) = P(O_{m_t} | m_t, \beta) P(O_\beta | m_t, \beta). \tag{5.4}$$

[3] Uncertainties on the jet energy scale (JES) often dominate the total uncertainty of the top-quark mass measurements. Performing an additional measurement of the JES within the phase space of the top-quark mass measurement (opposed to the phase space used in the jet calibration measurements) can often reduce the uncertainty in this parameter.

[4] The normalisation factor is obtained by integrating over all possible detector-level event configurations x:

$$N = \iint \mathrm{d}\sigma^{\text{LO}}(pp \to t\bar{t}, m_t) W(x, \vec{\beta}, \Phi_{t\bar{t}}) f_{\text{acc}}(x) \mathrm{d}x \tag{5.3}$$

Since only events with final state particles that lie within the detector acceptance and pass trigger and reconstruction selections are used in the measurement, i.e. in the numerator of Equation 5.2, an acceptance factor $f_{\text{acc}}(x)$ accounts for these effects in the normalisation. Specifically, $f_{\text{acc}}(x) = 0$ for any detector-level event x that does not satisfy these selection criteria.

[5] The last equality in Equation 5.4 assumes that O_{m_t} and O_β are independent random variables.

The probability density functions $P(O_X|m_t, \beta)$ are derived from simulation by generating the distributions of O_X for various values of m_t and β, the MC templates. Then, the dependence of O_X on m_t and β is parameterised using analytic functions. The different naming conventions arise from the choice of top-quark mass-sensitive observables O_{m_t} and the way the per-event likelihood is constructed. In the ideogram method, a kinematic fit [210, 211] is performed to determine the invariant top-quark mass distribution m_t^{fit}, which serves as the top-quark mass sensitive observable O_{m_t}. Different assignments of the jets as decay products of the intermediate W-boson and top quark, also called jet permutations, are possible in this kinematic fit. Typically, all m_t^{fit} values obtained from the different permutations are included in the per-event-likelihood, weighted with the goodness of fit-value P_{gof} derived during the kinematic m_t fit. To reduce the influence of poorly reconstructed events, the per-event likelihoods are weighted by the sum of the goodness-of-fit values of all permutations: $w_{\text{event}} = x \sum_{i=1}^{n_{\text{perm}}} P_{\text{gof}}(i)$, where x normalises the average w_{event} to 1. The per-event likelihood for the ideogram method is then expressed as:

$$\mathcal{L}_{\text{event},i}(O_{m_t}, O_\beta, m_t, \beta) = \left(\sum_i^{n_{\text{perm}}} P_{\text{gof}}(i) P(O^i_{m_t}, O^i_\beta | m_t, \beta) \right)^{w_{\text{event}}} . \tag{5.5}$$

As mentioned, for the parameter β often the JSF is used. The top-quark mass measurements in CMS using the ideogram method are classified as 1D, in which the JSF parameter is fixed to unity, as 2D, in which the JSF parameter is unconstrained in the fit, or as hybrid measurements, in which a Gaussian constraint centred at one is added for the JSF parameter [195]. The quite involved ideogram method achieves excellent statistical precision for a given dataset, which becomes increasingly less important though with the large data set collected at the LHC [195]. The construction of the per-event likelihood in the template method in ATLAS measurements is usually simpler. In ATLAS dilepton measurements, the invariant mass distribution of the charged lepton and the b-jet, $m_{\ell b}$, is used as a mass-sensitive observable in the fit. The ℓb-pairing is determined either by selecting the combination that yields the lowest average $m_{\ell b}$ value from the two possible ℓb-pairs [212, 213], or by using a neural network-based approach [124]. Per event, two $m_{\ell b}$-values are obtained per event. These are either combined by taking their average, which is used as the sensitive distribution [212, 213], or the invariant mass of the ℓb-pair with the highest $p_{\text{T}}^{\ell b}$ is selected [124]. In ATLAS single-lepton top-mass measurements [123, 212], the reconstructed top-quark invariant mass that maximises a kinematic likelihood [214] is used as the top-quark mass sensitive observable. Thereby, all possible invariant-mass combinations of reconstructed objects are evaluated. The likelihood is constructed from Breit-Wigner functions for the top quark and W boson, as well as transfer functions derived from simulation that map energies of the reconstructed objects to energies of parton-level objects. While the ATLAS measurements in the dilepton and the single-lepton decay channel discussed above use unbinned template fits, ATLAS top-quark mass measurements in the all-hadronic channel [215, 216] use binned template fits. In this case, the three-jet to di-jet mass ratio $R_{3/2}$ is utilised as top-quark mass sensitive observable. For each event, two $R_{3/2}$ values are obtained, one for each top-quark decay. The three-jet mass, corresponding to the reconstructed top-quark invariant mass, combines one b-jet and two non-b-tagged jets, while the di-jet mass is formed from the two

non-b-tagged jets from the three-jet pair and are assumed to originate from the W boson decay. Jet assignments are determined by minimizing a χ^2 based on the differences between the two di-jet masses per event and a fixed W-boson mass, and between the reconstructed three-jet masses, corresponding to $m_t = m_{\bar{t}}$. While in most of the ATLAS top-quark mass measurements from Figure 5.3 only a top-quark mass sensitive distribution is used in the fit (Template 1D), the ATLAS measurements in the single-lepton decay channel add additional distributions, which are sensitive to the JSF and the relative b-to-light-jet energy scale factor bJSF (Template 3D).

Analytic Matrix Weighting Technique Due to the presence of two neutrinos in the final state of a dileptonically decaying top-quark pair, there is no unique solution for the final state kinematics given the measurement results when no assumption on the top-quark mass is made.[6] While ATLAS top-quark mass measurements in the dilepton channel [124, 212, 213] often use $m_{\ell b}$ as an observable in an unbinned template fit, Run 1 CMS measurements in the same channel [217, 218] frequently employ the Analytic Matrix Weighting Technique (AMWT). In the AMWT, solutions for the final state kinematics are computed under the assumption of different top-quark mass values. Each m_t hypothesis is assigned a weight [219]

$$w_{\text{event}} = \left(\sum f(x_1) f(x_2)\right) p(E_{\ell^+}|m_t) p(E_{\ell^-}|m_t) \tag{5.6}$$

with

$$p(E|m_t) = \frac{4 m_t E (m_t^2 - m_b^2 - 2 m_t E)}{(m_t^2 - m_b^2)^2 + M_W^2 (m_t^2 - m_b^2) - 2 M_W^4}, \tag{5.7}$$

where x_i are the Bjorken-x values of the initial state partons, computed from the reconstructed final-state kinematics, and $f(x)$ denotes the parton distribution functions evaluated at $Q^2 = m_t^2$. The sum runs over different leading-order initial state parton configurations, e.g. $u\bar{u}$ or gg. For each event, the mass assumption m_t which maximises w_{event} is defined as the top-quark mass sensitive observable $O_{m_t} = m_{\text{AMWT}}$. This procedure is applied to both reconstructed data and simulated $t\bar{t}$ MC events with different input top-quark masses. The resulting m_{AMWT} distributions of the MC samples are used as templates in a binned maximum-likelihood fit — instead of a per-event likelihood — with the sample likelihood:

$$\mathcal{L}_{\text{sample}} = \prod_{i=1}^{N_{\text{bins}}} P(O_{m_t,i}|m_t). \tag{5.8}$$

[6] The $t\bar{t}$ system in dileptonic decay has a priori 24 degrees of freedom: the four-momenta of the six decay particles (two b-jets, two charged leptons, and two neutrinos). Fourteen of these 24 free parameters are measured (the three-momenta of the two charged leptons, the two b-jets, and two $E_{\text{T}}^{\text{miss}}$ components). The following nine constraints are imposed on the final state particles: the invariant mass of the W decay products needs to be equal to $m_W^{\text{world-average}}$ (two constraints), the masses of the six final state particles need to be equal to their measured world average values (six constraints) and the mass of the top and the anti-top quark need to be the same (one constraint). This leaves one free parameter: the top-quark mass.

CMS kinematic endpoint top-quark mass measurement A quite different analysis technique is used in the CMS top-quark mass measurement [220] in the dilepton channel at a center-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$ with an integrated luminosity of $5.0\,\text{fb}^{-1}$. Additionally to measuring the top-quark mass, in the so-called unconstrained mode, this measurement was used as a study of measuring the masses of the top quark, the W boson and the neutrino simultaneously, which is applicable to Beyond The Standard Model scenarios, in which several particles of the decay chain might be unknown. Thereby, the measurement used kinematic end-points of three minimax distributions (or strongly correlated distributions) constructed as

$$M_{T2} = \min_{\vec{p}_T^{\,a}+\vec{p}_T^{\,b}=\vec{p}_T^{\,\text{miss}}} \left\{ \max(M_T^a, M_T^b) \right\}, \tag{5.9}$$

where the decay of two identical particles, a and b, into decay products including at least one invisible daughter is considered. The $M_T^{a/b}$ variable is constructed as

$$m_{\text{invis}}^2 + m_{\text{vis}}^2 + 2(E_{\text{T}}^{\text{invis}} E_{\text{T}}^{\text{vis}} - \vec{p}_{\text{T}}^{\,\text{invis}} \cdot \vec{p}_{\text{T}}^{\,\text{vis}}), \tag{5.10}$$

where m, E_{T} and $\vec{p}_{\text{T}}$ are the invariant mass, the transverse energy $E_{\text{T}} = \sqrt{p_{\text{T}}^2 + m^2}$ and the transverse momentum of the invisible (invis) and the visible (vis) daughter particles, which may consist of more than one particle. If the measured missing transverse momentum could be correctly partitioned between the invisible decay products of the two decay chains a and b, the calculated M_T would have, for each decay chain individually, as an upper bound the mass of the decaying particle m. Therefore, even when taking the maximum of M_T^a and M_T^b, this upper bound still holds. Since an exact partitioning is not possible, the solution for the transverse momentum split that minimises $\max(M_T^a, M_T^b)$ is used, ensuring that $M_{T2} \leq m$ always holds. In the CMS top-quark mass measurement, the dependence of this upper bound on the top-quark mass is exploited, thereby relying only minimally on MC simulation. A competitive top-quark mass measurement of $m_t = 173.9 \pm 0.9(\text{stat.})^{+1.7}_{-2.1}(\text{syst})\,\text{GeV}$ was achieved with this method, with an uncertainty comparable to that of the CMS dilepton result at $\sqrt{s} = 7\,\text{TeV}$, listed in Figure 5.3, which reaches a precision of $172.5 \pm 1.6\,\text{GeV}$. This measurement thus demonstrates the strong top-quark mass sensitivity of kinematic end-points.

Evaluating systematic uncertainties through pseudo experiments In the top-quark mass measurements discussed above, systematic uncertainties are commonly estimated using pseudo-experiments. In contrast, the profile likelihood fit, employed in this thesis and in some of the recent top-quark mass analyses [121, 125, 198, 221], does not rely on this procedure, as will be explained in the next paragraph. It is nevertheless useful to outline the pseudo-experiment method, as it illustrates a key difference between the approaches. In this method, multiple pseudo data sets are generated by sampling from the background and signal distributions, while varying the parameters in the simulation associated with a given systematic uncertainty by ± 1 standard deviation. The average m_t fit value (and the value of the simultaneously fitted parameters, such as JSF) from the pseudo experiments using either the nominal MC events or the MC events with $\pm 1\sigma$ variations of a specific systematic uncertainty are compared, and the difference of the average fitted values taken as the uncertainty in m_t for this systematic. The uncertainties $\hat{\sigma}_{m_t}^{\alpha}$ of the different systematic uncertainty sources α are added in quadrature.

This, by construction, neglects possible correlations between the systematic uncertainty sources, which should be small, though, by representing each distinct uncertainty source by a separate systematic uncertainty.

Profile likelihood template fit The top-quark mass extraction presented in this thesis is based on a binned profile likelihood template fit, as is the case for several recent top-quark mass analyses [121, 125, 198, 221]. Therefore, the advantages, but also the difficulties with respect to the fit strategies presented above, should be highlighted. The currently most precise top-quark mass measurement was performed by CMS using a profile likelihood template fit in the single lepton channel [198], reaching a precision of 370 MeV. In total, five regions were used, providing sensitivity to the top-quark mass parameter and parameters related to the jet-energy scale of light and b-jets.[7] For the main top-quark mass sensitive observable, m_t^{fit}, an unbinned profile likelihood fit is used, which is obtained similarly as in the ideogram method, but only using the permutation with the best goodness-of-fit value P_{gof}, while the other four distributions are fitted in a normalised binned profile likelihood fit. The top-quark mass measurement [198] thus demonstrates the significant improvement in the top-quark mass uncertainty, that is achievable with a profile likelihood fit. While the details of the statistical fit procedure are described in Section 9, here only the difference of the profile likelihood fit to the traditional top-quark mass measurement fit strategies presented above should be highlighted. While the interpretation of the top-quark mass measured in direct measurements was discussed in the beginning of this section, the technical description of the previously discussed direct top-quark mass measurements, which do not use a profile likelihood fit, can be clearly formulated: It is the measurement of the top-quark mass parameter used in a specific Monte-Carlo generator with specific settings in, e.g., the parton-shower configuration. This interpretation is not changed by including the simultaneous fit of detector response parameters, such as e.g. the JSF. In a profile likelihood fit, systematic uncertainties are included in the likelihood by parameterising their influence on the observable through (nuisance) parameters and adding (Gaussian) constraints on these parameters, which quantify the prior knowledge of these uncertainties. The likelihood optimisation now incorporates the minimisation over these nuisance parameters, which also include uncertainties evaluated by modifying parameters in the MC generation, which cannot be predicted from first principles, the so-called modelling systematics. Thus, if the nuisance parameters after the fit differ from their input value, a top-quark mass value obtained in a profile likelihood fit does not correspond to the top-quark mass parameter in a given MC generator with specific settings, but would correspond to a mixture of theory models / configurations. This is one of the reasons why systematic uncertainties associated with the modelling of the signal process must be strictly scrutinised and well understood. A significant step toward the top-quark mass and decay width measurement in this thesis is thus the development and study of the modelling systematics for the $bb4\ell$ signal process, which is described in Section 3. Besides adjusting the best-fit value of the nuisance parameters during the profile likelihood fit, the uncertainty associated with a particular systematic variation can also be reduced (constrained) if the data shows enough discrimination power. Further, in the profile likelihood fit, correlations between systematic uncertainties

[7] The light jets are thereby defined as non-b-tagged jets, thus also including c-jets.

are taken into account through the fit method itself, since it jointly maximises the likelihood $\mathcal{L}(\mu, \vec{\theta})$ over all nuisance parameters $\vec{\theta}$. This joint profiling allows correlated variations of nuisance parameters induced by their common impact on the observables, which are reflected in the post-fit covariance matrix. The need to perform pseudo-data experiments to quantify the impact of each systematic uncertainty, as discussed previously, is thereby removed by the usage of a profile likelihood fit. Since the analysis presented in this thesis, uses a profile likelihood fit, the measured best-fit-values and the constraints on the nuisance parameters need to be carefully studied. This applies to both, the parameter extraction of the top-quark mass and the decay width, but was highlighted in the context of top-quark mass measurements, due to the subtle change in the meaning of the measured Monte-Carlo top-quark mass parameter when profiling systematic uncertainties or not profiling these, as the latter is done by most of the previous precision top-quark mass analyses.

5.2 Top-quark decay width measurements

The results of the most recent top-quark decay width measurements at the Tevatron and the LHC are summarised in Figure 5.4. As seen by the number of measurement results in Figure 5.4, the top-quark decay width is significantly less studied compared to the top-quark mass in the LHC experiments. The distinction between indirect and direct top-quark decay width measurements will be explained in the following. However, it can already be clearly seen in Figure 5.4 that the indirect measurements achieve considerably smaller uncertainties on the fitted top-quark decay width. The disadvantage of the indirect measurements, however, is that assumptions on the top-quark decay must be made, as detailed below, limiting the ability to interpret the results in BSM scenarios. The measurement presented in this thesis constitutes a direct top-quark decay width measurement.

Indirect measurements The top-quark decay width in the indirect measurements of D0 [222] and CMS [223] is extracted by combining two measurements: the t-channel single top quark production cross-section [224, 225] and the measurement of the ratio R of the branching fraction of $t \to Wb$ to the branching fraction of $t \to Wq$, where q denotes any down-type quark (d, s, b): $R = B(t \to Wb)/B(t \to Wq)$ [223, 226]. The R ratio is determined by measuring the observed b-jet multiplicities in $t\bar{t}$ events and comparing them to the expected b-jet multiplicities for different R. For example, if fewer $t \to Wb$ decays than in the SM are assumed, more events with smaller b-jet multiplicity are expected. The partial decay width can be expressed in terms of the t-channel single top-quark production cross-section

$$\Gamma(t \to Wb) = \sigma^{\text{measured}}_{\text{t-channel}} \frac{\Gamma^{\text{SM}}(t \to Wb)}{\sigma^{\text{SM}}_{\text{t-channel}}}. \tag{5.11}$$

The single top-quark t-channel production is chosen, because it is sensitive to the tWb coupling and has a larger production cross-section compared to the s-channel single top-quark production. Using the relation between the partial and the total top-quark decay width,

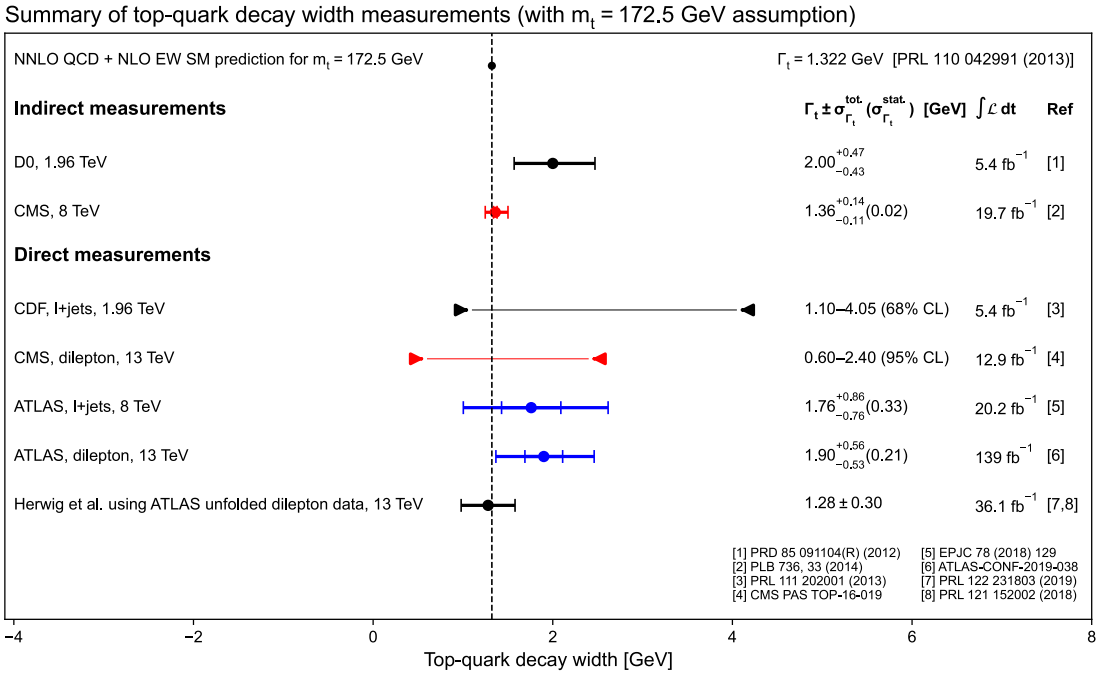


Figure 5.4: Summary of recent top-quark decay width measurements using a $m_{\text{top}} = 172.5\,\text{GeV}$ assumption.

$\Gamma(t \to Wb)$ and Γ_t,

$$\Gamma_t = \frac{\Gamma(t \to Wb)}{B(t \to Wb)}, \tag{5.12}$$

and assuming that the top-quark decays exclusively into a W-boson and a down-type quark ($\sum_q B(t \to Wq) = 1$), such that $R = B(t \to Wb)$, as well as that the g_{tWb}-coupling is the same in single top-quark production and top-quark decay, the total decay width can be written as

$$\Gamma_t = \frac{\sigma^{\text{measured}}_{\text{t-channel}}}{R^{\text{measured}}} \frac{\Gamma^{\text{SM}}(t \to Wb)}{\sigma^{\text{SM}}_{\text{t-channel}}} = \frac{\sigma^{\text{measured}}_{\text{t-channel}}}{B(t \to Wb)} \frac{\Gamma^{\text{SM}}(t \to Wb)}{\sigma^{\text{SM}}_{\text{t-channel}}}. \tag{5.13}$$

The assumption of a fixed top-quark mass is implicit in the calculation of $\Gamma^{\text{SM}}(t \to Wb)$, which is performed for $m_t = 172.5\,\text{GeV}$, as well as in the MC templates used for the R ratio and single top-quark t-channel production cross-section measurement, in both the D0 and the CMS measurements. While the indirect top-quark width measurements currently reach the smallest uncertainties, the direct measurements presented in the following do not rely on the restrictive assumption $\sum_q B(t \to Wq) = 1$.[8]

[8] This assumption would, for example, be explicitly violated in case of a top-quark decay to a BSM charged Higgs boson [227] or a flavour changing neutral current (FCNC) top-quark decay [41].

Direct measurements using $t\bar{t}$ predictions The direct measurements performed by the CDF [228], CMS [229], and ATLAS [230, 231] collaborations, as summarised in Figure 5.4, extract the top-quark decay width through a maximum likelihood fit to either the reconstructed top-quark mass (CDF) or a top-quark mass proxy $m_{\ell b}$ (CMS and ATLAS). In the ATLAS measurements, the lepton-b-jet pairing is performed either by selecting the ℓb-pair with the smallest angular separation ΔR [231] or using the assignment from a kinematic fit [230]. In contrast, the CMS measurement [229] considers all possible ℓb-pairings per event,[9] as this approach was found to improve the sensitivity of the measurement. The dilepton ATLAS measurement [231], which employs a profile likelihood fit, introduces the invariant mass of the two b-jets m_{bb}, in a dedicated control region as additional distribution in the fit to reduce the systematic uncertainties related to the b-jet reconstruction. Controlling jet-related systematics — particularly, the JES uncertainty — through inclusion of dedicated JES-sensitive distributions is a common strategy in single-lepton analyses, as previously discussed for top-quark mass measurements and also applied in the top-quark decay width measurements by CDF [228] and ATLAS [230] in the single-lepton channel. However, none of the so far discussed $t\bar{t}$ dilepton analyses included such a control region. All of these top-quark decay width measurements focus on the double-resonant $WWbb$ production, i.e. the $t\bar{t}$ production process. While CDF uses LO $t\bar{t}$ samples generated with Pythia and obtains templates for different Γ_t values by generating new LO $t\bar{t}$ Pythia samples with varied widths, the CMS and ATLAS measurements use NLO $t\bar{t}$ Powheg+Pythia 8 predictions, where variations of the top-quark decay width are modelled via per-event reweighting using a Breit-Wigner distribution. It was verified, that this reweighting procedure results in the same distribution as newly generated NLO $t\bar{t}$ samples with varied input top-quark decay widths. The ATLAS [230, 231] and CMS [229] measurements state several considerations regarding the range of the $m_{\ell b}$ distribution, which is used due to relying on MC predictions, which have NLO QCD accuracy in the on-shell $t\bar{t}$ production (with subsequent smearing of the top-quark invariant mass using a Breit-Wigner distribution) and LO QCD accuracy in top-quark decay description: The ATLAS measurement in the dilepton channel [231] restricts the $m_{\ell b}$ range to values below 150 GeV, motivated by the expected increasing impact of missing top-quark off-shell and NLO top-quark decay effects above this threshold [1]. The ATLAS measurement in the single-lepton channel [230] instead uses an extended $m_{\ell b}$ range up to 200 GeV, and tested explicitly that limiting the range to 150 GeV results or that reweighting $t\bar{t} + tW$ parton level kinematics to a fixed-order NLO QCD off-shell $WWbb$ calculation, resulted in a change of the measured top-quark decay width of less than 0.5 GeV. Thereby, the total uncertainty in the top-quark decay width measurement is $1.76^{+0.86}_{-0.76}$ GeV, thus emphasizing the benefit of using MC predictions, which include off-shell effects for future measurements. In contrast, the CMS measurement [229] uses a broader $m_{\ell b}$ range in the statistical fit of up to $m_{\ell b} = 300$ GeV. However, [229] states that in this regime the measurement probes both, missing higher order effects in the SM QCD calculation, as well as potential BSM contributions. Further, all the top-quark decay width measurements [228–231] are performed under the assumption of a fixed top-quark mass of $m_t = 172.5$ GeV. In the ATLAS measurement [231], the best-fit $\hat{\Gamma}_t$

[9] Thereby, all $m_{\ell b}$-combinations obtained with the two leading leptons and the leading or the two leading b-jets are considered — depending on the whether the observable is constructed in the analysis region requiring exactly one or more than two b-jets— without weighting the $m_{\ell b}$ combinations by any probability measure how likely a ℓb-combination is expected to originate from the same top-quark decay.

values are presented also for MC samples generated with alternative input top-quark masses around $\pm 0.5\,\text{GeV}$ of the nominal top-quark mass, leading to alternative $\hat{\Gamma}_t$ values, which are well within the uncertainty in the nominal top-quark decay width measurement. The CMS measurement [229] shows the variation of the width-sensitive observable $m_{\ell b}$ under both the top-quark decay width and top-quark mass variations. It is thereby stated in [229] that top-quark mass variations primarily affect the end-point region around $m_{\ell b} = 150\,\text{GeV}$, whereas the top-quark decay width variations predominantly influence the low $m_{\ell b}$-region.

Direct measurements using *WWbb* predictions The conclusions regarding the top-quark decay width sensitive phase space regions of the $m_{\ell b}$ distribution change significantly when using MC predictions for the full $WWbb$ production process. These predictions, generated with the $bb4\ell$ generator [1, 2] (see Section 2.5), include $t\bar{t}/tW$ interference, top-quark off-shell effects and NLO corrections in the top-quark decay. This was demonstrated in the top-quark decay width measurement presented in [232], which compared $bb4\ell$ predictions with varied input top-quark decay widths to an unfolded $m_{\ell b}^{\text{minimax}}$ distribution[10] from ATLAS [78], where the full $WWbb$ final-state — i.e. in the case of [78], the $t\bar{t} + tW$ process — was unfolded. The ATLAS unfolding measurement used $WWbb$ events in the dilepton decay channel and introduced a dedicated control region requiring at least three b-tagged jets. This control region was added to correct the MC prediction of $t\bar{t}$ events with additional heavy-flavour jets ($t\bar{t}$ +HF) to data. This background was found to dominate the high-$m_{\ell b}^{\text{minimax}}$ tail. The fitted normalisation factor for the MC $t\bar{t}$ +HF contribution was thereby determined to be $1.49 \pm 0.05(\text{stat.}) \pm 0.20(\text{syst.})$. The highest sensitivity to the top-quark decay width was observed in the tail of the $m_{\ell b}^{\text{minimax}}$ distribution, where the contribution from off-shell events and tWb-like events is enhanced. While the off-shell event fraction scales linearly with the top-quark decay width [233], the variation of the top-quark decay width influences the relative fraction of $t\bar{t}$ and tWb-like events and thus leads to shape variations of the distribution, when using off-shell $WWbb$ MC predictions [232, 234]. Using this analysis strategy, the most precise direct top-quark decay width measurement to date was achieved, with an uncertainty of 300 MeV. This constitutes a significant improvement compared to the previously most precise direct top-quark decay width measurement of $1.90^{+0.56}_{-0.53}\,\text{GeV}$ [231]. This result emphasizes the importance of using NLO QCD $WWbb$ MC predictions for future top-quark decay width measurements.

[10] The $m_{\ell b}^{\text{minimax}}$ distribution is discussed in detail in Section 7.2.

5.3 Lessons learned from previous measurements

Several important aspects of the analysis strategy presented in this thesis are derived from the experience of previous top-quark mass and decay width measurements. These are briefly summarised below:

- **Increased sensitivity from using $WWbb$ MC predictions** In particular, the sensitivity to top-quark decay width variations is significantly enhanced in the high-$m_{\ell b}^{\text{minimax}}$ tail, when modelling top-quark off-shell effects and $t\bar{t}/tW$ interference at NLO QCD accuracy, as implemented in the $bb4\ell$ MC generator. This MC generator was used in the currently most precise direct top-quark decay width measurement [232]. Furthermore, the uncertainty in the top-quark mass measurement is expected to be reduced when using $bb4\ell$ predictions, as the recently introduced top-quark lineshape uncertainty in ATLAS [90] — which accounts for uncertainties arising from the implementation of the Breit-Wigner smearing of the top-quark invariant mass in on-shell $t\bar{t}$ production simulations — can be avoided due to the inclusion of full NLO QCD top-quark off-shell effects in the $bb4\ell$ simulation.
- **Sensitive distribution in $WWbb$ dilepton decay channel** In previous top-quark mass and decay width measurements in the dilepton channel, the use of an $m_{\ell b}$-distribution — rather than attempting an explicit reconstruction of the top-quark invariant mass with the complication of two neutrinos in the final state — is a common practice. The top-quark decay width measurement in [232] demonstrated the gain in sensitivity achieved by including the off-shell region of the $m_{\ell b}^{\text{minimax}}$ distribution. The CMS top-quark mass measurement [220] showed that the kinematic end-point of such a mini-max observable also provides strong sensitivity to variations of the top-quark mass.
- **Top-quark decay width measurement depends on the assumed top-quark mass** Previous top-quark decay width measurements assume a fixed top-quark mass of $m_t = 172.5\,\text{GeV}$. Repeating these measurements with varied top-quark input masses can lead to shifts in the fitted top-quark decay width parameter. In all cases, the observed variations remain within the quoted total uncertainty in the decay width parameter. However, with increasing precision on the top-quark decay width measurement, this dependence might become more relevant. In the analysis presented in this thesis, the use of the $bb4\ell$ MC prediction is expected to reduce this dependence, as the sensitivity to the top-quark decay width arises primarily from the off-shell tail of the $m_{\ell b}^{\text{minimax}}$ distribution rather than from the shape of the reconstructed top-quark invariant mass peak.[11] More importantly, this issue does not arise in the measurement in this thesis, as both the top-quark mass and decay width are extracted simultaneously.

[11] See the estimation of the correlations between the top-quark mass and decay width in the simultaneous fit in Section 10.1, as well as the dependence of the sensitive distribution on variations in both parameters in Section 7.2.

- **Controlling the $WWbb$ + additional heavy-flavour contribution** In the ATLAS unfolding measurement [78] of the $m_{\ell b}^{\text{minimax}}$ distribution of the $WWbb$ production process, the dominant background in the high $m_{\ell b}^{\text{minimax}}$ tail was found to be the contribution of $t\bar{t}$ events with additional b-jets. This background was underestimated in the MC predictions and its normalisation was fitted to data in a dedicated control region, resulting in a substantial correction factor of $1.49 \pm 0.05(\text{stat.}) \pm 0.20(\text{syst.})$. Since the high $m_{\ell b}^{\text{minimax}}$ region is particularly sensitive to variations of top-quark decay width [232], controlling the $WWbb$ +additional heavy-flavour jet contribution will be important in the measurement presented in this thesis.
- **Constraining JES and bJES** Systematic uncertainties related to the JES and the bJES are often among the dominant sources of uncertainty in $t\bar{t}$ measurements across all decay channels. While both the top-quark mass and top-quark decay width measurements in the lepton+jets channel often include dedicated control regions to constrain these uncertainties, such an approach has only been implemented in the dilepton channel in [231], where the invariant mass of the two b-jets, m_{bb}, was used for this purpose.
- **Using a profile likelihood fit** The use of a profile-likelihood fit is expected to improve the precision on the measured top-quark mass and decay width parameters due to the inclusion of correlations between nuisance parameters, the adjustment of these to their best-fit values, as well as reducing systematic uncertainties if they can be constrained in the analysis phase space by the data. This approach was also followed in the currently most precise top-quark mass measurement by CMS [198]. But, the final fit result needs to be carefully scrutinised, especially regarding the treatment of modelling systematic uncertainties. Also, the recent direct ATLAS top-quark decay width measurement [231] was performed using a binned profile likelihood fit.

Therefore, in this thesis the top-quark mass and decay width parameters are measured simultaneously using a profile likelihood fit, $bb4\ell$ MC predictions, the top-quark mass and decay width sensitive $m_{\ell b}^{\text{minimax}}$ distribution and a dedicated treatment of the normalisation of $WWbb$ events with additional heavy-flavour jets, thereby controlling jet related systematic uncertainties by including an m_{bb} distribution in a control region.

CHAPTER 6

Data and Monte-Carlo simulation

The measurement presented in this thesis is performed with the data set of $140\,\text{fb}^{-1}$ of pp–collision data collected by the ATLAS detector in the full LHC Run 2 period (2015–2018) at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$. Observables build from these data events are compared to the corresponding observables obtained with simulated MC events, which describe the expected data for different theory models. In the case of this analysis, this corresponds to the $WWbb$ production process[1] for different top-quark mass and top-quark decay width values, as well as background processes, which are simulated as predicted by the SM. The first step in the event generation is the modelling of the hard process, the parton shower and hadronisation, which were discussed in detail in Section 2. In the next step, the interaction of the final-state particles, generated by the MC simulation, with the detector material is modelled with GEANT4 [235] for the nominal signal and background MC samples in this analysis. Especially, the simulation of particle showers in the calorimeter is a computational expensive task. For this reason, MC samples which are used to evaluate a systematic uncertainty are commonly simulated using ATLFAST 3 [236], in which the GEANT4 calorimeter shower simulation is replaced with parameterised shower shapes and generative machine learning algorithms. The detector simulation outputs are hits in the ID and in the muon spectrometer and energy deposits in the calorimeter subsystems, which are digitised into voltage and current signals, emulating the detector readout. After this step, both MC and real data events are processed with the same reconstruction chain, as well as with the same analysis-specific algorithms. In this analysis, reprocessed Run 2 data, as well as reprocessed MC samples are studied, using reconstruction algorithms which are also applied in Run 3 event reconstruction.

[1] More concretely, the NLO QCD prediction for the full $b\bar{b}\ell^+\ell'^-\nu_\ell\bar{\nu}_{\ell'}$ final state is considered through the $bb4\ell$ MC simulation, but for brevity this will be referred to as $WWbb$ production in the following.

6.1 Data events

The recorded data events considered in this analysis need to satisfy quality criteria [237] concerning the condition of the detector during data taking, as e.g. the detector subsystems need to be fully operational, and characteristics of the reconstructed event[2]. The integrated luminosity of data events per year in the Run 2 data taking period, which satisfy these quality criteria, and are therefore suitable for physics analyses, is listed in Table 6.1. As mentioned in Section 4.2, the mean number of pp interactions per bunch crossing $\langle\mu\rangle$ varies during the Run 2 data taking period, as listed in Table 6.1 and shown graphically in Figure 4.3. Additional pp interactions, i.e. not the hard scattering interaction of interest, are referred to as pile-up.

Year	Luminosity	$\langle\mu\rangle$
2015	3.24fb^{-1}	13.4
2016	33.40fb^{-1}	25.1
2017	44.63fb^{-1}	37.8
2018	58.79fb^{-1}	36.1
Total Run 2	140.07fb^{-1}	33.7

Table 6.1: Integrated luminosity per year and for the total Run 2 data taking period [238], as well as the mean number of interactions per bunch crossing $\langle\mu\rangle$. The $\langle\mu\rangle$ values are given for the full data set recorded by ATLAS in Run-2 [149].

6.2 Monte-Carlo events

To match the pile-up profile in the different data-taking years of Run 2, simulated inelastic interactions are overlaid on the hard-scatter MC events before the reconstruction step [239]. Thereby, the detector simulation and digitisation is run separately for the hard-scatter and the minimum-bias interactions, which are predominantly soft QCD processes that dominate the total pile-up, and the resulting simulated detector read-outs overlaid. These inelastic interactions are modelled with either Epos 2.0.1.4 [240] or Pythia 8.308 [44], using the Epos LHC tune [241] and the Pythia A3 tune [242] and the NNPDF2.3lo [243] PDF set, respectively. Thereby, the Pythia generated events are required to include either a high transverse momentum jet, a prompt photon, or a lepton originating from a B-hadron decay, while the Epos generator was used to model the remaining pile-up events.

When the hard-scatter process of the MC samples described in the following is simulated with either the Powheg method [46, 47] or with MadGraph5_aMC@NLO [60], the nominal sample prediction is obtained by interfacing the produced hard-scatter events with the Pythia 8 for parton showering and hadronisation. For these samples, the Pythia 8 A14 tune [95] is always used, derived with the NNPDF2.3lo PDF set [243]. For MC samples

[2] Data events satisfying these criteria are listed in the so-called *good run list*.

generated with SHERPA, the hadronisation and parton showering of the hard scatter events are performed in SHERPA itself, using the default parton shower tune for the showering provided by the SHERPA authors. Further, the decay of bottom and charm hadrons is simulated with EVTGEN [96], except for processes which were generated with SHERPA for which the heavy hadron decay is simulated within SHERPA. For the hard-process generation of MC events obtained with POWHEG or MADGRAPH5_AMC@NLO, the NNPDF3.0NLO PDF set [93] is used, while for events simulated with SHERPA instead the NNPDF3.0NNLO PDF set [93] is utilised.

6.2.1 Signal process

The development of the nominal ATLAS $bb4\ell$ MC sample, as well as the modelling uncertainty prescription, which should be applied in analyses using this sample to build their nominal prediction, were vital steps towards the measurement presented in this thesis. These studies are presented in detail in Section 3, while in this section only a brief overview of these MC samples is given.

The $WWbb$ production process in dileptonic decay is generated using the $bb4\ell$ process [1, 2] of the POWHEG BOX RES generator [89].[3] The top-quark mass is set to $m_{\text{top}} = 172.5\,\text{GeV}$, with a corresponding decay width of $\Gamma_{\text{top}} = 1.32\,\text{GeV}$. The POWHEG h_{damp} parameter is chosen as $1.5\,m_{\text{top}}$. The renormalisation (μ_R) and factorisation (μ_F) scales used in the matrix element calculation depend on the projected resonance history. For events projected onto a $t\bar{t}$ resonance history, the scales are defined as

$$[(m_t^2 + p_{\text{T},t}^2)(m_{\bar{t}}^2 + p_{\text{T},\bar{t}}^2)]^{1/4}, \tag{6.1}$$

while for events projected onto a tWb resonance history, they are set to

$$[(m_t^2 + p_{\text{T},t}^2)(m_b^2 + p_{\text{T},b}^2)]^{1/4}. \tag{6.2}$$

Here, m and p_{T} denote the invariant mass and the transverse momentum of the top and bottom quarks at the underlying Born level. Parton showering and hadronisation are performed using PYTHIA 8.312, with a dedicated parton shower interface for $bb4\ell$ events. The PYTHIA 8 recoil scheme is set to recoil-to-top, as described in Section 3.4.

The $bb4\ell$ MC sample used as prediction for the signal process is generated using the full GEANT4 detector simulation. Since alternative $bb4\ell$ MC samples described below are produced using the ATLFAST 3 detector simulation instead, also an additional nominal $bb4\ell$ MC sample is available produced with the ATLFAST 3 simulation. When evaluating systematic uncertainties with dedicated $bb4\ell$ samples or when parameterising the variation due to different top-quark masses and decay widths, this is performed by using the ATLFAST 3 version of the nominal $bb4\ell$ MC sample in order to stay consistent with the used ATLAS simulation software in this comparison. In the following, the list of alternative $bb4\ell$ MC samples is presented. The change of specific parameter settings to evaluate a systematic uncertainty or to allow for a

[3] The $bb4\ell$ process version used in this thesis is currently only available in [92].

parameterisation of the top-quark mass dependence are described, all other MC generation parameters are thereby left unchanged with respect to the nominal sample.

- **Alternative top-quark mass** Four $bb4\ell$ MC samples with a varied top-quark mass input parameter m_{top} are generated, which span the range $m_{\text{top}} = (171, 172, 173, 174)\,\text{GeV}$. The top-quark decay width at NLO QCD has been correspondingly adapted in the MC generation to $\Gamma_{\text{top}} = (1.28, 1.306, 1.333, 1.36)\,\text{GeV}$, respectively.
- **Alternative h_{damp}** The POWHEG h_{damp} parameter is set to $h_{\text{damp}} = 3\,m_{\text{top}}$ instead of the value of $h_{\text{damp}} = 1.5\,m_{\text{top}}$ used in the nominal MC sample generation.
- **Alternative $p_{\text{T}}^{\text{hard}}$** The PYTHIA 8 matching parameter is set to $p_{\text{T}}^{\text{hard}} = 1$ instead of the nominal $p_{\text{T}}^{\text{hard}} = 0$ setting [90].
- **Scale-resonance-veto** Instead of using the default dedicated $bb4\ell$ parton-shower interface for the matching of the parton shower in the top-quark decay subprocess, an alternative matching procedure is applied, using the scale-resonance-veto setting implemented in the same interface.
- **Underlying event variation** Parton shower parameters influencing the underlying event modelling are varied according to the A14Var1 uncertainty prescription [95].
- **Alternative colour reconnection model** The $bb4\ell$ MC sample obtained with the colour reconnection model CR0 is compared to the samples generated with the colour reconnection models CR1 and CR2, thereby using the appropriate ATLAS tunes [142].
- **Alternative parton shower Herwig 7** Instead of using the PYTHIA 8 parton shower, the $bb4\ell$ events from the hard process generation are interfaced with the HERWIG 7.3 parton shower [244], using thereby the default tune.
- **Alternative Pythia 8 recoil scheme** The PYTHIA 8 recoil scheme is changed from recoil-to-top, which is used in the nominal $bb4\ell$ MC sample production, to the recoil-to-colour setting.

***bb4ℓ* signal categories** The MC events generated with the $bb4\ell$ process are divided into two categories in this measurement: events with ($bb4\ell + \text{HF}$) and without ($bb4\ell$) additional particle-level heavy-flavour jets. This categorisation is performed by inspecting the number of particle-level jets, which are matched to a B-hadron with $p_{\text{T}} > 5\,\text{GeV}$ through an angular distance matching requiring $\Delta R < 0.3$.[4] These particle-level jets are thereby constructed by clustering all stable particles from simulation — having a mean lifetime of $\tau > 3 \cdot 10^{-11}\,\text{s}$ — using the anti-k_t algorithm with a radius parameter $R = 0.4$. At the analysis level, no kinematic selection was placed on these particle-level objects, but it was verified that the minimum transverse momentum is $p_{\text{T}} > 5\,\text{GeV}$. The number of particle-level b-jets originating from the hard process, which are not selected by these inclusive selection criteria is expected

[4] Technically, this is performed by using the centrally available classification variable `HadronConeExclTruthLabelID` used in ATLAS flavour tagging.

to be negligible. All MC events with more than two particle-level b-jets are classified as $bb4\ell + \text{HF}$ events.

6.2.2 Background process

Due to the specific final state composition of $WWbb$ production in the dileptonic decay channel and the stringent object selection criteria, the background contribution in this analysis is very small, amounting to less than 1 % of the total event yield in the signal region. While the object and event selection used to isolate the $WWbb$ events with a dileptonic $e\mu$ final state is described in detail in Section 7.1, the relative contributions of the individual background processes in the SR are already presented in the following. Specifically, the fractions of the total background yield arising from the different background processes are quoted to illustrate which sources dominate the residual background, despite their overall small contribution.

For the measurement, the considered background processes are grouped into the following categories: $t\bar{t}$+bosons, *LeptonFakes* and *Others*. These categories and the corresponding background MC samples are briefly described.

$t\bar{t}$ +boson The contribution of all events in the signal region of this analysis, which are grouped into the $t\bar{t}$ +boson category, namely the $t\bar{t}H$, $t\bar{t}Z$ and the $t\bar{t}W$ production process, amounts to roughly 70 % of the total background contribution. The $t\bar{t}H$ and $t\bar{t}Z$ production processes are modelled at NLO QCD with POWHEG BOX v2 [245, 246] and MADGRAPH5_aMC@NLO 3.5.1 [60], respectively. The $t\bar{t}W$ events are generated with SHERPA 2.2.14 [247], which provides NLO QCD accurate predictions for events with up to one jet and LO QCD accurate results for events with two jets. The total cross-sections of the $t\bar{t}$+boson production processes were corrected to the predicted cross-section values at NLO QCD and NLO EW accuracy [248] using MADGRAPH5_aMC@NLO of $0.88^{+0.09}_{-0.11}$ pb, $0.60^{+0.08}_{-0.07}$ pb and 507^{+35}_{-50} fb for $t\bar{t}Z$, $t\bar{t}W$ and $t\bar{t}H$, respectively. All three $t\bar{t}$+boson production processes contribute significantly to the total $t\bar{t}$+boson background categories, with approximate fractions of 40 % from $t\bar{t}W$, 35 % from $t\bar{t}Z$, and 25 % from $t\bar{t}H$.

LeptonFake Several background production processes are considered, which do not include two charged leptons in their final state and can only enter the event selection applied in the measurement, if another physics object is misidentified as a charged lepton. These processes are grouped in the so-called *LeptonFake* category,[5] which constitute around 20 % of the total

[5] In the composition of this background category, it is explicitly checked with a centrally available tool in the ATLAS Collaboration, the Isolation and Fake Forum TruthClassification Tool, if at least one muon or one electron, is not classified as a prompt muon or electron, but originates instead from photon conversion, hadron decay, was reconstructed with an inverted electric charge compared to the truth particle or cannot be classified. Electron and muons from a leptonic tau-decay are classified as prompt. The classification from muons from tau-decay with the current version of this tool is under investigation in the ATLAS Collaboration. For this reason, the TruthClassification Tool is not used to define the other background categories, i.e. no requirement of having only prompt muons is used in the other categories. Instead, it is applied for the defining the LeptonFake category, as the used MC samples for the $t\bar{t}$ and tW production

background contribution in the signal region of the analysis and is strongly dominated by the $t\bar{t} + tW$ processes in single-lepton decay.

- ***tt̄* and *tW* production in the single-lepton decay channel** Since a $bb4\ell$ MC sample including also the single-lepton decay channel of the $WWbb$ final-state [2] is not yet available within the ATLAS Collaboration, this contribution is instead estimated using simulated samples of top-quark pair production ($t\bar{t}$) and associated production of top-quarks with W bosons (tW). These MC events are generated using POWHEG and PYTHIA 8, with parameter settings consistent with those discussed in Section 3, which are briefly summarised here. The $t\bar{t}$ and tW MC events are generated with the POWHEG generator [59, 75] in the five-flavour number scheme. As for the $bb4\ell$ MC sample, the top-quark mass and decay width are set to $m_{\text{top}} = 172.5\,\text{GeV}$ and $\Gamma_{\text{top}} = 1.32\,\text{GeV}$, and the h_{damp} parameter is chosen as $1.5\,m_{\text{top}}$. The tW sample is produced using the diagram removal scheme [76]. The hard-scatter events are interfaced with PYTHIA 8.230 for the $t\bar{t}$ sample and PYTHIA 8.309 for the tW sample. For both samples, the PYTHIA 8 recoil-to-colour scheme is applied. The total $t\bar{t}$ cross-section is normalised to the NNLO prediction in QCD, including the resummation of next-to-next-to-leading logarithmic (NNLL) soft-gluon terms, calculated using TOP++ 2.0 [48, 249–254], yielding $\sigma(t\bar{t})_{\text{NNLO+NNLL}} = 832 \pm 51\,\text{pb}$. The inclusive tW cross-section is corrected to the theoretical prediction calculated at NLO in QCD with NNLL soft-gluon corrections [255, 256], $\sigma(tW)_{\text{NLO+NNLL}} = 71.7 \pm 3.8\,\text{pb}$.
- **Single top-quark production** Single top-quark production in the t-channel and s-channel is modelled using the POWHEG BOX v2 implementations of these processes [257, 258]. The t-channel process is generated in the four-flavour number scheme, while the s-channel process is generated in the five-flavour number scheme, both using the NNPDF3.0NLO [93] parton distribution function set. In both cases, the generated events are interfaced with PYTHIA 8.230. The inclusive cross-sections are scaled to the theoretical prediction calculated at NLO in QCD using HATHOR 2.1 [259, 260]. This results in cross-sections of $\sigma(t, t\text{-chan})_{\text{NLO}} = 136.02^{+5.40}_{-4.57}\,\text{pb}$ ($\sigma(\bar{t}, t\text{-chan})_{\text{NLO}} = 80.95^{+4.06}_{-3.61}\,\text{pb}$), and $\sigma(t, s\text{-chan})_{\text{NLO}} = 6.35^{+0.23}_{-0.20}\,\text{pb}$ ($\sigma(\bar{t}, s\text{-chan})_{\text{NLO}} = 3.97^{+0.19}_{-0.17}\,\text{pb}$) for single-top (single-anti-top) production.
- ***W*+jets** The W+jets production events are simulated with SHERPA v2.2.11 [247], with a matrix element calculation, which is NLO QCD accurate for up to two additional jets, and LO QCD accurate for up to five additional jets.
- **Diboson in single-lepton decay** The diboson production process in single-lepton decay is predicted with SHERPA v2.2.14. The matrix element calculation is performed at NLO QCD accuracy for up to one additional jet and at LO QCD accuracy for up to three additional jets.

process also contain events in the dileptonic decay channel. Since this contribution is already described by the $bb4\ell$ sample, it must be removed from the $t\bar{t}$ and tW MC sample. This is achieved by requiring at least one non-prompt electron or muon with the TruthClassification Tool. Addtionally, for the $t\bar{t}$ component it is required that not both top-quarks decay into leptons also at the level of the event record. It was checked, that for the other MC samples, applying this tool does not lead to a reduction of MC events passing the selection in the *LeptonFake* category.

Others Diboson production in multileptonic decay and Z+jets production are grouped into the *Others* background category, which accounts for approximately 10 % of the total background contribution in the signal region. Within this category, the background is dominated by Z+jets events, which constitute about 70 % of the *Others* contribution.

- ***Z*+jets** The Z+jet events are simulated with SHERPA v2.2.11 [247] for the $Z \to ee$ and $Z \to \mu\mu$ decay channels, and with SHERPA v2.2.14 for the $Z \to \tau\tau$ channel. The matrix element calculation is performed at NLO QCD accuracy for up to two additional jets, and at LO QCD accuracy for up to five additional jets. Off-shell photon contributions are included in the matrix-element calculation, and separate MC samples are produced for dilepton invariant masses of $m_{\ell\ell} > 40\,\text{GeV}$ and $10\,\text{GeV} < m_{\ell\ell} < 40\,\text{GeV}$. All corresponding Z+jets MC samples are used in this analysis, but the total predicted event yield is dominated by the $Z \to \tau\tau$ samples with $m_{\ell\ell} > 40\,\text{GeV}$, which contribute about 92 % of the full Z+jets background. This is plausible, as the $Z \to (\tau \to e)(\tau \to \mu)$ decays naturally produce the $e\mu$ final state targeted by the event selection.
- **Diboson in multileptonic decay** The diboson production process in multileptonic decay is predicted with SHERPA v2.2.12 or SHERPA v2.2.14, depending on the production process. The matrix element calculation is performed at NLO QCD accuracy for up to one additional jet and at LO QCD accuracy for up to three additional jets. The dominant contribution from this set of diboson MC samples in the signal region of the analysis stems from the $pp \to \ell\ell\nu\nu$ production process.

CHAPTER 7

Analysis strategy

In this analysis, the top-quark mass and decay width are measured using a profile-likelihood fit to the $m_{\ell b}^{\text{minimax}}$ distribution in the dilepton ($e\mu$) decay channel of the $WWbb$ final state. After describing the object and event selection applied in the definition of the analysis signal and control regions in Section 7.1, these regions are described in more detail in Section 7.2–Section 7.4.

7.1 Object definition and event selection

In this measurement, three analysis regions are studied, which provide sensitivity to the top-quark mass and width variations, to the modelling of the $bb4\ell + \text{HF}$ contribution and to the nuisance parameters parameterising jet energy scale uncertainties. These regions are defined with the following physics objects:

- **Electrons** Reconstructed electron objects are required to have $p_{\text{T}} > 28\,\text{GeV}$ and $|\eta| < 2.47$. Additionally, electrons in the pseudorapidity region of $1.37 < |\eta| < 1.52$ are excluded since this corresponds to the transition region between the barrel and the end-cap ECAL (refer to Section 4.3.4). The *Tight* identification and the *TightVarRad* isolation criteria need to be fulfilled (refer to Section 4.4.3).

- **Muons** The same transverse momentum selection of $p_{\text{T}} > 28\,\text{GeV}$ as for electron objects are applied for muons, while requiring the pseudorapidity to be $|\eta| < 2.5$. The *Medium* identification working point and the *PflowTightVarRad* isolation working point are applied (refer to Section 4.4.4).

- **Jets** The jets are reconstructed with the anti-k_t jet-clustering algorithm, with $R = 0.4$, using PFlow objects as the jet constituents (Section 4.4.1). Jets need to satisfy $p_{\text{T}} > 25\,\text{GeV}$ and $|\eta| < 2.5$.

- ***b*-tagged jets** A jet is classified as a b-jet in this analysis, if the tightest b-tagging working point of the GN2v01 flavour tagging algorithm, with 65 % efficiency, is satisfied (Section 4.4.2).

In order to reduce the amount of pile-up jets entering the analysis regions, the selected jets need to pass the *Tight* working-point of a neural-network-based Jet Vertex Tagger (NNJVT). This requirement is only applied for jets with $p_{\mathrm{T}} < 60\,\mathrm{GeV}$, as the pile-up contribution is only expected to be considerable for low-p_{T} jets. This pile-up jet tagging algorithm is an updated version of the JVT algorithm [261].

As the reconstruction of physics objects, e.g. muons, electrons and jets, is independent of each other, an overlap removal procedure is performed in order to avoid double-counting of final-state particles as multiple (different) reconstructed physics objects. The procedure from [262] is followed. For example, only one lepton object is retained if an electron and muon object share an ID track to avoid lepton duplicates. Which lepton object is retained depends on the specific characteristics of the reconstructed objects. If leptons and jets are within a ΔR cone of 0.2, the jet is removed to avoid including the lepton energy in the jet reconstruction. If the remaining jets and leptons are within a ΔR cone of 0.4, the lepton is removed since it is likely to originate from a hadron decay in the jet.

For each event, at least one of the standard ATLAS electron or muon triggers [263, 264] needs to be trigged and one of the reconstructed electron or muon objects must be matched to the lepton triggering these single lepton triggers.[1] The additional event selection criteria in the different regions are summarised in Figure 7.1 and are explained in more detail in the following.

Event selection in the signal region In the signal region of this analysis, events with a $WWbb$ final-state in dileptonic $e\mu$-decay are targeted by requiring exactly one electron and exactly one muon of opposite electric charges. Additionally, exactly two b-tagged jets (at the 65 % GN2v01 b-tagging efficiency working point) need to be present in the event. Further, a veto against a third b-tagged jet is applied at a looser working point with 85 % efficiency in order to reduce the contribution of $bb4\ell + \mathrm{HF}$ events in the signal region. The $m_{\ell b}^{\mathrm{minimax}}$ observable (see Section 7.2) is studied in the signal region, and all events passing the selection criteria of the signal region are considered, by including them in the over- and underflow bins of the $m_{\ell b}^{\mathrm{minimax}}$ distribution.

[1] Concretely, the 2015 single electron triggers `HLT_e24_lhmedium_L1EM20VH`, `HLT_e60_lhmedium`, `HLT_e120_lhloose` and the muon triggers `HLT_mu20_iloose_L1MU15`, `HLT_mu40` are used, as well as the 2016–2018 electron triggers `HLT_e26_lhtight_nod0_ivarloose`, `HLT_e60_lhmedium_nod0`, `HLT_e140_lhloose_nod0` and muon triggers `HLT_mu26_ivarmedium` and `HLT_mu50`.

Event selection in the $bb4\ell$ + HF control region To probe a potential mismodelling of the $bb4\ell$ + HF contribution in the MC prediction, the total number of events in a dedicated control region enriched in $bb4\ell$ + HF MC events is utilised. While with respect to the signal region selection, the b-jet requirement is modified to select events with at least three b-jets (at the GN2v01 65 % b-tagging working point), all other signal region selection criteria are applied also to the $bb4\ell$ + HF control region.

Event selection in the control region to constrain jet uncertainties To constrain jet energy scale uncertainties in the measurement, the invariant mass of two b-jets, m_{bb}, in an additional control region is used. This region uses the same b-jet requirements as the signal region, but different lepton-selection criteria. The same-lepton-flavour dilepton decay channels (ee and $\mu\mu$) are targeted by requiring exactly two electrons or exactly two muons with opposite electrical charge. To suppress events with leptons originating from hadron or Z-boson decays, the dilepton invariant mass is required to satisfy $m_{\ell\ell} > 15\,\text{GeV}$ and $m_{\ell\ell} \notin [80, 100]\text{GeV}$. To isolate the $WWbb$ production process, in which neutrinos from the W-boson decays lead to missing transverse momentum in the event, an additional requirement of $E_{\text{T}}^{\text{miss}} > 60\,\text{GeV}$ is imposed (compare to Section 4.4.5).[2] Finally, only events with an invariant mass of the b-jet pair in the range $m_{bb} \in [60, 300]\,\text{GeV}$ are considered, as motivated in Section 7.4.

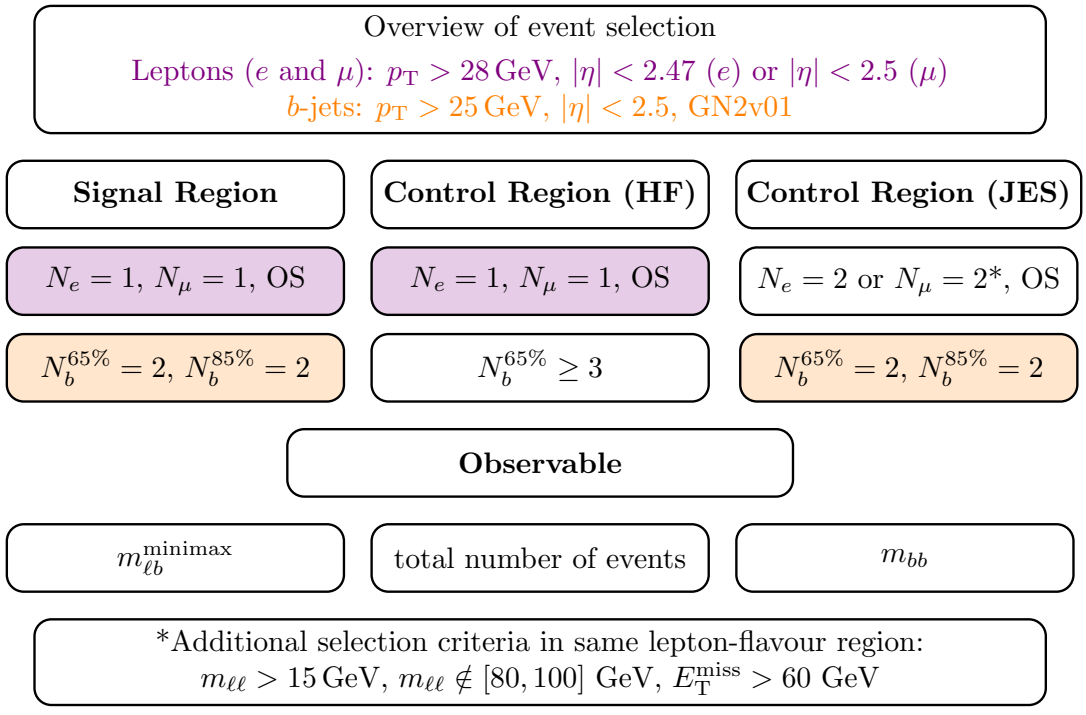


Figure 7.1: Sketch of the signal and control region definitions used in this measurement.

[2] The $E_{\text{T}}^{\text{miss}}$-selection requirement is added only in the same-lepton-flavour control region to especially reduce the Z+jets background contribution, which is found to be small in the signal region due to the different-lepton-flavour ($e\mu$) selection criterion.

After applying these object- and event-selection criteria, the total number of MC events per category and the number of Run 2 data events in the three measurement regions are determined and shown in Table 7.1. All three analysis regions are strongly dominated by the $bb4\ell$ and $bb4\ell + \text{HF}$ contributions, while the background contributions are significantly smaller or negligible. The $bb4\ell$ contribution dominates both the signal region and the control region targeting the reduction of JES uncertainties. In contrast, the $bb4\ell + \text{HF}$ contribution is the dominant process in the control region designed to constrain the modelling of additional heavy-flavour jet production, as expected. A contribution of $bb4\ell$ events is also observed in this control region. These events include either a c-jet or a light jet that is misidentified as a b-jet.

	SR	(%)	JES CR	(%)	HF CR	(%)
$bb4\ell$	$165\,000 \pm 15\,000$	95.4	$62\,600 \pm 6\,300$	94.9	149 ± 62	7.5
$bb4\ell + \text{HF}$	$7\,100 \pm 2\,300$	4.1	$2\,810 \pm 980$	4.3	$1\,720 \pm 420$	86.5
$t\bar{t}$+boson	560 ± 16	0.32	272 ± 9	0.41	116 ± 5	5.8
LeptonFakes	150 ± 130	0.08	41 ± 37	0.06	3.3 ± 3.1	0.17
Others	73 ± 37	0.04	250 ± 130	0.37	1.04 ± 0.58	0.05
Total	$172\,900 \pm 15\,100$	100	$66\,000 \pm 6\,500$	100	$1\,990 \pm 430$	100
Data	169 844		64 246		3 107	
MC / Data	1.02 ± 0.09		1.03 ± 0.10		0.64 ± 0.14	

Table 7.1: MC and data event yields in the signal region (SR), JES control region (JES CR) and $bb4\ell + \text{HF}$ control region (HF CR) defined in Figure 7.1. The MC event categories are introduced in Section 6.2. The quoted uncertainties in the number of events include all systematic uncertainties described in Section 8.

The implementation of the event selection, the application of correction factors and systematic uncertainties on the physics objects, and the creation of the final distributions used in the statistical fit was performed with the Athena [265], TopCPToolkit [266] and the FastFrames [267] software packages.

7.2 Event reconstruction in the signal region

The top-quark mass and decay width sensitive observable used in the signal region of this measurement is the $m_{\ell b}^{\text{minimax}}$ distribution [78]. It is defined as

$$m_{\ell b}^{\text{minimax}} = \min\{\max(m_{\ell_1 b_1}, m_{\ell_2 b_2}), \max(m_{\ell_1 b_2}, m_{\ell_2 b_1})\}. \tag{7.1}$$

In the signal-region event selection, exactly two leptons (one electron and one muon) and two b-jets are required. Therefore, any ordering of the leptons and b-jets in Equation 7.1 leads to the same $m_{\ell b}^{\text{minimax}}$ value. In the practical implementation, b_1 and ℓ_1 are defined as the leading p_{T} b-jet and lepton, while b_2 and ℓ_2 correspond to the remaining b-jet and lepton.

It was shown in [232], that the off-shell region and the interference region between $t\bar{t}$ and tW production in the $m_{\ell b}^{\text{minimax}}$ distribution are particularly sensitive to the top-quark decay width. Using the $m_{\ell b}^{\text{minimax}}$ observable, this region can be isolated in a way that is exact at LO. This can be understood by first considering the on-shell LO decay of one top quark $t \to Wb \to \ell\nu b$. At LO it can be shown that $m_{\ell b}^2 \leq m_t^2 - m_W^2$, where m_t and m_W denote the on-shell invariant masses of the top quark and the W boson. For the dileptonic decay of a top-antitop quark pair, and assuming that the decay products could be assigned unambiguously, the following upper bound would apply:

$$\max(m_{\ell_t b_t}, m_{\ell_{\bar{t}} b_{\bar{t}}}) \leq \sqrt{m_t^2 - m_W^2} \equiv m_{\ell b}^{\text{end}}, \tag{7.2}$$

where $\ell_t b_t$ and $\ell_{\bar{t}} b_{\bar{t}}$ denote the decay products from the top and anti-top quark, respectively. Experimentally, however, the correct assignment of the ℓb pairs originating from the same top-quark decay cannot be determined unambiguously. The construction of the $m_{\ell b}^{\text{minimax}}$ observables avoids the need for this assignment, while still leading to the same upper bound of $m_{\ell b}$. For illustration, it is assumed that all selected b-jets and leptons in the event are decay products of a top-quark pair and that, in the notation from Equation 7.1, the decay products of the top quark are ℓ_1 and b_1 and of the anti-top quark ℓ_2 and b_2. The first term $\max(m_{\ell_1 b_1}, m_{\ell_2 b_2})$ corresponds to the correct pairing and therefore has an upper bound of $m_{\ell b}^{\text{end}}$ at LO. In contrast, the maximum invariant mass of the wrong pairing, $\max(m_{\ell_2 b_1}, m_{\ell_1 b_2})$, is not constrained by the preservation of invariants related to the decaying top-quark resonance and can therefore exceed this bound. But, since $m_{\ell b}^{\text{minimax}}$ is defined as the minimum of these two quantities, $\max(m_{\ell_1 b_1}, m_{\ell_2 b_2})$ and $\max(m_{\ell_2 b_1}, m_{\ell_1 b_2})$, it follows that for an on-shell LO $t\bar{t}$ decay

$$m_{\ell b}^{\text{minimax}} \leq \sqrt{m_t^2 - m_W^2} \equiv m_{\ell b}^{\text{end}} \approx 150\,\text{GeV}. \tag{7.3}$$

The phase-space region above this kinematic end-point, of approximately 150 GeV, is particularly sensitive to the top-quark decay width Γ_t. In this phase-space region, the contribution from off-shell top quarks becomes significant and the fraction of non-double-resonant processes, such as tWb, increases. Since one of the ℓb combinations in these processes does not originate from a decaying top-quark resonance, the kinematic bound discussed above does not apply. The ratio between the $t\bar{t}$-dominated phase-space region and the region enhanced in $t\bar{t}/tW$ interference contributions is linearly sensitive to the top-quark decay width and independent of the coupling strength of the $t \to Wb$ vertex, g_{tWb}, as shown for example in [234]. This sensitivity arises from the different scaling of the double-resonant and the $t\bar{t}/tW$ interference terms with Γ_t. In the $bb4\ell$ MC generation, resonance-history projectors assign either a $t\bar{t}$ or a tWb resonance history to each generated event. In phase-space regions where $t\bar{t}/tW$ interference effects are enhanced, the fraction of events assigned to a tWb resonance history is expected to increase. Figure 7.2 illustrates how the fractions of $bb4\ell$ MC events assigned to $t\bar{t}$ and tWb resonance histories depend on the assumed value of the top-quark decay width. This dependence is particularly pronounced above the end-point around 150 GeV. In addition, the increasing contribution of the tWb component with increasing $m_{\ell b}^{\text{minimax}}$ can be observed. These features translate directly into a decay-width sensitivity of the high-$m_{\ell b}^{\text{minimax}}$ tail, which is shown for $bb4\ell$ MC events in the signal region in Figure 7.3(a). Thereby, the total event

yields of the $bb4\ell$ predictions with alternative top-quark decay widths are normalised to the nominal $bb4\ell$ event yield in the signal region. Since the end-point $m_{\ell b}^{\text{end}}$ is directly related to the top-quark mass parameter m_t, the region close to this boundary exhibits enhanced sensitivity to m_t. In addition, sensitivity to m_t is also observed in the threshold region and in the bulk of the $m_{\ell b}^{\text{minimax}}$ distribution, as illustrated in Figure 7.3(b). Thus, overall, the regions of the $m_{\ell b}^{\text{minimax}}$ distribution sensitive to variations of the top-quark decay width and mass are largely separated.

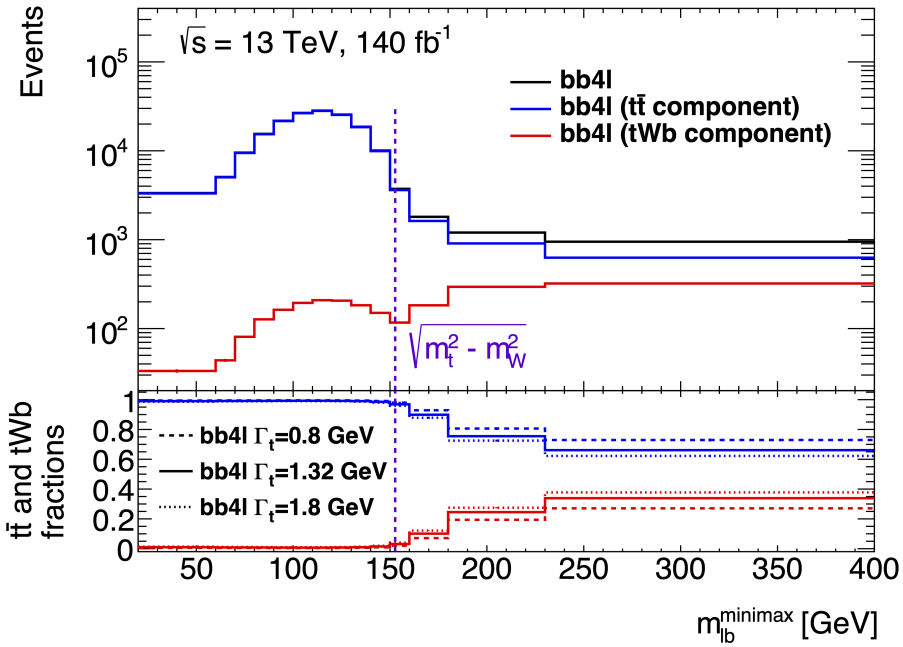
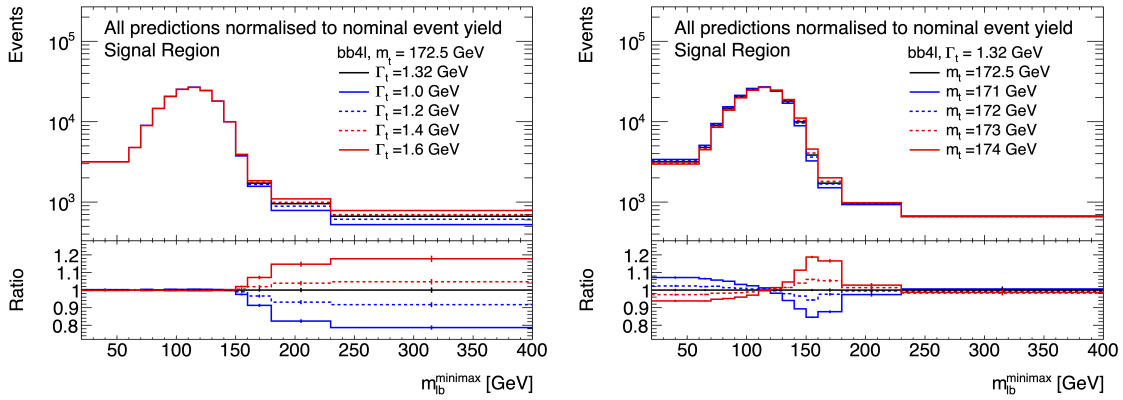


Figure 7.2: Number of $bb4\ell$ MC events in the fully inclusive $bb4\ell$ prediction (including both the $bb4\ell$ and $bb4\ell + \text{HF}$ contributions) that are projected onto a $t\bar{t}$ or a tWb resonance history. The fractions are shown in the lower ratio panel as a function of the $m_{\ell b}^{\text{minimax}}$ distribution in the signal region, and their variation with different assumptions for the top-quark decay width is illustrated. The kinematic end-point of the $m_{\ell b}^{\text{minimax}}$ distribution for a $t\bar{t}$ LO decay is given by $\sqrt{m_t^2 - m_W^2}$.

As the $bb4\ell$ MC sample is divided into two categories in this measurement ($bb4\ell$ and $bb4\ell$+HF), the $bb4\ell + \text{HF}$ MC prediction for different top-quark mass and decay widths is also studied. As shown in Figure 7.4, the sensitivity is reduced in this $bb4\ell + \text{HF}$ compared to the $bb4\ell$ contribution, but it remains non-negligible for variations of the top-quark mass and is therefore parametrised in the analysis.[3] In the definition of the $bb4\ell + \text{HF}$ contribution, the number of truth-particle jets passing loose kinematic cuts is considered. Consequently, these additional heavy-flavour truth jets may either not be reconstructed as a physics object, or the corresponding physics object may not pass the applied object selection criteria. In such cases, the two b-jets in the signal region still correspond to the two b-jets originating from top-quark

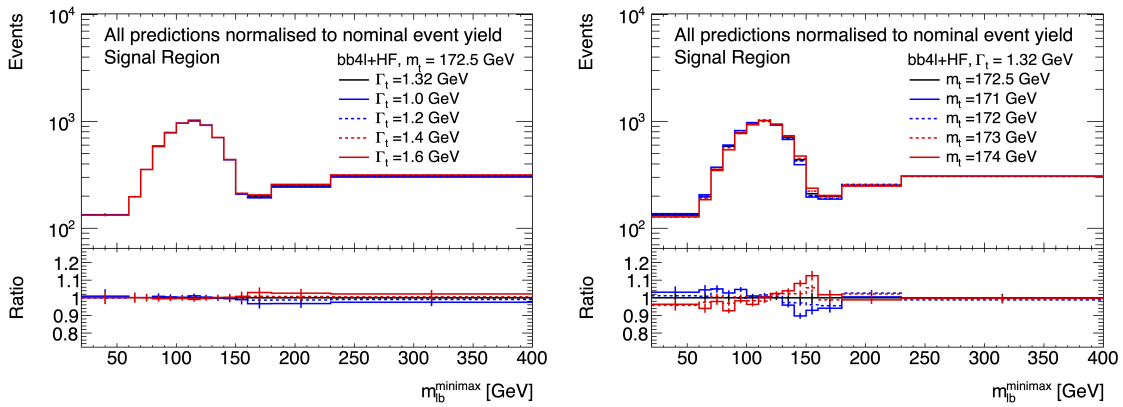
[3] The necessity to include the $bb4\ell + \text{HF}$ event yield parameterisation in terms of m_t and Γ_t is tested in Section 9.3.

decay and therefore exhibit the same dependence on the top-quark mass and decay width as the events in the $bb4\ell$ category.



(a) (b)

Figure 7.3: Distributions of the $m_{\ell b}^{\text{minimax}}$ observable in the signal region as predicted for the $bb4\ell$ contribution for different values of (a) the top-quark decay width and of (b) the top-quark mass. The total event yield in the signal region is thereby normalised to the total event yield of the nominal $bb4\ell$ sample with $m_t = 172.5\,\text{GeV}$ and $\Gamma_t = 1.32\,\text{GeV}$. The $bb4\ell$ predictions with varied values of the top-quark mass are shown in all cases for $\Gamma_t = 1.32\,\text{GeV}$, obtained in the $bb4\ell$ MC samples with $m_t \neq 172.5\,\text{GeV}$ via event reweighting, as described in the text.



(a) (b)

Figure 7.4: Same comparison as in Figure 7.3, but for the $bb4\ell + \text{HF}$ contribution.

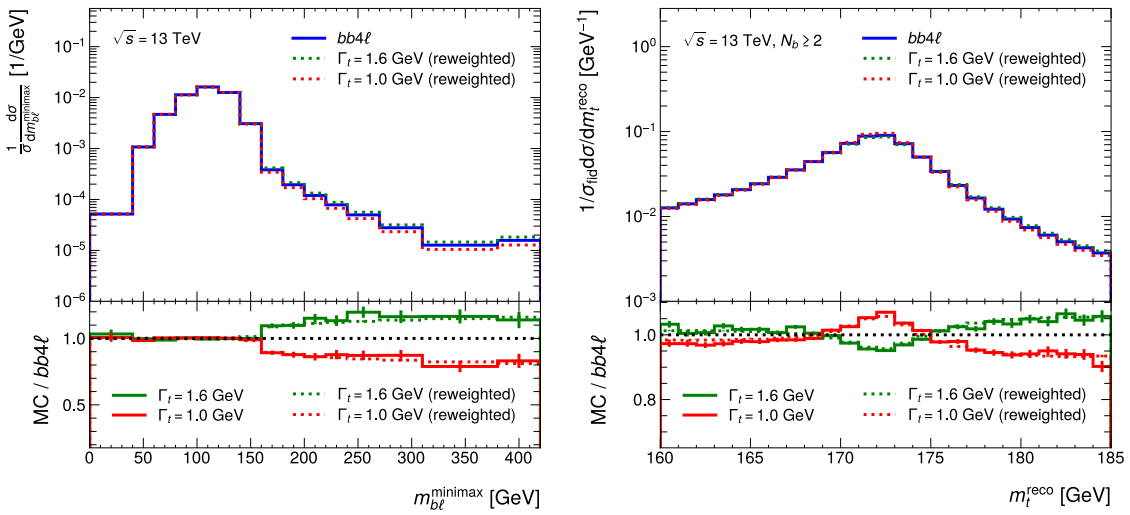
The dependence of the $m_{\ell b}^{\text{minimax}}$ distribution on the top-quark mass and decay width is evaluated using dedicated $bb4\ell$ MC samples generated with top-quark masses

$$m_t \in \{171, 172, 172.5, 173, 174\}\,\text{GeV}. \tag{7.4}$$

The variation of the top-quark decay width is evaluated in each of these samples using event weights calculated during the POWHEG generation process. With this procedure, predictions for the alternative decay widths

$$\Gamma_t \in \{0.5, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0, 2.2, 2.4, 2.6, 2.8, 3.0, 3.5, 4.0\}\,\text{GeV} \tag{7.5}$$

can be generated. The reweighting procedure is validated at NLO+PS level using two dedicated $bb4\ell$ MC samples generated with the nominal top-quark mass parameter $m_t = 172.5\,\text{GeV}$ and modified decay widths of $\Gamma_t = 1.0\,\text{GeV}$ and $\Gamma_t = 1.6\,\text{GeV}$. The resulting distributions are compared to those obtained from a reweighted nominal $bb4\ell$ MC sample, applying event weights determined during the POWHEG generation.[4] Good agreement between the distributions obtained from explicitly generated samples and those obtained with the reweighting procedure is observed, as shown in Figure 7.5 for two observables sensitive to the top-quark decay width. The use of event weights significantly reduces the computational resources required to evaluate $bb4\ell$ predictions for alternative top-quark decay widths.



(a) (b)

Figure 7.5: Comparison of $bb4\ell$ predictions with varied top-quark decay widths for (a) the $m_{\ell b}^{\text{minimax}}$ distribution from [78, 268] and (b) the reconstructed invariant top-quark mass, using the definitions from Section 3.1. The top-quark mass is fixed to $m_t = 172.5\,\text{GeV}$ in all $bb4\ell$ predictions. Results obtained from dedicated MC samples with explicitly set decay widths of $\Gamma_t = 1.0\,\text{GeV}$ and $\Gamma_t = 1.6\,\text{GeV}$ are compared to those derived using event weights applied to the nominal $bb4\ell$ MC sample to evaluate the same width variations (reweighted).

[4] These $bb4\ell$ MC samples with explicitly modified Γ_t are produced by varying the POWHEG input parameters `twidth` and `twidth_phsp`. The event weights to generate the Γ_t-varied MC samples through reweighting, were produced by varying the parameter `twidth` during the reweighting step (after the nominal event generation). Reweighting both the parameters `twidth` and `twidth_phsp` was also tested but resulted in significantly worse agreement between the explicitly generated and reweighted distributions and is therefore not used.

The comparison of the total MC prediction to data of the $m_{\ell b}^{\text{minimax}}$ distribution in the signal region is shown in Figure 7.6. The displayed uncertainty band includes all systematic uncertainties described in Section 8. A sizeable disagreement between data and the MC prediction is visible in the high-$m_{\ell b}^{\text{minimax}}$ tail, which is further discussed in the next section.

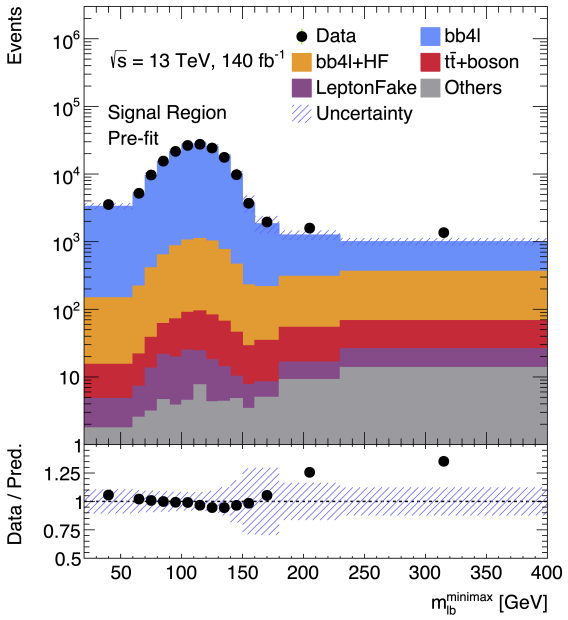


Figure 7.6: Comparison of the $m_{\ell b}^{\text{minimax}}$ distribution in the signal region as observed in data and as predicted by MC before the statistical fit. The uncertainty band corresponds to the full uncertainty in the MC predictions, including the systematic uncertainty estimates discussed in Section 8.

7.3 Controlling the additional heavy-flavour contribution

Since the $t\bar{t}$ contribution with additional heavy-flavour jets has been observed to be underestimated in several top-quark analyses [78, 269, 270], the modelling of the $bb4\ell + \text{HF}$ contribution is evaluated and corrected using data in a dedicated control region. In this region, the normalisation factor of the $bb4\ell + \text{HF}$ contribution is determined.[5] This correction is particularly important for the top-quark decay width measurement, as the relative contribution of $bb4\ell + \text{HF}$ events is enhanced in the high-$m_{\ell b}^{\text{minimax}}$ tail in the signal region, which is sensitive to variations of the top-quark decay width. Crucially, the shape of the $bb4\ell + \text{HF}$ fraction — i.e. $bb4\ell$ MC events with more than two particle-level b-jets (see Section 6.2.1) — follows the shape-variations expected from changes in the top-quark decay width. This behaviour is illustrated in Figure 7.7(a), which shows the fractions of $bb4\ell + \text{HF}$ MC events ($bb4\ell$+b-jets)

[5] The corresponding normalisation factor is free to vary in all analysis regions, with the strongest constraint expected to arise from the $bb4\ell + \text{HF}$ control region described in this section.

relative to the inclusive $bb4\ell$ prediction, meaning the sum of the $bb4\ell$ and $bb4\ell$+HF predictions. Additionally, the fraction of $bb4\ell$ events with at least one particle-level c-jet, but no additional particle-level b-jets, ($bb4\ell$+c-jets) relative to the inclusive $bb4\ell$ prediction is shown. In the following, the motivation for not including $bb4\ell$ MC events with additional particle-level c-jets in the $bb4\ell$ + HF event category is explained. While the contribution from $bb4\ell$+c-jets MC events in Figure 7.7(a) exhibits only a mild shape dependence and accounts for approximately 12 % of the inclusive $bb4\ell$ prediction in the signal region, the $bb4\ell$+b-jets contribution amounts to about 4 % in the bulk of the distribution and increases to approximately 30 % in the high-$m_{\ell b}^{\text{minimax}}$ tail. The strong increase of the $bb4\ell$+b-jets fraction at large $m_{\ell b}^{\text{minimax}}$ values is expected from the definition of the observable. If one of the two reconstructed b-jets does not originate from a top-quark decay, the corresponding $m_{\ell b}$ combination is no longer constrained by the kinematic end-point at $m_{\ell b}^{\text{end}}$, which is derived for LO top-quark pair decays. Consequently, even if the second $m_{\ell' b'}$ combination is correctly formed from a lepton and a b-jet originating from the same top-quark decay, the use of $\max(m_{\ell b}, m_{\ell' b'})$ will frequently select the $m_{\ell b}$ value that involves the additional heavy-flavour jet.

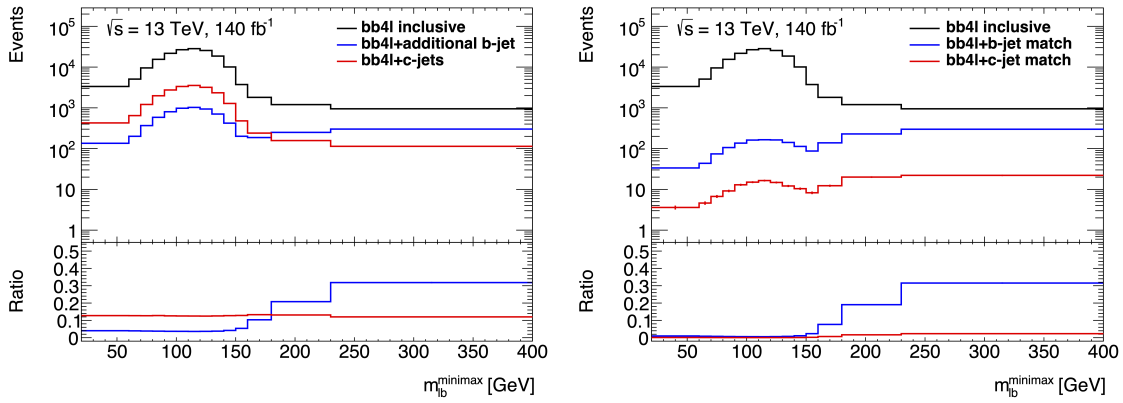


Figure 7.7: Comparison of the number of $bb4\ell$ MC events with additional particle-level heavy-flavour jets and the inclusive $bb4\ell$ MC prediction in the signal region. This comparison is shown using (a) for the fractions of $bb4\ell$ MC events with more than two particle-level b-jets ($bb4\ell$+additional b-jets) or exactly two particle-level b-jets and at least one particle-level c-jet ($bb4\ell$+c-jets). In this measurement, the $bb4\ell$ + HF contribution is defined as the $bb4\ell$+additional b-jets contribution. The comparison in (b) is performed by requiring that at least one of the reconstructed b-jets is not originating from the $WWbb$ process, whose truth type corresponds either to a b-jet ($bb4\ell$+b-jet match) or to a c-jet ($bb4\ell$+c-jet match).

It should be emphasized that the number of reconstructed b-jets in the signal region is exactly two and no explicit c-tagging information is used in the region definition. This means, that $bb4\ell$ events with additional particle-level heavy-flavour jets can only enter the signal region selection if either

- (I) the additional particle-level heavy-flavour jet is reconstructed and enters the b-jet object selection, but one of the b-jets from the $WWbb$ hard process does not or

- (II) both b-jets from the $WWbb$ hard process are reconstructed and enter the b-jet object selection, but the additional particle-level heavy-flavour jet does not.

To study whether the scenario (I) or (II) is more frequent, the fraction of $bb4\ell$ MC events is examined, in which at least one reconstructed and selected b-jet does not originate from the top-quark decay, or more precisely not from the $WWbb$ hard process.[6] This corresponds to the scenario (I). For this cross-check, reconstructed b-jets are matched to b-quarks from the hard process after final-state radiation within a cone of $\Delta R < 0.4$. The resulting fraction of $bb4\ell$ MC events in the signal region in which at least one selected (reconstructed) b-jet cannot be matched to a b-parton from the $WWbb$ hard process, but the truth type of the reconstructed jet corresponds to either a b-jet ($bb4\ell + b$-jet match) or a c-jet ($bb4\ell$+c-jet match),[7] is shown in Figure 7.7(b). Studying the $bb4\ell$+b-jet match and $bb4\ell$+c-jet match fractions allows the following conclusions:

- The fraction of events in which a particle-level c-jet is matched to a reconstructed b-jet (scenario (I)) is small, reaching only about 2 % in the high-$m_{\ell b}^{\text{minimax}}$ tail (Figure 7.7(b)). Consequently, for the majority of $bb4\ell + c$-jet MC events defined through the number of particle-level c-jets (Figure 7.7(a)), the reconstructed b-jets selected in the signal region do not originate from misidentified additional c-jets. In most cases, the additional particle-level c-jet is either not reconstructed, the reconstructed jet fails the kinematic selection requirements or is not misidentified as b-jet and therefore does not enter the b-jet object selection (scenario (II)). Consequently, and given the small shape dependence of the $bb4\ell$+c-jet fraction, this contributions is not included in the $bb4\ell + \text{HF}$ definition. Instead, an uncertainty of $\pm 50\,\%$ is assigned to $bb4\ell$ MC events containing exactly two particle-level b-jets and at least one particle-level c-jet.

- The $bb4\ell$+b-jet (Figure 7.7(a)) and the $bb4\ell$+b-jet match (Figure 7.7(b)) contribution in the high-$m_{\ell b}^{\text{minimax}}$ region is very similar. This indicates that indeed in this phase-space region one of the selected b-jets from the $bb4\ell + \text{HF}$ MC events, i.e. the $bb4\ell + b$-jet events, is not originating from the hard process. In the bulk of the distribution, however, significantly more $bb4\ell + b$-jet MC events than $bb4\ell + b$-jet match MC events are found. This can be explained similarly to the case of the $bb4\ell$+c-jets contribution, since in this phase-space region most additional particle-level b-jets appear not to be reconstructed or the reconstructed jet not to enter the object selection used in the analysis.

[6] The identification of b-partons from the hard process for $bb4\ell$ events projected on a tWb resonance history relies on the parton status codes in the event record and has been validated for the nominal Powheg+Pythia 8 $bb4\ell$ MC sample. For events projected onto a $t\bar{t}$ resonance history, the b-parton from the hard process is identified as the b-parton originating from the top-quark decay. Since this procedure depends strongly on the generator-specific event-record and status-code definitions, it is not used as the nominal $bb4\ell + \text{HF}$ classification scheme. This alternative classification requires a complete event record, corresponding in the $bb4\ell$ MC sample to either a $pp \to t\bar{t}$ or a $pp \to tWb$ resonance history. In the rare cases where the event record is incomplete — amounting to only 0.003 % of SR events — the default classification based on counting truth heavy-flavour jets is used instead.

[7] The truth-flavour classification of the reconstructed jets is obtained from a centrally provided variable used in the ATLAS flavour tagging group, which matches truth heavy-flavour hadrons with $p_{\text{T}} > 5\,\text{GeV}$ to reconstructed jets within a ΔR cone of 0.3. If both, a D-hadron and a B-hadron are matched to the same reconstructed jet, the jet is labelled as a truth b-jet.

In the $bb4\ell$ + HF control region, defined by the selection criteria in Section 7.1, the relative contribution of $bb4\ell$ MC events with additional particle-level heavy-flavour jets with respect to the inclusive $bb4\ell$ prediction is also evaluated. These fractions amount to 92 % for events with additional particle-level b-jets and 6 % for events with additional particle-level c-jets. The same fractions are obtained when requiring explicitly that at least one of the reconstructed b-jets is not matched to a b-parton from the hard process, but its truth flavour corresponds either to a b-jet or c-jet ($bb4\ell$+b-jet match and $bb4\ell$+c-jet match). Since the selection in this control region requires at least three reconstructed b-jets, it is plausible that the same results are obtained when either requiring additional particle-level heavy flavour jets to be present in the event or requiring that these particle-level jets are matched to at least one of the reconstructed b-jets.[89] The comparison between the nominal MC prediction and data in the control region is presented in Figure 7.8. A significant discrepancy between data and MC is observed, which is accounted for by introducing a free-floating normalisation factor for the $bb4\ell$ + HF contribution.

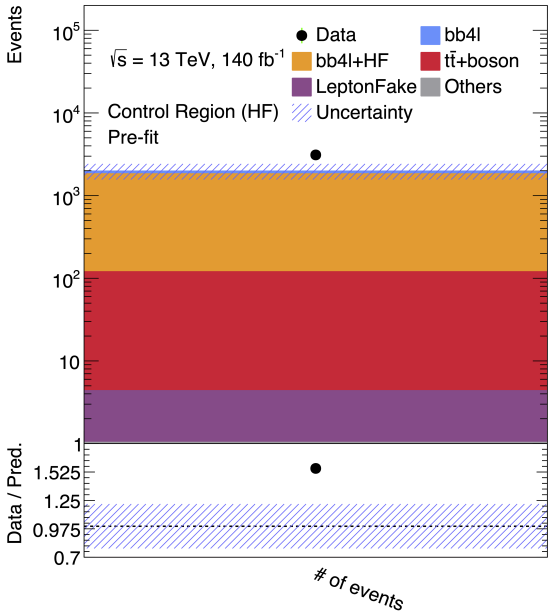


Figure 7.8: Same comparison as in Figure 7.6, but for the number of events entering the $bb4\ell$ + HF control region of the measurement.

[8] These fractions could differ if e.g. an additional particle-level heavy flavour jet is not reconstructed or not selected, but instead a light jet is misidentified as b-jet. The fraction of these events is expected to be negligble.

[9] This comparison should highlight the impact of the $bb4\ell$ + HF truth classification defintion, as the $bb4\ell$ + HF normalisation factor is predominantly determined in the $bb4\ell$ + HF control region, where using either the $bb4\ell$+b-jet (Figure 7.7(b)) or the $bb4\ell$+b-jet match (Figure 7.7(b)) definitions lead to nearly identical results, while the corresponding fractions in the bulk of the $m_{\ell b}^{\text{minimax}}$ distribution in the signal region will differ significantly.

As this control region targets only the overall scaling of the $bb4\ell$ + HF fraction, no shape information is used. Instead, the total number of events in the region is considered. No dependence on the top-quark mass or decay width is assumed for the event yield, as only shape differences introduced by m_t and Γ_t variations are considered in the analysis.[10]

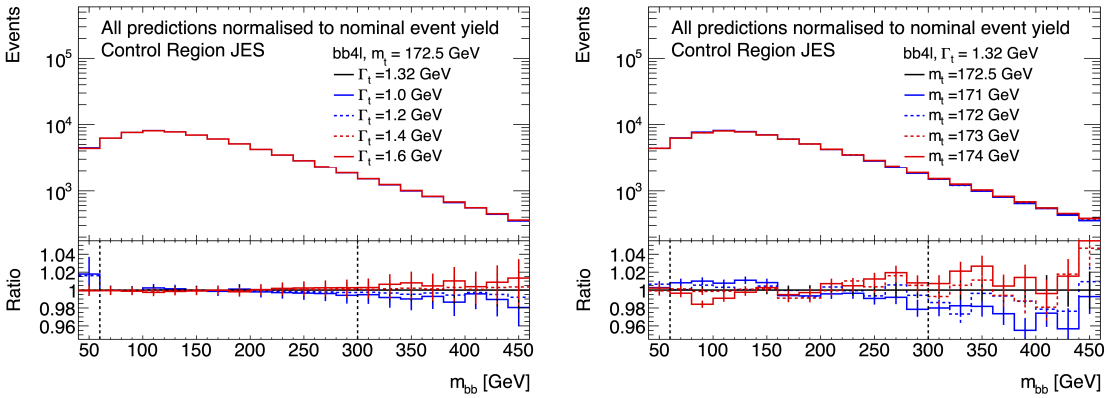
7.4 Controlling jet energy scale uncertainties

An additional control region is introduced to constrain systematic uncertainties related to the jet energy scale (JES), which were found to have a significant impact on the sensitivity of top-quark mass measurement, as will be discussed in Section 10. To avoid any statistical overlap with the signal region instead of selecting events in the different-lepton-flavour ($e\mu$) decay channel, the same-lepton-flavour (ee and $\mu\mu$) channels are used, as described in Section 7.1. To still probe jets with similar kinematic properties, the b-jet requirements are kept identical to those used in the signal region definition. As an observable sensitive to the jet energy scale, the invariant mass of the two b-jets, m_{bb}, is studied in the control region. An important design criterion for this region is that it should be largely independent of the top-quark mass and decay width. This requirement is fulfilled for the chosen range $m_{bb} \in [60, 300]\,\text{GeV}$. As shown in Figure 7.9, only a very small dependence on the top-quark decay width is observed at values of $m_{bb} > 300\,\text{GeV}$.[11] Consequently, no dependence of the event yields per m_{bb}-bins on m_t and Γ_t is considered in the statistical fit.[12] Figure 7.10 shows good agreement between data and the nominal MC prediction in this control region within systematic uncertainties.

[10] If changing the top-quark mass or the top-quark decay width parameter would lead to differences in the $bb4\ell$ + HF modelling and thus to different fractions of events in the signal and the control region, this constant parameterisation would not be appropriate. That these fractions do not depend on the top-quark mass and decay width was explicitly checked.

[11] The dependence on the top-quark mass is evaluated using separate MC samples and is therefore more strongly affected by statistical fluctuations.

[12] In Section 9.3, it is explicitly shown that including or not including the m_t and Γ_t event yield parameterisation in the CRs has a negligible influence on the performed closure tests.



(a) (b)

Figure 7.9: Same comparison as in Figure 7.3, but for the predicted $bb4\ell$ contribution in the m_{bb} distribution of the JES control region. The m_{bb} range of $[60, 300]$ GeV in-between the dotted vertical lines in the ratio panel is used as control region.

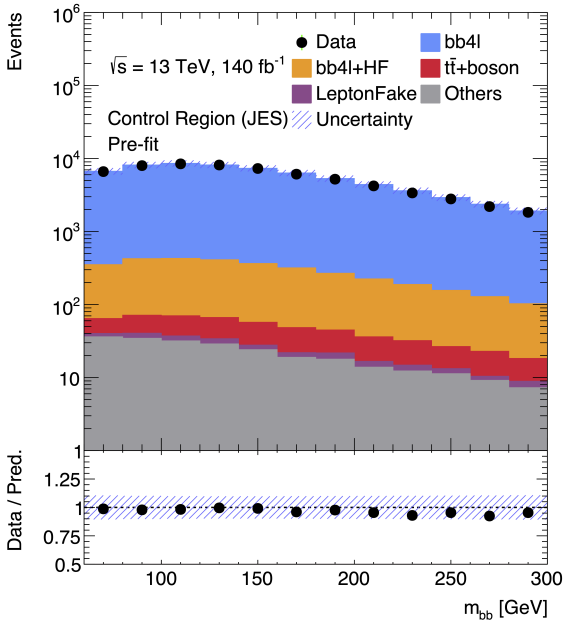


Figure 7.10: Same comparison as in Figure 7.6, but for the m_{bb} distribution in the JES control region.

CHAPTER 8

Systematic uncertainties

Several sources of systematic uncertainties are considered in this analysis, which can be grouped into experimental uncertainties and modelling uncertainties. The experimental uncertainties are related to the calibration of physics objects and the luminosity measurement, while the modelling uncertainties are evaluated by variations of parameters in the MC generation that cannot be predicted from first principles, as well as uncertainties in the background normalisations. These systematic uncertainty sources are discussed in Section 8.1 and Section 8.2, respectively. The systematic uncertainties affect both the shape and the normalisation of the total MC prediction, which consists of the signal and background process predictions.[1] Also in the case of background normalisation uncertainties both the shape and the normalisation of the total MC prediction is affected, since variations of the relative number of MC events per category, which are described in Section 6.2, lead to shape variations in the overall prediction.

8.1 Experimental uncertainties

The experimental systematic uncertainties are determined for both, the signal and background MC predictions and are treated as correlated between the different event categories and analysis regions. These uncertainties are associated with the correction factors obtained from the calibration measurements of physics objects as described in the following. Additionally, a total uncertainty in the calibrated luminosity of the full Run 2 data set of 0.83 % [238] is adopted.

[1] The exception is the luminosity uncertainty, which only affects the normalisation of the total MC prediction.

8.1.1 Jet uncertainties

The calibration of PFlow jets [271] is performed through a sequence of steps. First, jets are calibrated using simulation-based corrections, which include pile-up corrections and the absolute MC-based jet energy scale (JES) corrections. After these steps, the jets are defined at the MCJES scale. This is followed by the final MC-based correction, the global sequence calibration (GSC), which reduces the jet energy scale dependence on the flavour of the parton initiating a jet and fluctuations in the particle shower development. The simulation-based corrections are followed by in-situ measurements to account for differences in the jet energy scale between data and MC and to measure the jet energy resolution (JER). Since systematic uncertainties associated to the jet reconstruction are among the dominant contributions in the top-quark mass measurement, the origin and evaluation of the JES and JER uncertainties derived in [271] are discussed in detail in the following. The impact of using the uncertainties provided by the central recommendations on the signal region in this analysis is illustrated for a few cases.

The goal of the JES calibration, discussed first, is to correct the energy of the reconstructed jets such that it corresponds to the energy of the associated truth jet. These truth jets are constructed by including all simulated stable particles, except for muons, neutrinos and particles from pile-up interactions, in the anti-k_t jet clustering algorithm with $R = 0.4$.

Pile-up corrections The first step in the jet energy scale calibration is the pile-up subtraction based on the jet transverse-momentum density (p_{T}/A). The jet area A entering this density definition is determined from the number of ghost particles[2] clustered into a jet using the k_t-algorithm [272] with a radius parameter $R = 0.4$.[3] The per-event pile-up density ρ is defined as the median jet-p_{T} density, i.e. the median of all values $p_{\mathrm{T}}^{\mathrm{jet},i}/A_{\mathrm{jet},i}$ for jets with $|\eta| < 2$ in a given event. Only jets in the central region are considered, as ρ decreases nearly to zero in the forward region due to the lower occupancy in this detector region with coarser segmentation [261]. Jets originating from the hard-scattering process are expected to populate the high-ρ tail of this distribution and therefore have only a limited impact on the median value. Consequently, ρ predominantly reflects the average transverse-momentum contribution per unit jet area from pile-up interactions. The pile-up corrected transverse momentum of each jet is then given by $p_{\mathrm{T}}^{\mathrm{corr}} = p_{\mathrm{T}}^{\mathrm{reco}} - \rho \times A$. This correction is applied to all jets, not only the central jets used in the determination of ρ, by scaling the jet four-momentum by the ratio $p_{\mathrm{T}}^{\mathrm{corr}}/p_{\mathrm{T}}^{\mathrm{reco}}$ while leaving the jet direction unchanged. An uncertainty associated with the ρ-based pile-up subtraction is derived by comparing differences in the calculated ρ values between different event topologies (Z+jets, γ+jets and dijets) in MC and data. For example,

$$\Delta_\rho = (\rho^{Z+j} - \rho^{\gamma+j})_{\mathrm{MC}} - (\rho^{Z+j} - \rho^{\gamma+j})_{\mathrm{data}}. \tag{8.1}$$

[2] See the definition under Figure 4.14.

[3] While the anti-k_t jet clustering algorithm is typically used in data analyses, the k_t-algorithm is employed here exclusively for the p_{T}-density-based pile-up subtraction due to its increased sensitivity to soft radiation, see also Section 4.4.1.

This difference is translated into an uncertainty in the calibrated jet transverse momentum according to

$$\Delta_{p_{\mathrm{T}}} = \max(|\Delta_\rho|) \times C_{p_{\mathrm{T}}}^{\mathrm{JES}} \times \pi R^2, \tag{8.2}$$

where $C_{p_{\mathrm{T}}}^{\mathrm{JES}}$ is a correction factor used to define the jet p_{T} at the MCJES scale discussed below, and R is the jet radius parameter. This jet uncertainty constitutes the dominant systematic uncertainty in the top-quark mass measurement when the jet pre-recommendation described below are excluded and is discussed further in the corresponding Section B.2.

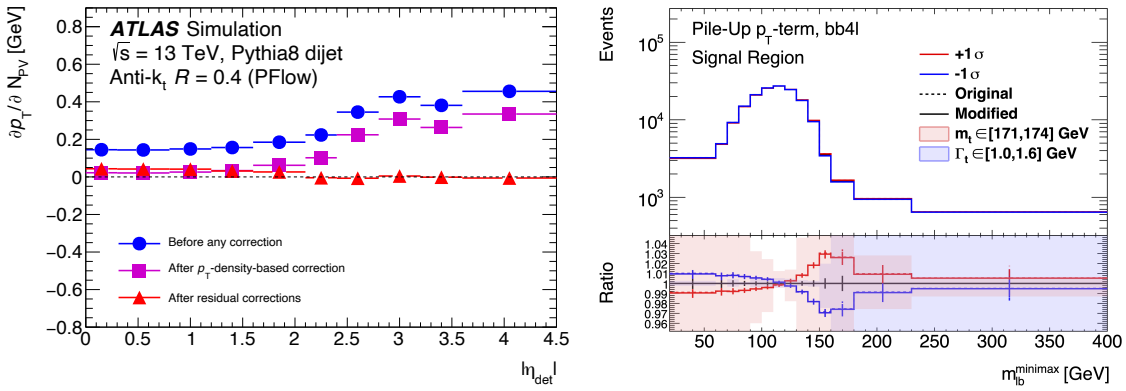
After the pile-up density correction, a residual dependence of the difference between the reconstructed jet p_{T} and the truth jet p_{T} on the number of primary vertices (N_{PV}) and the mean number of interactions per bunch crossing (μ) is observed. This behaviour reflects the sensitivity to in-time and out-of-time pile-up,[4] respectively, since N_{PV} is reconstructed using tracks, which are only sensitive to in-time pile-up [273]. The residual dependence is especially pronounced for jets at large pseudorapidities ($|\eta| > 2$), as shown in Figure 8.1(a). This is plausible, since the ρ-based pile-up subtraction is derived only in the central regions of the calorimeter. Including both the p_{T}-density-based and the residual pile-up corrections, the corrected jet p_{T} is given by

$$p_{\mathrm{T}}^{\mathrm{corr}} = p_{\mathrm{T}}^{\mathrm{reco}} - \rho \times A - \alpha \times (N_{\mathrm{PV}} - 1) - \beta \times \mu. \tag{8.3}$$

The coefficients α and β are determined from independent linear fits in bins of the truth jet transverse momentum $p_{\mathrm{T}}^{\mathrm{truth}}$, using geometrical matching between the truth and reconstructed jets with $\Delta R < 0.3$, and the absolute pseudorapidity $|\eta_{\mathrm{det}}|$, defined with respect to the geometric centre of the detector. A p_{T}-dependent systematic uncertainty is assigned to these residual corrections by fitting the observed logarithmic dependence of the α and β-coefficients on $p_{\mathrm{T}}^{\mathrm{truth}}$. The nominal α and β values in each $|\eta_{\mathrm{det}}|$ region are taken from the logarithmic fit evaluated at $p_{\mathrm{T}}^{\mathrm{truth}} = 25\,\mathrm{GeV}$, where pile-up effects are most relevant. The difference on the residual pile-up correction in dependence of p_{T} between using this nominal parameterisation or the full logarithmic dependence of α and β on $p_{\mathrm{T}}^{\mathrm{truth}}$ is assigned as a systematic uncertainty. The impact of this uncertainty on the $m_{\ell b}^{\mathrm{minimax}}$ distribution of the $bb4\ell$ contribution in the signal region is illustrated in Figure 8.1(b).

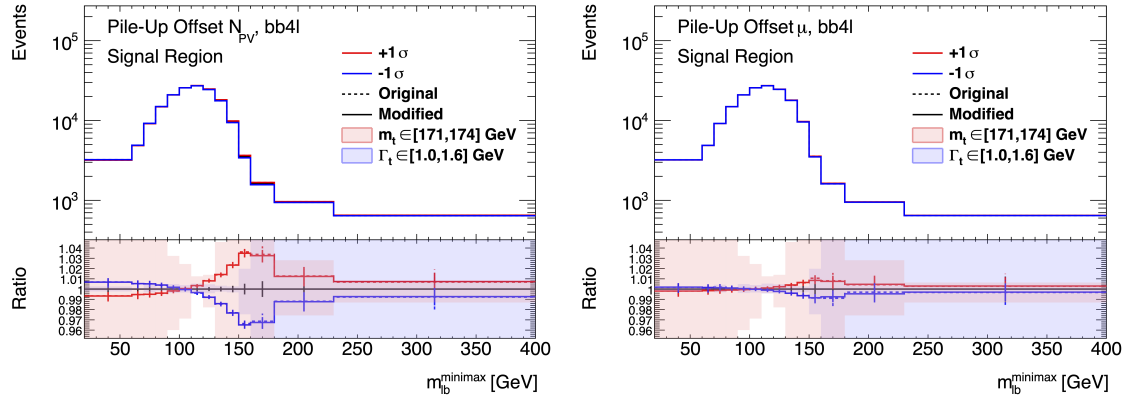
In addition, two further uncertainties are considered for the pile-up p_{T} correction, based on in-situ measurements of the jet transverse momentum dependence on N_{PV} and μ. These in-situ studies use either jets only reconstructed from tracks, which are largely insensitive to pile-up, or Z+jets events, exploiting the p_{T} balance between a reconstructed jet and the Z boson. The small observed non-closure in the pile-up dependence on N_{PV} and μ between MC and data is taken as a systematic uncertainty. The impact of these uncertainties on the $bb4\ell$ contribution in the SR of this analysis is shown in Figure 8.2. The shape of the impact of these two uncertainties is similar, while the uncertainty associated with the in-time pile-up dependence correction is larger in magnitude.

[4] See Section 4.2 for the definition of in-time and out-of-time pile-up.



(a) (b)

Figure 8.1: (a) Dependence of the reconstructed jet p_{T} on in-time pile-up — parameterised through the number of reconstructed primary vertices N_{PV} — with and without pile-up corrections as a function of $|\eta_{\mathrm{det}}|$ [271]. (b) Impact of the p_{T}-dependent uncertainty on the JES arising from the residual pile-up subtraction on the $m_{\ell b}^{\mathrm{minimax}}$ distribution of the $bb4\ell$ contribution in the signal region. The $\pm 1\sigma$ variations are shown as dashed lines, while the output of the smoothing algorithm described in Section 9.1, used to reduce the impact of statistical fluctuations in the templates for systematic variations, is shown as solid lines. Additionally, the effects of top-quark mass and decay width variations are illustrated by the red and blue shaded bands, respectively.



(a) (b)

Figure 8.2: Same comparison as in Figure 8.1(b), but for the JES pile-up uncertainties parameterising the non-closure of the (a) N_{PV} and (b) μ jet p_{T} dependence found between data and MC in in-situ measurements.

Absolute jet energy scale and η calibration In the next step, the reconstructed jet energy E^{reco} and reconstructed pseudorapidity η^{reco} are corrected to the corresponding truth quantities E^{truth} and η^{truth}, as predicted by PYTHIA 8 dijet MC simulations, after the pile-up corrections described above have been applied. This calibration accounts for the non-compensating calorimeter response,[5] energy losses in passive material, out-of-cone radiation, and biases in the jet pseudorapidity reconstruction. These latter biases arise mainly due to transition regions between calorimeter subsystems and from differences in calorimeter read-out granularity. To correct for the absolute jet energy scale, the jet energy response $\mathcal{R}$ is determined in MC as the mean of a Gaussian fit to the $E^{\text{reco}}/E^{\text{truth}}$ distribution in bins of energy and pseudorapidity.[6] The jet energy calibration factor is then defined as the inverse of the jet energy response $\mathcal{R}$, and is applied to the reconstructed jet energy. An additional bias in the reconstructed jet pseudorapidity is observed, in particular in the detector transition regions of the barrel and endcap calorimeters ($|\eta^{\text{reco}}| \sim 1.4$) and the endcap and forward calorimeters ($|\eta^{\text{reco}}| \sim 3.1$). This bias is removed by applying a second correction based on $\eta^{\text{reco}} - \eta^{\text{truth}}$, derived in bins of energy and pseudorapidity. The calibration procedure described above only modifies the jet p_{T} and the jet η, not the full jet four-momentum. Reconstructed jets, after these corrections are applied, are defined at the MCJES scale.

Global sequential calibration Even after the application of the previous JES calibrations steps, the jet energy response can still vary for a given jet p_{T} and η due to differences in the properties of quark- and gluon-initiated jets. Quark-initiated jets often contain hadrons carrying a high momentum fraction of the jet momentum and therefore traverse large parts of the calorimeter system. In contrast, gluon initiated jets tend to consist of a larger number of lower-p_{T} particles, resulting in broader shower shapes. To account for these differences, the global sequential calibration applies a sequence of multiplicative corrections based on observables sensitive to the jet shape and constituent composition. In total, six variables are used to parameterise the jet response. For each observable, a correction to the jet four-momentum is derived in bins of p_{T} and η, similar to the method in the MCJES calibration. The GCS does not change the average jet energy response, but instead improves the jet energy resolution (JER) $\sigma_{\mathcal{R}}$. The resolution is defined as the standard deviation of a Gaussian fit to the jet p_{T} response, which is given by $p_{\text{T}}^{\text{reco}}/p_{\text{T}}^{\text{truth}}$.

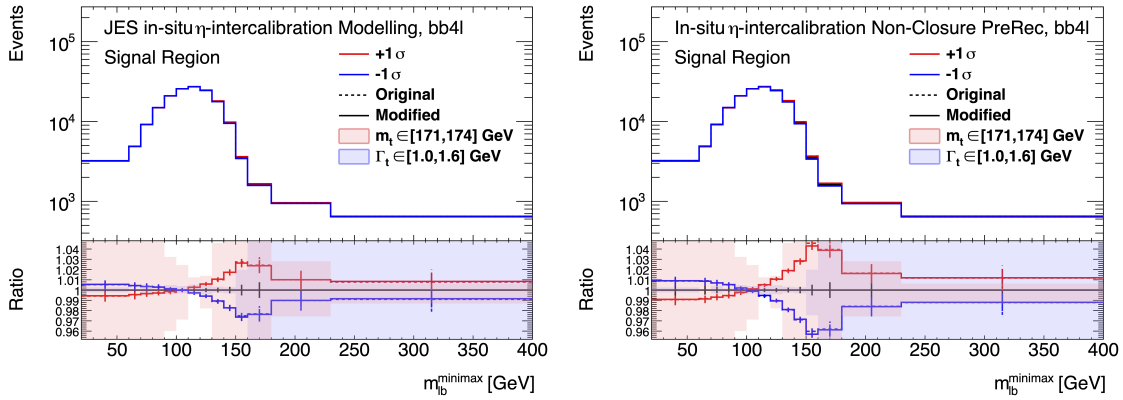
Jet energy scale in-situ calibration After the MCJES and GSC calibration steps, the in-situ jet calibration corrects for differences in the jet energy response in data and the calibrated simulation. This is achieved by correcting the jet energy scale in data using measurements that compare the jet p_{T} to that of well-calibrated recoiling reference objects or system. The jet

[5] A non-compensating calorimeter response means that the calorimeter does not measure hadronic and electromagnetic shower components equally, leading to a reduced response for hadrons compared to electromagnetic energy deposits. Consequently, the fraction of energy carried by the hadronic component is underestimated. See also the discussion in Section 4.3.5.

[6] Instead of deriving $\mathcal{R}$ in bins of the reconstructed energy and pseudorapidity, $\mathcal{R}$ is derived in bins of the truth particle energy and reconstructed pseudorapidity, as the response follows a Gaussian distribution in the latter but not in the former. A numerical inversion procedure is applied to obtain $\mathcal{R}$ in bins of the reconstructed quantities.

response measured in data and MC is compared via $c = \mathcal{R}^{\text{data}}_{\text{in-situ}}/\mathcal{R}^{\text{MC}}_{\text{in-situ}}$, where the in-situ calibration consists of applying the derived correction factor c^{-1} formulated in terms of the jet p_{T} (and jet η where applicable) to data. The analysis presented in this thesis uses reprocessed Run 2 data, for which no dedicated in-situ jet calibrations are available yet. Instead, the Run 2 in-situ calibration derived with the previous ATLAS reconstruction software release is applied. Additional preliminary, conservative (*pre-recommendation*) uncertainties [274] are included to cover differences between the two reconstruction formats.

Three in-situ measurements are performed sequentially. First, the η intercalibration corrects the energy scale of forward jets to match that of the more precisely calibrated central jets using the p_{T} balance in dijet events. Systematic uncertainties are derived to account for mismodelling of physics objects, detector and event topology effects impacting the dijet momentum balance. These uncertainties are parameterised in bins of p_{T} and η and combined into a single systematic uncertainty per bin. The dominant contribution to the η intercalibration uncertainty arises from modelling uncertainties and is estimated by comparing the calibration results obtained with Powheg+Pythia 8 and Sherpa dijet samples. Statistical uncertainties are considered due to the finite size of the data and MC samples. In addition, a non-closure uncertainty is assigned by comparing the jet response in data and MC after applying the η intercalibration. Since the η intercalibration derived with the previous reconstruction software version is also used for the reprocessed Run 2 data, an additional *pre-recommendation* non-closure uncertainty is included to account for differences between the two reconstruction versions [274]. The impact of the modelling and pre-recommendation non-closure uncertainties on the $m^{\text{minimax}}_{\ell b}$ distribution of the $bb4\ell$ contribution in the signal region is shown in Figure 8.3.



(a) (b)

Figure 8.3: Same comparison as in Figure 8.1(b), but for (a) the JES in-situ η intercalibration modelling and (b) the η intercalibration non-closure pre-recommendation uncertainty.

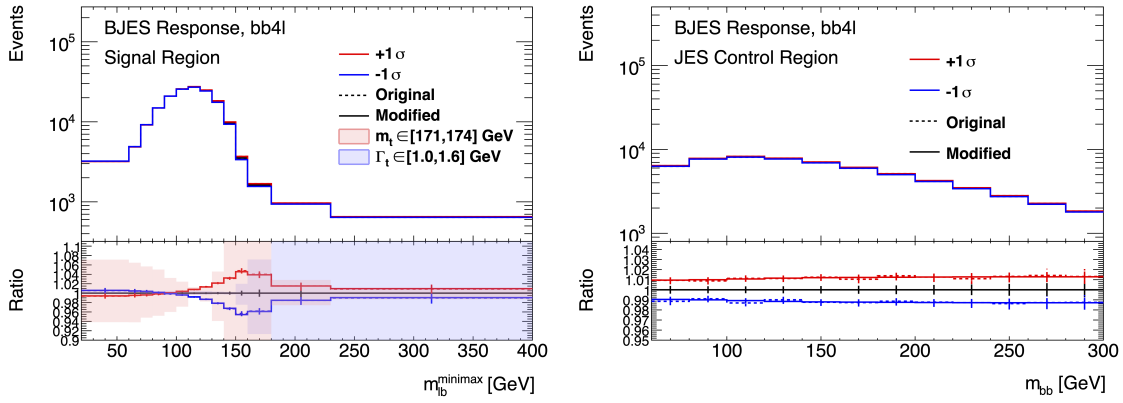
The second in-situ measurement uses the missing-E_{T} projection fraction (MPF) method in Z/γ+jets events. This method measures the full hadronic recoil of an event relative to the p_{T} of a calibrated photon or leptonically decaying Z boson. By using the full hadronic recoil instead of a single jet, the sensitivity to pileup, the underlying event — as on average the

total p_{T} of these contributions cancels out directionally — and jet definitions is reduced. This technique only uses central jets, but through the η intercalibration the derived corrections can also be applied to forward jets. Thereby, the calibration is derived for jets in a p_{T} range of $17\,\mathrm{GeV} < p_{\mathrm{T}} < 1.2\,\mathrm{TeV}$. The third and final in-situ measurement targets high-p_{T} jets by comparing the p_{T} of a single high-p_{T} jet to that of recoiling low-p_{T} jets, which have already been well calibrated by the preceding steps. This method extends the jet p_{T} calibration to 2.4 TeV. For both of these in-situ measurements, modelling uncertainties are estimated through generator comparisons, comparing the corrections derived either with Powheg+Pythia 8 or Pythia 8, depending on the physics process, to the corrections obtained using the Sherpa generator. Additional systematic uncertainties account for the dependence on analysis selection cuts, event topology, statistical uncertainties and calibration uncertainties of the reference physics objects, such as leptons and photons. The independent systematic uncertainties from the in-situ calibrations, each parameterised as a function of jet p_{T}, are combined using an eigenvector decomposition [275] in the category reduction scheme based on their origin (detector, statistical, modelling or mixed). Among these uncertainty components, the eigenvector associated to the generator modelling uncertainty in the in-situ measurements is the most important in the analysis presented in this thesis and is discussed further in Section B.2.

Additional pre-recommendation uncertainties are considered for the in-situ jet systematic uncertainties when using reprocessed Run-2 data [274]. Compared to the earlier reconstruction software version, the reprocessed data use updated topo-cluster noise thresholds and a modified primary-vertex reconstruction algorithm (see Section 4.3.3). To account for these changes, systematic uncertainties are derived from differences in jet response when varying only the noise thresholds or only the vertexing algorithm in the reprocessed data. Furthermore, in-situ non-closure uncertainties are applied by comparing the jet-energy response between samples reconstructed with the two software versions. These pre-recommendation uncertainties currently represent the dominant systematic uncertainty in the top-quark mass measurement and are discussed further in Section 10.1.

Finally, three additional uncertainties are included to account for the dependence of the jet response on the flavour content of a jet: the flavour composition, the flavour response uncertainty, and a dedicated uncertainty considered for b-quark initiated jets. The flavour composition uncertainty accounts for differences in the jet response between gluon and quark-initiated jets by considering an uncertainty in the simulated gluon-to-quark jet fraction. The flavour response uncertainty captures differences in the jet response of gluon jets simulated with Pythia and Herwig, while no dedicated uncertainty is assigned for light-quark jets due to their similar response when varying the parton shower model. For heavy-flavour jets (b, c), these two uncertainties are replaced by a dedicated uncertainty accounting for response differences for light (u, d, s) and heavy-flavour jets. While the light-quark and gluon jet uncertainties have negligible impact on the measurement presented in this thesis, the uncertainty in the b-jet energy scale response is significant. While in the signal region its influence on the $bb4\ell$ prediction induces a shape effect, similar to the shape effect expected from top-quark mass variations, in the JES control region the b-jet energy scale response uncertainty leads mainly to a normalisation effect and cannot be strongly constrained, as

shown in Figure 8.4.



(a) (b)

Figure 8.4: Same comparison as in Figure 8.1(b), but for the impact of the b-jet energy scale (BJES) response uncertainty on the $bb4\ell$ prediction in (a) the signal and (b) the JES control region.

The origin of the flavour uncertainties has been studied in [276] and an alternative evaluation scheme was introduced in [277]. This was found to reduce the flavour uncertainty in a recent top-quark mass measurement in ATLAS [125]. Applying this new scheme to evaluate the flavour uncertainties in the jet-calibrations of reprocessed Run 2 data is expected to further reduce the uncertainty in the measurement presented in this analysis. Finally, additional uncertainties in high-p_{T} jets [278] and jets with *punch-through* [279] are considered, which do not have a large impact in this analysis.

Jet energy resolution In simulation, the JER $\sigma_{\mathcal{R}}$ is corrected to match the resolution observed in data with in-situ measurements. Correction factors depend on the jet transverse momentum and pseudorapidity. A precise determination of the JER is important for data analyses involving jets or missing transverse momentum. The JER is measured using the *dijet balance* method relying on an approximate scalar balance between the transverse momenta of the two leading, i.e. highest p_{T}, jets in dijet events is assumed ($p_{\mathrm{T}}^{\mathrm{jet1}} \approx p_{\mathrm{T}}^{\mathrm{jet2}}$). The JER is determined both in MC and data in bins of p_{T} and η. If the JER measured in data is larger than in MC, the jets in the MC sample are additionally smeared until their average resolution matches the one observed in data. If the JER in MC is larger than in data, no smearing is applied, as the data should not be modified. Instead, the full difference $\sigma_{\mathrm{syst}} = \sigma^{\mathrm{data}} - \sigma^{\mathrm{MC}}$ is taken as a systematic uncertainty. Systematic uncertainties in the JER determination are estimated by varying the dijet selection criteria, using alternative MC generators and by modifying the noise term that contributes to the transverse momentum dependence of the resolution $\sigma(p_{\mathrm{T}})/p_{\mathrm{T}}$ as $1/p_{\mathrm{T}}$. The noise term is determined using a separate in-situ method and varied within uncertainties evaluated with this method. These uncertainties are propagated to physics analyses by applying an additional Gaussian smearing to the jets p_{T}

with the width σ_{smear}, defined as

$$\sigma^2_{\text{smear}} = \left(\sigma_{\text{nominal}} + |\sigma_{\text{syst}}|\right)^2 - \sigma^2_{\text{nominal}}, \tag{8.4}$$

where σ_{nominal} corresponds to the nominal JER determination, and σ_{syst} denotes the variation in resolution when considering a given systematic. If $\sigma_{\text{syst}} > 0$, the smearing is applied to jets in simulation. If $\sigma_{\text{syst}} < 0$, this is instead performed in data. As this procedure is only used to evaluate the systematic uncertainty, the nominal data sample remains unchanged. In practice, instead of smearing the data sample, a pseudo-data sample constructed from simulation can be smeared to evaluate the effect of the uncertainty. This latter recommendation is used in the present analysis. In total, 13 independent uncertainties related to the JER measurement are included in this analysis. Thereby, an eigenvector decomposition similar to the case of the JES uncertainties is applied to reduce the number of final independent uncertainties, which need to be considered. Additionally, a JER pre-recommendation uncertainty is added, which takes into account the changed noise threshold of the software version used in reprocessed Run 2 data and the reconstruction software, which was used to perform these JER in-situ measurements [274]. In general, the impact of the JER uncertainties on the top-quark mass and decay width measurement is found to be small.

NNJVT tagger The calibration of the NNJVT hard-scatter jet efficiency is performed based on the procedure used for the JVT tagger [261]. The efficiency measured in simulation is compared to that observed in data using $Z(\rightarrow \mu\mu)$+jets events, and the resulting scale factors are applied to simulated events. Systematic uncertainties in the calibration measurement related to pile-up modelling, variation of selection criteria and different generator choices are accounted for by varying the scale factor applied to simulation.

8.1.2 Flavour tagging uncertainties

Due to imperfect modelling in the MC event generation and in the simulation of the detector response, differences arise between data and simulation in the input variables to the GN2v01 flavour-tagging algorithm, and consequently in the classifier response. These differences are accounted for by deriving correction factors for each GN2v01 working point [188], which are applied to simulated events to align the tagging efficiencies of b-jets and mis-tagging efficiencies of c-jets and light jets between data and simulation [280–282]. These correction factors are derived in bins of jet p_{T}. The calibration of b-jets and c-jets is performed using $t\bar{t}$ events in dilepton and single-lepton ($W \rightarrow cs$) final states, respectively. In both calibration measurements, the dominant systematic uncertainty arises from the modelling of $t\bar{t}$ production. The calibration of light jets is carried out using Z+jets events. Due to the strong background rejection of light jets by the flavour-tagging algorithm, the fraction of light jets in the calibration regions would be small compared to the contributions from b-jets and c-jets, making the calibration challenging. To address this, an alternative tagger based on transformed track variables is employed, which reduces the b-jet and c-jet efficiencies while leaving light-jet efficiencies approximately unchanged. The mis-tagging rate of light jets is derived using this alternative tagger and dedicated uncertainties are introduced to account for

the extrapolation to the nominal GN2v01 tagger. These extrapolation uncertainties lead to the most dominant source of systematic uncertainties in the light jet calibration. To reduce the number of systematic uncertainties related to the flavour tagging efficiency calibration in data analyses, an eigenvector decomposition is performed, similar to the case of the JES and JER uncertainties, with 129 uncorrelated systematic uncertainty components: 55 from the b-jet calibration, 24 from the c-jet calibration, 42 from the light jet calibration and 8 from the extrapolation of the calibration results to high transverse momentum jets. Additionally, MC-to-MC scale factors are determined to align the efficiencies in alternative generators to the nominal generator Powheg+Pythia 8 used in the calibration. The uncertainty corresponding to the first eigenvector of the systematic uncertainty decomposition of the b-jet efficiency is further discussed in Section 10.2.

8.1.3 Lepton uncertainties

Systematic uncertainties associated with the reconstructed lepton objects include uncertainties in the determination of the trigger, reconstruction, identification and isolation efficiencies, as well as uncertainties related to the lepton energy or momentum scale and resolution. The impact of these systematic uncertainties on the extraction of the top-quark mass and the decay width in this thesis is generally found to be small.

Electron calibration and uncertainties The calibration of the electron energy scale and resolution is performed using $Z \to ee$ events [190]. The thereby derived energy scale correction is applied to data, while a resolution correction is applied to simulated events. Systematic uncertainties in the derivation of these correction factors arise from pile-up modelling, calorimeter-layer inter-calibration, and an uncertainty estimated by applying an alternative correction factor extraction method. These uncertainties are propagated as systematic variations on the electron energy scale and resolution. Scale factors correcting the trigger, reconstruction, identification and isolation efficiencies [190, 263] are measured by comparing data and MC simulation in $Z \to ee$ and $J/\psi \to ee$ events and are applied to simulated events. Systematic uncertainties associated with the determination of these correction factors are accounted for in the analysis presented in this thesis by applying the provided scale factor variations.

Muon calibration and uncertainties The muon momentum scale and resolution in simulated events are corrected as derived from the comparison of $J/\psi \to \mu\mu$ and $Z \to \mu\mu$ decays in data and in simulation [283]. Uncertainties in the correction factors are considered in this analysis and include uncertainties related to the event selection, observable definitions, background contributions, and modelling of the decay and final state radiation in the calibration procedure. In addition, residual detector misalignment effects, which introduces a charge-dependent bias on the muon momentum measurement, are corrected in data events, while simulated events assume ideal alignment. Scale factors correcting the reconstruction, identification, isolation, vertex association and trigger efficiencies of simulation events [192, 264] are derived by comparing efficiencies in simulation and data using $Z \to \mu\mu$ and $J/\psi \to \mu\mu$ events. Uncertainties in

the determination of these scale factors are accounted for by applying alternative correction factors to the muon objects in the analysis presented in this thesis.

8.1.4 Missing transverse momentum uncertainties

Systematic uncertainties in the $E_{\mathrm{T}}^{\mathrm{miss}}$ determination [193] are derived from the uncertainty in the p_{T} calculation for both the hard term, calculated with the reconstructed physics objects, and the soft term reconstructed with charged tracks (see Section 4.4.5). The systematic uncertainties described above, related to the energy or momentum scale of the lepton and jet objects, are propagated to the $E_{\mathrm{T}}^{\mathrm{miss}}$ calculation. Dedicated systematic uncertainties for the soft term are derived using $Z \to \mu\mu$ events with no additional jets, in which the total transverse momentum of the hard process $(\vec{p}_{\mathrm{T}}^{\,Z})$[7] should be balanced by the soft term $(\vec{p}_{\mathrm{T}}^{\,\mathrm{soft}})$ in case of perfect detector resolution and full phase space coverage. Systematic uncertainties in the $E_{\mathrm{T}}^{\mathrm{miss}}$ response and the resolution are derived by comparing data to MC simulation. Three systematic uncertainty components are derived, considering the offset of the soft p_{T} component w.r.t. the axis of the hard (Z) p_{T} component, as well as a resolution smearing along both the longitudinal and the perpendicular axis of this projection.

8.2 Modelling uncertainties

Systematic uncertainties estimated with variations in the MC generator settings are evaluated for the $WWbb$ signal process, while normalisation uncertainties are assigned to the background processes. These uncertainties are described in Section 8.2.1 and Section 8.2.2, respectively.

8.2.1 Signal modelling uncertainties

As the definition of the modelling uncertainties in the $bb4\ell$ MC simulation was discussed in detail in Section 3, only a brief description of the considered uncertainties is given here. The modelling uncertainties of the $bb4\ell$ and $bb4\ell + \mathrm{HF}$ contributions are treated as uncorrelated in the statistical fit. The modelling uncertainties with the largest impact on the top-quark mass fit are the colour reconnection uncertainty and the μ_R matrix element variation of the $bb4\ell$ contribution. For the top-quark decay width measurement, modelling uncertainties dominate the total systematic uncertainty. Among these, the most important are the alternative Herwig 7 parton-shower uncertainty for both the $bb4\ell$ and $bb4\ell + \mathrm{HF}$ contributions, as well as the μ_R and μ_F matrix-element scale variations of the $bb4\ell$ contribution. The impact of the dominant modelling uncertainties in the signal region is discussed in more detail in Section 10.

[7] The transverse momentum vector $\vec{p}_{\mathrm{T}}$ is defined as (p_x, p_y).

Hard-process generation: matrix-element scale variation The uncertainty due to missing higher-order corrections in the perturbative calculation is estimated by varying the central scale μ_0 of the renormalisation (μ_R) and factorisation (μ_F) scales by factors of 0.5 and 2. The variations are performed independently for the two scales. The μ_R uncertainty is evaluated with predictions obtained with $(\mu_R, \mu_F) \in \{(0.5, 1.0), (2.0, 1.0)\}\,\mu_0$, while the μ_F uncertainty is evaluated with predictions for which $(\mu_R, \mu_F) \in \{(1.0, 0.5), (1.0, 2.0)\}\,\mu_0$. These variations are evaluated using event weights.

Hard-process generation: h_{damp} variation Changing the POWHEG h_{damp} parameter affects the modelling of the hard emission generated in the POWHEG simulation. The variation to a larger h_{damp} value of $3\,m_{\mathrm{top}}$, compared to the nominal value of $1.5\,m_{\mathrm{top}}$ used in the $bb4\ell$ MC sample, leads to a larger fraction of POWHEG real-emission events generated according to the POWHEG formula of Equation 2.60 rather than according to the real-emission matrix element squared.

Hard-process generation: PDF variation PDF uncertainties are evaluated using the PDF4LHC21 PDF set [113]. The 30 eigenvector variations are propagated via event weights.

Parton-shower matching: $p_{\mathrm{T}}^{\mathrm{hard}}$ variation The uncertainty associated with the matching of the hard production process to the parton shower is evaluated by varying the parton-shower matching scale through the PYTHIA 8 $p_{\mathrm{T}}^{\mathrm{hard}}$ parameter, which is varied from its nominal value $p_{\mathrm{T}}^{\mathrm{hard}} = 0$ to $p_{\mathrm{T}}^{\mathrm{hard}} = 1$. The $p_{\mathrm{T}}^{\mathrm{hard}} = 1$ variation introduces holes and overlaps in the emission phase-space which is covered by the POWHEG hard process generation and the PYTHIA 8 parton-shower emissions.

Parton-shower matching: scale-resonance-veto Similarly, the matching uncertainty in the top-quark decay subprocess is evaluated by using the PYTHIA 8 p_{T} definitions in the parton-shower vetoing, instead of the (correct) POWHEG p_{T} definition. This again results in a mismatch in the emission phase space covered by the two generators.

Parton shower: Herwig 7 The prediction obtained with the nominal POWHEG+PYTHIA 8 $bb4\ell$ MC sample is compared to the prediction obtained when using the alternative HERWIG 7 parton shower. It should be noted, that for the HERWIG 7.3.0 version used in the $bb4\ell$ MC generation, no dedicated b-tagging MC-to-MC-scale factors are available. Instead, the corresponding scale factors derived for the HERWIG 7.2.3 version are applied. The impact of the absence of dedicated scale factors requires further investigation in future studies.

Parton shower: Recoil uncertainty An uncertainty in the modelling of parton-shower emissions from the top-quark decay is assessed by changing the Pythia 8 recoil scheme from the nominal recoil-to-top configuration to the recoil-to-colour option. This modifies the assignment of the colour-connected recoil partner for gluons emitted in the top-quark decay subprocess, which would be colour-connected to the top quark, and affects the emission pattern from the top-quark decay.

Parton shower: Pythia 8 splitting kernel variation Uncertainties in the modelling of ISR and FSR emissions in the Pythia 8 parton shower are evaluated by varying the scale at which the strong coupling constant in the Pythia 8 splitting kernels is evaluated by factors of 0.5 and 2, as well as varying the non-singular terms in the splitting kernel parameterisation (c_{NS}) from their nominal value of zero to ± 2. These variations are evaluated via event weights. Currently, no dedicated $bb4\ell$ MC sample[8] is available within the ATLAS Collaboration describing the $\alpha_S(M_Z)$ variations in the Powheg Sudakov. As discussed in Section 3.4, this uncertainty is closely related to the FSR μ_R $Q \to Qg$ splitting kernel variations. Since these variations are not found to have a large impact in the present analysis, the Powheg $\alpha_S(M_Z)$ variation uncertainty is not expected to constitute a large systematic uncertainty. This should be verified once the corresponding $bb4\ell$ MC samples are available.

Parton shower: Underlying event The underlying event modelling in the Pythia 8 A14 tune is varied according to the A14Var1 prescription [95].

Parton shower: Colour reconnection Three different Pythia 8 colour reconnection models [142] are compared: the nominal CR0 model and the alternative CR1 and CR2 models. Dedicated ATLAS Pythia 8 tunes are applied to ensure the validity of the comparison.

Additional uncertainty: $bb4\ell + c$ contribution An additional uncertainty is assigned to the $bb4\ell$ contribution containing truth c-jets. Events containing exactly two truth b-jets and at least one truth c-jet are assigned a normalisation uncertainty of $\pm 50\,\%$. This uncertainty is introduced since the normalisation of the $bb4\ell + c$ contribution is not constrained using data in a dedicated control region in this analysis.

[8] While a MC sample with a relatively small number of events at NLO+PS level is available, which was studied in Section 3.4, no large MC samples including detector simulation and reconstruction have been produced so far, as for this uncertainty the production of new LHE requires significant computational resources, it is expected not to be dominant in this data analysis and in collaboration with the authors of [2] new recommendations for the nominal MC generation settings are being developed. With the new recommendations, more number of events of the Powheg $\alpha_S(M_Z)$ $bb4\ell$ variations will be generated including also the detector simulation and reconstruction.

8.2.2 Background normalisation

Background normalisation uncertainties are considered following the conservative uncertainty estimates used in recent ATLAS measurements in the $t\bar{t}$ dilepton channel at $\sqrt{s} = 13\,\text{TeV}$ [102, 284]. The background categories are described in Section 6.2.2.

$t\bar{t}$+boson normalisation A normalisation uncertainty of 13.3 % is considered for the $t\bar{t}$+boson templates, following the prescription used in recent ATLAS measurements in the $t\bar{t}$ dilepton final state at $\sqrt{s} = 13\,\text{TeV}$ [102, 284]. This corresponds to the largest theoretical uncertainty associated with the total cross-section predictions of the $t\bar{t}$+boson processes.

Other normalisation The *Other* background category includes diboson and Z+jets events. A conservative normalisation uncertainty of 50 % is assumed. Previous ATLAS $t\bar{t}$ dilepton analyses have employed 40 % [284] or 50 % [102] uncertainties to the diboson contribution to cover possible mismodelling of the diboson normalisation in events with additional jets [285].

LeptonFake normalisation The MC prediction of the *LeptonFake* contribution is not constrained in this analysis using a dedicated control region. Therefore, a conservative normalisation uncertainty of 90 % is applied. In the recent ATLAS $WWbb$ unfolding measurement in the dilepton channel at $\sqrt{s} = 13\,\text{TeV}$ [102] a 50 % normalisation uncertainty was applied for this background. It has been verified in [102] that the data–MC disagreement is indeed covered by this uncertainty, by studying a validation region defined with similar kinematic selections as the signal region in the measurement, but requiring two same-sign leptons instead of opposite-sign leptons to increase the fraction of MC events with fake leptons.

CHAPTER 9

Fit strategy

In this section, first the likelihood function used in the profile likelihood fit is described in Section 9.1, followed by a brief introduction into the numerical parameter optimisation methods and uncertainty estimation techniques used in this measurement in Section 9.2. Finally, the parameterisation of the event yields per bin of the sensitive $m_{\ell b}^{\text{minimax}}$ distribution in the signal region in terms of the top-quark mass and decay width is described in Section 9.3.

9.1 The likelihood function

In this analysis, a binned profile likelihood template fit is used. First, the likelihood function is discussed, when only including the sensitive distribution $m_{\ell b}^{\text{minimax}}$ in the signal region. Later, additional terms will be introduced to account for the $bb4\ell + \text{HF}$ and JES control regions. The $m_{\ell b}^{\text{minimax}}$ histogram can be regarded as the result of a single measurement — namely, in this analysis the full ATLAS Run 2 data taking — which measures an N_{bins}-dimensional random vector [286] with N_{bins} being the number of bins in the $m_{\ell b}^{\text{minimax}}$ histogram. The joint probability density function (p.d.f.) for an observed bin count of $(N_{\text{bin1}}^{\text{data}}, \cdots, N_{\text{binN}}^{\text{data}})$ given the model parameter θ, is constructed as the product of the Poisson distributions $\text{Pois}(N_i^{\text{data}}|N_i^{\text{pred}}(\theta))$ in each bin i. The model parameter $\hat{\theta}$, which maximizes this joint likelihood given the measured data, is called the maximum likelihood estimator (MLE) and corresponds to the central parameter estimate in this analysis, as will be further detailed below. The profile likelihood fit model in this thesis is build using the HistFactory [287] and TRExFitter [288] software packages.

Parameters of interest If the sensitive distribution is well-chosen, the expected number of events per bin N_i^{pred} will have a significant dependence on the parameters of interest: the top-quark mass m_t and top-quark decay width Γ_t in this case. The likelihood L, i.e. fixing the measured event counts N_i^{data} in the joint p.d.f., is then given by

$$L(m_t, \Gamma_t) = \prod_{i=1}^{N_{\text{bins}}} \text{Pois}(N_i^{\text{data}} | N_i^{\text{pred}}(m_t, \Gamma_t)) = \prod_{i=1}^{N_{\text{bins}}} e^{-N_i^{\text{pred}}(m_t,\Gamma_t)} \frac{N_i^{\text{pred}}(m_t, \Gamma_t)^{N_i^{\text{data}}}}{N_i^{\text{data}}!}. \quad (9.1)$$

The expectation values N_i^{pred} are obtained with MC simulation. Additionally to the $WWbb$ production process (the $bb4\ell$ MC simulation), also other production processes lead to background events which are included in the measurement as described in Section 6. The expected number of events in bin i can thus be written as

$$\begin{aligned} N_i^{\text{pred}} &= \sum_s N_{s,i}^{\text{pred}} \\ &= N_{bb4\ell,i}^{\text{pred}} + N_{bb4\ell+\text{HF},i}^{\text{pred}} + N_{t\bar{t}+\text{bosons},i}^{\text{pred}} + N_{\text{Others},i}^{\text{pred}} + N_{\text{LeptonFakes},i}^{\text{pred}}, \end{aligned} \quad (9.2)$$

where $N_{s,i}^{\text{pred}}$ denotes the predicted number of events in bin i per MC event category s, which are defined in Section 6.2. The parameterisation of the top-quark mass and decay width dependence of $N_i^{\text{pred}}(m_t, \Gamma_t)$ is discussed in Section 9.3. Only an m_t- and Γ_t-dependence of the $bb4\ell$ and $bb4\ell$ +HF templates is considered.

Nuisance parameters In a profile likelihood fit, systematic uncertainties are directly included in the likelihood via nuisance parameters $\vec{\alpha}$ with Gaussian constraints, such that the expectation values read $N_i^{\text{pred}}(m_t, \Gamma_t, \vec{\alpha})$. This dependence is again evaluated via MC simulation: For each systematic uncertainty α, a MC template, i.e. in this case the $m_{\ell b}^{\text{minimax}}$ histogram in the signal region, is constructed, with the parameters in the simulation associated to this systematic uncertainty varied by $\pm 1\sigma_\alpha$,[1] while all other parameters are kept at their nominal value. For each systematic uncertainty α, there are therefore three MC templates: the nominal template, the $+1\sigma_\alpha$ and $-1\sigma_\alpha$ templates. But to maximize the likelihood function w.r.t. to these parameters, a continuous parameterisation of the dependence of the expected number of events on these systematic variations is needed. This is achieved by interpolating between these three points: In HistFactory, this is technically done by first splitting the effect of the systematic uncertainty into a normalisation and a shape component. The normalisation component is derived for the full distribution, interpolating between the total event yield from the down to the up variation. The shape component is then parameterised by interpolating between the bin content in each bin, when the total yield of the $(+1\sigma_\alpha)$ and

[1] These variations are not always available, in particular for two-point modelling systematics: One example is the parton-shower uncertainty, which is evaluated by comparing a MC template generated with the nominal Pythia 8 or alternative Herwig 7 parton shower. This uncertainty is included in the profile likelihood fit by defining it as a $+1\sigma_\alpha$ uncertainty and symmetrizing it, i.e. adding the same variation with opposite sign as the $-1\sigma_\alpha$ uncertainty. This is not justified from a statistical point of view, but is meant as a proxy to capture variations of further samples that have not been considered. Recently, the possibility to derive statistically meaningful theory-nuisance parameters is being discussed, see e.g. [289].

$(-1\sigma_\alpha)$ variations are normalised to the nominal yield. For the normalisation component an exponential interpolation is used, and thus avoiding negative normalisations, while the shape components are interpolated piece-wise linear. These different interpolation options are illustrated in Figure 9.1.

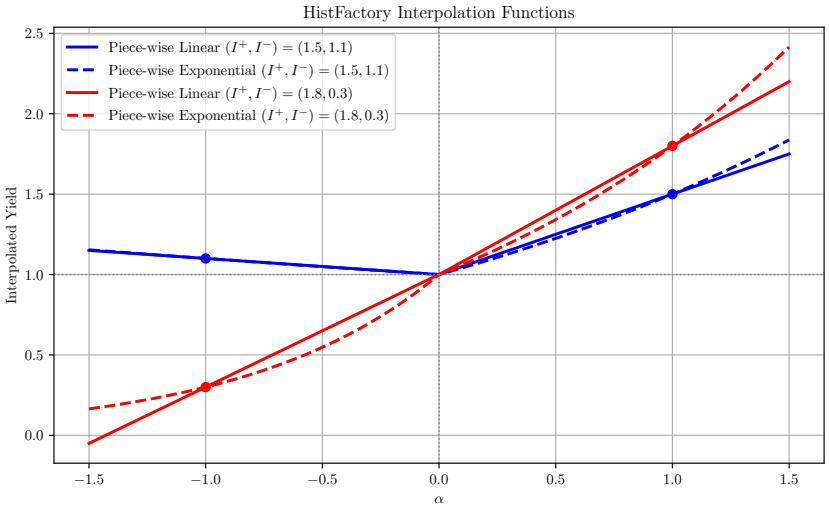


Figure 9.1: Visualisation of the piece-wise linear and piece-wise exponential interpolation options for nuisance parameter parameterisation in HistFactory. These two interpolation options are shown for two different input values $I(\alpha = 1) = I^+$ and $I(\alpha = -1) = I^-$. The parameterisation is scaled such that the nominal template $I(\alpha = 0)$ always equals 1. Graphic similar to visulation in [287].

The estimation of the uncertainty associated with a given systematic uncertainty is (ideally) obtained from a different measurement, e.g. a jet energy scale calibration measurement, see Section 8.1.1. Instead of including the full statistical model of these measurements into the likelihood used in this analysis, a Gaussian constraint term $G(a|\alpha, \sigma_\alpha)$ is added to describe the auxiliary measurement a. Conventionally, these parameters are scaled such that before the fit, $a = \alpha_{\text{nominal}} = 0$ and the $\pm 1\sigma$ variations correspond to $\alpha = \pm 1$. The likelihood including nuisance parameters α and their Gaussian constraints is given by

$$L(m_t, \Gamma_t, \vec{\alpha}) = \prod_{i=1}^{N_{\text{bins}}} \text{Pois}(N^{\text{data}}_{\text{SR},i} | N^{\text{pred}}_{\text{SR},i}(m_t, \Gamma_t, \vec{\alpha})) G(\vec{0}|\vec{\alpha}, \vec{1}) \tag{9.3}$$

$$= \prod_{i=1}^{N_{\text{bins}}} \text{Pois}(N^{\text{data}}_{\text{SR},i} | N^{\text{pred}}_{\text{SR},i}(m_t, \Gamma_t, \vec{\alpha})) \prod_j \frac{1}{\sqrt{2\pi}} e^{-\frac{\alpha_j^2}{2}}. \tag{9.4}$$

Not only the auxiliary measurements, but also the $m_{\ell b}^{\text{minimax}}$ bin yields are used to further constrain the nuisance parameters and reduce the corresponding systematic uncertainty. Since the physics object calibrations are performed in a wide phase space region to be applicable to most physics analyses, the estimation of systematic uncertainties can be improved in the concrete physics analysis phase space by the inclusion of the nuisance parameters in the likelihood.

Smoothing The histograms to parameterise the effect of systematic uncertainties $\vec{\alpha}$ are obtained from a finite set of MC generated events and thus are subject to statistical fluctuations. Consider the case of a systematic uncertainty α, in which one of the bins of the histogram corresponding to the $+1\sigma_\alpha$ variation shows a significant statistical fluctuation. When performing the profile likelihood fit with this nuisance parameter α, the uncertainty associated with this parameter would be constrained through this particular bin, since there is a strong disagreement of the $+1\sigma_\alpha$ variation w.r.t. data. This constraint is correct for this specific bin, but since it is propagated to the full distribution, it would lead to an overconstraint of this systematic uncertainty, leading to a smaller total uncertainty in the parameters of interest and an unreliable uncertainty estimate. To prevent these kinds of overconstraints, the MC templates corresponding to systematic variations are modified, smoothed, to reduce the impact of statistical fluctuations before the parameterisation of the bin content dependence on α is calculated. Several smoothing algorithms exist, and they are chosen in this analysis by comparing the smoothed histograms with the original input histograms. Statistical uncertainties in the MC templates, which are used to evaluate the systematic uncertainties, are not taken into account in the fit model.

Monte Carlo statistical uncertainties For the nominal signal and background MC templates, uncertainties due to statistical fluctuations in the MC predictions can be taken into account. This is done in the HistFactory framework by introducing a nuisance parameter γ for each bin to account for the statistical uncertainty associated with the total event yield of the bin. This approach is a simplification w.r.t. the original Barlow and Beeston method [290], in which a nuisance parameter in each bin is assigned to each individual sample, which would lead to a significant increase of parameters in the fitting procedure. As the MC statistical uncertainty in all bins considered in this analysis is well below 1 %, no dedicated uncertainty term is included in the likelihood.

Unconstrained normalisation factors Lastly, factors modifying the total event yield obtained from the $bb4\ell$ and the $bb4\ell + \text{HF}$ sample are introduced: the normalisation factors $\phi_{bb4\ell}$ and $\phi_{bb4\ell+\text{HF}}$. In this case, no constraint term is added, such that in the likelihood maximisation no penality term is included when deviating from $\phi = 1$. These normalisation factors are thus free to vary. Instead of correcting the $WWbb$ production process total cross-section to higher order calculations, the normalisation factor is determined from data. In data analyses focussing on the $t\bar{t}$ on-shell production process, it was found that the $t\bar{t}$ +HF prediction is underestimated [78, 132]. Thus, a separate normalisation factor is introduced for this contribution and measured through a dedicated $bb4\ell + \text{HF}$ control region, as described in Section 7.3. The full likelihood function $L(m_t, \Gamma_t, \vec{\alpha}, \vec{\phi})$ used in this analysis is thus given by

$$L(m_t, \Gamma_t, \vec{\alpha}, \vec{\phi}) = \prod_{i=1}^{N_{\text{bins}}} \text{Pois}(N_i^{\text{data}} | N_i^{\text{pred}}(m_t, \Gamma_t, \vec{\alpha}, \vec{\phi})) G(\vec{0} | \vec{\alpha}, \vec{1}) \tag{9.5}$$

with

$$N_i^{\text{pred}}(m_t, \Gamma_t, \vec{\alpha}, \vec{\phi}) = \sum_s N_{s,i}^{\text{pred}}(m_t, \Gamma_t, \vec{\alpha}, \vec{\phi}) \tag{9.6}$$

$$= \sum_s N_{s,i}^{\text{pred,nom}} \cdot \delta_{s,i}(m_t, \Gamma_t) \prod_j \sigma_i(\alpha_j)\eta(\alpha_j)\phi_s \tag{9.7}$$

where N_{bins} now also includes the single bin in the $bb4\ell + \text{HF}$ control region and the bins in the m_{bb} distribution of the JES control region, which was described in Section 7.4. The function $\delta_{s,i}(m_t, \Gamma_t)$ describes the modification of the nominal event yield of bin i for the samples $s \in \{bb4\ell, bb4\ell + \text{HF}\}$ as a function of m_t and Γ_t and will be discussed in Section 9.3. The functions $\sigma_i(\alpha_j)$ and $\eta(\alpha_j)$ describe the modification of the nominal event yield due to the shape impact of the uncertainty corresponding to α_j in the bin i and due to the overall normalisation impact of the uncertainty α_j, respectively. The normalisation factors ϕ_s only apply for the $bb4\ell$ and $bb4\ell + \text{HF}$ contribution.

9.2 Parameter optimisation and uncertainty estimation

After having defined the likelihood function $L(m_t, \Gamma_t, \vec{\alpha}, \vec{\phi})$ in the previous section, this section summarises the methods used to determine its local maximum and to estimate the uncertainties of the corresponding best-fit parameters.

Negative log-likelihood minimisation The best-fit parameters $(\hat{m}_t, \hat{\Gamma}_t, \hat{\vec{\alpha}}, \hat{\vec{\phi}})$ are obtained by maximising the likelihood function with respect to all parameters

$$(\hat{m}_t, \hat{\Gamma}_t, \hat{\vec{\alpha}}, \hat{\vec{\phi}}) = \underset{m_t, \Gamma_t, \vec{\alpha}, \vec{\phi}}{\text{argmax}}\, L(m_t, \Gamma_t, \vec{\alpha}, \vec{\phi}). \tag{9.8}$$

In the TRExFitter framework, this is achieved through the numerical minimisation of the negative log-likelihood using the Migrad minimiser of Minuit [291]. Starting from an initial parameter vector $\vec{x}_0 = (m_{t,0}, \Gamma_{t,0}, \vec{\alpha}_0, \vec{\phi}_0)$, the gradient $G = \Delta_{\vec{x}}(-\ln L(\vec{x}_0))$ is numerically computed. A line search along the direction opposite to the gradient is performed to approach the minimum of $-\ln L$. Parameter correlations are accounted for via the covariance matrix V, which is the inverse of the Hessian matrix H^{-1} built from the second derivatives of the negative log-likelihood function:

$$H_{ij} = \frac{\partial^2(-\ln L(\vec{x}))}{\partial x_i \partial x_j}. \tag{9.9}$$

Initially, the Hessian is set to the identity matrix. Assuming the Hessian matrix at the point $\vec{x}_i$ is known, the change in gradient can be expressed as $G_{i+1} = G_i + H_i(\vec{x}_{i+1} - \vec{x}_i)$. Setting this expression to zero — since $G_{i+1} = 0$ is the target — leads to the step

$$G_i + H_i(\vec{x}_{i+1} - \vec{x}_i) = 0 \Leftrightarrow \vec{x}_{i+1} = \vec{x}_i - H_i^{-1} G_i = \vec{x}_i - V_i G_i. \tag{9.10}$$

Since the Hessian is initially approximate, the actual step is determined via a line search along this direction by finding an α that minimises

$$-\ln L(\vec{x}_{i+1}) = -\ln L(\vec{x}_i - \alpha V \times G). \tag{9.11}$$

Computing the change in gradient ΔG and the change in the parameter values Δx, the estimation of the covariance matrix V is updated, as the estimated curvature must be able to explain the gradient change seen after updating the parameters $\vec{x}$ through the line search step. This procedure is repeated until the estimated distance to the minimum (EDM) is smaller than 0.001. The EDM can be calculated by using the Taylor expansion of $-\ln L$ around its minimum, which is given by

$$-\ln L(\vec{x}) = -\ln L(\vec{x}_{\text{min}}) + \frac{1}{2}(\vec{x} - \vec{x}_{\text{min}})^T H(\vec{x} - \vec{x}_{\text{min}}), \tag{9.12}$$

if the negative log-likelihood $-\ln L(x)$ is well approximated by a quadratic form around the minimum, where in the Taylor expansion the first derivate is zero. Expressing the quadratic term in terms of the gradient G and the covariance matrix V, the EDM for $F = -\ln L(\vec{x})$ is given by

$$F(\vec{x}) - F(\vec{x}_{\text{min}}) = \frac{1}{2}(\vec{x} - \vec{x}_{\text{min}})^T H(\vec{x} - \vec{x}_{\text{min}}) = \frac{1}{2} G^T V G. \tag{9.13}$$

Pruning Performing the numerical minimisation of the negative log-likelihood in the $\mathcal{O}(100)$-dimensional parameter space is computationally expensive. To reduce computation time and improve fit stability, nuisance parameters α associated to systematic uncertainties that have negligible effect on the distributions considered in the statistical fit can be neglected, i.e., pruned. The pruning is performed by evaluating the effect of each nuisance parameter α on both the shape and normalisation of the distributions:

- **Shape variation:** The histograms corresponding to $\alpha = \pm 1\sigma$ are first normalised to the nominal event yield (corresponding to $\alpha = 0$). The rescaled varied histograms are then compared to the nominal histogram bin by bin. If in any bin, the deviation between nominal and rescaled varied histogram exceeds 0.01 %, the corresponding shape systematic σ_α is retained; otherwise, it is removed from the likelihood function.
- **Normalisation variation:** The total event yield of the nominal and the varied histograms are compared. The normalisation systematic η_α is dropped from the likelihood function, if the difference between the nominal event yield and the yield from the varied histogram is smaller than 0.1 %.

This pruning procedure is applied separately for each sample and each analysis region (the signal and the two control regions). Applying these pruning thresholds, the number of considered systematic uncertainties is reduced from 388 to 274.

Uncertainty estimation in a profile likelihood fit The uncertainties on the best-fit parameters are estimated using the properties of the profile-likelihood ratio. The profile likelihood ratio is defined as

$$\lambda(\mu) = \frac{L(\mu, \hat{\hat{\theta}}_\mu)}{L(\hat{\mu}, \hat{\theta})}, \tag{9.14}$$

where the denominator is constant and corresponds to the maximum of the likelihood function L defined above. Instead of writing out the parameters of interests and the nuisance parameters separately, these are grouped into the factors μ and θ, respectively. The numerator is obtained by maximising the likelihood function with respect to θ for a fixed value of μ. The thereby determined nuisance parameters are denoted as $\hat{\hat{\theta}}$. By construction, the maximum of $\lambda(\mu)$ is one, which occurs at $\mu = \hat{\mu}$. The reason to introduce the profile likelihood ratio is that, while the likelihood function L has arbitrary normalisation and no probabilistic meaning by itself, $\lambda(\mu)$ has a known asymptotic distribution. According to Wilk's theorem, the logarithm of the profile-likelihood ratio asymptotically follows a χ^2 distribution [292]:

$$-2\ln\lambda(\mu) = \frac{(\mu - \hat{\mu})^2}{\sigma^2} + \mathcal{O}(1/\sqrt{N}), \tag{9.15}$$

which allows the construction of confidence intervals, with $\Delta(-2\ln\lambda) = 1$ corresponding to a 68 % (1σ) confidence interval on μ. The uncertainty σ can be evaluated in two ways, which are equivalent if $-2\ln\lambda(\mu)$ is exactly quadratic around the minimum. Expanding the logarithm of the profile-likelihood ratio around the minimum gives

$$\ln\lambda(\mu)|_{\hat{\mu}} = \underbrace{[\ln L(\mu, \hat{\hat{\theta}}_\mu)|_{\mu=\hat{\mu}} - \ln L(\hat{\mu}, \hat{\theta})]}_{=0} + \frac{1}{2}\frac{\partial^2 \ln L(\mu, \hat{\hat{\theta}}_\mu)}{\partial \mu^2}|_{\hat{\mu}}(\mu - \hat{\mu})^2, \tag{9.16}$$

where the first derivative vanishes at the minimum. Comparing to Equation 9.15, the identification

$$\frac{1}{\sigma^2} = -\frac{\partial^2 \ln L}{\partial \mu^2}|_{(\hat{\mu},\hat{\theta})} \tag{9.17}$$

can be made. For multiple parameters $\vec{x}$, this can be generalised in terms of the covariance matrix

$$V_{ij} = \left(-\frac{\partial^2 \ln L}{\partial\theta_i \partial\theta_j}|_{\hat{\theta}}\right)^{-1}. \tag{9.18}$$

The diagonal elements V_{ii} correspond to the variance of the parameters θ_i. The off-diagonal elements V_{ij} describe correlations between parameters θ_i and θ_j. If the likelihood factorises with respect to all parameters, the mixed derivates of V_{ij} vanish and the covariance matrix becomes diagonal. Non-zero off-diagonal elements indicate that variations in a parameter θ_i can be partially compensated by variations in another parameter θ_j, leading to tilted likelihood contours in the (θ_i, θ_j)-plane. The strength of this correlation is described by the correlation coefficient $\rho_{ij} = V_{ij}/\sqrt{V_{ii}V_{jj}}$. The second method to evaluate uncertainty estimates on the

best-fit values is to calculate the parameters μ' for which

$$-2\ln\lambda(\mu') = 1, \tag{9.19}$$

as this would correspond to the values of $(\mu' - \hat{\mu})^2/\sigma^2 = 1 \Leftrightarrow \mu' = \hat{\mu} \pm \sigma$. In this analysis, the uncertainty in the nuisance parameters $\vec{\alpha}$ and on the normalisation factors $\vec{\phi}$ are evaluated according to Equation 9.17 as implemented in MINUIT with HESSE, which per definition leads to symmetric estimated uncertainties $\hat{\sigma}_{\hat{\theta}}$. For the parameters of interest, the top quark mass m_t and the top quark decay width Γ_t, the formula from Equation 9.19 is used, implemented in MINUIT with MINOS, which can produce asymmetric uncertainty intervals.

9.3 Parameterising the top-quark mass and decay width sensitivity

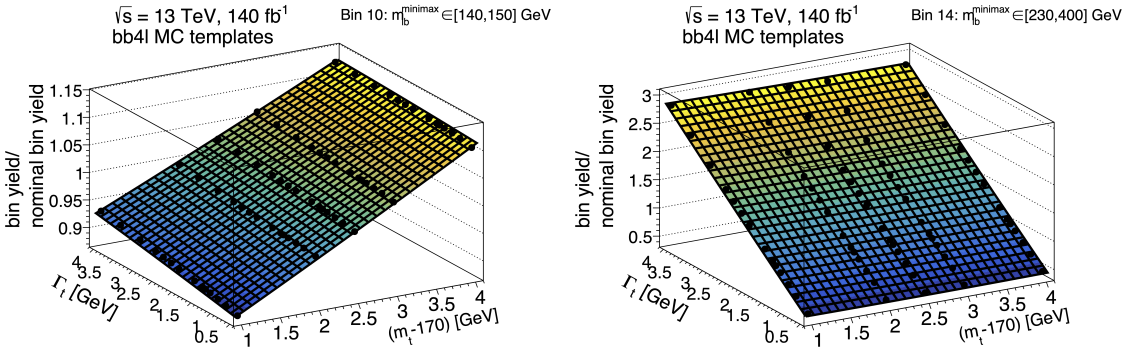
While in Section 7.2 the variation of the $bb4\ell$ MC predictions for different top-quark mass and decay width values of the $m_{\ell b}^{\text{minimax}}$ distribution in the signal region was discussed, a continuous prediction of the event yield per bin as a function of the top-quark mass and decay width is required for the statistical fit. This is achieved by fitting a two-dimensional quadratic polynomial $\delta_i^s(m_t, \Gamma_t)$ to the template yields in each bin i, obtained from the $bb4\ell$ and $bb4\ell + \text{HF}$ MC predictions for varied m_t and Γ_t, i.e. $s \in \{bb4\ell, bb4\ell + \text{HF}\}$.[2,3] The number of events per bin is then given by $N_i^s(m_t, \Gamma_t) = N_i^{s,\text{nominal}} \delta_i^s(m_t, \Gamma_t)$.[4] For this procedure, all $bb4\ell$ MC predictions are normalised to the total event yield of the nominal $bb4\ell$ prediction in the signal region, as only shape differences induced by the parameter variation are considered. Since the absolute $m_{\ell b}^{\text{minimax}}$ distribution is used in the fit, a free-floating normalisation factor is added, as discussed previously. Furthermore, as the mass-varied $bb4\ell$ samples are only available with the ATLFAST 3 detector simulation, these samples are compared, for the purpose of deriving the parameterisation $N_i^s(m_t, \Gamma_t)$, to the nominal $bb4\ell$ MC sample ($m_t = 172.5\,\text{GeV}$, $\Gamma_t = 1.32\,\text{GeV}$) also simulated with ATLFAST 3.[5] This procedure is performed separately for both the $bb4\ell$ and for the $bb4\ell + \text{HF}$ contribution. In Figure 9.2, two examples of the $N_i^{bb4\ell}(m_t, \Gamma_t)$ parameterisation of the event yield in the $m_{\ell b}^{\text{minimax}}$ bins of (a) $m_{\ell b}^{\text{minimax}} \in [140, 150]\,\text{GeV}$ and of (b) $m_{\ell b}^{\text{minimax}} \in [230, 400]\,\text{GeV}$ are shown. These particular bins exhibit high sensitivity to the top-quark mass and decay width, respectively.

[2] This fitting procedure is implemented in the TREXFITTER framework [288], where the method `Fit` of the class `TGraph2DErrors` provided by the ROOT framework [293] is used. This corresponds to a χ^2 fitting procedure using the MINUIT minimiser. The statistical MC uncertainties on the template event yields in each bin are thereby taken into account. It should be noted, that the mass-varied $bb4\ell$ predictions are obtained by separate MC sample generations and thus have uncorrelated MC statistical errors, while the decay-width varied $bb4\ell$ predictions are obtained through reweighting and therefore have correlated statistical uncertainties for predictions with the same top-quark mass parameter.

[3] No smoothing procedure is thereby applied to the $bb4\ell$ or $bb4\ell + \text{HF}$ template distributions with varied m_t and Γ_t.

[4] Please note that $\delta_i^s(m_t, \Gamma_t)$ does not need to equal 1 for the nominal parameters of interest $m_t = 172.5\,\text{GeV}$ and $\Gamma_t = 1.32\,\text{GeV}$.

[5] The implicit assumption made in this procedure is that the top-quark mass and decay width dependence of the $m_{\ell b}^{\text{minimax}}$ distribution is identical for distributions obtained with ATLFAST 3 and GEANT4 simulations.



(a) (b)

Figure 9.2: Comparison of the fitted $N_i^{bb4\ell}(m_t, \Gamma_t)$ dependence and the template yields of the mass- and width-varied $bb4\ell$ predictions, normalised to the total event yield of the nominal $bb4\ell$ MC prediction in the signal region. These template yields are used as inputs to derive the $\delta_i^{bb4\ell}(m_t, \Gamma_t)$ parameterisation. The comparison is shown for bin $i = 10$ ($m_{\ell b}^{\text{minimax}} \in [140, 150]\,\text{GeV}$) and bin $i = 14$ ($m_{\ell b}^{\text{minimax}} \in [230, 400]\,\text{GeV}$). The y-axis displays the ratio of the $N_i(m_t, \Gamma_t)$ predictions to the nominal $bb4\ell$ template yield ($m_t = 172.5\,\text{GeV}$ and $\Gamma_t = 1.32\,\text{GeV}$).

To test the quality of the mass and decay width parameterisation, pseudo-data tests are performed: A pseudo-data set is constructed by adding the $bb4\ell$ and $bb4\ell + \text{HF}$ MC predictions with varied top-quark mass and decay width parameters[6] and the nominal background predictions. A fit including only statistical uncertainties is then performed to this pseudo-data set to test whether the input top-quark masses and decay widths can be correctly extracted when applying the $N_i^s(m_t, \Gamma_t)$ parameterisation described above. If the $N_i^{bb4\ell}(m_t, \Gamma_t)$ and $N_i^{bb4\ell+\text{HF}}(m_t, \Gamma_t)$ parameterisation would exactly reproduce the input template yields of the $bb4\ell$ and $bb4\ell + \text{HF}$ predictions for all m_t and Γ_t, perfect closure of this test would be expected. In practice, the parameterisation does not reproduce the template input points exactly. This behaviour is expected due to the finite number of MC events, which leads to statistical fluctuations in the predicted event yield per bin relative to the underlying theoretical prediction that would be obtained with unlimited number of generated MC events.

The results of these pseudo-data tests are shown in Figure 9.3. The cross-markers indicate the input values of the top-quark mass and decay width used to produce the pseudo data, while the points indicate the corresponding fitted values obtained from the pseudo-data test. The error bars represent the uncertainty estimate from the statistical fit without including systematic uncertainties to the extracted m_t and Γ_t values. All tested pseudo-data mass and decay widths are recovered within the uncertainty of the fitted parameter. It should be noted, that the parameterisation is derived by shifting the top-quark mass parameter by

[6] If the $bb4\ell$ mass-varied predictions were available with the GEANT4 simulation, these predictions could be directly used to build the pseudo-data set. Instead, as these MC samples are only available in ATLFAST 3, the ratios of the event yields per bin in the $m_{\ell b}^{\text{minimax}}$ distribution of the mass-varied ATLFAST 3 $bb4\ell$ prediction and the nominal ATLFAST 3 $bb4\ell$ prediction are used for each bin in the nominal $bb4\ell$ GEANT4 prediction. This procedure is performed again for both the $bb4\ell$ and the $bb4\ell + \text{HF}$ contributions separately.

$(m_t - 170)$ GeV in the parameterisation, leading to a similar order of magnitude for the m_t and Γ_t parameters. The improvement of the parameterisation due to this choice and further tests of alternative parameterisation of $N_i^s(m_t, \Gamma_t)$ are presented in Appendix A.

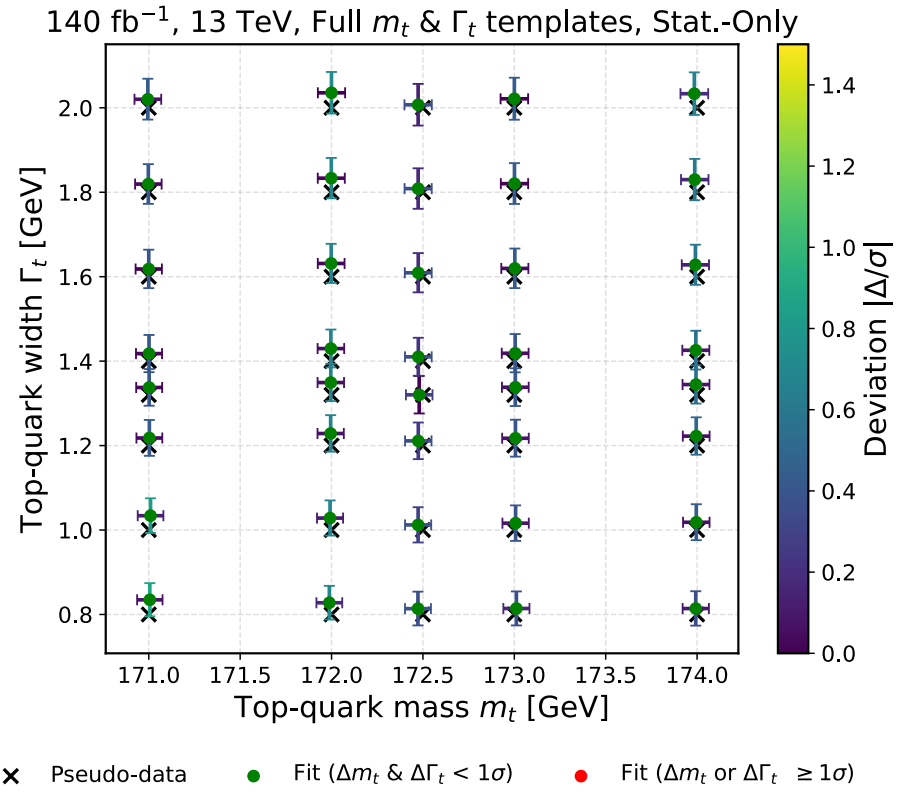


Figure 9.3: Closure test of the parameterisation. The input top-quark mass and decay width of the pseudo data (cross markers) are compared to the fitted values (point markers). The fit includes only statistical uncertainties, using the nominal two-dimensional parameterisation $N_i(m_t, \Gamma_t)$. The colour of the error bars on the fitted parameters indicates the corresponding values of $\frac{|m_t^{\text{fit}} - m_t^{\text{in}}|}{\sigma_{m_t}^{\text{stat.-only}}}$ and $\frac{|\Gamma_t^{\text{fit}} - \Gamma_t^{\text{in}}|}{\sigma_{\Gamma_t}^{\text{stat.-only}}}$.

CHAPTER 10

Estimating the top-quark mass and decay width sensitivity

In this section, the sensitivity of the top-quark mass and decay width measurement is evaluated. This is achieved by testing the profile likelihood fit model using a data set representative of the MC simulation, the Asimov data set, in Section 10.1, and by examining the best-fit normalisation factors and nuisance parameter values obtained from the fit to ATLAS Run 2 data in Section 10.2. All systematic uncertainties described in Section 8 are taken into account, including in particular the $bb4\ell$ modelling uncertainties discussed in detail in Section 3. The development of these modelling uncertainties was crucial to obtain a realistic estimate of the achievable precision in this measurement.

10.1 Profile likelihood fit using the Asimov data set

The statistical fit model described in Section 9 is studied using an Asimov data set N_i^{data} [292] constructed from the nominal signal and background predictions. The nominal signal prediction corresponds to the event yields per bin obtained with the parameterisation $N_i^{\text{pred}}(m_t = 172.5\,\text{GeV}, \Gamma_t = 1.32\,\text{GeV})$. As discussed in Section 9.3, this prediction does not need to correspond exactly to the nominal $bb4\ell$ signal MC templates. The use of the Asimov data set allows the evaluation of the expected uncertainty in the top-quark mass and decay width. This evaluation is performed for the case when all nuisance parameters are evaluated to be exactly zero, i.e. they correspond to their nominal values, the $bb4\ell$ and $bb4\ell$ + HF normalisation factors are unity, and the input parameters of interest are retrieved. The uncertainties in the nuisance parameters can be reduced with respect to their pre-fit value, reflecting the constraining power of the data if it were the Asimov data set. This is taken as an estimate for the constraining power expected in the ATLAS Run 2 data set.

First, the impact of introducing the control regions used to correct the normalisation of the $bb4\ell + \text{HF}$ contribution and to constrain JES uncertainties on the expected sensitivity to the top-quark mass and decay width is evaluated by comparing three different Asimov fits:[1]

(a) 1D Asimov fit: Profile likelihood fit of the $m_{\ell b}^{\text{minimax}}$ distribution in the signal region.

(b) 2D Asimov fit: Profile likelihood fit corresponding to (a), additionally including the total number of events in the heavy-flavour control region. This region is expected to reduce the uncertainty in the $bb4\ell + \text{HF}$ normalisation factor and the top-quark decay width.

(c) 3D Asimov fit: Profile likelihood fit corresponding to (b), additionally including the m_{bb} distribution of the JES control region, which is expected to constrain in particular the JES pre-recommendation uncertainties and therefore improve the sensitivity to the top-quark mass parameter.

The uncertainty estimates in the top-quark mass and decay width, as well as in the normalisation factors of the $bb4\ell$ and $bb4\ell + \text{HF}$ contributions, obtained from the three different Asimov fits, are shown in Figure 10.1. All systematic uncertainties discussed in Section 8 are included. The sensitivity to the top-quark mass and decay width is found to be dominated by systematic uncertainties. When not including systematic uncertainties in the 3D Asimov fit, the expected uncertainties amount to 0.07 GeV for the top-quark mass and 0.04 GeV for the top-quark decay width.

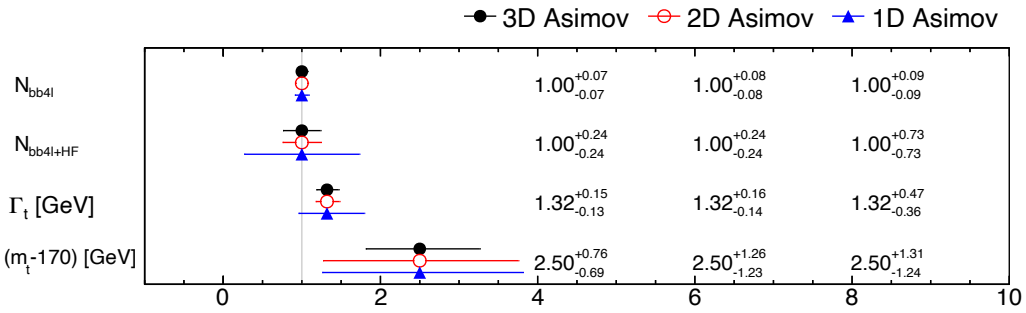


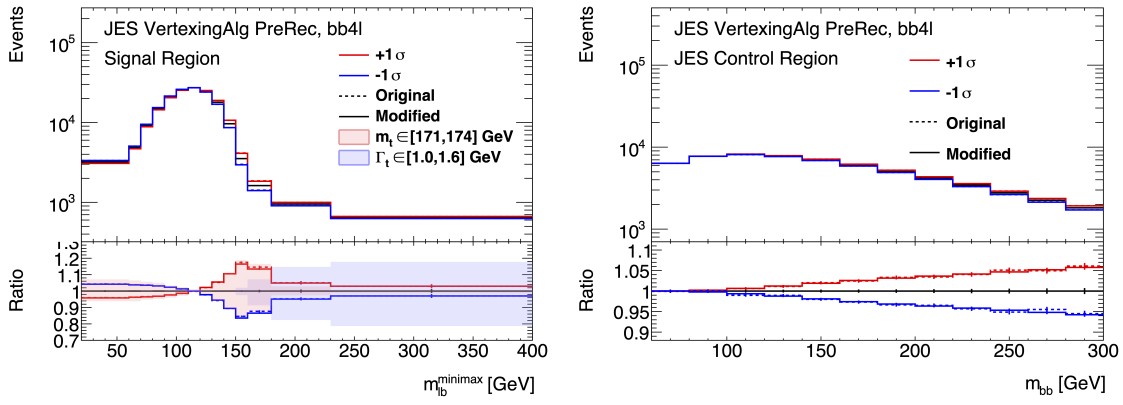
Figure 10.1: Estimated uncertainties in the normalisation factors of the $bb4\ell$ and $bb4\ell$+HF contributions ($N_{bb4\ell}$ and $N_{bb4\ell+\text{HF}}$), the top-quark mass ($m_t - 170\,\text{GeV}$), and the top-quark decay width (Γ_t), obtained from the three Asimov fits discussed in the text. The post-fit parameter values and uncertainties are listed below the fit labels (3D Asimov, 2D Asimov, and 1D Asimov).

The inclusion of the $bb4\ell + \text{HF}$ control region significantly reduces the uncertainty in the $bb4\ell + \text{HF}$ normalisation factor and in the top-quark decay width, while having only a minor impact on the top-quark mass and the $bb4\ell$ normalisation. Prior to introducing the $bb4\ell + \text{HF}$ control region, the parameters $N_{bb4\ell+\text{HF}}$ and Γ_t are strongly correlated. From the correlation matrix obtained in the negative log-likelihood minimisation of the 1D Asimov

[1] While in the following, the nomenclature 1D, 2D and 3D is used, this corresponds to the profile likelihood fit to either one, two or three one-dimensional histograms, instead of to a one-, two- or three-dimensional observable.

fit, the correlation coefficient is estimated to be approximately $-88\,\%$. This behaviour can be understood from the shape of the $bb4\ell + \text{HF}$ contribution in the signal region, shown in Figure 7.7(a), which demonstrates that variations in the $bb4\ell + \text{HF}$ normalisation factor have a similar effect to changes in the top-quark decay width. By including additional information about the $bb4\ell + \text{HF}$ contribution through the dedicated control region, the correlation between $N_{bb4\ell+\text{HF}}$ and Γ_t is reduced drastically to approximately $2\,\%$. This is achieved since the $bb4\ell + \text{HF}$ control region is highly sensitive to $N_{bb4\ell+\text{HF}}$, while having no sensitivity to variations in Γ_t.

The introduction of the JES control region, on the other hand, has negligible impact on the uncertainty in $N_{bb4\ell+\text{HF}}$ and in the top-quark decay width, but a significant effect on the top-quark mass measurement. This behaviour is partly due to a reduction in the correlation between the top-quark mass parameter and the nuisance parameter describing the JES *VertexingAlg* pre-recommendation uncertainty, which is associated with the change in the primary vertexing algorithm between the reconstruction software versions used in the in-situ JES calibration and used in this measurement. In the 2D Asimov fit, the correlation coefficient between these parameters is approximately $72\,\%$, which decreases to about $36\,\%$ after introducing the JES control region. Figure 10.2 illustrates the impact of the JES *VertexingAlg* pre-recommendation uncertainty. In (a), the effect of this uncertainty in the signal region is compared to variations of the top-quark mass and decay width, demonstrating the strong correlation with m_t. In contrast, (b) shows that the JES control region is sensitive to the JES *VertexingAlg* pre-recommendation uncertainty. As discussed in Section 7.4, the JES control region is not sensitive to variations of m_t.



(a) (b)

Figure 10.2: Impact of the JES *VertexingAlg* pre-recommendation uncertainty on the $bb4\ell$ prediction in (a) the signal region and (b) the JES control region. The MC predictions corresponding to the $\pm 1\sigma$ variations of this uncertainty are shown as dashed lines, while the output of the smoothing algorithm is shown as solid lines. Additionally, in (a) the influence of the top-quark mass and decay width variations in the signal region are illustrated by the red and blue bands, respectively.

The constraining power estimated in the Asimov fit for a subset[2] of the considered nuisance parameters — those which are most relevant for this measurement — evaluated with the three Asimov fits is shown in Figure 10.3. After the introduction of the JES control region, additionally to reducing the correlation of the JES *VertexingAlg* pre-recommendation uncertainty and the top-quark mass variation, the corresponding nuisance parameter uncertainty can be more strongly constrained. While slightly increased constraints in the 3D compared to the 1D Asimov fit are also observed in the $bb4\ell$ modelling systematic uncertainties, the dominant expected constraining power is indeed originating from the signal region. This is also the case for the significantly constrained $bb4\ell$ +Herwig 7 nuisance parameter, and will be further discussed in Section 10.2.

The correlation coefficients, evaluated in the 3D Asimov fit, between nuisance parameters, the parameters of interest (the top-quark mass and the decay width) and the normalisation factors are shown in Figure 10.4. Only quantities that exhibit at least one correlation coefficient above 20 % with another quantity are retained. As shown in this figure, the normalisation factor $N_{bb4\ell}$ is strongly correlated with the uncertainty associated to $bb4\ell$ MC events with additional particle-level c-jets ($bb4\ell$ soft c-jets in Figure 10.4). This is expected, as in Figure 7.7(a) the fraction of $bb4\ell$ events with additional c-jets was found to be approximately flat across the full $m_{\ell b}^{\text{minimax}}$ distribution, corresponding nearly to a pure normalisation effect with negligible impact on the shape. For the $bb4\ell + \text{HF}$ normalisation factor, the strongest correlations are observed with the nuisance parameters describing variations of the FSR $g \to q\bar{q}$ splitting kernel. This modelling uncertainty directly affects the number of events with additional b-jets through variations in the description of $g \to b\bar{b}$ splittings, and is therefore expected to have a significant impact on $N_{bb4\ell+\text{HF}}$.

While the largest correlation coefficients between nuisance parameters θ_i and the parameters of interest $\mu \in \{m_t, \Gamma_t\}$ are also shown in Figure 10.4, the dominant sources of systematic uncertainties affecting μ can be identified with the uncertainty breakdown based on the covariance matrix [294] shown in Figure 10.5.[3] Thereby, the impact of each systematic uncertainty in μ is evaluated in the Gaussian approximation through

$$\Delta_{\theta_i}^{\mu} = \sigma_{\mu}^{\text{tot}} \rho_{\theta_i,\mu} \sigma_{\theta_i}, \qquad \rho_{\theta_i,\mu} = V_{\theta_i,\mu} / \sqrt{V_{\theta_i,\theta_i} V_{\mu,\mu}} \tag{10.1}$$

where $\rho_{\theta_i,\mu}$ is the correlation coefficient built with the covariance matrix V, σ_{θ_i} the (post-fit) uncertainty in the nuisance parameter θ_i, and $\sigma_{\mu}^{\text{tot}}$ the total uncertainty in the parameter of interest μ. In case of asymmetric uncertainties — which are possible when performing the uncertainty estimation using Minos— in Equation 10.1 double-sided Gaussian distributions $\sigma^{\text{up}} \neq \sigma^{\text{down}}$ are used instead, which can thus lead to asymmetric impacts. The final fit model is studied by using the covariance matrix, as the covariance matrix is built at the best-fit point in the full likelihood parameter space (meaning the parameters of interest, normalisation factors and nuisance parameters) and all systematic uncertainty sources are considered. These

[2] This is not shown for all 274 nuisance parameters entering the fit model after pruning for better visibility, instead the focus lies on the nuisance parameters related to the $bb4\ell$ and $bb4\ell + \text{HF}$ modelling, jet, $E_{\text{T}}^{\text{miss}}$ and flavour-tagging b-jet efficiency uncertainties in this figure.

[3] It was verified that the conclusions from Figure 10.4 do not change when using either Minos to evaluate all parameter uncertainties in the likelihood or the global shifted observables approach [294].

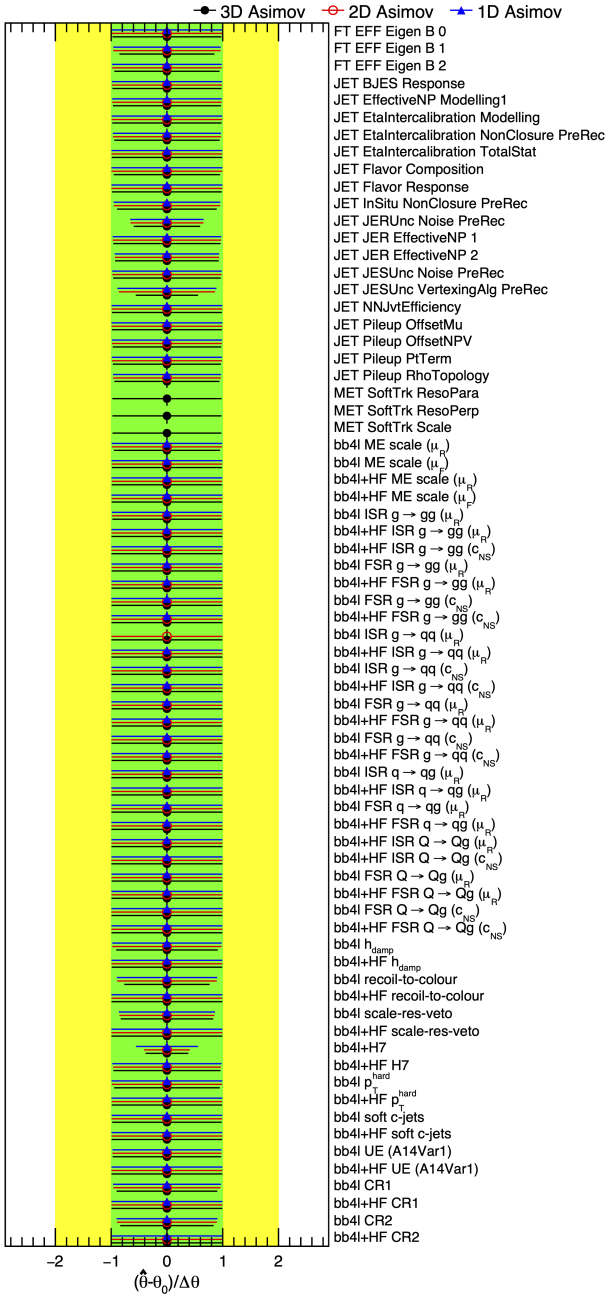


Figure 10.3: Post-fit values and uncertainties for a subset of nuisance parameters ($bb4\ell$ and $bb4\ell + \text{HF}$ modelling, $E_{\text{T}}^{\text{miss}}$, and most impactful jet and flavour tagging b-jet efficiency uncertainties) used in the 3D, 2D and 1D Asimov fits. The best-fit and the pre-fit values for the nuisance parameters are given by $\hat{\theta}$ and θ_0, respectively. For the Asimov data set, it is always $\hat{\theta} = \theta_0$. The pre-fit uncertainty is given by $\Delta\theta$ and is shown — applied to the pre-fit value θ_0 — by the green band. In the case of the 1D and 2D Asimov fit (which do not include the JES control region), the nuisance parameters related to the $E_{\text{T}}^{\text{miss}}$ uncertainties are pruned and thus are not visible in the figure.

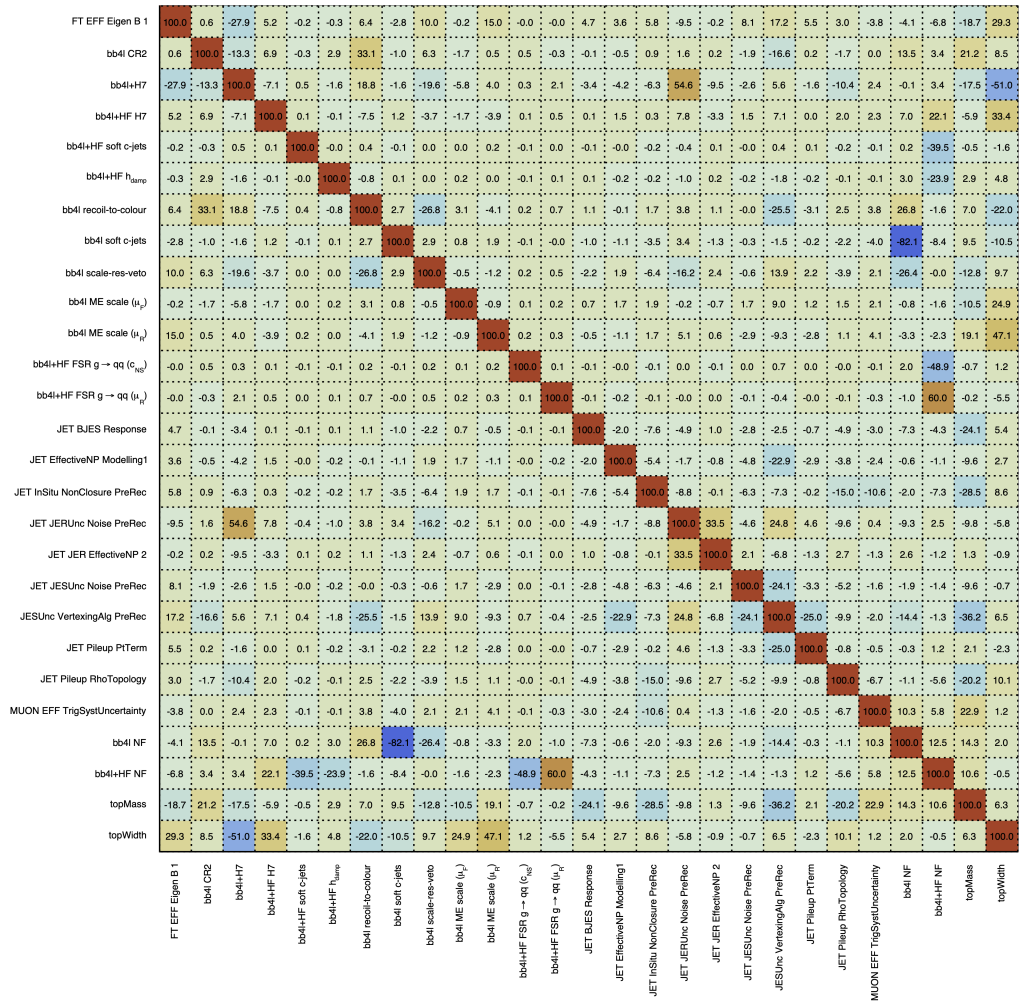
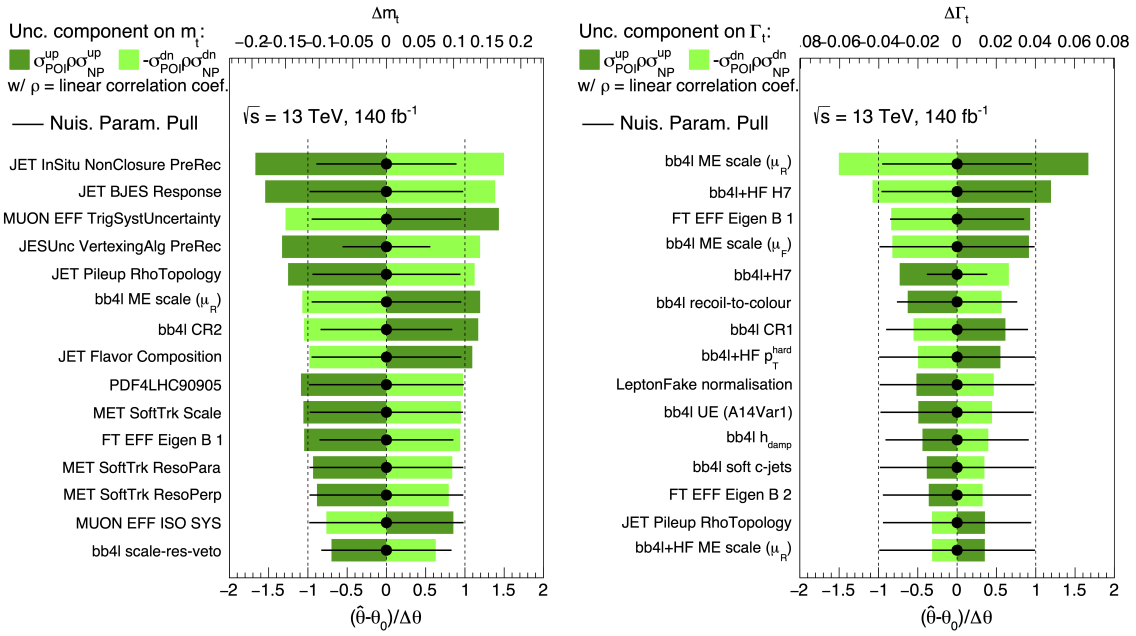


Figure 10.4: Correlation coefficients between nuisance parameters, the top-quark mass and decay width parameters, and the normalisation factors with at least one coeffiecent $> 20\,\%$ as determined in the minimisation procedure of the 3D Asimov fit.

impacts of systematic uncertainty contributions can be added in quadrature to obtain the total systematic uncertainty estimate on m_t and Γ_t.

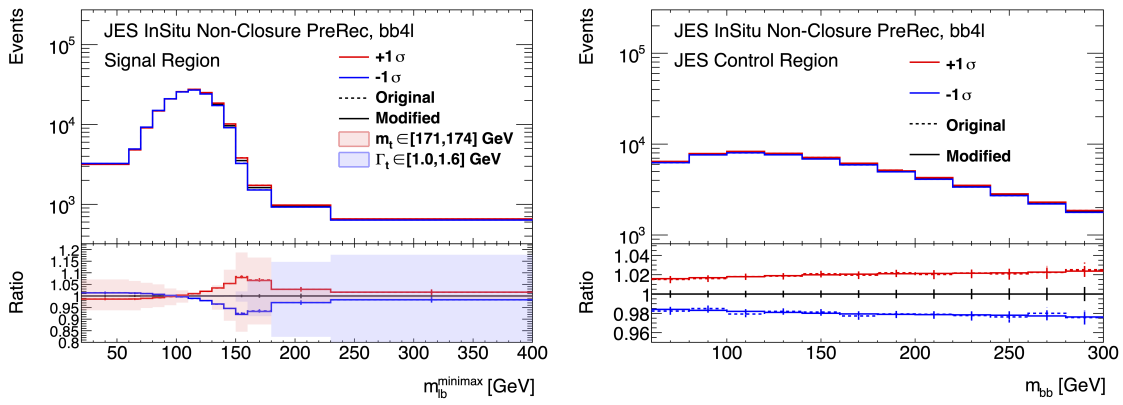


(a) (b)

Figure 10.5: Impact of systematic uncertainties evaluated with the uncertainty decomposition based on the covariance matrix for (a) the top-quark mass and (b) the decay width. The 15 systematic uncertainties with the highest impact on the respective parameter of interest in decreasing order of their impact are shown. The numerical values of these impacts on $\Delta\hat{m}_t$ and $\Delta\hat{\Gamma}_t$ can be read from the upper x-axis.

The impact of the JES *VertexingAlg* pre-recommendation uncertainty on the mass measurement is strongly reduced after the introduction of the JES control region, which is by the dominant uncertainty component in the case of the 2D Asimov fit. In the 3D Asimov fit, a further JES pre-recommendation uncertainty is important in the mass fit, which accounts for the non-closure of the in-situ corrections, when using the reconstruction software version in this measurement or in the JES in-situ calibration. Its impact on the $bb4\ell$ prediction in both the signal region and the JES control region is shown in Figure 10.6. In contrast to the JES uncertainty related to the primary vertexing algorithm, the in-situ non-closure uncertainty induces only a small variation in the JES control region and can therefore not be constrained as effectively. As a consequence of introducing the JES control region, uncertainties related to the muon trigger efficiency and to the missing transverse momentum enter the fifteen most impactful systematic uncertainties affecting m_t, which instead are not important in the 2D Asimov fit. The dominant jet-related uncertainties on the mass measurement are further discussed in Appendix B.2.

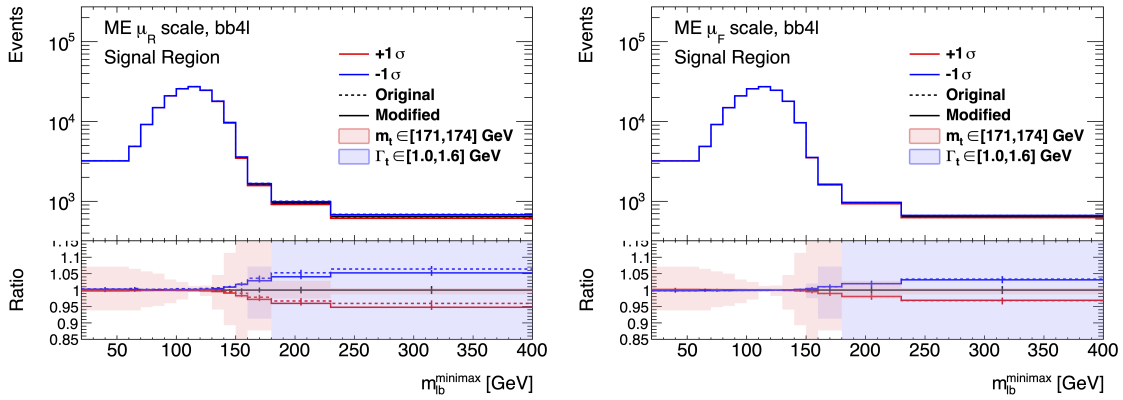
The systematic uncertainty contributions to the top-quark decay width measurement are generally dominated by modelling uncertainties on the $bb4\ell$ and $bb4\ell + \text{HF}$ contributions. In particular, the most important systematic uncertainty sources are the matrix-element scale



(a) (b)

Figure 10.6: Same comparison as in Figure 10.2, but for the JES pre-recommendation uncertainty covering the JES in-situ non-closure.

variations in the $bb4\ell$ contribution, as well as the parton-shower uncertainty in the $bb4\ell$ and $bb4\ell + \text{HF}$ predictions. The impact of matrix-element renormalisation (μ_R) and factorisation (μ_F) scale variations on the $bb4\ell$ prediction in the signal region are shown in Figure 10.7. The resulting variations closely resemble the shape changes induced by variations of the top-quark decay width, explaining their significant influence on the Γ_t measurement.



(a) (b)

Figure 10.7: Same comparison as in Figure 10.2, but for the matrix-element scale variation μ_R (a) and μ_F (b) variation uncertainty in the $bb4\ell$ prediction in the SR.

The impact of the parton-shower uncertainty on both, the $bb4\ell$ and $bb4\ell + \text{HF}$ contributions, in the signal and the JES control region is shown in Figure 10.8. While the effect of the parton-shower uncertainty in the $bb4\ell$ prediction in the signal region differs from variations of Γ_t in the bulk of the $m_{\ell b}^{\text{minimax}}$ distribution, the pronounced shape variation of the $bb4\ell$+Herwig 7 uncertainty in the tail of the distribution resembles the variations expected for a change in the Γ_t parameter — when shifting the overall normalisation. The correlation coefficient between

the top-quark decay width variation and the parton-shower uncertainty is thereby significant, −51 %, as seen in Figure 10.4. Furthermore, the parton-shower uncertainty evaluated for the $bb4\ell$ + HF contribution corresponds to a large normalisation change of approximately 30 %, as observed also in the JES control region. Interestingly, the parton-shower uncertainty in the $bb4\ell$ + HF prediction decreases at large values of $m_{\ell b}^{\text{minimax}}$. This indicates that, while the modelling of additional soft b-jets differs significantly between $bb4\ell$ + Pythia 8 and $bb4\ell$ + Herwig 7, these differences are less pronounced for additional b-jets with higher transverse momenta. This interpretation is further supported by the observation that the normalisation difference due to the parton-shower uncertainty in the $bb4\ell$ + HF contribution is reduced to about 6 % in the $bb4\ell$ + HF control region, in which at least one additional b-jet is required to have substantial transverse momentum. Further studies, comparing the simultaneous and independent top-quark mass and decay width profile likelihood fits and the impact of excluding the jet pre-recommendation uncertainties, which are expected to be reduced significantly with the upcoming JES and JER recommendations, can be found in Appendix B.1 and Appendix B.2, respectively.

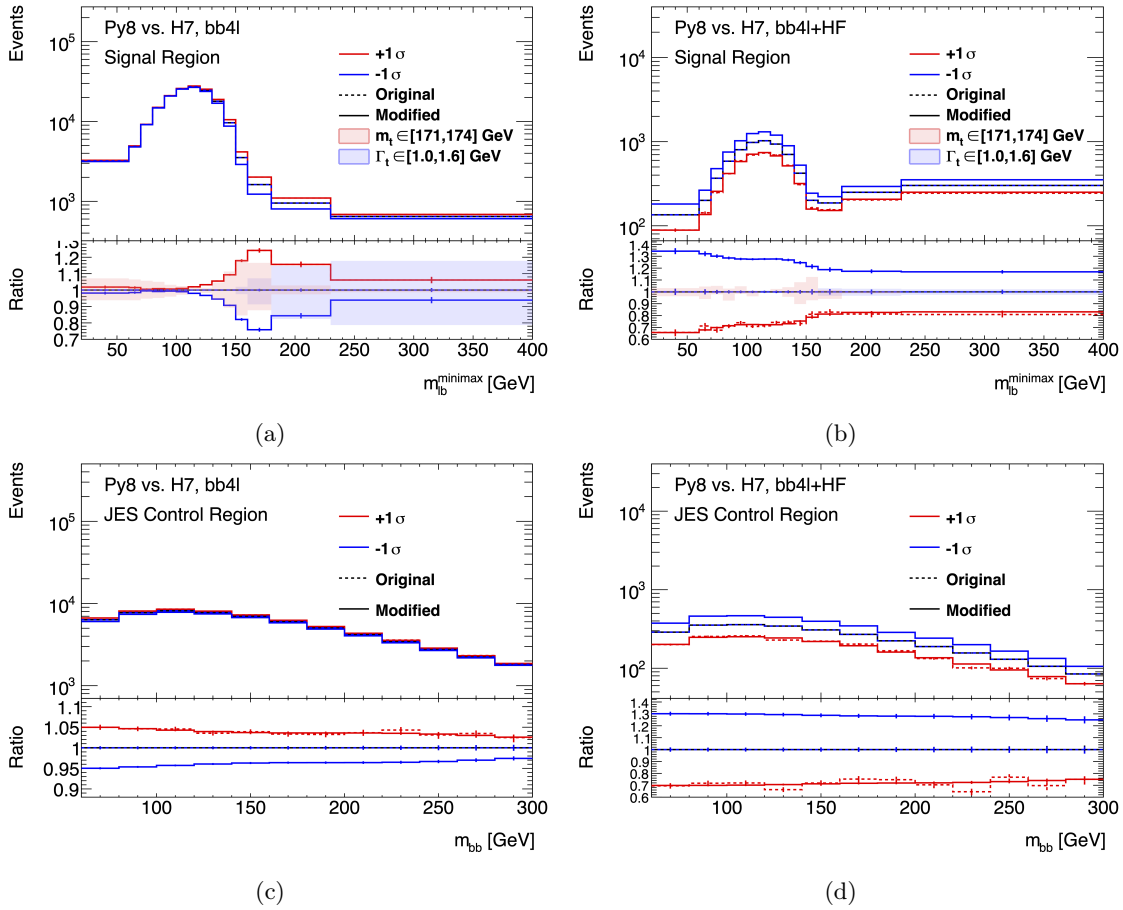


Figure 10.8: Same comparison as in Figure 10.2, but for the $bb4\ell$+Herwig 7 uncertainty evaluated for the $bb4\ell$ ((a),(c)) and $bb4\ell$ + HF ((b), (d)) contribution.

10.2 The blinded fit

To further study the fit model, ATLAS Run 2 data is used in the statistical fit, but the best-fit values of the parameters of interest — the top-quark mass and the decay width — are concealed, i.e. the fit is *blinded*. This approach is useful to study the behaviour of the model in the fit to data, in particular the post-fit values and uncertainties of the nuisance parameters, without biasing the choices made in the analysis set-up due to known best-fit values of the parameters of interest for different fit configurations.

The estimated uncertainty in the blinded parameters of interest and the measured value for the $bb4\ell$ and $bb4\ell + \text{HF}$ normalisation factors in the fit to Run 2 ATLAS data are given by

$$\hat{m}_t = \hat{m}_t{}^{+0.61}_{-0.65}\,\text{GeV}, \;\; \hat{\Gamma}_t = \hat{\Gamma}_t{}^{+0.23}_{-0.20}\,\text{GeV}, \;\; \hat{N}_{bb4\ell+\text{HF}} = 1.73 \pm 0.41, \;\; \hat{N}_{bb4\ell} = 0.96 \pm 0.07. \quad (10.2)$$

Compared to the expected uncertainties evaluated with the Asimov data set, the uncertainty in the top-quark mass parameter is reduced, while the uncertainty in the top-quark decay width parameter increased in the fit to ATLAS Run 2 data. As expected from the large data–MC discrepancy observed in the $bb4\ell + \text{HF}$ control region prior to the fit, the normalisation factor $\hat{N}_{bb4\ell+\text{HF}}$ differs substantially from unity. Its value is consistent with the comparison of the $bb4\ell$ predictions to unfolded data shown in Figure 3.17(b). The normalisation factor of the $bb4\ell$ contribution is compatible with unity.

Taking into account the fitted normalisation factors for both, the $bb4\ell$ and $bb4\ell + \text{HF}$ contributions, the inclusive NLO+PS $bb4\ell$ prediction in the signal region is effectively scaled by approximately $-2\,\%$. As the inclusive K-factor $\sigma_{\text{NNLO}}/\sigma_{\text{NLO}}$ for off-shell $t\bar{t}$ production was calculated to be approximately $10.8\,\%$ [295], this might be surprising. However, it should be noted, that the inverse width corrections discussed in [2] are not applied to the nominal $bb4\ell$ MC prediction in this analysis. Including these corrections leads to a reduction of the $bb4\ell$ normalisation of approximately $6\,\%$, as described in Section 3.2. Their application has been recommended to the ATLAS Collaboration in private communication with the authors of [2] and will be included in future MC productions of $bb4\ell$ samples, both for the nominal prediction and for systematic variations.

To assess whether the post-fit model $L_{\text{post-fit}}$ provides an adequate description of the data, a goodness-of-fit test based on the saturated model [296] is performed. The saturated model L_{sat} introduces one free parameter per data point, allowing it to reproduce the data perfectly. The test statistic,

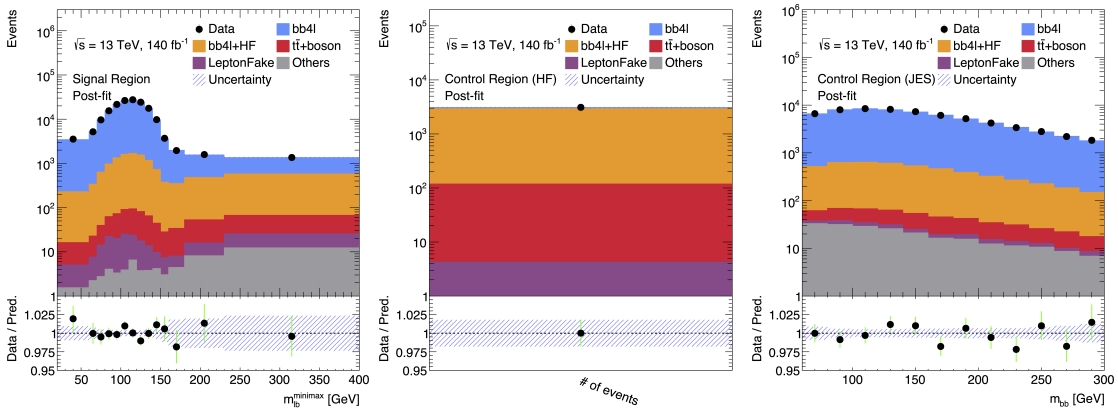
$$\chi^2 = -2\ln\lambda = -2\ln\frac{L_{\text{post-fit}}}{L_{\text{sat}}}, \quad (10.3)$$

follows asymptotically a χ^2 distribution according to Wilks' theorem. The corresponding p-value,

$$p = p(\chi^2 \geq \chi^2_{\text{obs}}) = 1 - p(\chi^2 \leq \chi^2_{\text{obs}}) = 1 - F(\chi^2_{\text{obs}}, N_{\text{dof}}), \quad (10.4)$$

is computed using the cumulative distribution function F of the χ^2 distribution with $N_{\text{dof}} = N_{\text{data}} - N_{\text{par}} = 27 - 4$ degrees of freedom, where N_{par} includes the two parameters of interest

and the two free-floating normalisation factors. This p-value quantifies the probability of observing a χ^2 value, which is at least as large as the one obtained, assuming the post-fit model is correct. A goodness-of-fit probability of 43.7 % is obtained using the saturated model, indicating good agreement between the post-fit model and the data. The good post-fit description of the ATLAS Run 2 data in the analysis regions is shown in Figure 10.9.

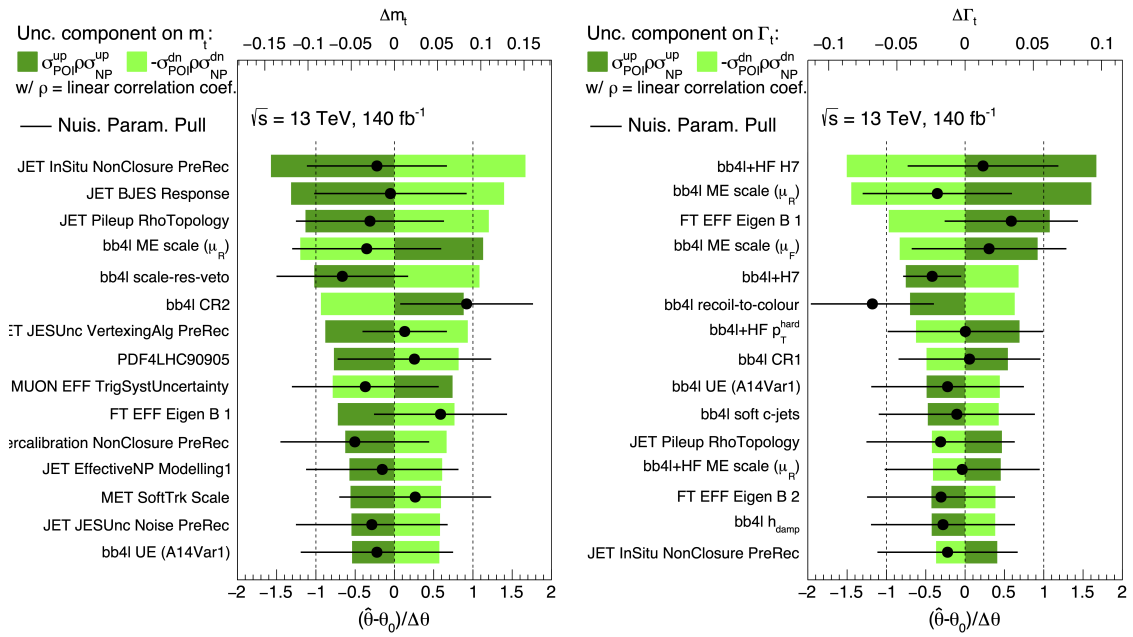


(a) (b) (c)

Figure 10.9: Comparison of the ATLAS Run 2 data in (a) the signal region, (b) the $bb4\ell + \text{HF}$ control region and the (c) JES control region with the total MC prediction in these regions, in which the best-fit parameters of interest $\hat{m}_t$, $\hat{\Gamma}_t$, the best-fit normalisation factors $\hat{N}_{bb4\ell}$ and $\hat{N}_{bb4\ell+\text{HF}}$, as well as the best-fit nuisance parameter values $\hat{\theta}$ and corresponding uncertainties $\Delta\hat{\theta}$ as obtained from the blinded fit to data are employed to obtain the predicted event yields per bin. The blue uncertainty band shows the post-fit uncertainty in the MC prediction, and the green error bars on the data points represent the statistical uncertainty in these.

In Figure 10.10, the impact of the highest ranked systematic uncertainties in the best-fit top-quark mass and decay width value in the blinded fit to ATLAS Run 2 data is shown. Significant deviations between the post-fit and pre-fit nuisance parameter values are observed for several $bb4\ell$ modelling uncertainties, including the PYTHIA 8 recoil-scheme uncertainty, the colour reconnection CR2 uncertainty and the matching uncertainty in the top-quark decay (scale-resonance-veto). As expected from the Asimov fit studies in Section 10.1, the HERWIG 7 uncertainty is strongly constrained. Moreover, in the blinded fit, the corresponding nuisance parameter is shifted such that it remains compatible with the nominal parton shower model at slightly above 1σ.

For the estimation of these modelling systematics, the nominal prediction is compared to a single alternative prediction, which defines the $+1\sigma$ variation in the pre-fit model. The corresponding -1σ variation is obtained by symmetrising the uncertainty, i.e. by applying a mirrored shift with respect to the nominal prediction. For the recoil, scale-resonance-veto, and HERWIG 7 uncertainties, the fit does not prefer the alternative prediction itself, but rather the mirrored variation. A preference in the profile likelihood fit for the mirrored variation therefore does not correspond to a concrete physics model, but instead more strongly disfavours the alternative prediction.



(a) (b)

Figure 10.10: Same comparison as in Figure 10.5, but for the systematic uncertainties with the largest impact on the best-fit $\hat{m}_t$ and $\hat{\Gamma}_t$ value in the blinded fit to ATLAS Run 2 data.

To investigate which analysis regions have the largest impact on these shifts, the corresponding systematic uncertainties are decorrelated across the signal region, the JES control region, and the $bb4\ell + \text{HF}$ control region, introducing a separate nuisance parameter for each region. The pre- and post-fit values and uncertainties when correlating or decorrelating these nuisance parameters across analysis regions are compared in Figure 10.11.

The largest shifts in the post-fit nuisance parameters for the $bb4\ell$ recoil, scale-resonance-veto, and colour reconnection (CR2) uncertainties originate from the signal region, and the reduction in the Herwig 7 uncertainty is also predominantly driven by the signal region. For all of these uncertainties, the post-fit trends are similar for nuisance parameters in the signal and JES control region, whereas the $bb4\ell + \text{HF}$ control region does not constrain or shift these parameters. Thereby, differences between data and MC predictions in the single-bin $bb4\ell + \text{HF}$ control region, lead to a change of the $bb4\ell + \text{HF}$ normalisation factor $\hat{N}_{bb4\ell+\text{HF}}$, rather than leading to a shift of a nuisance parameter, which would include a likelihood penalty.

With the future consolidated JES and JER recommendations for reprocessed Run 2 data and MC, it may become possible to remove the JES control region, eliminating the need for these decorrelation studies. Still, since the dominant changes between pre-fit and post-fit nuisance parameter values corresponding to these uncertainties arise from the signal region, they must be scrutinised further. The effects of these uncertainties on the $bb4\ell$ distribution in the signal region are shown in Figure 10.12.

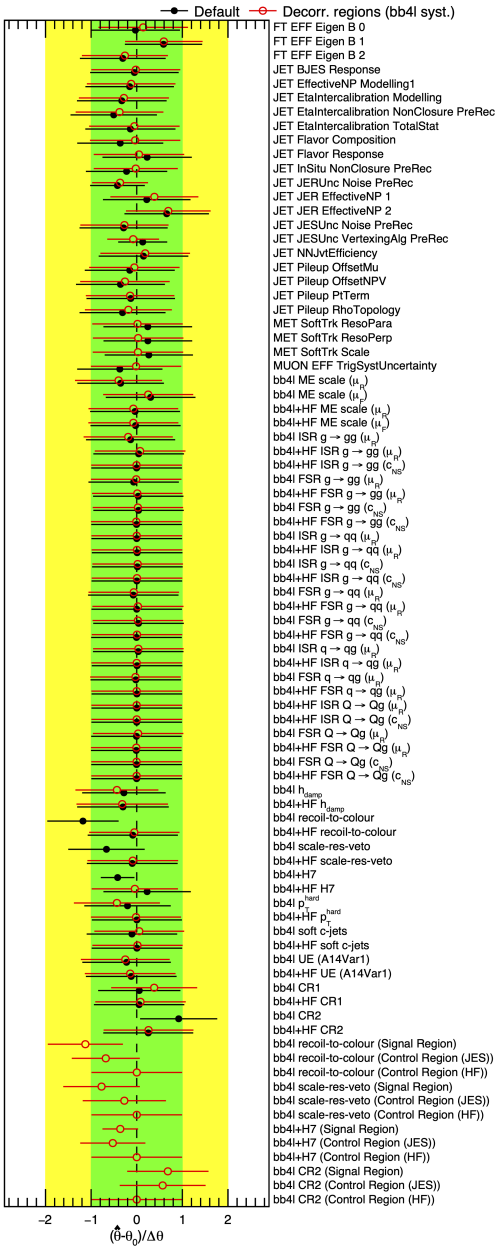


Figure 10.11: Same comparison as in Figure 10.3, but for the blinded fit to data when either correlating (default) or decorrelating (decorr. regions $bb4\ell$ syst) systematic uncertainties in the $bb4\ell$ prediction, whose nuisance parameters show the largest deviations between their pre-fit and post-fit values, per region. These systematic uncertainties are the $bb4\ell$ recoil, colour reconnection (CR2), scale-resonance-veto and HERWIG 7 uncertainties. Therefore, in the default fit, one nuisance parameter for these uncertainties is defined ($bb4\ell$ recoil-to-colour, $bb4\ell$ CR2, $bb4\ell$ scale-resonance-veto, $bb4\ell$+H7) and shown in the figure, while for the fit *decorr. regions* $bb4\ell$ *syst* three separate nuisance parameters are introduced per uncertainty ($bb4\ell$ recoil-to-colour (region), $bb4\ell$ CR2 (region), $bb4\ell$ scale-resonance-veto (region), $bb4\ell$+H7 (region)) with region $\in$ {Signal Region, Control Region (JES), Control Region (HF)}.

An interesting observation concerns the recoil uncertainty: its impact on the top-quark mass measurement is small in this analysis, despite being one of the dominant uncertainties in the most recent ATLAS top-quark mass measurement in the dilepton channel [124]. The measurement used an $m_{\ell b}$ distribution up to 140 GeV as the mass-sensitive observable. In the bulk of the $m_{\ell b}^{\text{minimax}}$ distribution, the recoil uncertainty affects the shape of the observable similarly to mass variations, but near the endpoint around 150 GeV these two shape variations differ distinctly, allowing the recoil uncertainty and the mass variations to be largely uncorrelated. In the current analysis, this results in a correlation coefficient of only 7 % in the profile likelihood fit to Asimov data, as shown in Figure 10.4.

The matching uncertainty in the top-quark decay subprocess (scale-resonance-veto) is expected to be replaced by a more targeted uncertainty definition, as the current implementation introduces a mismatch between the scales used in POWHEG and PYTHIA 8 to quantify this uncertainty. As indicated by the corresponding post-fit nuisance parameter value, the alternative prediction is indeed disfavoured by the data in the signal region. Nevertheless, the associated degree of freedom is still actively used in the fit. Therefore, understanding the impact of an improved definition of the matching uncertainty will be important for this analysis.

Finally, the colour-reconnection uncertainty is mostly influencing the low $m_{\ell b}^{\text{minimax}}$ region, while its evaluation in the high-$m_{\ell b}^{\text{minimax}}$ tail is limited by MC statistical uncertainties, as seen in Figure 10.12(c). It should be noted however, that this uncertainty does not have a large impact on the top-quark decay width measurement, as seen in Figure 10.10. With the current uncertainty estimate, the data in the signal region favours the alternative CR2 model over the CR0 colour reconnection model.[4]

It is explicitly verified in the blinded fit that the top-quark mass and decay width measurements are largely independent. In Figure 10.13, the blinded best-fit values of m_t and Γ_t obtained from the simultaneous fit are compared to those from separate fits extracting only the top-quark mass or decay width, as described in detail in Appendix B.1. Excellent agreement is observed between the best-fit values of the parameter of interest across the different fit models. Further, the comparison of MC predictions and data in distributions, which were not used in the profile likelihood fit, can be found in Appendix B.3.

[4] In a recent ATLAS top-quark mass measurement [125], which employed a profile likelihood fit, the data likewise favoured the CR2 model. However, the use of the single-lepton $t\bar{t}$ channel and the focus on high transverse-momentum top quarks in [125], prevents a direct comparison with the measurement presented in this thesis.

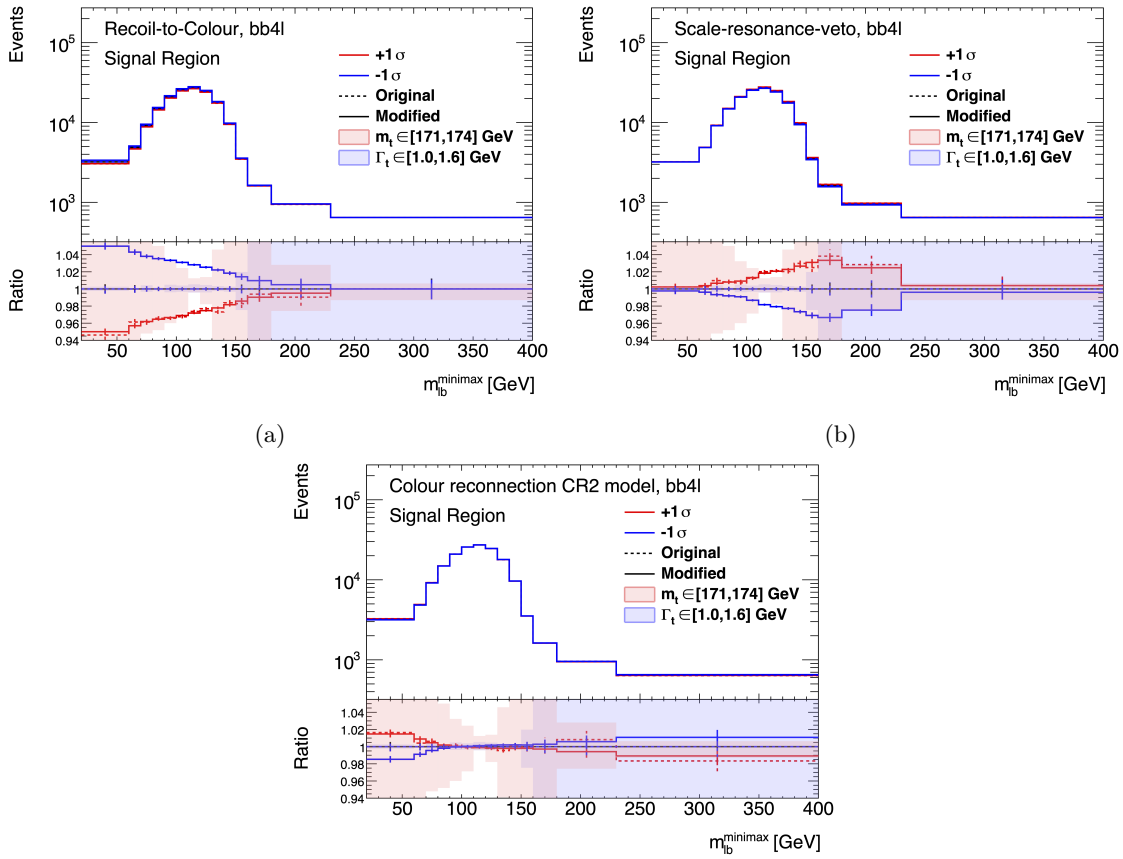
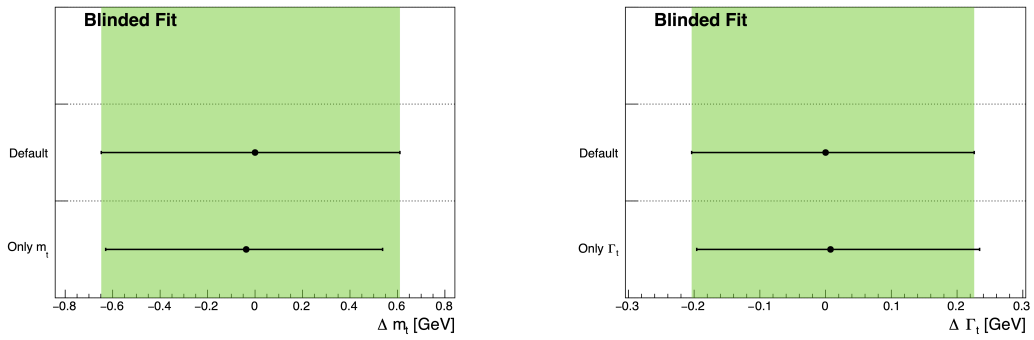


Figure 10.12: Same comparison as in Figure 10.2, but for the impact of the (a) PYTHIA 8 recoil uncertainty, (b) the matching uncertainty in the top-quark decay (scale-resonance-veto) and (c) the colour reconnection uncertainty CR2 on the $bb4\ell$ prediction in the signal region.

10.3 Future developments

Some aspects of the analysis presented in this thesis can be further improved. For the top-quark mass measurement, the updated JES and JER in-situ calibrations, derived using reprocessed Run 2 data and simulation, are expected to reduce the associated uncertainties. To estimate the expected improvement, the top-quark mass uncertainty is re-evaluated without taking the jet pre-recommendation uncertainties into account. This leads to a reduction on the uncertainty in the top-quark mass of approximately 100 MeV with respect to the uncertainty obtained in the nominal fit, which includes the jet pre-recommendation uncertainties. Once these calibrations are applied, the need for a dedicated JES control region will be reassessed, potentially simplifying the analysis strategy and the understanding of constraints and post-fit shifts of the $bb4\ell$ modelling uncertainties. In collaboration with the authors of [2], the nominal MC generation settings for the $bb4\ell$ sample are under final review. One anticipated change is the adoption of a lower central renormalisation and factorisation scale in the NLO QCD



(a) (b)

Figure 10.13: Comparison of the blinded best-fit values for (a) the top-quark mass and (b) decay width as determined in the simultaneous (m_t, Γ_t)-fit (Default) or by independent fits of either the top-quark mass or decay width, as described in Section B.1. The green band indicates the uncertainty in the blinded parameter of interest in the *default* fit.

matrix-element calculation, which is particularly relevant since the matrix-element scale-variation uncertainty was found to significantly affect the top-quark decay-width measurement. Additionally, the systematic-uncertainty prescription is being refined, including potentially an alternative evaluation of the matching uncertainties. The top-quark decay-width measurement will benefit from a more detailed understanding of the observed discrepancy between data and MC in the number of additional b-jets. Studies are ongoing to validate the $bb4\ell + \text{HF}$ truth definition and the application of the normalisation factor for the $bb4\ell + \text{HF}$ contribution in the signal region, which is largely determined from data in the corresponding $bb4\ell + \text{HF}$ control region. The MC modelling of additional heavy-flavour jets is also being actively investigated in collaboration with the authors of [2].

CHAPTER 11

Conclusions

In this thesis, the achievable precision in a simultaneous top-quark mass and decay width measurement using the ATLAS Run 2 dataset has been studied. Using simulated data, the expected sensitivity is found to be 0.73 GeV for the top-quark mass and 0.14 GeV for the top-quark decay width. This constitutes a significant improvement in the top-quark decay width sensitivity compared to the currently most precise measurement with an uncertainty of 0.3 GeV [232]. This improvement is driven by several factors, including the use of the full Run 2 data set, advanced object reconstruction techniques and, in particular, an analysis strategy that exploits the enhanced sensitivity provided by improved MC predictions generated with the $bb4\ell$ MC generator [1, 2]. A substantial part of this work focused on making these MC predictions suitable as a nominal signal model within an ATLAS measurement. This required the detailed study and implementation of all relevant modelling systematic uncertainties. As many top-quark measurements in ATLAS are limited by modelling uncertainties, their reliable estimation is essential for a robust uncertainty evaluation. The nominal $bb4\ell$ ATLAS MC sample developed in this work therefore provides a central contribution to future top-quark analyses within the ATLAS Collaboration. Indeed, MC samples produced during this thesis have already been utilised in several publications [102, 104, 297]. The studies presented here demonstrate the significant impact of the modelling of additional heavy-flavour jets on the top-quark decay width measurement, as well as the influence of jet pre-recommendation uncertainties in the top-quark mass determination. The latter is expected to be significantly reduced with the availability of updated jet calibrations. In contrast, the former requires a more detailed understanding of the modelling of additional heavy-flavour contributions in the $bb4\ell$ MC, which were found to differ from those in the $t\bar{t}$ MC simulation. Further studies are ongoing in collaboration with the authors of [2] to investigate these effects in more detail and to refine the systematic uncertainty model. These efforts are expected to further improve the robustness of future measurements.

APPENDIX A

Testing alternative parameterisations of the top-quark mass and decay width dependence

Several alternative parameterisations of the top-quark mass and decay width parameterisation per bin $N_i(m_t, \Gamma_t)$ are presented in this appendix, while the nominal parameterisation used in the measurement is described in Section 9.3. These alternative parameterisations are used to study the impact of several general choices made in the alternative parameterisation, as well as the impact of using a different functional for $N_i(m_t, \Gamma_t)$.

The following general choices are made in the nominal parameterisation: First, the dependence of both the $bb4\ell$ and the $bb4\ell + \text{HF}$ contribution is parameterised, resulting in two parameterisations $N_i^{bb4\ell}(m_t, \Gamma_t)$ and $N_i^{bb4\ell+\text{HF}}(m_t, \Gamma_t)$, which are used in the statistical fit. Second, the top-quark mass parameter is shifted by $(m_t - 170)\,\text{GeV}$ in the parameterisation in order to obtain a similar order of magnitude for the m_t and Γ_t parameters. Finally, the m_t and Γ_t dependence of the event yields per bin in the control regions is not parameterised. The impact of these choices is shown in Table A.1.

To compare the different parameterisation choices, the absolute difference between the fitted top-quark mass and decay width values (m_t^{fit}, Γ_t^{fit}) and the values used to generate the pseudo-data set (m_t^{in}, Γ_t^{in}) is divided by the uncertainty in the fitted parameter values obtained in the fit without systematic uncertainties,

$$(\Delta/\sigma)_{m_t} = \frac{|m_t^{\text{fit}} - m_t^{\text{in}}|}{\sigma_{m_t}^{\text{stat.-only}}} \text{ and } (\Delta/\sigma)_{\Gamma_t} = \frac{|\Gamma_t^{\text{fit}} - \Gamma_t^{\text{in}}|}{\sigma_{\Gamma_t}^{\text{stat.-only}}}. \tag{A.1}$$

The average and the maximum values of $(\Delta/\sigma)_{m_t}$ and $(\Delta/\sigma)_{\Gamma_t}$ are evaluated for input top-quark masses $m_t \in \{171, 172, 172.5, 173, 174\}\,\text{GeV}$ and input top-quark decay widths of $\Gamma_t \in \{0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0\}\,\text{GeV}$. In addition, the maximum deviations in $\Delta m_t = |m_t^{\text{fit}} - m_t^{\text{in}}|$ and $\Delta\Gamma_t = |\Gamma_t^{\text{fit}} - \Gamma_t^{\text{in}}|$ within the same m_t^{in} and Γ_t^{in} range are compared in Table A.1.

	$(\frac{\Delta}{\sigma})^{\text{avg}}_{m_t}$	$(\frac{\Delta}{\sigma})^{\text{avg}}_{\Gamma_t}$	$(\frac{\Delta}{\sigma})^{\text{max}}_{m_t}$	$(\frac{\Delta}{\sigma})^{\text{max}}_{\Gamma_t}$	$(\Delta)^{\text{max}}_{m_t}$ [GeV]	$(\Delta)^{\text{max}}_{\Gamma_t}$ [GeV]
default	0.13	0.47	0.38	0.89	0.02	0.04
no $bb4\ell$ + HF param.	0.42	0.54	0.74	1.39	0.06	0.07
incl. CR param.	0.13	0.47	0.39	0.88	0.03	0.04
no m_t shift	0.14	0.49	0.36	1.15	0.03	0.05

Table A.1: Comparison of the average absolute differences $(\Delta/\sigma)_{m_t} = |m_t^{\text{in}} - m_t^{\text{fit}}|/\sigma$ and $(\Delta/\sigma)_{\Gamma_t} = |\Gamma_t^{\text{in}} - \Gamma_t^{\text{fit}}|/\sigma$ between the input values $(m_t^{\text{in}}, \Gamma_t^{\text{in}})$ and the fitted values $(m_t^{\text{fit}}, \Gamma_t^{\text{fit}})$ obtained from the fits including only statistical uncertainties to pseudo-data sets. The averages are shown for different choices in the derivation of the parameterisation $N_i(m_t, \Gamma_t)$. In the default parameterisation, the $bb4\ell$ + HF contribution is parameterised like the $bb4\ell$ contribution, the control-region yields are not parameterised and the top-quark mass parameter is shifted to $m_t - 170\,\text{GeV}$. The impact of these choices is tested by altering the described parameterisation choices in the tests labelled (no $bb4\ell$ + HF param.), (incl. CR param.), and (no m_t shift), respectively. The average and maximum values of $(\Delta/\sigma)_{m_t}$ and $(\Delta/\sigma)_{\Gamma_t}$, as well as the maximum deviations Δm_t and $\Delta\Gamma_t$, are computed over the ranges $m_t \in \{171, 172, 172.5, 173, 174\}$ GeV and $\Gamma_t \in \{0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0\}$ GeV.

First, it was explicitly tested whether neglecting the m_t and Γ_t dependence in the $bb4\ell$ + HF contribution affects the closure of the fit (no $bb4\ell$ + HF param.). For this test, only the $bb4\ell$ component is parameterised in terms of m_t and Γ_t, $N_i^{bb4\ell}(m_t, \Gamma_t)$, while the $bb4\ell$ + HF parameterisation $N_i^{bb4\ell+\text{HF}}(m_t, \Gamma_t)$ is fixed to unity for all m_t and Γ_t values. For the nominal fit model, which includes the m_t and Γ_t parameterisation of the $bb4\ell$ + HF contribution, an average deviation of 0.13σ between the fitted and input mass values is observed. When the $bb4\ell$ + HF parameterisation is omitted, this deviation increases to 0.42σ. The effect is less pronounced for the width extraction: the average deviation amounts to 0.47σ in the nominal fit model and increases slightly to 0.54σ when the $bb4\ell$ + HF dependence is not parameterised. This behaviour is expected, as the $bb4\ell$ + HF contribution predominantly exhibits sensitivity to variations of the top-quark mass, as shown in Figure 7.4.

It is further verified that not parameterising the m_t- and Γ_t-dependence in the control regions has a negligible influence on the closure of the pseudo-data tests (incl. CR param). Only negligible differences in the metrics shown in Table A.1 are observed between the default parameterisation, which does not include the control-region m_t- and Γ_t-parameterisation, and the alternative parameterisation where the control-region yields are also parameterised. It should be emphasised that in these pseudo-data tests the pseudo data are generated from the same MC prediction that is used to derive the parameterisation. Consequently, the statistical fluctuations in the pseudo data correspond exactly to the fluctuations present in the MC templates used to construct the parameterisation.

Finally, the influence of parameterising $N_i(m_t, \Gamma_t)$ in terms of m_t or $m_t - 170\,\text{GeV}$ is studied (no m_t shift), while the latter choice is used in the nominal parameterisation. While no significant change in the average deviations $(\Delta/\sigma)^{\text{avg}}_{m_t}$ and $(\Delta/\sigma)^{\text{avg}}_{\Gamma_t}$ is observed, omitting the mass shift leads to more pronounced deviations for pseudo-data points with $m_t = 171\,\text{GeV}$ and $\Gamma_t < 1.2\,\text{GeV}$, where $(\Delta/\sigma)_{\Gamma_t} > 1$ is observed.

The results of a different kind of pseudo-data test are shown in Figure A.1. In this test, instead of using the nominal parameterisation $N_i(m_t, \Gamma_t)$, which is derived from all available template distributions of varied top-quark masses and decay widths, the event yield parameterisation $N_i'(m_t, \Gamma_t)$ is derived by omitting all template distributions that have the same top-quark mass or decay width (or both) as the pseudo-data set. This pseudo-data fit therefore tests whether top-quark mass and decay width values that are not used in the parameterisation can still be correctly recovered. As seen in Figure A.1, this is indeed the case if the (m_t, Γ_t)-dependence is interpolated. In contrast, extrapolating to masses outside the range of template values — this is the case in the pseudo-data tests with $m_t = 171\,\text{GeV}$ and $m_t = 174\,\text{GeV}$ — leads instead to a larger discrepancy between fitted and input mass and decay width values, as expected.

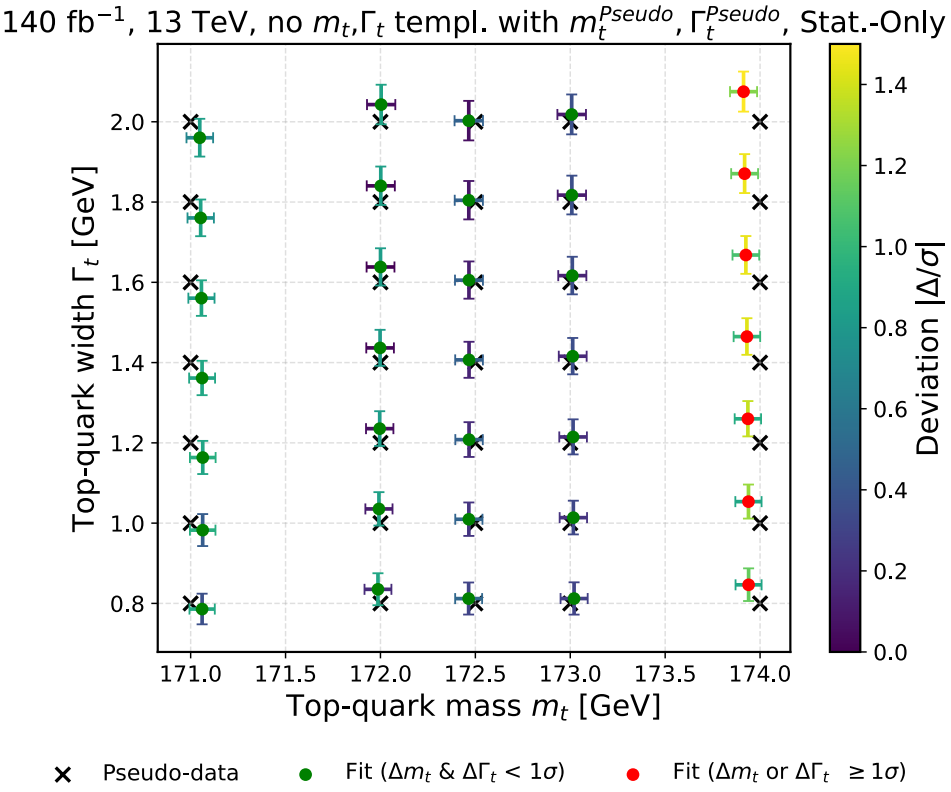


Figure A.1: Same comparison as in Figure 9.3, but using the parameterisation $N_i'(m_t, \Gamma_t)$ instead of the nominal $N_i(m_t, \Gamma_t)$. For each tested pseudo-data point (m_t, Γ_t), $N_i'(m_t, \Gamma_t)$ is rederived by excluding all template points with the same top-quark mass or decay width as the pseudo data. For example, when testing $m_t = 172\,\text{GeV}$ and $\Gamma_t = 2.0\,\text{GeV}$, all $bb4\ell$ predictions with $m_t = 172\,\text{GeV}$ (for all Γ_t) and with $\Gamma_t = 2.0\,\text{GeV}$ (for all m_t) are excluded during the derivation of $N_i'(m_t, \Gamma_t)$.

Alternative parameterisations of $N_i(m_t, \Gamma_t)$ were also tested using linear, cubic, or quartic polynomials instead of the default quadratic form. Similar comparisons as made in Table A.1 for different parameterisation choices are also performed for the different polynomial orders in Table A.2. Using a lower-order polynomial reduces agreement between input and fitted

m_t and Γ_t values when all template distributions are used (full templates), but improves the agreement when the parameterisation must be extrapolated (removed templates). It was confirmed, that the worse agreement in the average and the maximum $(\Delta/\sigma)_{m_t}$ and $(\Delta/\sigma)_{\Gamma_t}$, in the test where templates were removed, is indeed dominated by the extrapolated mass points $m_t = 171\,\text{GeV}$ and $m_t = 174\,\text{GeV}$. The quadratic parameterisation is chosen as a default, since the improvement from a linear to quadratic parameterisation in the closure tests — when using all template values — is sizeable, while going to even higher-order polynomials the improvement is less pronounced, especially in the case of a cubic parameterisation. Furthermore, the quadratic parameterisation is more robust compared to the cubic and quartic parameterisation in the extrapolation tests. Although extraction of a parameter outside the range of template m_t and Γ_t values would require the generation of additional template points to ensure correct parameterisation, the improved extrapolation behaviour of the quadratic form compared to higher-order polynomials indicates that it is a more robust parameterisation.

	Full templates				Removed templates			
	$(\frac{\Delta}{\sigma})^{\text{avg}}_{m_t}$	$(\frac{\Delta}{\sigma})^{\text{avg}}_{\Gamma_t}$	$(\frac{\Delta}{\sigma})^{\text{max}}_{m_t}$	$(\frac{\Delta}{\sigma})^{\text{max}}_{\Gamma_t}$	$(\frac{\Delta}{\sigma})^{\text{avg}}_{m_t}$	$(\frac{\Delta}{\sigma})^{\text{avg}}_{\Gamma_t}$	$(\frac{\Delta}{\sigma})^{\text{max}}_{m_t}$	$(\frac{\Delta}{\sigma})^{\text{max}}_{\Gamma_t}$
linear	0.22	0.48	0.48	1.19	0.27	0.61	0.55	1.44
quadratic	0.13	0.47	0.38	0.89	0.52	0.69	1.22	1.51
cubic	0.13	0.47	0.37	0.86	6.12	1.19	15.20	3.51
quartic	0.09	0.48	0.20	0.79	6.02	1.42	18.19	4.37

Table A.2: Comparison of different functional forms of the parameterisation $N_i(m_t, \Gamma_t)$ — linear, quadratic, cubic, or quartic — similar to Table A.1. In the first pseudo-data test (Full templates), all available MC templates of m_t and Γ_t are used, and the average deviations between fitted and input values are computed over the ranges $m_t \in \{171, 172, 172.5, 173, 174\}\,\text{GeV}$ and $\Gamma_t \in \{0.8, 1.0, 1.2, 1.4, 1.6, 1.8, 2.0\}\,\text{GeV}$. In a second test (Removed templates), all MC templates with $m_t = m_t^{\text{in}}$ and $\Gamma_t = \Gamma_t^{\text{in}}$ are excluded when deriving the alternative parameterisation $N_i'(m_t, \Gamma_t)$, and the averages are computed over the same parameter ranges as in the (Full templates) test.

APPENDIX B

Further studies of the top-quark mass and decay width measurement

B.1 Simultaneous or independent extraction of the top-quark mass and decay width

An important question in this analysis is whether the measurement of the top-quark mass and decay width significantly influence each other in a simultaneous fit. As discussed in Section 7.2, the phase-space regions sensitive to variations in m_t and Γ_t only slightly overlap. Therefore, the correlation coefficient between these two parameters of interest in the simultaneous fit in Figure 10.4 is found to be small with approximately 6 %. To test this explicitly, independent Asimov fits of the top-quark mass and decay width are compared to the simultaneous Asimov fit. In the independent fits, either only the top-quark mass or top-quark decay width dependence of the $m_{\ell b}^{\text{minimax}}$ bin yields in the signal region is parameterised, which corresponds to the new parameterisations $N_i(m_t)$ and $N_i(\Gamma_t)$, respectively, while setting $\Gamma_t = 1.32\,\text{GeV}$ or $m_t = 172.5\,\text{GeV}$. Therefore, this approach not only tests the independence of the m_t and Γ_t extraction, but also compares the one-dimensional parameterisation of the event yield per bin, which is only dependent on either m_t or Γ_t, against the two-dimensional $N_i(m_t, \Gamma_t)$ parameterisation. In all cases, a quadratic polynomial is used to derive these N_i parameterisations.

In Figure B.1, the uncertainties on the top-quark mass and decay width, as well as on the normalisation factors $N_{bb4\ell}$ and $N_{bb4\ell+\text{HF}}$, are compared between the independent and the simultaneous m_t- and Γ_t-Asimov fits. Only a marginal change in the m_t sensitivity is observed. This difference might be due to the use of the one-dimensional parameterisation $N_i(m_t)$ in the independent fit, as opposed to the two-dimensional parameterisation $N_i(m_t, \Gamma_t)$ employed in the simultaneous fit.

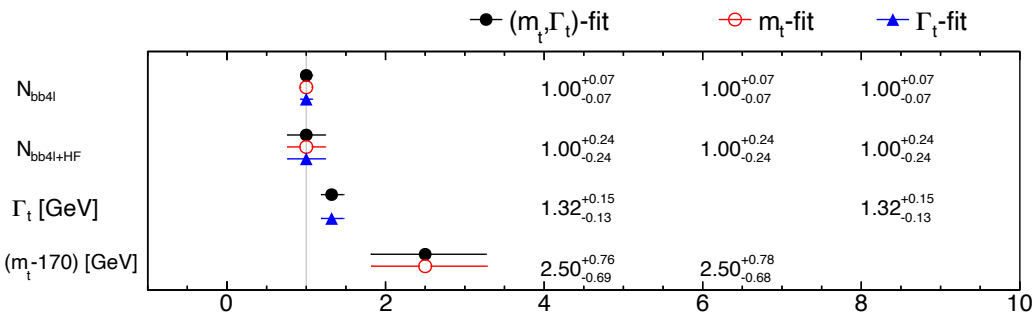
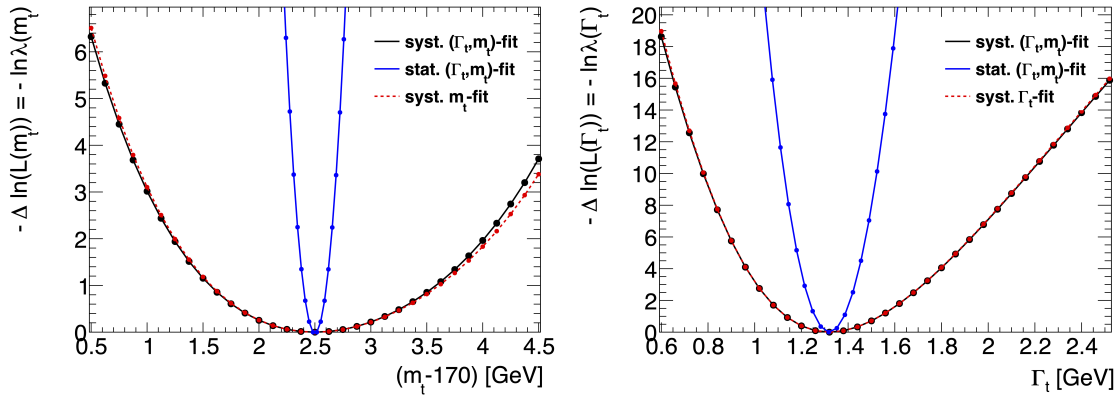


Figure B.1: Same comparison as in Figure 10.1, but for the Asimov fit including the signal, $bb4\ell$+HF and JES control region when extracting simultaneously the top-quark mass and decay width parameter, or when extracting either the top-quark mass or top-quark decay width in separate Asimov fits. Thereby, also the parameterisations $N_i(m_t, \Gamma_t)$, $N_i(m_t)$ and $N_i(\Gamma_t)$ are separately derived in each Asimov fit.

The good agreement in the uncertainty estimate between the simultaneous and the independent Asimov fits of the top-quark mass and decay width is also reflected in the shape of the negative logarithm of the profile likelihood ratio around its minimum of $\hat{m}_t$ and $\hat{\Gamma}_t$, as shown in Figure B.2. These distributions, commonly referred to as *profiled likelihood scans*, are obtained by profiling all other free parameters in the likelihood for fixed values of m_t or Γ_t. The values of m_t and Γ_t at which $-\ln\lambda(m_t) = 0.5$ and $-\ln\lambda(\Gamma_t) = 0.5$ define the 1σ confidence intervals, as discussed in Section 9.2. Good agreement is found when comparing the likelihood scans obtained from the simultaneous and independent fits, both including all systematic uncertainties. In addition, the likelihood scan for the simultaneous fit without including systematic uncertainties is shown, illustrating the strongly reduced uncertainty in the statistical-uncertainty-only Asimov fit, particularly in the case of the top-quark mass measurement. It was explicitly verified, that the impacts of the most dominant systematic uncertainties in the simultaneous or independent measurements of the top-quark mass and decay width are very similar.

The independence of the top-quark mass and decay width measurements is further assessed with the two-dimensional likelihood scan obtained from the simultaneous Asimov fit, shown in Figure B.3. The shape of the likelihood in the (m_t, Γ_t)-plane exhibits no significant correlation between the two parameters, as indicated by the absence of a pronounced tilt in the confidence contours. This observation is consistent with the previous results and confirms that the simultaneous extraction of m_t and Γ_t does not introduce a notable mutual dependence. It is therefore demonstrated that using the off-shell $WWbb$ predictions, which introduce a top-quark decay width sensitivity in the high $m_{\ell b}^{\text{minimax}}$ tail, the dependence of the top-quark decay width measurement on the assumed top-quark mass could be significantly reduced compared to measurements utilising $t\bar{t}$ MC to evaluate the signal predictions (refer to Section 5.2).



(a) (b)

Figure B.2: Negative logarithm of the profile likelihood ratio, $-\ln\lambda$, as a function of (a) the top-quark mass and (b) the decay width. The likelihood scans are shown for the simultaneous fit including all systematic uncertainties and only statistical uncertainties, as well as for the independent Asimov fits including all systematic uncertainties. The position where $-\ln\lambda = 0.5$ corresponds to the 1σ confidence interval.

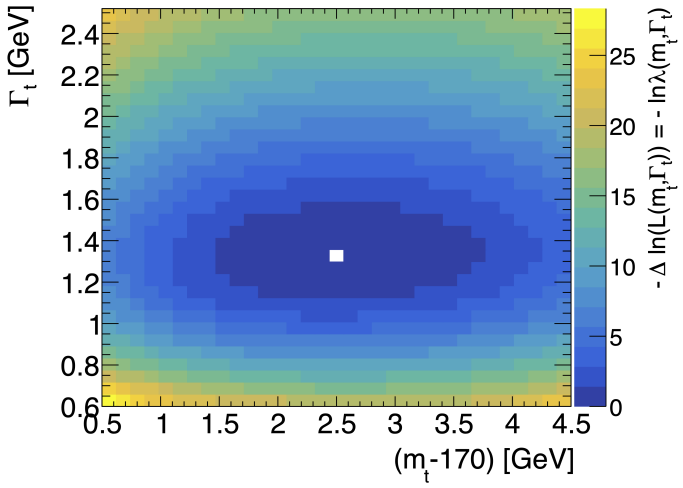


Figure B.3: Two-dimensional distribution of the negative logarithm of the profile likelihood ratio, $-\ln\lambda$, as a function of the top-quark mass m_t and decay width Γ_t, obtained from the simultaneous (m_t, Γ_t) Asimov fit including all systematic uncertainties. The shape of the likelihood contours indicates no significant correlation between the two parameters.

B.2 Influence of jet pre-recommendation uncertainties

As discussed in Section 10.1, the precision of the top-quark mass measurement is strongly affected by the jet pre-recommendation uncertainties. These uncertainties are currently considered for reprocessed Run 2 data, as the in-situ JES and JER calibrations derived with previous reconstruction algorithms are still used. With the upcoming jet uncertainty recommendations, which will include newly derived in-situ JES and JER calibrations, these pre-recommendations are expected to be reduced significantly. Therefore, the sensitivity of the measurement is evaluated using the Asimov data set and removing these uncertainties from the fit model.

In Figure B.4, the uncertainties on the top-quark mass and decay width, as well as on the normalisation factors, are compared between fits including the jet pre-recommendation uncertainties and those excluding them (NoPreRec). This comparison is performed both for Asimov fits including the signal region, the $bb4\ell + \text{HF}$ control region, and the JES control region (3D), and for Asimov fits excluding the JES control region (2D).

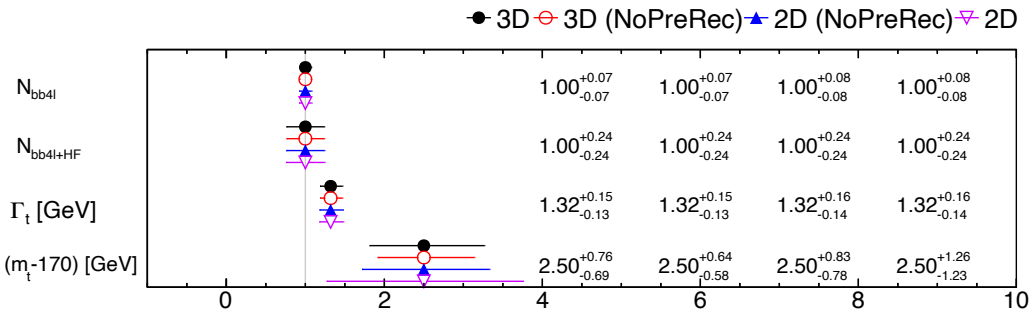


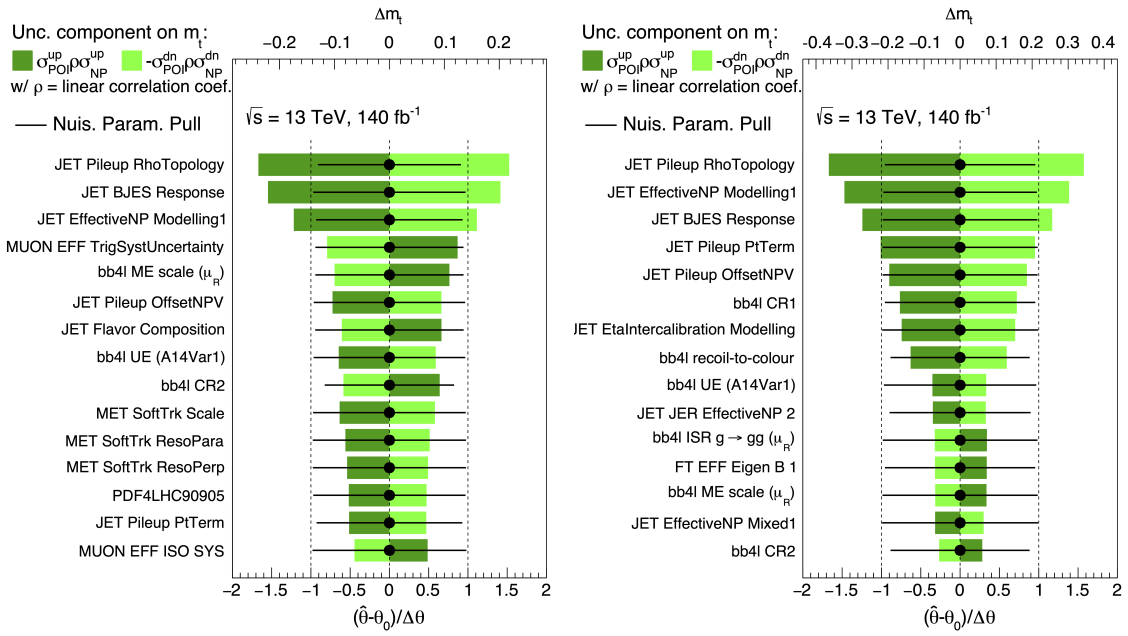
Figure B.4: Same comparison as in Figure 10.1, but for Asimov fits either including or excluding (NoPreRec) the JES pre-recommendation uncertainties and either using the signal region, the $bb4\ell+\text{HF}$ and the JES control region (3D) or only the signal and the $bb4\ell + \text{HF}$ control region (2D).

The exclusion of the jet pre-recommendation uncertainties, as well as the inclusion or omission of the JES control region, has a negligible impact on the uncertainties of the normalisation factors and only a minor effect on the top-quark decay width. This is expected, as these uncertainties are neither strongly correlated with the normalisation factors nor among the dominant contributions to the decay width uncertainty, which is instead driven primarily by $bb4\ell$ and $bb4\ell + \text{HF}$ modelling uncertainties. A slight reduction in the decay width uncertainty is observed in the 3D fit compared to the 2D fit, likely due to the stronger constraint on the $bb4\ell$ modelling uncertainties when including the JES control region.

In contrast, the uncertainty in the top-quark mass is significantly affected by the jet pre-recommendation uncertainties. This effect is most pronounced in the 2D fit, where the JES control region is not included, and the pre-recommendation uncertainties remain weakly constrained by data while being strongly correlated with variations of m_t. In this case, removing the pre-recommendation uncertainties reduces the uncertainty in the top-quark

mass by approximately 35 %. In the 3D fit, where the JES control region provides additional constraints, the reduction is smaller, around 15 %.

When the jet pre-recommendation uncertainties are included, the JES control region is crucial for the top-quark mass measurement, improving the precision on $\hat{m}_t$ by about 42 %. This improvement decreases to roughly 24 % when the pre-recommendation uncertainties are excluded. The most dominant systematic uncertainties in the top-quark mass measurement, when excluding the jet pre-recommendation uncertainties, are shown in Figure B.5, when either (a) including or (b) excluding the JES control region in the Asimov fit.



(a) (b)

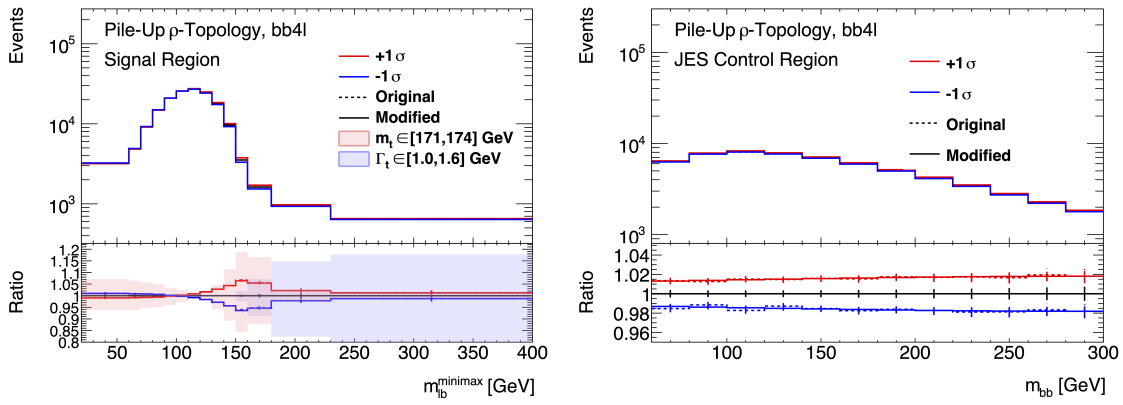
Figure B.5: Same comparison as in Figure 10.5, but for uncertainty estimate on the top-quark mass obtained in the (m_t, Γ_t)-Asimov fit, in which the jet pre-recommendation uncertainties are excluded and (a) all analysis regions are considered or (b) only the signal and the $bb4\ell + \text{HF}$ control region are considered in the profile likelihood fit.

In both the 3D and 2D fits, the leading contributions arise from JES uncertainties related to pile-up subtraction, in-situ calibration, and flavour response. When including the JES control region, additional uncertainties — primarily associated with the normalisation in this region, such as muon trigger efficiency and soft-term calibration affecting the missing transverse momentum — become relevant.

While the introduction of the JES control region could drastically decrease the correlation between the top-quark mass and the JES *VertexingAlg* pre-recommendation uncertainty, and further constrain the latter additionally through data in the JES control region, this is not the case for the most dominant JES uncertainties, when excluding the pre-recommendation uncertainties. As illustrated in Figure B.6 and Figure B.7, the variations induced by the JES ρ-topology pile-up subtraction uncertainty and the JES in-situ calibration modelling

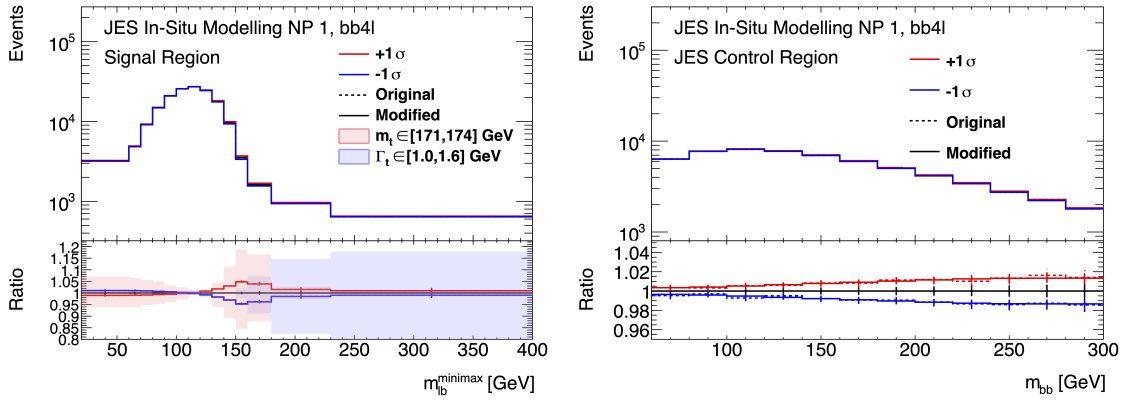
uncertainty in the signal region closely resemble those from changes in m_t, while their impact in the JES control region is comparatively weak. Consequently, the JES control region provides limited additional constraint on these dominant systematic uncertainty sources.

These observations suggest that, once updated in-situ jet calibrations for reprocessed Run 2 data become available, the current JES control region could be replaced by an observable with enhanced sensitivity to the dominant JES uncertainties, or potentially be removed. This would simplify the analysis setup and avoid influencing the nuisance parameter best-fit values related to $bb4\ell$ modelling uncertainties through data in this JES control region, as further discussed in Section 10.2.



(a) (b)

Figure B.6: Same comparison as in Figure 10.2, but for the JES ρ-topology uncertainty from the pile-up subtraction.



(a) (b)

Figure B.7: Same comparison as in Figure 10.2, but for the nuisance parameter 1 from the eigenvector decomposition of sources to the JES in-situ modelling uncertainty.

B.3 Compatibility in validation regions

The post-fit model discussed in Section 10.2 is further studied by comparing both the pre-fit and the post-fit predictions to ATLAS Run 2 data in observables that were not directly used in the profile likelihood fit. To construct the post-fit model, the fitted parameters of interest, normalisation factors and nuisance parameter values are applied to the MC prediction. This allows testing whether the derived corrections are specific to the observables used in the analysis regions or whether they represent more general improvements to the MC modelling. The pre-fit comparisons additionally provide insight into the general agreement between the nominal ATLAS $bb4\ell$ MC prediction (as well as smaller background contributions) with Run 2 data.

In Figure B.8, the transverse momentum distribution of the leading b-jet and the charged lepton in the signal region are shown. For both the pre-fit and the post-fit models, the spectra in simulation are harder than those observed in data. These transverse momentum distributions are sensitive to the modelling of the reconstructed top-quark p_{T}, for which several ATLAS and CMS measurements of the $t\bar{t}$ process have reported a similar mismodelling [298–300]. This data-MC discrepancy is generally attributed to missing higher-order corrections, as applying NNLO QCD reweighting based on the invariant mass and p_{T} of the $t\bar{t}$ system [57, 90, 91] improves the agreement with data for NLO+PS $t\bar{t}$ predictions. Currently, no differential NNLO QCD predictions are available for the $pp \to b\bar{b}\ell^{+}\bar{\nu}_{\ell}\ell^{-}\nu_{\ell}$ process. The question of whether and how NNLO QCD corrections derived in the on-shell $t\bar{t}$ process can be consistently applied to the $bb4\ell$ MC simulation remains under active discussion. Consequently, no NNLO-based reweighting is applied to the nominal $bb4\ell$ prediction in this analysis. Nevertheless, within uncertainties, both the pre-fit and post-fit models show reasonable agreement with the data.

More inclusive event properties, such as the jet multiplicity and the scalar sum of jet transverse momenta are shown in Figure B.9. Also for these observables, good agreement between data and MC predictions is observed within uncertainties, both before and after the fit. Studying such global event properties — without direct sensitivity to the discriminating observable $m_{\ell b}^{\mathrm{minimax}}$— is particularly important for validating the new $bb4\ell$ ATLAS MC sample.

Lastly, the angular distance ΔR between the lepton-b-jet pair with the smallest separation, as well as the missing transverse momentum, are shown in Figure B.8. Since the $\Delta R(\ell, b)$ distribution has a more direct impact on the shape of the $m_{\ell b}$ observable, the post-fit model yields a slightly improved description of the data. For the JES control region, a requirement of $E_{\mathrm{T}}^{\mathrm{miss}} > 60\,\mathrm{GeV}$ is imposed — therefore, the modelling of the missing transverse momentum is also validated, and good agreement between data and MC is observed.

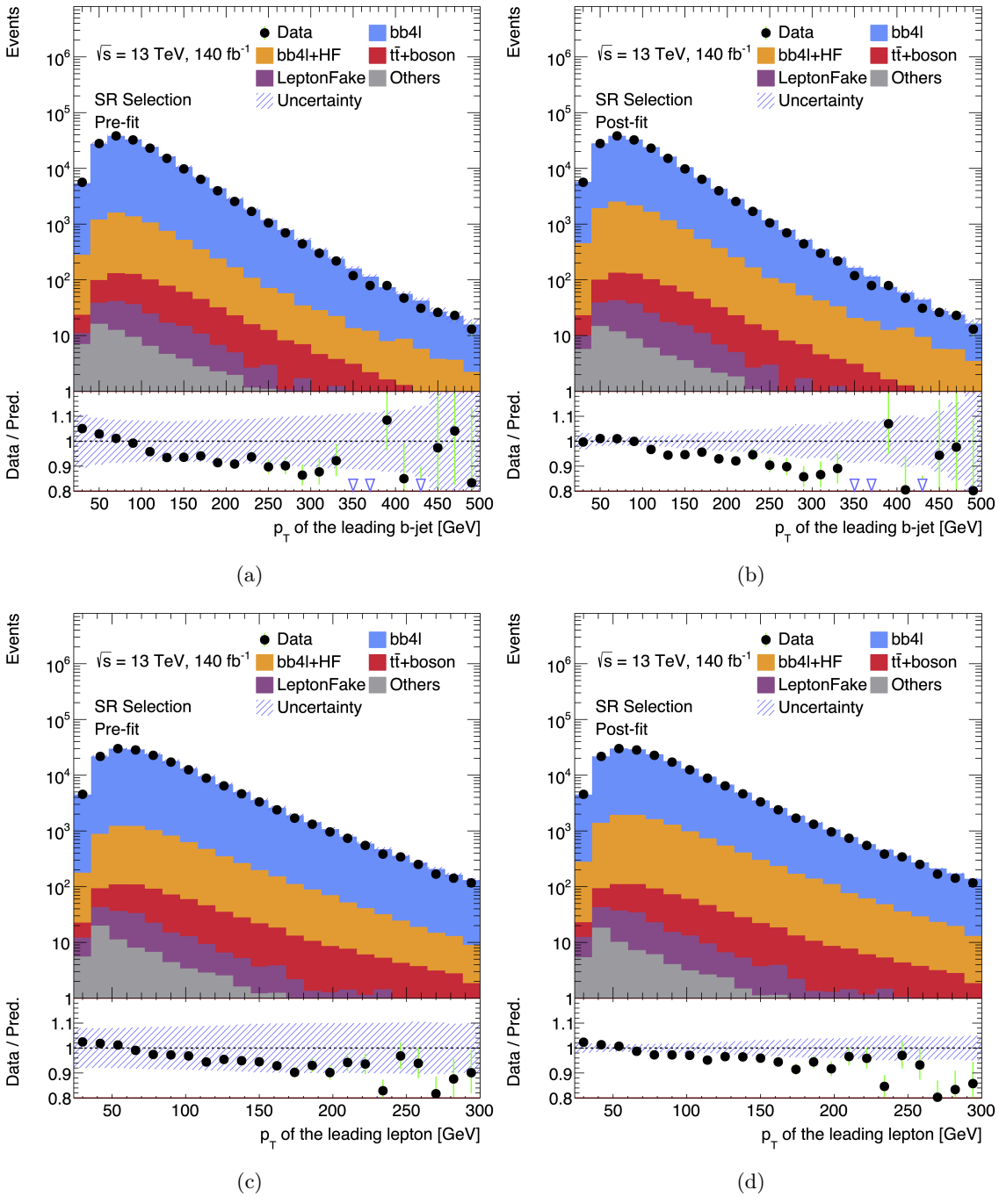


Figure B.8: Comparison of the transverse momentum of the leading — i.e. the highest p_{T}— b-jet (top) and charged lepton (bottom) in the signal region of ATLAS Run 2 data with the total MC prediction, using either the nominal model (Pre-fit, left) or the model, in which the best-fit values $\hat{m}_t$, $\hat{\Gamma}_t$, $\hat{N}_{bb4\ell}$, $\hat{N}_{bb4\ell+\mathrm{HF}}$, and the best-fit nuisance parameter values $\hat{\theta}$ with corresponding uncertainties $\Delta\hat{\theta}$ are applied as determined from the blinded fit (right). The blue uncertainty band shows the post-fit uncertainty in the MC prediction, and the green error bars on the data points represent the statistical uncertainty in these.

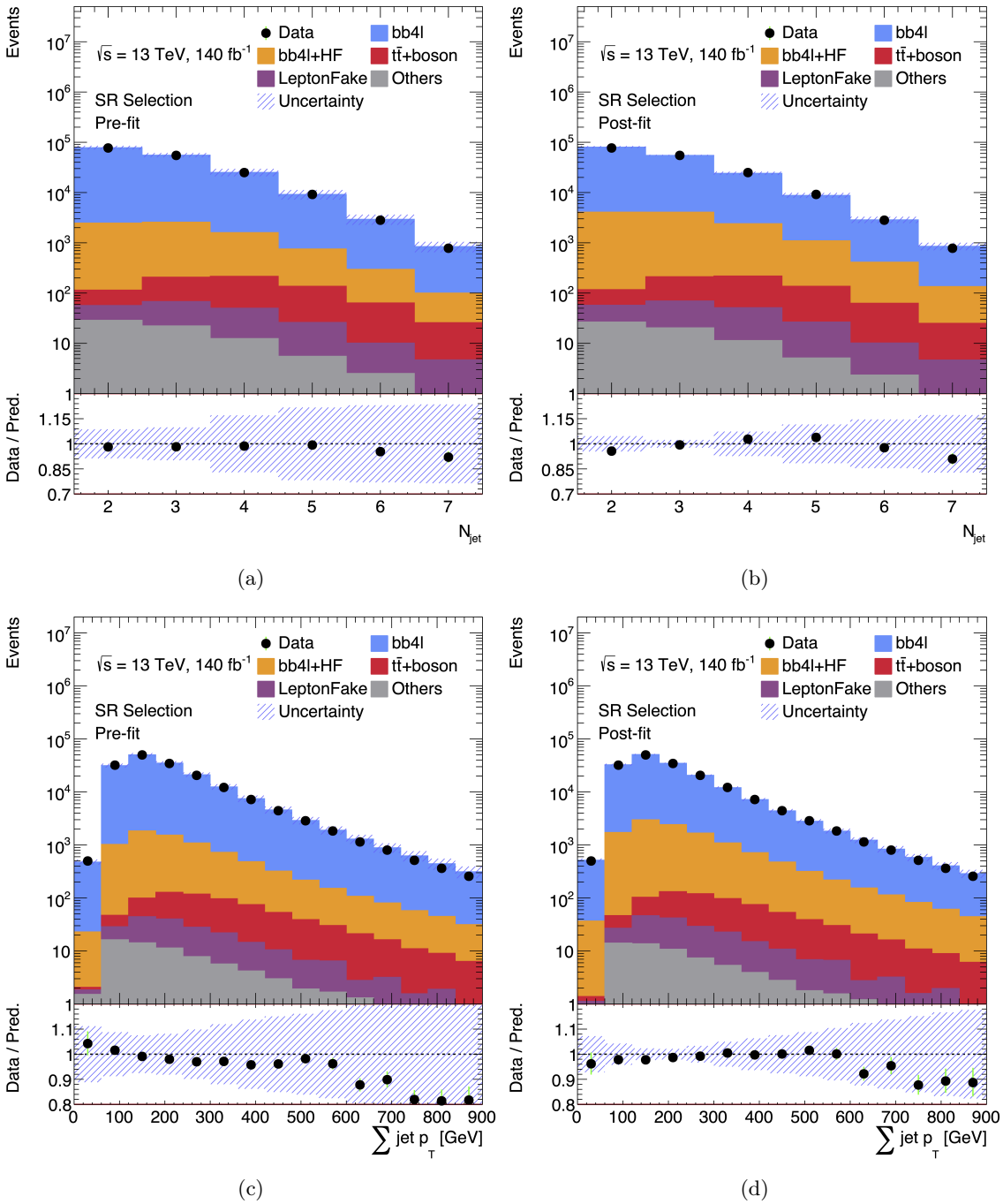


Figure B.9: Same comparison as in Figure B.8, but the jet multiplicity (top) and the scalar sum of all jet transverse momenta (bottom) in the signal region.

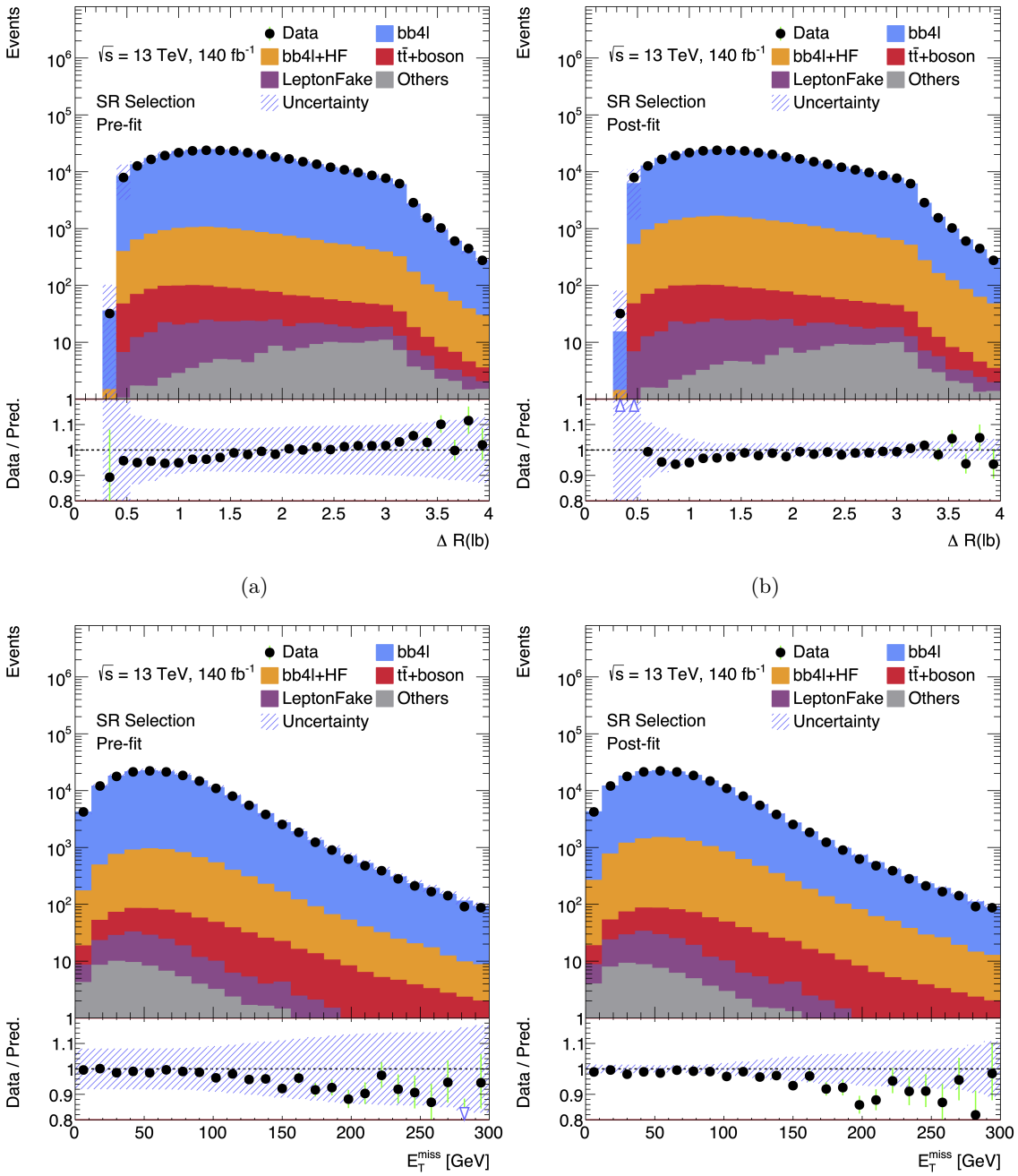


Figure B.10: Same comparison as in Figure B.8, but for the angular distance ΔR between the charged lepton and the b-jet pairs, which are closest in ΔR (top), and the transverse missing momentum (bottom) in the signal region.

APPENDIX C

High-p_{T} jet tagging using a Boosted Decision Tree

The efficiency of flavour-tagging algorithms to correctly classify b-jets significantly decreases as the transverse momentum of the jet ($p_{\mathrm{T}}^{\mathrm{jet}}$) increases. With the extensive data set collected by the ATLAS experiment and the anticipated increase of the data set in the High-Luminosity LHC, high-$p_{\mathrm{T}}^{\mathrm{jet}}$ physics searches are becoming feasible to probe Beyond the Standard Model theories. For these physics analyses, a b-tagging algorithm optimized for the high-$p_{\mathrm{T}}^{\mathrm{jet}}$ region is highly desirable. But, two main difficulties diminish the b-tagging efficiency at high jet p_{T}. As shown in Figure C.1, the number of fragmentation tracks per b-jet grows with increasing $p_{\mathrm{T}}^{\mathrm{jet}}$. Thereby, all reconstructed tracks, which were produced by light hadrons in the detector, are classified as fragmentation tracks, whether the light hadrons originate from the hard process directly or are generated during the parton shower evolution. Furthermore, the track reconstruction efficiency is reduced at high $p_{\mathrm{T}}^{\mathrm{jet}}$ and fewer tracks originating from B-hadron decay are reconstructed, even though the true number of heavy-flavour decay tracks stays constant due to Lorentz-invariance. Consequently, the influence of tracks from B-hadron decay in b-jets at high $p_{\mathrm{T}}^{\mathrm{jet}}$ is reduced in those jet variables, which are reconstructed using the properties of the whole track ensemble of the jet.

To mitigate the problem of the increasing multiplicity of fragmentation tracks at high $p_{\mathrm{T}}^{\mathrm{jet}}$, the strategy presented in this study is to select only those tracks as input to the b-tagging algorithm, which are most likely to originate from a B-hadron decay. This is achieved through the Track Classification Tool (TCT), which assigns to each track a weight, indicating whether the track originates either from a heavy-flavour hadron decay, from fragmentation or from other sources as pile-up or material interaction. The optimisation of the TCT as a tool, which is used during jet classification is described in Section C.2, while details about the used reconstructed objects and truth definitions are given in Section C.1. Only those three reconstructed tracks of a jet, which are assigned the highest heavy-flavour hadron decay weight from the TCT, are used as input to the Classified Track Tagger (CTT) b-tagging algorithm. This procedure is further described in Section C.3, where also a study regarding the number of input tracks to

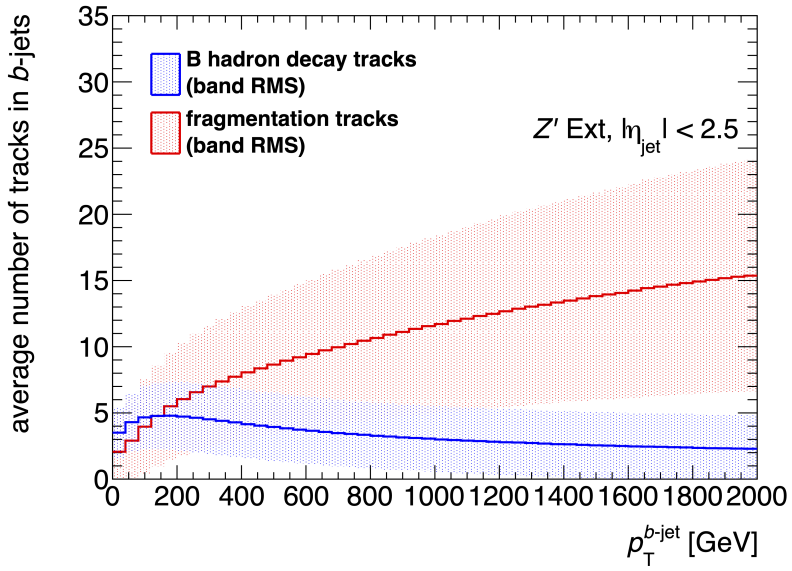


Figure C.1: Average number of B-hadron decay tracks (blue) and fragmentation tracks (red) in b-jets as a function of b-jet p_{T}. The corresponding root mean squared error (RMS) is shown with the shaded region.

the CTT is presented. Both the TCT and the CTT use as underlying multivariate machine learning tool a Boosted Decision Tree (BDT). A reconstructed physics object, in case of the TCT, a reconstructed track and in case of the CTT, a reconstructed jet, is represented to the BDT through several input variables, which are found to have discrimination power for the classification problem at hand. In a single decision tree, a sequence of one-dimensional cuts on the sensitive distributions is applied, to achieve the best separation into the classification categories, e.g. for the CTT the separation into the signal (b-jets) and background (light jets) class. The optimal cut values are determined during the BDT training. In a BDT, not only one decision tree is trained, but instead multiple. The general idea is to focus on those events in a new decision-tree training, which where misclassified in the previous one. This is done by assigning a higher weight to these misclassified events. The final BDT score is calculated by using a weighted sum of the output of the individual decision trees, where higher weights are given to the output of decision trees, which have the lowest classification error. The BDT training performed in this study, for both the TCT and the CTT BDT, was done with the TMVA software package [301]. The final CTT b-tagging performance and a comparison to the performance of the ATLAS b-tagging algorithm DL1r [302], as well as the influence of b-jets without any reconstructed heavy-flavour decay tracks on the CTT classification performance is shown in Section C.4. Lastly, the Conclusions are presented. In the following, a few quantitive statements will be made, while the focus lies in the qualitative study of important input features for high p_{T} b-tagging.

C.1 Monte-Carlo samples and object definitions

Monte-Carlo samples For the b-tagging studies presented in this Appendix, Monte-Carlo samples are needed which provide a source of light and b-jets over a wide range of transverse momenta. For the lower $p_{\mathrm{T}}^{\mathrm{jet}}$ range of $0.02 - 0.5\,\mathrm{TeV}$, these jets are mainly provided by a Powheg+Pythia 8 NLO $t\bar{t}$ sample [59, 85, 303], in which the top quark decays either in the dilepton or single lepton channel. In the generation of this sample, the A14 Pythia 8 tune[1] [95] was used and the Powheg parameter influencing the generation of the hardest emission h_{damp} was set to $1.5\,m_{\mathrm{top}}$ with the ATLAS default setting of $m_{\mathrm{top}} = 172.5\,\mathrm{GeV}$. The PDF set NNPDF3.0NLO [93] was used in the MC sample generation. Jets with higher transverse momenta are obtained from a Z' simulation sample. These Z' events are simulated at LO in QCD with Pythia 8 also using the A14 tune and the NNPDF2.3LO PDF set [243]. In order to obtain a broad p_{T} spectrum of the jets produced in the decay of the Z' resonance, an event-by-event reweighting is applied to increase the natural width of the resonance. Further, to ensure a balanced amount of light, b- and c-jets in the MC sample, the branching fractions of the Z' to bb, cc and to two light jets are each set to one-third [302].

Selection criteria on reconstructed objects The jets analysed in this study are formed by clustering particle-flow objects using the anti-k_t jet clustering algorithm with a radius parameter of $R = 0.4$ (see Section 4.4.1). To meet the selection criteria, a jet must have a transverse momentum p_{T} larger than 20 GeV and absolute value of pseudorapidity $|\eta|$ less than 2.5. Additionally, the jet must not be identified as pile-up jet by the Jet Vertex Tagger (JVT) [304]. Jets are also excluded in the selection, if an electron or muon[2] falls within a ΔR cone of 0.2 of the jet's direction. Further, the jet is required to have at least three associated tracks.[3] Also, the reconstructed tracks need to pass selection cuts based on the number of hits in the ID, the impact parameters and the quality of the track reconstruction, as specified in Table C.1. As listed in this table, the transverse momentum of the track needs to be larger than 700 MeV and is not allowed to be larger than the jet p_{T}. This corresponds to a rather loose track selection, especially compared to the selection applied in the DL1r [302] and graph neural network taggers [188] in ATLAS.

[1] For both the $t\bar{t}$ and the Z' MC sample described in the following, the A14 Pythia 8 tune, which was derived using the NNPDF2.3LO PDF set [243] was used.

[2] Both, the reconstructed electron and the reconstructed muon need to pass the identification working point *Medium* [190, 192]. Compare also to Section 4.4.3 and Section 4.4.4.

[3] In this study, reconstructed tracks are matched to a jet, if they are within a $\Delta R = 0.4$ cone around the jet direction. At least three associated tracks are required per jet, since the CTT will need three input tracks. Also passing default values as input to the CTT for missing tracks (i.e. in the case of jets with less than three tracks) is possible, but would lead to spikes in the CTT discriminant distribution, of which the origin is well understood though. In fact, during the CTT BDT training with higher number of input tracks in Section C.3, the jet selection criterion still requires only more than three reconstructed tracks per jet.

$N_{\text{pixel hits}}$	> 0
$N_{\text{SCT hits}}$	> 3
$N_{\text{pixel + SCT hits}}$	> 6
$N_{\text{shared hits}}$	< 2
z_0	$-15\,\text{mm} < z_0 < 15\,\text{mm}$
d_0	$-5\,\text{mm} < d_0 < 5\,\text{mm}$
$p_{\mathrm{T}}^{\text{track}}$	$700\,\text{MeV} < p_{\mathrm{T}}^{\text{track}} < p_{\mathrm{T}}^{\text{jet}}$
χ^2	$\chi^2/\text{nof} \leq 5$
σ_{3D}	> 1
$\sigma_Z^2 + \sigma_R^2$	> 1

Table C.1: Analysis cuts applied on reconstructed tracks. Selection criteria are based on the number of hits in the inner detector pixel ($N_{\text{pixel hits}}$) and the silicon strip detector ($N_{\text{SCT hits}}$), as well as to the sum of the hits in pixel and strip detector ($N_{\text{pixel + SCT hits}}$). Further, the number of shared hits between two reconstructed tracks is also considered. Requirements on the longitudinal (z_0) and transverse (d_0) impact parameters are set, as well as on the track transverse momentum (p_{T}). Loose requirements on the track quality are placed by setting a threshold on the χ^2 obtained in the track reconstruction fit. Finally, selection criteria on the signed impact parameter significances are applied, which reject tracks, which are likely to originate from the primary vertex.

Jet truth definitions The true jet flavour is inferred by a ΔR matching of truth heavy-flavour hadrons to reconstructed jets. If a b quark or a B-hadron is found within a ΔR cone of 0.3 around the direction of a reconstructed jet, this jet is declared a b-jet. The same D-hadron-to-jet-matching procedure is performed for c-jets, if the corresponding jet has not been declared as a b-jet. This ensures that b-jets, in which a B-hadron decays into a D-hadron are still included in the b-jet category. All other jets are classified as light jets.

Track truth definitions Reconstructed tracks are divided into three categories in this study: B-hadron decay tracks (also called in the following heavy-flavour (HF) tracks), fragmentation and *garbage* tracks. A reconstructed track is classified as B-hadron decay track, if it is associated to a truth hadron in the event record that has a B-hadron as ancestor. Thereby, the decay chain is traced back for four steps. The broader class of *garbage* tracks consists of tracks, which have been produced by GEANT4 [235], which includes tracks from long-lived particle decays and tracks from particles that have been produced in the interaction of particles from the pp collision with the detector material. Additionally, pile-up tracks are included in the *garbage* category, which are identified by not being associated to the hard-process primary vertex. Tracks that do not fall in the B-hadron decay or *garbage* track categories are classified as fragmentation tracks.

C.2 Track Classification Tool

The aim of the Track Classification Tool (TCT) is to determine the physics process from which a reconstructed track originates. Underlying this multivariate analysis tool is a multi-class BDT, which assigns to each track a weight of stemming either from B-hadron decay, from fragmentation or from the *garbage* category. The interpretation of these weights, respectively called `wgtB`, `wgtL` and `wgtG`, as probabilities is natural, since they add up to one. As input to the three-category BDT, ten input variables are used, which describe properties of the track and of the jet to which the track is assigned. These variables are summarised in Table C.2.

input variable	description
σ_{3D}	signed 3D impact parameter significance
σ_R	signed transverse impact parameter significance
σ_z	signed longitudinal impact parameter significance
d_0	transverse impact parameter
prb_P	transformed transverse track momentum $p_{\text{T}}^{\text{track}}$ w.r.t. transverse jet momentum $p_{\text{T}}^{\text{jet}}$ $\text{prb}_P = \frac{p_{\text{frac}}}{1+p_{\text{frac}}}$, $p_{\text{frac}} = \frac{p_{\text{T}}^{\text{track}} - p_{\text{T}}^{\text{cut}}}{\sqrt{p_{\text{T}}^{\text{jet}}}}$ with $p_{\text{T}}^{\text{cut}} = 700\,\text{MeV}$
$p_{\text{T}}^{\text{rel}}$	transverse component of p^{track} w.r.t. p^{jet}
$p_{\text{T}}^{\text{jet}}$	transverse momentum of the jet to which the track was assigned (jet-track assignment through ΔR cone of 0.4)
$N_{\text{hits IBL}}$	number of hits in the innermost B-layer
$N_{\text{hits BL}}$	number of hits in the next-to-innermost B-layer
η_{track}	pseudorapidity of the track

Table C.2: Input variables in the BDT training of the Track Classification Tool.

For the purpose of incorporating the TCT track-classification output into the CTT jet-classification algorithm, in this study, the multi-class BDT of the TCT was retrained. Compared to the BDT training of the previous TCT model, the retraining is performed with increased training statistics. Additionally, rather than training a single BDT over the entire $p_{\text{T}}^{\text{jet}}$ range as in the previous approach, five separate BDTs were trained, each optimized for one of five distinct $p_{\text{T}}^{\text{jet}}$ regions. The number of training N_{train} and testing events N_{test} per track category used in these $p_{\text{T}}^{\text{jet}}$ regions is listed in Table C.3.[4] Additionally, in Table C.3 the TCT classification power of heavy-flavour decay tracks against the other two track categories is shown through the area under the ROC curve (AUC) value. The ROC (Receiver Operating Characteristic) curve thereby indicates how the true positive and the false positive rate depend on the decision threshold x of e.g. the heavy-flavour decay track classifier output `wgtB`.

[4] Thereby, it was ensured that the $p_{\text{T}}^{\text{jet}}$ distribution is identical for the different track categories by using the down-sampling method, i.e. selecting the same amount of HF, fragmentation and *garbage* tracks in each bin of a finely binned $p_{\text{T}}^{\text{jet}}$ histogram. This is important in order to ensure that the track classification is not dependent on the jet p_{T}, which may differ in the MC samples used in the BDT training between light and b-jets, of which only the latter is a source of HF tracks.

Thereby, a reconstructed track with a BDT output of `wgtB` $> x$ is classified as an HF track. While an AUC value of 1 corresponds to perfect classification, an AUC value of 0.5 indicates a random classifier. As expected, for increasing values of $p_{\mathrm{T}}^{\mathrm{jet}}$ the separation power between the different track categories decreases since the dense jet core at high p_{T} degrades the track reconstruction.

$p_{\mathrm{T}}^{\mathrm{jet}} \in$	$[0.02, 0.2]$ TeV	$[0.2, 0.5]$ TeV	$[0.5, 1]$ TeV	$[1, 2]$ TeV	$[2, 7]$ TeV
$N_{\mathrm{train}} = N_{\mathrm{test}}$	3.3 M	5.2M	4.1 M	6.7 M	6.3M
AUC	0.937	0.927	0.912	0.898	0.876

Table C.3: AUC value (HF vs. other tracks) and number of training and testing events per track category in the BDT training of the retrained TCT model.

This can also be seen when comparing the ROC curves[5] in Figure C.2 directly. Thereby, the heavy-flavour track efficiency is shown against the other track rejection, which includes both fragmentation and *garbage* tracks for the different $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges. Besides the performance of the newly trained BDT, also the comparison to the previous TCT BDT model is shown in this figure. For all $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges an improvement in the classification power is found.

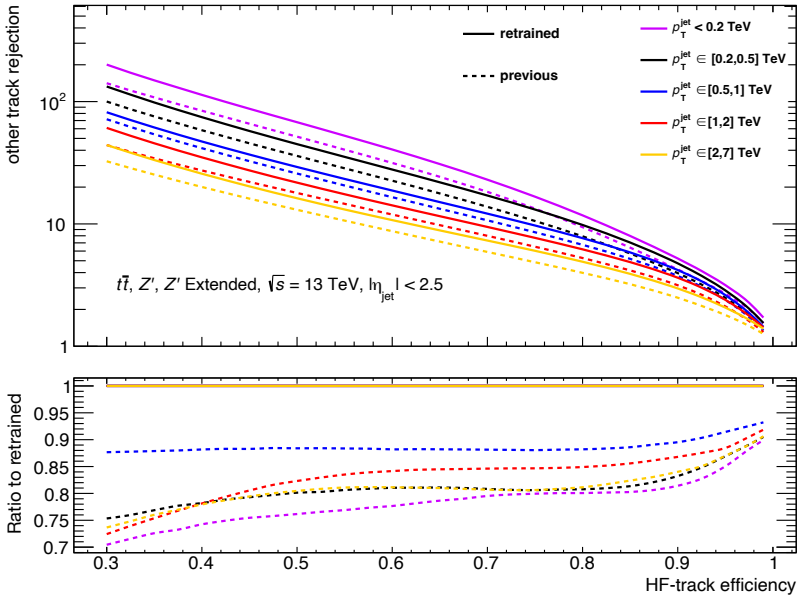


Figure C.2: Comparison of the ROC curves of the previous TCT BDT and the newly retrained TCT BDT model in different $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges. The ROC curves presented in this figure are, as in the remainder of this Section, obtained by evaluating the BDT only on the testing events.

[5] In this Appendix, several figures will be described as displaying ROC curves, while thereby the standard procedure in the ATLAS community is followed of showing as rejection (on the y-axis) the inverse of the false positive rate. This rejection power can be easily interpreted, as e.g. a rejection power of 100 indicates that 1 out of 100 tracks are misidentified.

To gain deeper insights into the most critical input features for track classification, an importance ranking of the input variables is generated using the TMVA framework. This ranking is determined by analysing how frequently a decision tree splits based on a specific input variable and assessing the significance of each split in distinguishing between signal and background [301]. For all $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges, the highest ranked variable is the transformed track momentum prb_P. Further, for all $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges the three following highest ranked variables contain the transverse component of the track w.r.t. the jet three-momentum $p_{\mathrm{T}}^{\mathrm{rel}}$, the transverse impact parameter d_0 and the signed transverse impact parameter significance σ_R.[6] The distributions of these variables for an example $p_{\mathrm{T}}^{\mathrm{jet}}$ range of $p_{\mathrm{T}}^{\mathrm{jet}} \in [1, 2]\,\mathrm{TeV}$ is shown in Figure C.3 for heavy-flavour (HF, blue), fragmentation (red) and *garbage* tracks (green) truth matched either to a light (dashed) or a b-jet (solid).

The highest ranked variable, the transformed track momentum prb_P, describes how much of the total transverse momentum of the jet is carried by the reconstructed track (Figure C.3(a)). The prb_P distribution obtained from reconstructed tracks from B-hadron decay shows a clear peak toward large values of this variable. This is understood since a b quark, being a heavy quark, radiates less collinear partons due to the dead-cone effect and thus retains more of its initial momentum after parton showering. Furthermore, b quarks typically originate from the primary interaction vertex and initiate a b-jet, thereby carrying a large fraction of the jet momentum. In contrast, $g \to b\bar{b}$ splittings in the parton shower — which would result in smaller values of prb_P — are relatively rare. In Figure C.3(a), it is additionally shown, that the fragmentation tracks in light and b-jets show a different behaviour in the prb_P distribution.[7] Due to the large B-hadron mass, the three-momenta of the B-hadron decay products can have a sizable transverse component $p_{\mathrm{T}}^{\mathrm{rel}}$ w.r.t. the jet three-momentum also after boosting from the B-hadron rest frame to the laboratory frame. This effect is shown in Figure C.3(b), while fragmentation and *garbage* tracks tend to smaller values of $p_{\mathrm{T}}^{\mathrm{rel}}$. Further, in the lower two plots in this figure the distributions of the transverse impact

[6] The transverse impact parameter d_0 is defined as the transverse distance of the point of closest approach of the reconstructed track to the primary vertex. The signed impact parameter significance is defined as the impact parameter divided by its estimated uncertainty. The location of a possible secondary vertex w.r.t. the primary vertex is indicated by the sign of σ_R. While a secondary vertex after the primary vertex — in the direction of flight of the particle — leads to a positive sign, a negative sign of σ_{R} would signify a secondary vertex before the primary vertex. While HF decay tracks (and *garbage* tracks from material interaction) will tend to have positive signed impact parameter significance values, fragmentation tracks have a symmetric distribution of σ_{R} around zero due to resolution effects. See e.g. [305] for a nice graphic representation.

[7] This difference is not completely unexpected, since in light jets, the hadron tracks labelled as fragmentation tracks can originate, i.e. can retain a large momentum fraction from the primary parton produced in the hard process. If fragmentation tracks from light and b-jets have considerably distinct TCT input features, a decision has to be made during the TCT training, if fragmentation tracks only from light or from light and b-jets should be used. This was tested by trying to separate fragmentation tracks from light and from b-jets through a BDT, using the same input parameters as for the TCT BDT. Thereby, only small separation power between the two classes was found (AUC values between 0.54 and 0.64 depending on the $p_{\mathrm{T}}^{\mathrm{jet}}$ range). Further, two trainings of the jet-classification algorithm CTT were compared using either a TCT training with fragmentation tracks from light jets or from both light and b-jets. As anticipated, no large differences were found between the CTT training results, amounting only to a few percent differences between the rejection curves over the entire b-jet efficiency range. Thus, the default TCT training, in which the fragmentation tracks in the BDT training are obtained from both light and b-jets, is left unchanged.

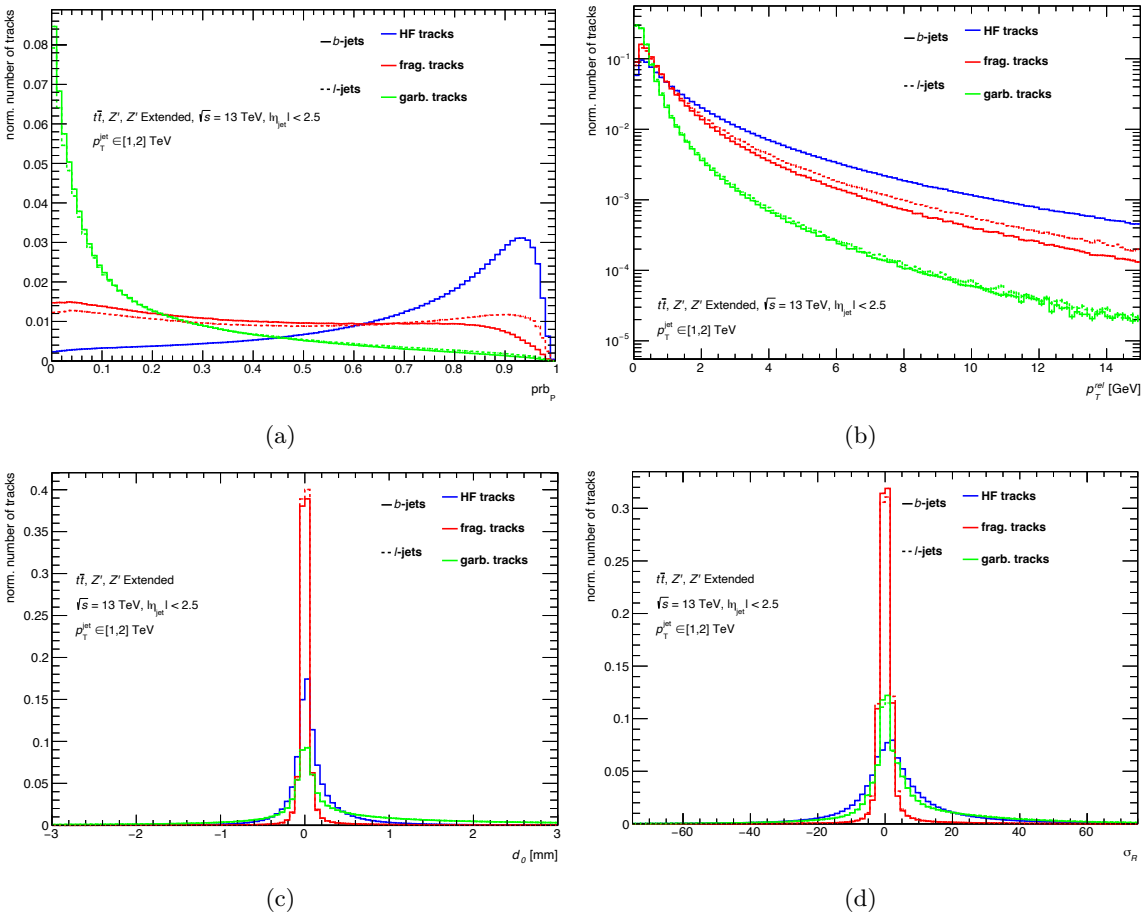


Figure C.3: Distributions of the four highest ranked variables $\mathrm{prb_P}$ (a), $p_{\mathrm{T}}^{\mathrm{rel}}$ (b), d_0 (c) and σ_R (d) in the TCT BDT training for $p_{\mathrm{T}}^{\mathrm{jet}} \in [1, 2]$ TeV.

parameter d_0 (Figure C.3(c)) and the signed transverse impact parameter significance σ_{R} (Figure C.3(d)) are shown. As expected, the HF and *garbage* tracks tend to larger impact parameters compared to fragmentation tracks, as only the latter originate from the primary vertex. Two additional input variables to the TCT were studied, the primary vertex weight, which is determined during the primary vertex fitting procedure (compare to Section 4.3.3) and is a measure of how compatible a given track is with the reconstructed primary vertex of the hard process, and the minimum distance $\min_{j \neq i} \Delta R(i, j)$ between the track i to any other track $j \neq i$ in the same jet. The primary vertex weight has a large correlation to the impact parameter distributions, which are already included in the TCT BDT training and its inclusion as input variable to the TCT BDT did not lead to an improvement neither in the TCT nor the CTT performance, when using a retrained TCT model including the primary vertex weight. Some improvement in the TCT classification was found using additionally the $\min_{i \neq j} \Delta R(i, j)$ input variable. The motivation to include this per-track-variable was that b quarks were expected to radiate less collinear partons due to the dead-cone effect and thus the $\min_{i \neq j} \Delta R(i, j)$ distribution would tend to larger values than e.g. for fragmentation

tracks. Indeed, the opposite was found. This could be explained with a high correlation of this variable with prb_P, meaning that the jet direction follows closely the B-hadron track direction and thus the B-hadron track is often located in the dense jet core. Even though, some improvement in the TCT track classification was found including the $\min_{i\neq j} \Delta R(i,j)$ variable, it did not translate to a significant improvement in the CTT jet classification and thus this variable was not used in the default TCT BDT model. As an example, the BDT discriminant distribution for the three categories of HF tracks `wgtB` (a), fragmentation tracks `wgtL` (b) and *garbage* tracks `wgtG` (c) for tracks from the different track categories is shown in Figure C.4.

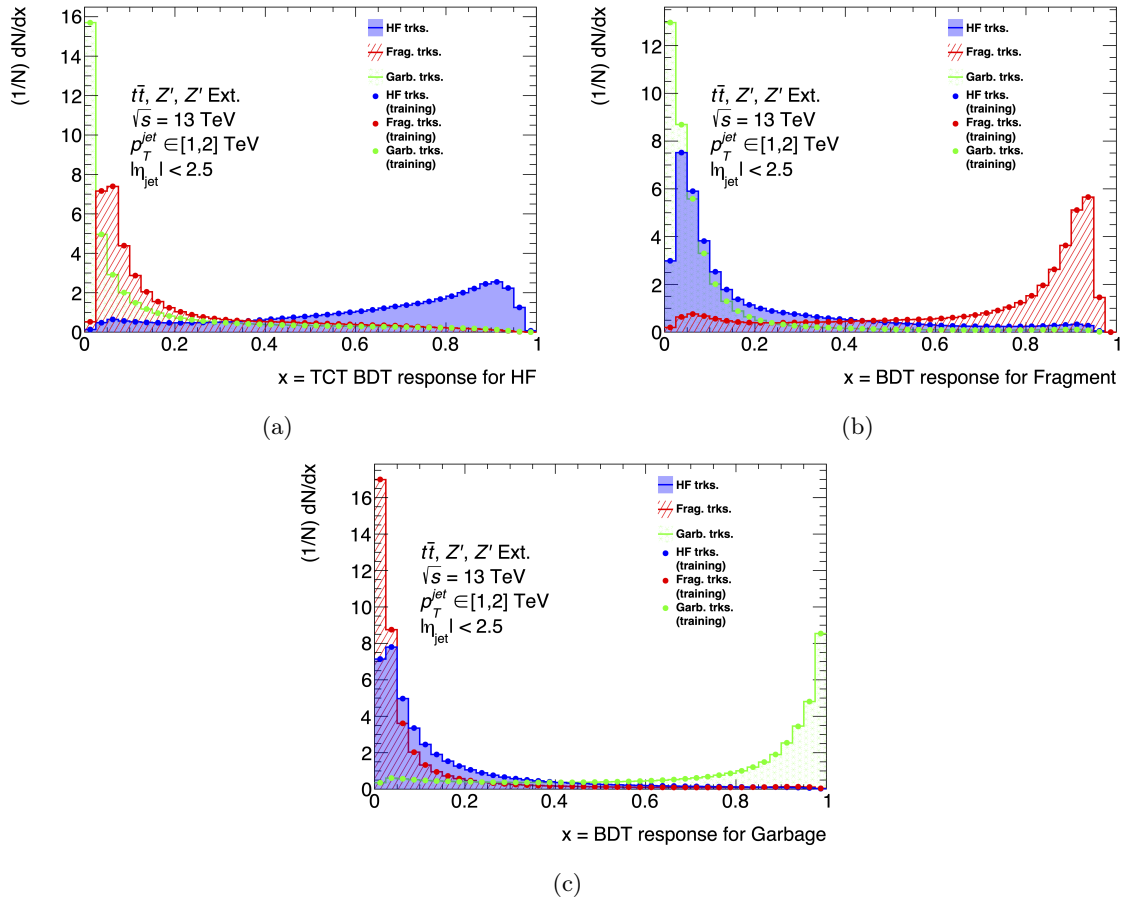


Figure C.4: BDT discriminant distributions for classification as heavy-flavour (HF) (`wgtB`, (a)), fragmentation (`wgtL`,(b)) or *garbage* (`wgtG`,(c)) track for the different track categories from jets with $p_{\mathrm{T}}^{\mathrm{jet}} \in [1,2]$ TeV. No overtraining of the BDT is observed, as the discriminant distributions for both training and testing events are very similar.

Thereby only those tracks are considered in the training and testing of the TCT model shown in Figure C.4, which are assigned to a jet which satisfies $p_{\mathrm{T}}^{\mathrm{jet}} \in [1,2]$ TeV. Figure C.4 demonstrates, that the BDT can separate better fragmentation and *garbage* tracks from the other track categories than it can separate heavy-flavour (HF) tracks from fragmentation

and *garbage* tracks. In conclusion, it was shown that the performance of the TCT could be improved by increasing the training statistics and through training dedicated BDT models in five different $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges. Both the retrained and the previous TCT BDT model are used in the following to study the performance of the CTT. In this way, the propagation of the performance improvement of the retrained TCT to the CTT can be studied.

C.3 Structure of the Classified Track Tagger

As described in the Introduction, the general idea behind the Classified Track Tagger (CTT) is to consider only those tracks in the jet-classification algorithm, which have the highest probability of originating from a B-hadron decay. This probability is quantified by the `wgtB` output of the TCT described in Section C.2. In the default set-up of the CTT, the three tracks inside a jet with the highest values of `wgtB` are used. Of these three input tracks, the three weights (`wgtB`, `wgtL`, `wgtG`) obtained from the TCT are used as input variables for the underlying CTT BDT, namely the weights of the track being a B-hadron decay, fragmentation or *garbage* track. Additionally, the track multiplicity information is added through the input variable $N_{\mathrm{track}}/p_{\mathrm{T}}^{\mathrm{jet}}$, since it was found to improve the jet-classification performance. Thereby, the most important, i.e. the highest ranked input variable, to the CTT BDT is the highest `wgtB` value of all tracks inside the jet. The structure of the CTT BDT is graphically represented in Figure C.5. A single BDT over the full $p_{\mathrm{T}}^{\mathrm{jet}}$ range is trained for the CTT jet-classification algorithm. Still, in the following, the CTT classification performance will be presented in separate $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges to emphasize the difficulty of high $p_{\mathrm{T}}^{\mathrm{jet}}$ tagging and to compare the jet properties of high and low p_{T} jets.

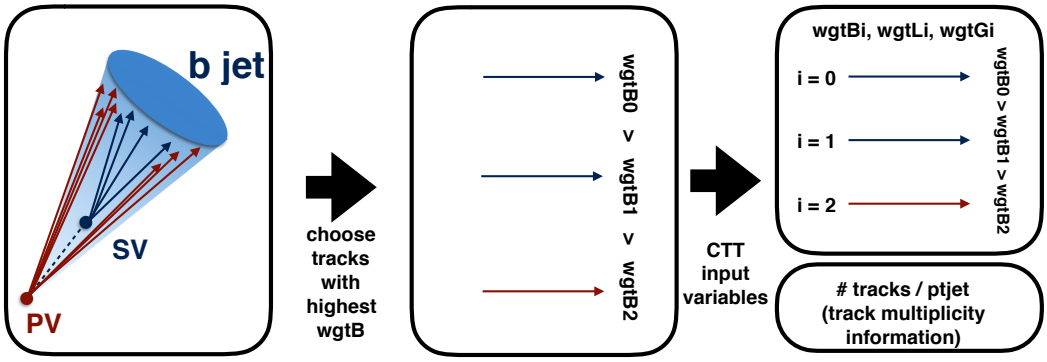
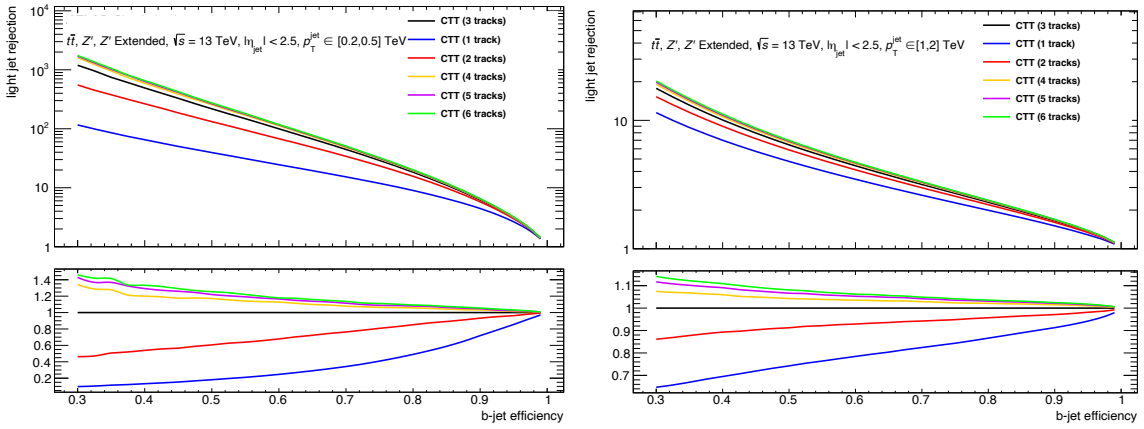


Figure C.5: Graphic representation of the CTT algorithm. From each jet those three tracks with the highest `wgtB` TCT output are selected and the three TCT weights of each of the three tracks used as input variables in the TCT training. As an additional variable the track multiplicity information inside the jet is added in the BDT training.

Number of input tracks A fundamental design choice concerning the CTT BDT structure is the number of input tracks used for the jet classification. To this end, in Figure C.6 the performance of the CTT trained with different number of input tracks is shown for jets in two different $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges.



(a) (b)

Figure C.6: ROC curves obtained with the CTT trained and tested with jets satisfying $p_{\mathrm{T}}^{\mathrm{jet}} \in [0.2, 0.5]$ TeV (a) and $p_{\mathrm{T}}^{\mathrm{jet}} \in [1, 2]$ TeV (b) and different number of input tracks in the CTT BDT.

As expected, the jet classification is improved when using a higher number of input tracks, since more information is provided to the BDT. On the other hand, the relative increase in performance decreases with a growing number of input tracks. This can be well understood, when considering Figure C.1 in the Introduction, as only a limited amount of tracks from the B-hadron decay are actually reconstructed, especially when focusing on the high $p_{\mathrm{T}}^{\mathrm{jet}}$ region. This is reflected in the `wgtB` distributions shown in Figure C.7, in which the `wgtB` value is shown for the highest (left) or second highest ranked(right) `wgtB` track inside either a light (red) or a b-jet (blue).

For jets with $p_{\mathrm{T}}^{\mathrm{jet}}$ in the range [1,2] TeV (Figure C.7, bottom), a reduction in separation power between light and b-jets becomes apparent when considering lower-ranked `wgtB` tracks, such as those with the second-highest `wgtB` value within a jet. In contrast, for jets in lower $p_{\mathrm{T}}^{\mathrm{jet}}$ intervals (Figure C.7, top), the `wgtB` distributions of these lower-ranked tracks continue to show pronounced shape differences between tracks originating from light jets or b-jets. This difference leads to a larger performance gain in jet classification in the lower $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges, when including more input tracks, as illustrated in Figure C.6, where the ROC curves of the CTT, trained on jets within the $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges of [0.2,0.5] TeV (a) and [1,2] TeV (b), are compared for different numbers of input tracks. A 20 % increase in light jet rejection is found at a b-jet efficiency of 70 % in the $p_{\mathrm{T}}^{\mathrm{jet}} \in [0.2, 0.5]$ TeV case, when using six instead of three input tracks in the BDT training and testing, while for jets with $p_{\mathrm{T}}^{\mathrm{jet}} \in [1, 2]$ TeV this increase of light jet rejection at the same b-jet efficiency amounts only to ~ 6 %. Even though the BDT training is still feasible even with a larger number of input tracks, a reduced number might be beneficial in view of systematic variations of the input features to the BDT. In order to make a more

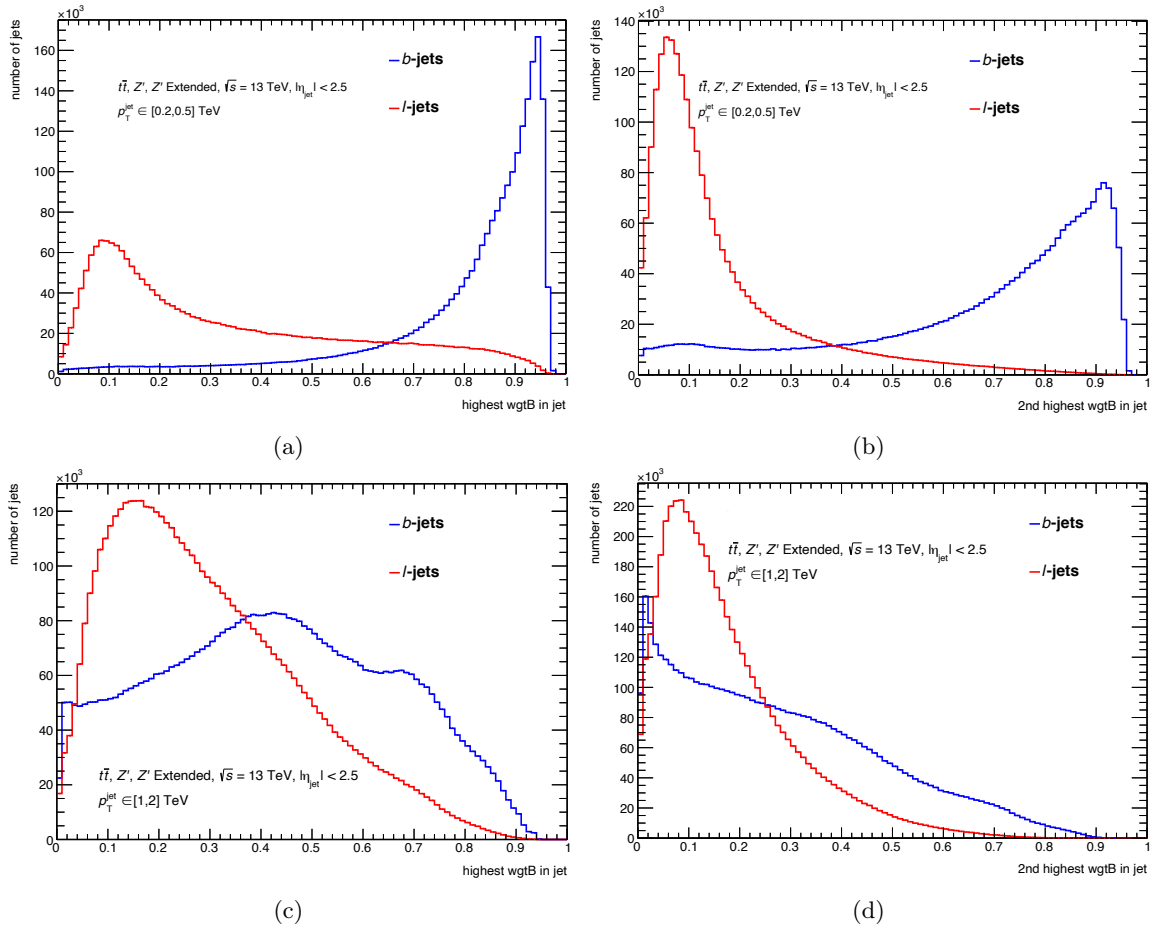


Figure C.7: Distribution of `wgtB` of the first (left, (a) and (c)) and the second (right, (b) and (d)) track inside a jet, when ordered according to `wgtB` for light jets (red) and b-jets (blue) within the range $p_{\mathrm{T}}^{\mathrm{jet}} \in [0.2, 0.5]$ TeV (top, (a) and (b)) and $p_{\mathrm{T}}^{\mathrm{jet}} \in [1, 2]$ TeV (bottom, (c) and (d)).

informed choice in this matter, a study of the systematic uncertainties would be needed. Since the CTT should in particular be aimed at b-tagging at high $p_{\mathrm{T}}^{\mathrm{jet}}$ and the improvement in performance through the inclusion of additional input tracks is reduced in this kinematic region, the number of input tracks is fixed to three tracks.

C.4 Results

In the following, the comparison of the CTT performance using either the previous or the retrained TCT model is discussed, as well as the comparison to the ATLAS b-tagging algorithm DL1r. Since a significant fraction of b-jets without any reconstructed heavy-flavour tracks was found at high $p_{\mathrm{T}}^{\mathrm{jet}}$, the CTT jet-classification performance on these jets is also described.

Improvement through TCT retraining First it is validated, that the increase in TCT performance through retraining of the BDT described in Section C.2 propagates to an improvement of the CTT jet-classification power. In Figure C.8, the light jet rejection vs. b-jet efficiency for different $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges is shown, using as input variables to the CTT BDT the weights obtained with the previous or the retrained TCT model.

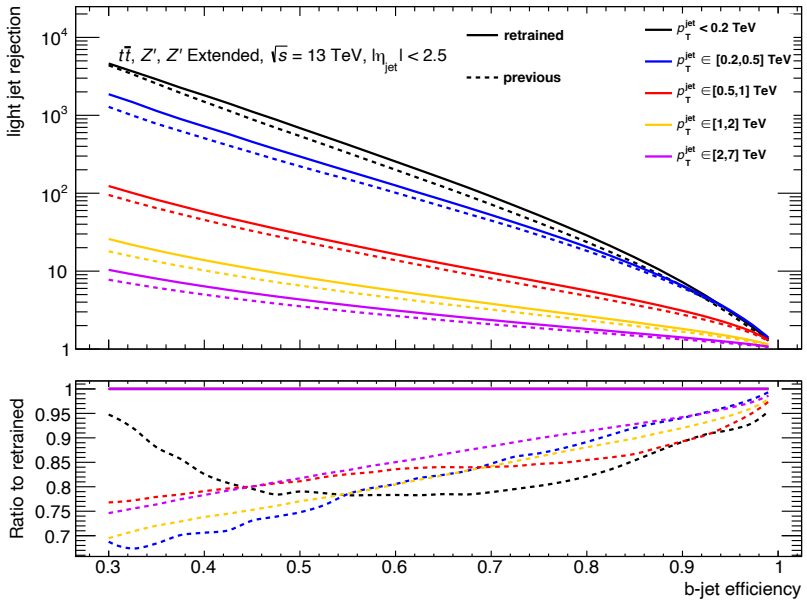


Figure C.8: The light jet rejection vs. b-jet efficiency obtained with the CTT using as input variables either the TCT weights from the previous TCT BDT or the retrained model.

In fact, the CTT performance can be improved by using the retrained TCT model in all $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges. This is expected, since as seen in Section C.2 the TCT performance also improved through retraining for all $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges. As discussed in Section C.3, in lower $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges the TCT performance is improved compared to the high $p_{\mathrm{T}}^{\mathrm{jet}}$ range. This is directly reflected also in the CTT BDT discriminant distributions, which are shown in Figure C.9 for the two $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges $p_{\mathrm{T}}^{\mathrm{jet}} \in [0.2, 0.5]$ TeV and $p_{\mathrm{T}}^{\mathrm{jet}} \in [1, 2]$ TeV.

Comparison with DL1r In Figure C.10, the comparison of the light-jet rejection vs. the b-jet efficiency obtained with the CTT using either the output of the retrained or the previous TCT BDT or with the DL1r b-tagging algorithm is shown. For the lower $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges, the CTT is clearly not competitive to the DL1r algorithm.[8] On the other hand, in the higher $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges, the performance is comparable, even though the CTT structure is much simpler. The neural network which underlies the DL1r uses as input variables the output of several low level taggers, which focus on secondary vertex finding and track origin classification [302]. The training statistics and the time for training this machine learning model are considerable

[8] Note that the lowest $p_{\mathrm{T}}^{\mathrm{jet}}$ range shown in Figure C.10 is between [0.5,1] TeV.

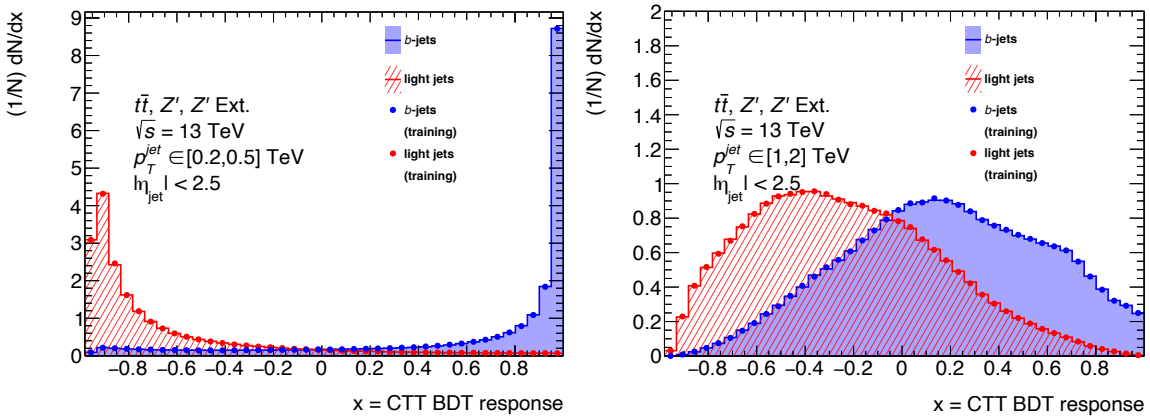


Figure C.9: The CTT BDT discriminant distributions for jets with $p_{\mathrm{T}}^{\mathrm{jet}} \in [0.2, 0.5]\,\mathrm{TeV}$ (a) or $p_{\mathrm{T}}^{\mathrm{jet}} \in [1, 2]\,\mathrm{TeV}$ (b).

and further, retuning of the low-level taggers requires ideally also a retuning or retraining of the neural network of the DL1r. The BDT underlying the CTT depends only on the TCT and further through the relatively simple BDT structure used, the training statistics and the training time is moderate.

Influence of *b*-jets without reconstructed heavy-flavour tracks At high *b*-jet p_{T}, a considerable fraction of *b*-jets without any reconstructed heavy-flavour (HF) decay tracks was found, as shown in Figure C.11. For *b*-jets at 2 TeV, this fraction amounts to approximately 28 %. Since the strategy of the CTT builds upon identifying tracks from the *B*-hadron decay, the influence of these *HF-trackless* *b*-jets on the jet-classification performance is investigated.

In Figure C.12(a), the CTT performance is shown for all *b*-jets, as well as for *b*-jets with and without reconstructed tracks from *B*-hadron decay. Since the fraction of HF-trackless *b*-jets grows with $p_{\mathrm{T}}^{\mathrm{jet}}$, only *b*-jets with high transverse momenta of $p_{\mathrm{T}}^{\mathrm{jet}} \in [1, 2]\,\mathrm{TeV}$ are considered in Figure C.12(a). Clearly, the classification power of the CTT is reduced for *b*-jets without *B*-hadron decay tracks, but a (small) separation from light jets is still possible. This is because the features of fragmentation tracks in *b*-jets differ from those in light jets, and the CTT algorithm is able to exploit these differences. To illustrate this, Figure C.12(b) shows the highest ranked input variable in the CTT BDT training — the maximum `wgtB` value among all tracks in a jet — for *b*-jets without any reconstructed heavy-flavour tracks and for light jets. This can be compared to the corresponding distribution of all *b*-jets in the same $p_{\mathrm{T}}^{\mathrm{jet}} \in [1, 2]\,\mathrm{TeV}$ range in Figure C.7. Fragmentation tracks in light jets are more likely to be classified as *B*-hadron decay tracks than fragmentation tracks in *b*-jets, which is consistent with the discussion of differences in fragmentation tracks between light and *b*-jets in Section C.2. These differences appear to be captured by the CTT BDT. It was verified that training a BDT specifically on *b*-jets without heavy-flavour tracks improves the classification

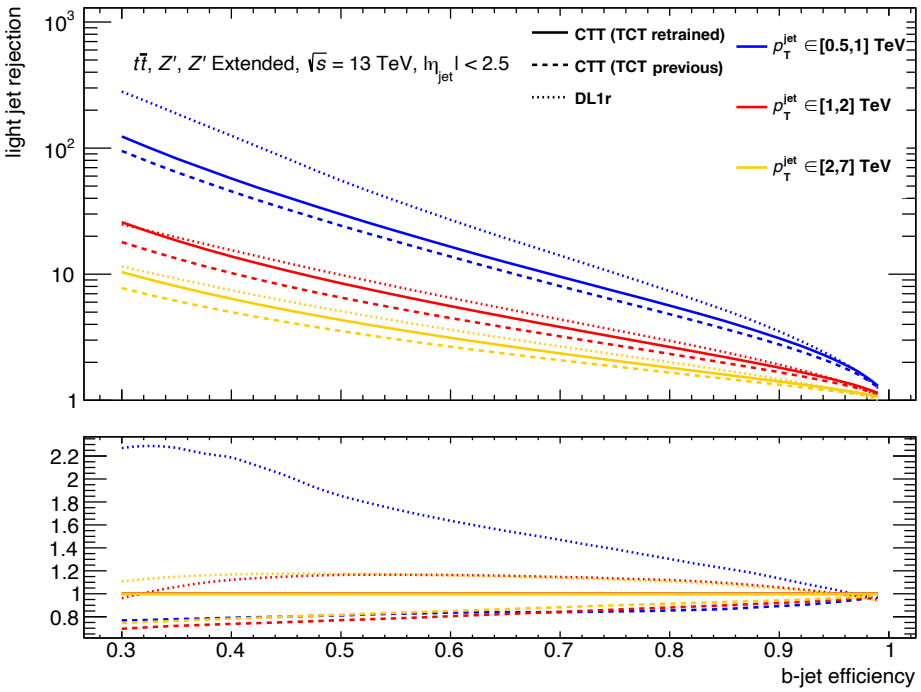


Figure C.10: Comparison of the b-jet-efficiency vs. light jet rejection using either the CTT (with underlying previous (dashed) or retrained (solid) TCT model) or the DL1r (dotted) b-tagging algorithm for three different $p_{\mathrm{T}}^{\mathrm{jet}}$ ranges.

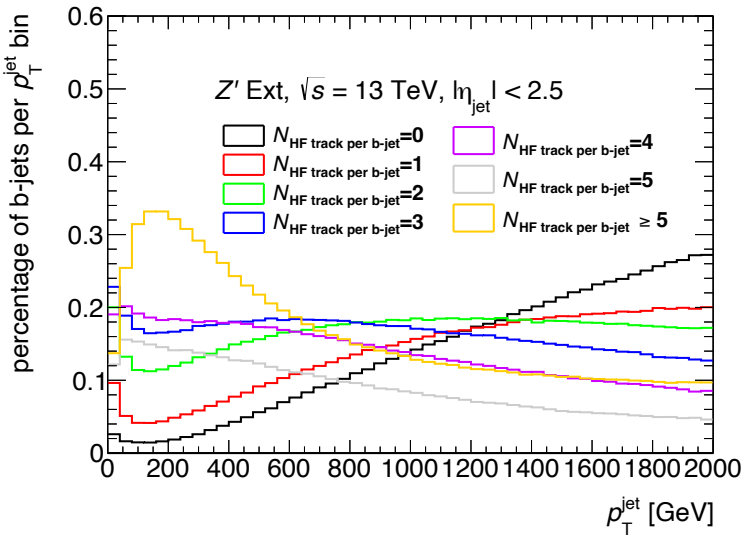
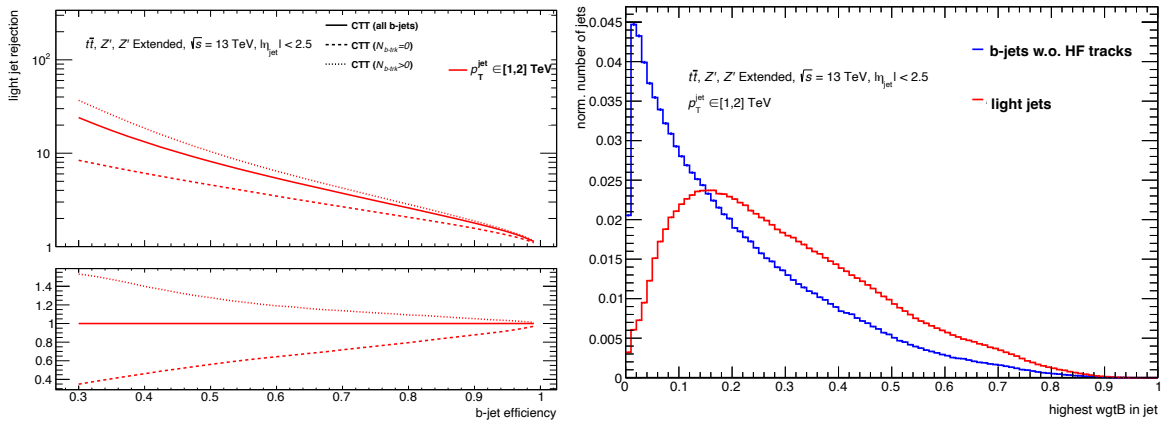


Figure C.11: Percentage of b-jets with different number of reconstructed tracks from heavy-flavour decay as a function of p_{T} of the b-jet.

performance on these jets.[9]



(a) (b)

Figure C.12: (a) Comparison of b-jet-efficiency vs. light jet rejection using the CTT b-tagging algorithm, which is trained on all b-jets, but evaluated either on all b-jets (solid) or only on b-jets without (dashed) or with (dotted) reconstructed tracks from B-hadron decay products in a $p_{\mathrm{T}}^{\mathrm{jet}} \in [1, 2]$ TeV range. (b) Distribution of the highest `wgtB` TCT score of all reconstructed tracks in light (red) and b-jets without any reconstructed B-hadron decay track (blue) within the range $p_{\mathrm{T}}^{\mathrm{jet}} \in [1, 2]$ TeV.

C.5 Conclusion

A new b-tagging algorithm, the Classified Track Tagger, was developed with particular focus on the classification of jets at high p_{T}. For this purpose the Track Classification Tool was retrained, which is used to select interesting tracks for the jet classification and additionally provides the sensitive input distributions to the CTT BDT. The improved TCT performance due to increased training statistics and the BDT training in five separate $p_{\mathrm{T}}^{\mathrm{jet}}$ bins was demonstrated in Section C.2. The improvement in the track classification was shown to be directly beneficial for the CTT b-tagging algorithm in Section C.4. The number of input tracks used in the CTT was fixed to three tracks, as especially for the high $p_{\mathrm{T}}^{\mathrm{jet}}$ range, the inclusion of more input tracks does not lead to a significant improvement in b-tagging performance. Compared to the b-tagging algorithm DL1r used in ATLAS, the CTT is not reaching the performance of the former for the low $p_{\mathrm{T}}^{\mathrm{jet}}$ range. But, in the high $p_{\mathrm{T}}^{\mathrm{jet}}$ range the performance of the two b-tagging algorithms is comparable. This is the case, even though the DL1r algorithm is much more complex. The simplicity of the CTT with respect to DL1r could be beneficial regarding systematics and in retraining the algorithm, which could be necessary e.g. due to changes in the ATLAS track reconstruction. On the other hand, the new generation of graph neural network flavour tagging algorithms in ATLAS do not depend on input of

[9] To make a more quantitive statement, the AUC values are compared: While the default CTT BDT training for jets with $p_{\mathrm{T}}^{\mathrm{jet}} \in [1, 2]$ TeV leads to an AUC value of 0.76 when tested on all b-jets and an AUC value of 0.65 when tested only on jets without HF tracks, an AUC value of 0.82 could be achieved for b-jets without HF tracks with a dedicated CTT BDT training only including b-jets without HF tracks.

lower-level taggers and can be easily retrained to suit particular physics purposes, as e.g. high p_{T} b-tagging [188, 306]. Additionally, b-jets without any reconstructed B-hadron decay tracks were investigated, and it was studied, that they can be to some extent distinguished from light jets due to different properties of their associated fragmentation tracks.

Acknowledgements

I want to thank my supervisor Markus Cristinziani for his continued support and advice during my time as PhD student in Siegen and during this thesis. I am truly grateful for all your time and effort.

Thanks to my great colleagues from the ATLAS collaboration: Andrea, Dominic, Federica and Nedaa-Alexandra for their help and advice in the modelling studies and Andrej, Marcel, Miguel and Tomas in the data analysis. I really appreciate the friendly and supportive environment that you created in these projects.

Many thanks to my lovely colleagues and friends in Siegen, which made my time in this city so valuable to me. I want to thank Diptaparna, Elisabeth, Vadim and Wolfgang for reviewing parts of this thesis, but also for their help in physics and technical questions during my time in Siegen. Especially I want to thank my dear friends Beatrice and Gabriel, of whose loving support and friendship I could always be sure of and with whom the sometimes rainy Siegen felt so quickly like a sunny and friendly home to me.

Thank you to my kind, funny and loving family, which always supported me unconditionally and will always give me strength in live, as they were also during this period as a PhD student.

I want to thank Ties, who can always make me laugh and with whom on my side any problem seems less big to me.

Bibliography

[1] T. Ježo, J. M. Lindert, P. Nason, C. Oleari and S. Pozzorini, *An NLO+PS generator for $t\bar{t}$ and Wt production and decay including non-resonant and interference effects*, Eur. Phys. J. C **76** (2016) 691, arXiv: 1607.04538 [hep-ph].

[2] T. Ježo, J. M. Lindert and S. Pozzorini, *Resonance-aware NLOPS matching for off-shell $t\bar{t}$ + tW production with semileptonic decays*, JHEP **10** (2023) 008, arXiv: 2307.15653 [hep-ph].

[3] F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, John Wiley And Sons, 1984, ISBN: 978-0-471-88741-6.

[4] S. Navas et al., *Review of particle physics*, Phys. Rev. D **110** (2024) 030001.

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

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

[7] G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585.

[8] T. Kinoshita, *Mass singularities of Feynman amplitudes*, J. Math. Phys. **3** (1962) 650.

[9] T. D. Lee and M. Nauenberg, *Degenerate Systems and Mass Singularities*, Phys. Rev. **133** (1964) B1549, ed. by G. Feinberg.

[10] A. Denner, S. Dittmaier, S. Kallweit and S. Pozzorini, *NLO QCD corrections to off-shell top-antitop production with leptonic decays at hadron colliders*, JHEP **10** (2012) 110, arXiv: 1207.5018 [hep-ph].

[11] G. Leibbrandt and J. D. Williams, *Two-Loop Quark Self-Energy in a New Formalism: (II) Renormalization of the Quark Propagator in the Light-Cone Gauge*, Nucl. Phys. B **566** (2000) 373, arXiv: hep-th/9911207.

[12] R. Tarrach, *The Pole Mass in Perturbative QCD*, Nucl. Phys. B **183** (1981) 384.

[13] P. Marquard, L. Mihaila, J. H. Piclum and M. Steinhauser, *Relation between the pole and the minimally subtracted mass in dimensional regularization and dimensional reduction to three-loop order*, Nucl. Phys. B **773** (2007) 1, arXiv: hep-ph/0702185.

[14] T. Muta, *Foundations of quantum chromodynamics. Second edition*, vol. 57, 1998.

[15] A. H. Hoang, *What is the Top Quark Mass?*, Ann. Rev. Nucl. Part. Sci. **70** (2020) 225, arXiv: 2004.12915 [hep-ph].

[16] J. Erler and M. Schott, *Electroweak Precision Tests of the Standard Model after the Discovery of the Higgs Boson*, Prog. Part. Nucl. Phys. **106** (2019) 68, arXiv: 1902.05142 [hep-ph].

[17] J. Haller et al., *Update of the global electroweak fit and constraints on two-Higgs-doublet models*, Eur. Phys. J. C **78** (2018) 675, arXiv: 1803.01853 [hep-ph].

[18] F. Bezrukov, M. Y. Kalmykov, B. A. Kniehl and M. Shaposhnikov, *Higgs Boson Mass and New Physics*, JHEP **10** (2012) 140, arXiv: 1205.2893 [hep-ph].

[19] G. Degrassi et al., *Higgs mass and vacuum stability in the Standard Model at NNLO*, JHEP **08** (2012) 098, arXiv: 1205.6497 [hep-ph].

[20] S. Alekhin, A. Djouadi and S. Moch, *The top quark and Higgs boson masses and the stability of the electroweak vacuum*, Phys. Lett. B **716** (2012) 214, arXiv: 1207.0980 [hep-ph].

[21] A. V. Bednyakov, B. A. Kniehl, A. F. Pikelner and O. L. Veretin, *Stability of the Electroweak Vacuum: Gauge Independence and Advanced Precision*, Phys. Rev. Lett. **115** (2015) 201802, arXiv: 1507.08833 [hep-ph].

[22] Y. Tang, *Vacuum Stability in the Standard Model*, Mod. Phys. Lett. A **28** (2013) 1330002, arXiv: 1301.5812 [hep-ph].

[23] 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, arXiv: 1905.02302 [hep-ex].

[24] CMS Collaboration, *Measurement of the top quark pole mass using $t\bar{t}$+jet events in the dilepton final state in proton–proton collisions at $\sqrt{s} = 13$ TeV*, (2022), arXiv: 2207.02270 [hep-ex].

[25] ATLAS and CMS Collaborations, *Highlights of the HL-LHC physics projections by ATLAS and CMS*, ATL-PHYS-PUB-2025-018, 2025, URL: https://cds.cern.ch/record/2928907.

[26] M. Jezabek and J. H. Kühn, *QCD Corrections to Semileptonic Decays of Heavy Quarks*, Nucl. Phys. B **314** (1989) 1.

[27] G. Burdman, *Quantum Field Theory I*, Lecture notes, University of Sao Paulo's Physics Institute, 2024, URL: http://fma.if.usp.br/~burdman/QFT1/lecture_26.pdf.

[28] M. E. Peskin and D. V. Schroeder, *An Introduction to quantum field theory*, Reading, USA: Addison-Wesley, 1995.

[29] K. G. Chetyrkin, R. Harlander, T. Seidensticker and M. Steinhauser, *Second order QCD corrections to* $\Gamma(t \to Wb)$, Phys. Rev. D **60** (1999) 114015, arXiv: hep-ph/9906273.

[30] A. Czarnecki and K. Melnikov, *Two loop QCD corrections to top quark width*, Nucl. Phys. B **544** (1999) 520, arXiv: hep-ph/9806244.

[31] L.-B. Chen, H. T. Li, J. Wang and Y. Wang, *Analytic result for the top-quark width at next-to-next-to-leading order in QCD*, Phys. Rev. D **108** (2023) 054003, arXiv: 2212.06341 [hep-ph].

[32] J. H. Kühn, "Theory of top quark production and decay", *23rd Annual SLAC Summer Institute on Particle Physics: The Top Quark and the Electroweak Interaction (SSI 95)*, 1996 1, arXiv: hep-ph/9707321.

[33] T. Dado, *Direct measurement of the top-quark decay width with the ATLAS detector*, PhD thesis: Comenius University Bratislava and Georg-Augustus-Universität Göttingen, 2019, URL: https://repository.cern/records/efgwf-qhm73.

[34] A. Dabelstein, W. Hollik, C. Junger, R. A. Jimenez and J. Sola, *Strong supersymmetric quantum effects on the top quark width*, Nucl. Phys. B **454** (1995) 75, arXiv: hep-ph/9503398.

[35] J. I. Guasch, *Supersymmetric radiative corrections to top quark and Higgs boson physics*, PhD thesis: Universitat Autonoma de Barcelona, 1999, arXiv: hep-ph/9906517.

[36] ATLAS Collaboration, *Search for a scalar partner of the top quark in the jets plus missing transverse momentum final state at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, JHEP **12** (2017) 085, arXiv: 1709.04183 [hep-ex].

[37] ATLAS Collaboration, *Search for top-squark pair production in final states with one lepton, jets, and missing transverse momentum using* $36\,fb^{-1}$ *of* $\sqrt{s} = 13$ *TeV pp collision data with the ATLAS detector*, JHEP **06** (2018) 108, arXiv: 1711.11520 [hep-ex].

[38] ATLAS Collaboration, *Search for dark matter and other new phenomena in events with an energetic jet and large missing transverse momentum using the ATLAS detector*, JHEP **01** (2018) 126, arXiv: 1711.03301 [hep-ex].

[39] ATLAS Collaboration, *Search for supersymmetry in final states with charm jets and missing transverse momentum in* 13 *TeV pp collisions with the ATLAS detector*, JHEP **09** (2018) 050, arXiv: 1805.01649 [hep-ex].

[40] F. Canelli, A. Knue and F. Maltoni, *Top Quark review in Review of particle physics from S. Navas et al.*, Phys. Rev. D **110** (2024) 030001, and 2025 update.

[41] F. Larios, R. Martinez and M. A. Perez, *New physics effects in the flavor-changing neutral couplings of the top quark*, Int. J. Mod. Phys. A **21** (2006) 3473, arXiv: hep-ph/0605003.

[42] S. Höche, *Introduction to parton-shower event generators*, (2014), Lecture notes of Theoretical Advanced Study Institute in Elementary Particle Physics (TASI), arXiv: 1411.4085 [hep-ph].

[43] G. Heinrich, *Colourful Loops: Introduction to Quantum Chromodynamics and Loop Calculations*, Lecture notes, Max Planck Institute for Physics, Munich, 2018, URL: https://wwwth.mpp.mpg.de/members/gudrun/members_gudrun__files/teaching/SS2018/QCD.pdf.

[44] C. Bierlich et al., *A comprehensive guide to the physics and usage of PYTHIA 8.3*, SciPost Phys. Codeb. **2022** (2022) 8, arXiv: 2203.11601 [hep-ph].

[45] J. Campbell, J. Huston and F. Krauss, *The Black Book of Quantum Chromodynamics : a Primer for the LHC Era*, Oxford University Press, 2018, ISBN: 978-0-19-965274-7.

[46] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004) 040, arXiv: hep-ph/0409146.

[47] S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations with parton shower simulations: the POWHEG method*, JHEP **11** (2007) 070, arXiv: 0709.2092 [hep-ph].

[48] 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, arXiv: 1303.6254 [hep-ph].

[49] M. Czakon, D. Heymes and A. Mitov, *High-precision differential predictions for top-quark pairs at the LHC*, Phys. Rev. Lett. **116** (2016) 082003, arXiv: 1511.00549 [hep-ph].

[50] S. Catani et al., *Top-quark pair hadroproduction at next-to-next-to-leading order in QCD*, Phys. Rev. D **99** (2019) 051501, arXiv: 1901.04005 [hep-ph].

[51] 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, arXiv: 1906.06535 [hep-ph].

[52] W. Bernreuther, M. Fucker and Z.-G. Si, *Weak interaction corrections to hadronic top quark pair production: Contributions from quark-gluon and b anti-b induced reactions*, Phys. Rev. D **78** (2008) 017503, arXiv: 0804.1237 [hep-ph].

[53] J. H. Kühn, A. Scharf and P. Uwer, *Electroweak effects in top-quark pair production at hadron colliders*, Eur. Phys. J. C **51** (2007) 37, arXiv: hep-ph/0610335.

[54] W. Hollik and D. Pagani, *The electroweak contribution to the top quark forward-backward asymmetry at the Tevatron*, Phys. Rev. D **84** (2011) 093003, arXiv: 1107.2606 [hep-ph].

[55] C. Gütschow, J. M. Lindert and M. Schönherr,
Multi-jet merged top-pair production including electroweak corrections,
Eur. Phys. J. C **78** (2018) 317, arXiv: 1803.00950 [hep-ph].

[56] R. Frederix, I. Tsinikos and T. Vitos,
Probing the spin correlations of $t\bar{t}$ production at NLO QCD+EW,
Eur. Phys. J. C **81** (2021) 817, arXiv: 2105.11478 [hep-ph].

[57] M. Czakon et al., *Top-pair production at the LHC through NNLO QCD and NLO EW*,
JHEP **10** (2017) 186, arXiv: 1705.04105 [hep-ph].

[58] S. Frixione, P. Nason and B. R. Webber,
Matching NLO QCD and parton showers in heavy flavour production,
JHEP **08** (2003) 007, arXiv: hep-ph/0305252.

[59] S. Frixione, G. Ridolfi and P. Nason,
A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction,
JHEP **09** (2007) 126, arXiv: 0707.3088 [hep-ph].

[60] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*,
JHEP **07** (2014) 079, arXiv: 1405.0301 [hep-ph].

[61] S. Hoeche et al., *Next-to-leading order QCD predictions for top-quark pair production with up to two jets merged with a parton shower*, Phys. Lett. B **748** (2015) 74,
arXiv: 1402.6293 [hep-ph].

[62] K. Cormier, S. Plätzer, C. Reuschle, P. Richardson and S. Webster,
Parton showers and matching uncertainties in top quark pair production with Herwig 7,
Eur. Phys. J. C **79** (2019) 915, arXiv: 1810.06493 [hep-ph].

[63] J. Mazzitelli et al.,
Next-to-Next-to-Leading Order Event Generation for Top-Quark Pair Production,
Phys. Rev. Lett. **127** (2021) 062001, arXiv: 2012.14267 [hep-ph].

[64] J. Mazzitelli et al., *Top-pair production at the LHC with MINNLO$_{PS}$*,
JHEP **04** (2022) 079, arXiv: 2112.12135 [hep-ph].

[65] P. F. Monni, P. Nason, E. Re, M. Wiesemann and G. Zanderighi,
MiNNLO$_{PS}$: a new method to match NNLO QCD to parton showers,
JHEP **05** (2020) 143, [Erratum: JHEP 02, 031 (2022)], arXiv: 1908.06987 [hep-ph].

[66] P. F. Monni, E. Re and M. Wiesemann,
MiNNLO$_{PS}$: optimizing $2 \to 1$ hadronic processes, Eur. Phys. J. C **80** (2020) 1075,
arXiv: 2006.04133 [hep-ph].

[67] S. Frixione, E. Laenen, P. Motylinski and B. R. Webber, *Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations*,
JHEP **04** (2007) 081, arXiv: hep-ph/0702198.

[68] P. Artoisenet, R. Frederix, O. Mattelaer and R. Rietkerk,
Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations,
JHEP **03** (2013) 015, arXiv: 1212.3460 [hep-ph].

[69] J. M. Campbell, R. K. Ellis, P. Nason and E. Re,
Top-Pair Production and Decay at NLO Matched with Parton Showers,
JHEP **04** (2015) 114, arXiv: 1412.1828 [hep-ph].

[70] J. Gao, C. S. Li and H. X. Zhu,
Top Quark Decay at Next-to-Next-to Leading Order in QCD,
Phys. Rev. Lett. **110** (2013) 042001, arXiv: 1210.2808 [hep-ph].

[71] M. Brucherseifer, F. Caola and K. Melnikov,
$\mathcal{O}(\alpha_s^2)$ *corrections to fully-differential top quark decays*, JHEP **04** (2013) 059,
arXiv: 1301.7133 [hep-ph].

[72] J. Gao and A. S. Papanastasiou,
Top-quark pair-production and decay at high precision, Phys. Rev. D **96** (2017) 051501,
arXiv: 1705.08903 [hep-ph].

[73] A. Behring, M. Czakon, A. Mitov, A. S. Papanastasiou and R. Poncelet,
Higher order corrections to spin correlations in top quark pair production at the LHC,
Phys. Rev. Lett. **123** (2019) 082001, arXiv: 1901.05407 [hep-ph].

[74] M. Czakon, A. Mitov and R. Poncelet,
NNLO QCD corrections to leptonic observables in top-quark pair production and decay,
JHEP **05** (2021) 212, arXiv: 2008.11133 [hep-ph].

[75] E. Re, *Single-top Wt-channel production matched with parton showers using the POWHEG method*, Eur. Phys. J. C **71** (2011) 1547, arXiv: 1009.2450 [hep-ph].

[76] S. Frixione, E. Laenen, P. Motylinski, C. White and B. R. Webber,
Single-top hadroproduction in association with a W boson, JHEP **07** (2008) 029,
arXiv: 0805.3067 [hep-ph].

[77] E. Bothmann, F. Krauss and M. Schönherr,
Single top-quark production with SHERPA, Eur. Phys. J. C **78** (2018) 220,
arXiv: 1711.02568 [hep-ph].

[78] ATLAS Collaboration,
Probing the Quantum Interference between Singly and Doubly Resonant Top-Quark Production in pp Collisions at $\sqrt{s} = 13$ TeV with the ATLAS Detector,
Phys. Rev. Lett. **121** (2018) 152002, arXiv: 1806.04667 [hep-ex].

[79] F. Cascioli, S. Kallweit, P. Maierhöfer and S. Pozzorini,
A unified NLO description of top-pair and associated Wt production,
Eur. Phys. J. C **74** (2014) 2783, arXiv: 1312.0546 [hep-ph].

[80] R. Frederix, *Top Quark Induced Backgrounds to Higgs Production in the $WW^{(*)} \to ll\nu\nu$ Decay Channel at Next-to-Leading-Order in QCD*,
Phys. Rev. Lett. **112** (2014) 082002, arXiv: 1311.4893 [hep-ph].

[81] G. Bevilacqua, M. Czakon, A. van Hameren, C. G. Papadopoulos and M. Worek,
Complete off-shell effects in top quark pair hadroproduction with leptonic decay at next-to-leading order, JHEP **02** (2011) 083, arXiv: 1012.4230 [hep-ph].

[82] A. Denner, S. Dittmaier, S. Kallweit and S. Pozzorini,
NLO QCD corrections to WWbb production at hadron colliders,
Phys. Rev. Lett. **106** (2011) 052001, arXiv: 1012.3975 [hep-ph].

[83] G. Heinrich, A. Maier, R. Nisius, J. Schlenk and J. Winter,
NLO QCD corrections to $W^+W^-b\bar{b}$ production with leptonic decays in the light of top quark mass and asymmetry measurements, JHEP **06** (2014) 158,
arXiv: 1312.6659 [hep-ph].

[84] A. Denner and M. Pellen,
Off-shell production of top-antitop pairs in the lepton+jets channel at NLO QCD,
JHEP **02** (2018) 013, arXiv: 1711.10359 [hep-ph].

[85] 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, arXiv: 1002.2581 [hep-ph].

[86] E. Boos et al., "Generic user process interface for event generators",
Proceedings, Physics at TeV colliders, Euro Summer School
(Les Houches, France, 21st May–1st June 2001), arXiv: hep-ph/0109068 [hep-ph],
URL: http://lss.fnal.gov/archive/preprint/fermilab-conf-01-496-t.shtml.

[87] J. Alwall et al., *A Standard Format for Les Houches Event Files*,
Comput. Phys. Commun. **176** (2007) 300, arXiv: hep-ph/0609017 [hep-ph].

[88] A. Buckley, F. Krauss, S. Navas and P. Richardson, *Monte Carlo Numbering Scheme*,
https://pdg.lbl.gov/2024/mcdata/mc_particle_id_contents.html,
Accessed: May, 2025.

[89] T. Ježo and P. Nason, *On the Treatment of Resonances in Next-to-Leading Order Calculations Matched to a Parton Shower*, JHEP **12** (2015) 065,
arXiv: 1509.09071 [hep-ph].

[90] 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, 2023, URL: https://cds.cern.ch/record/2872787.

[91] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section in the lepton+jets channel at $\sqrt{s} = 13$ TeV with the ATLAS experiment*,
Phys. Lett. B **810** (2020) 135797, arXiv: 2006.13076 [hep-ex].

[92] T. Ježo, J. M. Lindert and S. Pozzorini, *powheg-box-res-bb4l-sl-beta source code*,
https://gitlab.cern.ch/tjezo/powheg-box-res-bb4l-sl-beta,
Accessed: October 2024.

[93] The NNPDF Collaboration, *Parton distributions for the LHC run II*,
JHEP **04** (2015) 040, arXiv: 1410.8849 [hep-ph].

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

[95] ATLAS Collaboration, *ATLAS Pythia 8 tunes to 7 TeV data*,
ATL-PHYS-PUB-2014-021, 2014, URL: https://cds.cern.ch/record/1966419.

[96] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152.

[97] J. Lindert, *Private communication*, 2025.

[98] C. Bierlich et al., *Robust Independent Validation of Experiment and Theory: Rivet version 3*, SciPost Phys. **8** (2020) 026, arXiv: 1912.05451 [hep-ph].

[99] M. Cacciari, G. P. Salam and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063, arXiv: 0802.1189 [hep-ph].

[100] M. Cacciari, G. P. Salam and G. Soyez, *FastJet user manual*, Eur. Phys. J. C **72** (2012) 1896, arXiv: 1111.6097 [hep-ph].

[101] M. Cacciari, G. P. Salam and G. Soyez, *The Catchment Area of Jets*, JHEP **04** (2008) 005, arXiv: 0802.1188 [hep-ph].

[102] ATLAS Collaboration, *Measurements of differential cross-sections of $WbWb$ production in the dilepton channel in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, JHEP **02** (2026) 153, arXiv: 2506.14700 [hep-ex].

[103] ATLAS Collaboration, *Measurement of the top-quark Yukawa coupling from $t\bar{t}$ production in the lepton+jets final state using pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **01** (2026) 117, arXiv: 2509.16039 [hep-ex].

[104] ATLAS Collaboration, *Observation of a cross-section enhancement near the $t\bar{t}$ production threshold in $\sqrt{s}$ =13 TeV pp collisions with the ATLAS detector*, Reports on Progress in Physics **89** (2026) 057801, arXiv: 2601.11780 [hep-ex].

[105] CMS Collaboration, *Observation of quantum entanglement in top quark pair production in proton–proton collisions at $\sqrt{s} = 13$ TeV*, Rept. Prog. Phys. **87** (2024) 117801, arXiv: 2406.03976 [hep-ex].

[106] CMS Collaboration, *Observation of a pseudoscalar excess at the top quark pair production threshold*, Rept. Prog. Phys. **88** (2025) 087801, arXiv: 2503.22382 [hep-ex].

[107] CMS Collaboration, *Measurements of $t\bar{t}$ differential cross sections in proton–proton collisions at $\sqrt{s} = 13$ TeV using events containing two leptons*, JHEP **02** (2019) 149, arXiv: 1811.06625 [hep-ex].

[108] CMS Collaboration, *Rivet CMS_2018_I1703993*, https://rivet.hepforge.org/analyses/CMS_2018_I1703993.html, Accessed: October 2024.

[109] S. Ferrario Ravasio, T. Ježo, P. Nason and C. Oleari, *A theoretical study of top-mass measurements at the LHC using NLO+PS generators of increasing accuracy*, Eur. Phys. J. C **78** (2018) 458, [Addendum: Eur.Phys.J.C 79, 859 (2019)], arXiv: 1906.09166 [hep-ph].

[110] M. R. Masouminia and P. Richardson,
Hadronization and decay of excited heavy hadrons in Herwig 7, JHEP **07** (2024) 278, arXiv: 2312.02757 [hep-ph].

[111] ATLAS Collaboration, *Studies of $t\bar{t}/tW$ interference effects in $b\bar{b}\ell^{+}\ell^{-'}\nu\bar{\nu}'$ final states with Powheg and MadGraph5_aMC@NLO setups*, ATL-PHYS-PUB-2021-042, 2021, URL: https://cds.cern.ch/record/2792254.

[112] ATLAS Collaboration, *Measurement of the top-quark pole mass in dileptonic $t\bar{t}+1$-jet events at $\sqrt{s}=13$ TeV with the ATLAS experiment*, JHEP **12** (2025) 023, arXiv: 2507.02632 [hep-ex].

[113] PDF4LHC Working Group,
The PDF4LHC21 combination of global PDF fits for the LHC Run III,
J. Phys. G **49** (2022) 080501, arXiv: 2203.05506 [hep-ph].

[114] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*,
J. Phys. G **43** (2016) 023001, arXiv: 1510.03865 [hep-ph].

[115] 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, arXiv: 1412.3989 [hep-ph].

[116] S. Dulat et al.,
New parton distribution functions from a global analysis of quantum chromodynamics,
Phys. Rev. D **93** (2016) 033006, arXiv: 1506.07443 [hep-ph].

[117] S. Alioli, P. Nason, C. Oleari and E. Re,
NLO Higgs boson production via gluon fusion matched with shower in POWHEG,
JHEP **04** (2009) 002, arXiv: 0812.0578 [hep-ph].

[118] Rivet, *MC_HFDECAYS*,
https://rivet.hepforge.org/analyses/MC_HFDECAYS.html,
Accessed: September 2024.

[119] ATLAS Collaboration,
Towards a Systematic Uncertainty Model for the Powheg bb4ℓ Monte Carlo sample,
ATL-PHYS-PUB-2025-036, 2025, URL: https://cds.cern.ch/record/2942776.

[120] ATLAS Collaboration,
Observation of quantum entanglement with top quarks at the ATLAS detector,
Nature **633** (2024) 542, arXiv: 2311.07288 [hep-ex].

[121] 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, arXiv: 2209.00583 [hep-ex].

[122] ATLAS Collaboration, *Measurements of jet observables sensitive to b-quark fragmentation in $t\bar{t}$ events at the LHC with the ATLAS detector*,
Phys. Rev. D **106** (2022) 032008, arXiv: 2202.13901 [hep-ex].

[123] ATLAS Collaboration, *Measurement of the top quark mass in the $t\bar{t}\to$ lepton+jets channel from $\sqrt{s}=8$ TeV ATLAS data and combination with previous results*,
Eur. Phys. J. C **79** (2019) 290, arXiv: 1810.01772 [hep-ex].

[124] ATLAS Collaboration, *Measurement of the top-quark mass in $t\bar{t} \to$ dilepton events with the ATLAS experiment using the template method in 13 TeV pp collision data*, ATLAS-CONF-2022-058, 2022, URL: https://cds.cern.ch/record/2826701.

[125] ATLAS Collaboration, *Measurement of the top quark mass with the ATLAS detector using $t\bar{t}$ events with a high transverse momentum top quark*, Phys. Lett. B **867** (2025) 139608, arXiv: 2502.18216 [hep-ex].

[126] H. Brooks, C. T. Preuss and P. Skands, *Sector Showers for Hadron Collisions*, JHEP **07** (2020) 032, arXiv: 2003.00702 [hep-ph].

[127] S. Mrenna and P. Skands, *Automated parton-shower variations in PYTHIA 8*, Phys. Rev. D **94** (2016) 074005, arXiv: 1605.08352 [hep-ph].

[128] ATLAS Collaboration, *Measurement of jet activity produced in top-quark events with an electron, a muon and two b-tagged jets in the final state in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **77** (2017) 220, arXiv: 1610.09978 [hep-ex].

[129] ATLAS Collaboration, *Rivet ATLAS_2017_I1495243*, https://rivet.hepforge.org/analyses/ATLAS_2017_I1495243.html, Accessed: October 2024.

[130] ATLAS Collaboration, *Measurements of inclusive and differential fiducial cross-sections of $t\bar{t}$ production with additional heavy-flavour jets in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **04** (2019) 046, arXiv: 1811.12113 [hep-ex].

[131] ATLAS Collaboration, *Rivet ATLAS_2018_I1705857*, https://rivet.hepforge.org/analyses/ATLAS_2018_I1705857, Accessed: September 2024.

[132] ATLAS Collaboration, *Measurement of $t\bar{t}$ production in association with additional b-jets in the $e\mu$ final state in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **01** (2025) 068, arXiv: 2407.13473 [hep-ex].

[133] ATLAS Collaboration, *Rivet ATLAS_2024_I2809112*, https://rivet.hepforge.org/analyses/ATLAS_2024_I2809112.html, Accessed: October 2025.

[134] J. R. Christiansen and P. Z. Skands, *String Formation Beyond Leading Colour*, JHEP **08** (2015) 003, arXiv: 1505.01681 [hep-ph].

[135] UA1 Collaboration, *A study of the general characteristics of proton-antiproton collisions at $\sqrt{s}$ =0.2 to 0.9 TeV*, Nucl. Phys. B **335** (1990) 261.

[136] CDF Collaboration, *Soft and hard interactions in $p\bar{p}$ collisions at $\sqrt{s}$ =1800 and 630 GeV*, Phys. Rev. D **65** (7 2002) 072005, URL: https://link.aps.org/doi/10.1103/PhysRevD.65.072005.

[137] CDF Collaboration, *Measurement of Particle Production and Inclusive Differential Cross Sections in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Rev. D **79** (2009) 112005, [Erratum: Phys.Rev.D 82, 119903 (2010)], arXiv: 0904.1098 [hep-ex].

[138] CMS Collaboration, *Charged Particle Multiplicities in pp Interactions at $\sqrt{s}$ =0.9, 2.36, and 7 TeV*, JHEP **01** (2011) 079, arXiv: 1011.5531 [hep-ex].

[139] ATLAS Collaboration, *Charged-particle multiplicities in pp interactions measured with the ATLAS detector at the LHC*, New J. Phys. **13** (2011) 053033, arXiv: 1012.5104 [hep-ex].

[140] ALICE Collaboration, *Multiplicity dependence of the average transverse momentum in pp, p-Pb, and Pb-Pb collisions at the LHC*, Phys. Lett. B **727** (2013) 371, arXiv: 1307.1094 [nucl-ex].

[141] S. Argyropoulos and T. Sjöstrand, *Effects of color reconnection on $t\bar{t}$ final states at the LHC*, JHEP **11** (2014) 043, arXiv: 1407.6653 [hep-ph].

[142] ATLAS Collaboration, *A study of different colour reconnection settings for Pythia8 generator using underlying event observables*, ATL-PHYS-PUB-2017-008, 2017, URL: https://cds.cern.ch/record/2262253.

[143] 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, arXiv: 2209.07874 [hep-ex].

[144] ATLAS Collaboration, *Rivet ATLAS_2022_I2152933*, https://rivet.hepforge.org/analyses/ATLAS_2022_I2152933.html, Accessed: September 2024.

[145] *LHC The Guide FAQ*, URL: https://home.cern/resources/brochure/knowledge-sharing/lhc-facts-and-figures (visited on 28/11/2024).

[146] *Overall view of the LHC*, URL: https://cds.cern.ch/record/1708849 (visited on 29/11/2024).

[147] C. Bracco, L. Jorat, R. Noulibos, P. van Trappen and W. Weterings, *Measurements with the stripping foil test stand in the Linac4 Transfer Line*, EPJ Web Conf. **229** (2020) 01003, URL: https://doi.org/10.1051/epjconf/202022901003.

[148] L. Evans, *The Large Hadron Collider: a marvel of technology*, 2nd ed., Fundamental Sciences, EPFL Press, 2019, ISBN: 978-2-88914-482-2, URL: https://cds.cern.ch/record/2645935/files/Evans2018.pdf.

[149] *Public ATLAS Luminosity Results for Run-2 of the LHC*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2 (visited on 28/11/2024).

[150] *ATLAS Public Event Displays UNSG-2021-64*, URL: https://atlaspo.cern.ch/public/event_display/ (visited on 28/11/2024).

[151] ATLAS Collaboration, *The ATLAS experiment at the CERN Large Hadron Collider: a description of the detector configuration for Run 3*, JINST **19** (2024) P05063, arXiv: 2305.16623 [physics.ins-det].

[152] *Computer generated image of the ATLAS detector*, URL: https://cds.cern.ch/record/1095924 (visited on 29/11/2024).

[153] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[154] ATLAS Collaboration, *ATLAS pixel detector electronics and sensors*, JINST **3** (2008) P07007, URL: https://cds.cern.ch/record/1119279.

[155] B. Abbott et al., *Production and integration of the ATLAS Insertable B-Layer*, JINST **13** (2018) T05008, arXiv: 1803.00844 [physics.ins-det].

[156] ATLAS Collaboration, *Operation and performance of the ATLAS semiconductor tracker in LHC Run 2*, JINST **17** (2021) P01013, arXiv: 2109.02591 [physics.ins-det].

[157] ATLAS Collaboration, *Performance of the ATLAS Transition Radiation Tracker in Run 1 of the LHC: tracker properties*, JINST **12** (2017) P05002, arXiv: 1702.06473 [hep-ex].

[158] ATLAS Collaboration, *Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2*, Eur. Phys. J. C **77** (2017) 673, arXiv: 1704.07983 [hep-ex].

[159] Y. Takubo, *ATLAS IBL operational experience*, ATL-INDET-PROC-2016-012, 2017, URL: https://cds.cern.ch/record/2235541.

[160] *ATLAS Tracking Software Tutorial*, URL: https://atlassoftwaredocs.web.cern.ch/internal-links/tracking-tutorial/index.html (visited on 06/12/2024).

[161] B. Mindur, *ATLAS Transition Radiation Tracker (TRT): Straw tubes for tracking and particle identification at the Large Hadron Collider*, ATL-INDET-PROC-2016-001, 2017, URL: https://cds.cern.ch/record/2139567.

[162] C. Grupen and B. Shwartz, *Particle Detectors*, 2nd ed., Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, Cambridge University Press, 2023, ISBN: 978-1-009-40151-7.

[163] A. Bingül, *The ATLAS TRT and its Performance at LHC*, J. Physics: Conf. Ser. **347** (2012) 012025, URL: https://dx.doi.org/10.1088/1742-6596/347/1/012025.

[164] ATLAS Collaboration, *Software Performance of the ATLAS Track Reconstruction for LHC Run 3*, Comput. Softw. Big Sci. **8** (2024) 9, arXiv: 2308.09471 [hep-ex].

[165] R. L. Glückstern, *Uncertainties in track momentum and direction, due to multiple scattering and measurement errors*, Nucl. Instrum. Meth. **24** (1963) 381.

[166] ATLAS Collaboration, *Reconstruction of primary vertices at the ATLAS experiment in Run 1 proton–proton collisions at the LHC*, Eur. Phys. J. C **77** (2017) 332, arXiv: 1611.10235 [hep-ex].

[167] ATLAS Collaboration, *Development of ATLAS Primary Vertex Reconstruction for LHC Run 3*, ATL-PHYS-PUB-2019-015, 2019, URL: https://cds.cern.ch/record/2670380.

[168] ATLAS Collaboration, *ATLAS Liquid Argon Calorimeter: Technical Design Report*, ATLAS-TDR-2; CERN-LHCC-96-041, 1996, URL: https://cds.cern.ch/record/331061.

[169] ATLAS Collaboration, *ATLAS Calorimeter Performance: Technical Design Report*, ATLAS-TDR-1; CERN-LHCC-96-040, 1996, URL: https://cds.cern.ch/record/331059.

[170] J. E. Brau, J. A. Jaros and H. Ma, *Advances in Calorimetry*, Ann. Rev. Nucl. Part. Sci. **60** (2010) 615.

[171] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using LHC Run 1 data*, Eur. Phys. J. C **74** (2014) 3071, arXiv: 1407.5063 [hep-ex].

[172] C. W. Fabjan and F. Gianotti, *Calorimetry for Particle Physics*, Rev. Mod. Phys. **75** (2003) 1243, URL: https://cds.cern.ch/record/692252.

[173] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **96** (2017) 072002, arXiv: 1703.09665 [hep-ex].

[174] M. Groh et al., *Dependence of the space-to-drift-time relationship of monitored drift-tube chambers on the magnetic field in the ATLAS muon spectrometer*, IEEE Nuclear Science Symposium Conference Record, 2005 **2** (2005) 796, URL: https://api.semanticscholar.org/CorpusID:12386786.

[175] ATLAS Collaboration, *ATLAS Muon Spectrometer: Technical Design Report*, ATLAS-TDR-10; CERN-LHCC-97-022, CERN, 1997, URL: https://cds.cern.ch/record/331068.

[176] ATLAS Collaboration, *Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics*, (2009), arXiv: 0901.0512 [hep-ex].

[177] ATLAS Collaboration, *Muons in the Calorimeters: Energy Loss Corrections and Muon Tagging*, ATL-PHYS-PUB-2009-009, 2009, URL: https://cds.cern.ch/record/1169055.

[178] *ATLAS Approved DAQ Plots*, URL: https://twiki.cern.ch/twiki/bin/view/AtlasPublic/ApprovedPlotsDAQ (visited on 07/01/2025).

[179] 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, arXiv: 1606.02625 [hep-ex].

[180] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, Eur. Phys. J. C **77** (2017) 317, arXiv: 1611.09661 [hep-ex].

[181] E. Govorkova et al., *LHC physics dataset for unsupervised New Physics detection at 40 MHz*, Sci. Data **9** (2022) 118, arXiv: 2107.02157 [physics.data-an].

[182] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77** (2017) 466, arXiv: 1703.10485 [hep-ex].

[183] ATLAS Collaboration, *Transformer networks for constituent-based b-jet calibration with the ATLAS detector*, ATL-PHYS-PUB-2024-015, 2024, URL: https://cds.cern.ch/record/2905688.

[184] ATLAS Collaboration, *Tracking studies for b-tagging with* 900 *GeV collision data with the ATLAS detector*, ATLAS-CONF-2010-003, 2010, URL: https://cds.cern.ch/record/1273190.

[185] HLAV Collaboration, *Averages of b-hadron, c-hadron, and τ-lepton properties as of 2023*, Phys. Rev. D **113** (2026) 012008, arXiv: 2411.18639 [hep-ex].

[186] BaBar Collaboration, *Measurement of the B^0 and B^+ meson lifetimes with fully reconstructed hadronic final states*, Phys. Rev. Lett. **87** (2001) 201803, arXiv: hep-ex/0107019.

[187] ATLAS Collaboration, *Precision measurement of the B^0 meson lifetime using $B^0 \to J/\psi K^{*0}$ decays with the ATLAS detector*, Eur. Phys. J. C **85** (2025) 736, arXiv: 2411.09962 [hep-ex].

[188] ATLAS Collaboration, *Transforming jet flavour tagging at ATLAS*, Nature Commun. **17** (2026) 541, arXiv: 2505.19689 [hep-ex].

[189] *ATLAS Jet Flavour Tagging With GN1 and DL1d*, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/FTAG-2023-01/ (visited on 07/01/2025).

[190] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, JINST **14** (2019) P12006, arXiv: 1908.00005 [hep-ex].

[191] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79** (2019) 639, arXiv: 1902.04655 [hep-ex].

[192] 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, arXiv: 2012.00578 [hep-ex].

[193] 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, arXiv: 1802.08168 [hep-ex].

[194] ATLAS Collaboration, *Top cross section summary plots - May 2025*, ATL-PHYS-PUB-2025-024, 2025, URL: https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2025-024/.

[195] CMS Collaboration, *Review of top quark mass measurements in CMS*, Phys. Rept. **1115** (2025) 116, arXiv: 2403.01313 [hep-ex].

[196] P. Azzi et al., *Report from Working Group 1: Standard Model Physics at the HL-LHC and HE-LHC*, CERN Yellow Rep. Monogr. **7** (2019) 1, arXiv: 1902.04070 [hep-ph].

[197] CMS Collaboration, *Measurement of $t\bar{t}$ normalised multi-differential cross sections in pp collisions at $\sqrt{s} = 13$ TeV, and simultaneous determination of the strong coupling strength, top quark pole mass, and parton distribution functions*, Eur. Phys. J. C **80** (2020) 658, arXiv: 1904.05237 [hep-ex].

[198] 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, arXiv: 2302.01967 [hep-ex].

[199] A. Pathak et al., *Precision Top Mass Measurement Using Energy Correlators*, Proceedings of the European Physical Society Conference on High Energy Physics (EPS-HEP2023) (2024) 312, arXiv: 2311.14389 [hep-ph].

[200] J. Holguin et al., *Top quark mass extractions from energy correlators: a feasibility study*, JHEP **04** (2025) 072, arXiv: 2407.12900 [hep-ph].

[201] 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, arXiv: 1807.06617 [hep-ph].

[202] LHCTopWG, *LHCTopWG Summary Plots*, 2024, URL: https://twiki.cern.ch/twiki/bin/view/LHCPhysics/LHCTopWGSummaryPlots.

[203] D0 Collaboration, *Measurement of the Top Quark Mass Using the Matrix Element Technique in Dilepton Final States*, Phys. Rev. D **94** (2016) 032004, arXiv: 1606.02814 [hep-ex].

[204] D0 Collaboration, *Precision measurement of the top-quark mass in lepton+jets final states*, Phys. Rev. D **91** (2015) 112003, arXiv: 1501.07912 [hep-ex].

[205] CDF Collaboration, *Measurement of the top-quark mass in the $t\bar{t}$ dilepton channel using the full CDF Run II data set*, Phys. Rev. D **92** (2015) 032003, arXiv: 1505.00500 [hep-ex].

[206] CDF Collaboration, *Top-Quark Mass Measurement in Events with Jets and Missing Transverse Energy Using the Full CDF Data Set*, Phys. Rev. D **88** (2013) 011101, arXiv: 1305.3339 [hep-ex].

[207] CDF and D0 Collaborations, *Combination of CDF and D0 results on the mass of the top quark using up 9.7 fb^{-1} at the Tevatron*, (2016), arXiv: 1608.01881 [hep-ex].

[208] T. Martini, T. Nuraliyev and P. Uwer,
Determination of the top-quark mass from top-quark pair events with the matrix element method at next-to-leading order: Potential and prospects,
Phys. Rev. D **107** (2023) 076013, arXiv: 2301.03280 [hep-ph].

[209] F. Fiedler, A. Grohsjean, P. Haefner and P. Schieferdecker, *The Matrix Element Method and its Application in Measurements of the Top Quark Mass*,
Nucl. Instrum. Meth. A **624** (2010) 203, arXiv: 1003.1316 [hep-ex].

[210] D0 Collaboration, *Direct measurement of the top quark mass at D0*,
Phys. Rev. D **58** (1998) 052001, arXiv: hep-ex/9801025.

[211] J. D'Hondt et al.,
Fitting of Event Topologies with External Kinematic Constraints in CMS, tech. rep.,
CERN, 2006, URL: https://cds.cern.ch/record/926540.

[212] ATLAS Collaboration, *Measurement of the top quark mass in the $t\bar{t} \to$ lepton+jets and $t\bar{t} \to$ dilepton channels using $\sqrt{s} = 7$ TeV ATLAS data*,
Eur. Phys. J. C **75** (2015) 330, arXiv: 1503.05427 [hep-ex].

[213] ATLAS Collaboration, *Measurement of the top quark mass in the $t\bar{t} \to$ dilepton channel from $\sqrt{s} = 8$ TeV ATLAS data*, Phys. Lett. B **761** (2016) 350,
arXiv: 1606.02179 [hep-ex].

[214] J. Erdmann et al., *A likelihood-based reconstruction algorithm for top-quark pairs and the KLFitter framework*, Nucl. Instrum. Meth. A **748** (2014) 18,
arXiv: 1312.5595 [hep-ex].

[215] ATLAS Collaboration, *Measurement of the top-quark mass in the fully hadronic decay channel from ATLAS data at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. C **75** (2015) 158,
arXiv: 1409.0832 [hep-ex].

[216] ATLAS Collaboration, *Top-quark mass measurement in the all-hadronic $t\bar{t}$ decay channel at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **09** (2017) 118,
arXiv: 1702.07546 [hep-ex].

[217] CMS Collaboration, *Measurement of the top-quark mass in $t\bar{t}$ events with dilepton final states in pp collisions at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. C **72** (2012) 2202,
arXiv: 1209.2393 [hep-ex].

[218] CMS Collaboration,
Measurement of the top quark mass using proton–proton data at $\sqrt{s} = 7$ and 8 TeV,
Phys. Rev. D **93** (2016) 072004, arXiv: 1509.04044 [hep-ex].

[219] R. H. Dalitz and G. R. Goldstein, *Decay and polarization properties of the top quark*,
Phys. Rev. D **45** (1992).

[220] CMS Collaboration, *Measurement of masses in the $t\bar{t}$ system by kinematic endpoints in pp collisions at $\sqrt{s} = 7$ TeV*, Eur. Phys. J. C **73** (2013) 2494,
arXiv: 1304.5783 [hep-ex].

[221] CMS Collaboration,
Measurement of the $t\bar{t}$ production cross section, the top quark mass, and the strong coupling constant using dilepton events in pp collisions at $\sqrt{s} = 13$ TeV,
Eur. Phys. J. C **79** (2019) 368, arXiv: 1812.10505 [hep-ex].

[222] D0 Collaboration, *An improved determination of the width of the top quark*,
Phys. Rev. D **85** (2012) 091104, arXiv: 1201.4156 [hep-ex].

[223] CMS Collaboration,
Measurement of the ratio $\mathcal{B}(t \to Wb)/\mathcal{B}(t \to Wq)$ in pp collisions at $\sqrt{s} = 8$ TeV,
Phys. Lett. B **736** (2014) 33, arXiv: 1404.2292 [hep-ex].

[224] D0 Collaboration, *Model-independent measurement of t-channel single top quark production in $p\bar{p}$ collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Lett. B **705** (2011) 313.

[225] CMS Collaboration, *Measurement of the single-top-quark t-channel cross section in pp collisions at $\sqrt{s} = 7$ TeV*, JHEP **12** (2012) 035, arXiv: 1209.4533 [hep-ex].

[226] D0 Collaboration,
Precision measurement of the ratio $B(t \to Wb)/B(t \to Wq)$ and Extraction of V_{tb},
Phys. Rev. Lett. **107** (2011) 121802, arXiv: 1106.5436 [hep-ex].

[227] D. P. Roy, *Looking for the charged Higgs boson*, Mod. Phys. Lett. A **19** (2004) 1813, arXiv: hep-ph/0406102.

[228] CDF Collaboration, *Direct Measurement of the Total Decay Width of the Top Quark*,
Phys. Rev. Lett. **111** (2013) 202001, arXiv: 1308.4050 [hep-ex].

[229] CMS Collaboration, *Bounding the top quark width using final states with two charged leptons and two jets at $\sqrt{s} = 13$ TeV*, CMS-PAS-TOP-16-019, 2016,
URL: https://cds.cern.ch/record/2218019.

[230] ATLAS Collaboration, *Direct top-quark decay width measurement in the $t\bar{t}$ lepton+jets channel at $\sqrt{s} = 8$ TeV with the ATLAS experiment*, Eur. Phys. J. C **78** (2018) 129, arXiv: 1709.04207 [hep-ex].

[231] ATLAS Collaboration, *Measurement of the top-quark decay width in top-quark pair events in the dilepton channel at $\sqrt{s} = 13$ TeV with the ATLAS detector*,
ATLAS-CONF-2019-038, 2019, URL: https://cds.cern.ch/record/2684952.

[232] C. Herwig, T. Ježo and B. Nachman,
Extracting the Top-Quark Width from Nonresonant Production,
Phys. Rev. Lett. **122** (2019) 231803, arXiv: 1903.10519 [hep-ex].

[233] N. Kauer and D. Zeppenfeld,
Finite width effects in top quark production at hadron colliders,
Phys. Rev. D **65** (2002) 014021, arXiv: hep-ph/0107181.

[234] S. Liebler, G. Moortgat-Pick and A. S. Papanastasiou, *Probing the top-quark width through ratios of resonance contributions of $e^+e^- \to W^+W^-b\bar{b}$*, JHEP **03** (2016) 099, arXiv: 1511.02350 [hep-ph].

[235] S. Agostinelli et al., *GEANT4 — a simulation toolkit*,
Nucl. Instrum. Meth. A **506** (2003) 250.

[236] ATLAS Collaboration, *AtlFast3: The Next Generation Of Fast Simulation in ATLAS*, Comput. Softw. Big Sci. **6** (2021) 7, arXiv: 2109.02551 [hep-ex].

[237] ATLAS Collaboration, *ATLAS data quality operations and performance for 2015–2018 data-taking*, JINST **15** (2020) P04003, arXiv: 1911.04632 [physics.ins-det].

[238] 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, arXiv: 2212.09379 [hep-ex].

[239] ATLAS Collaboration, *Emulating the impact of additional proton–proton interactions in the ATLAS simulation by pre-sampling sets of inelastic Monte Carlo events*, Comput. Softw. Big Sci. **6** (2021) 3, arXiv: 2102.09495 [hep-ex].

[240] K. Werner, F.-M. Liu and T. Pierog, *Parton ladder splitting and the rapidity dependence of transverse momentum spectra in deuteron–gold collisions at the BNL Relativistic Heavy Ion Collider*, Phys. Rev. C **74** (2006) 044902, arXiv: hep-ph/0506232.

[241] T. Pierog, I. Karpenko, J. M. Katzy, E. Yatsenko and K. Werner, *EPOS LHC: Test of collective hadronization with data measured at the CERN Large Hadron Collider*, Phys. Rev. C **92** (2015) 034906, arXiv: 1306.0121 [hep-ph].

[242] 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, 2016, URL: https://cds.cern.ch/record/2206965.

[243] NNPDF Collaboration, *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244, arXiv: 1207.1303 [hep-ph].

[244] G. Bewick et al., *Herwig 7.3 release note*, Eur. Phys. J. C **84** (2024) 1053, arXiv: 2312.05175 [hep-ph].

[245] H. B. Hartanto, B. Jäger, L. Reina and D. Wackeroth, *Higgs boson production in association with top quarks in the POWHEG BOX*, Phys. Rev. D **91** (2015) 094003, arXiv: 1501.04498 [hep-ph].

[246] M. Ghezzi, B. Jäger, S. L. P. Chavez, L. Reina and D. Wackeroth, *Hadronic production of top-quark pairs in association with a pair of leptons in the powheg box framework*, Phys. Rev. D **106** (2022) 014001, arXiv: 2112.08892 [hep-ph].

[247] E. Bothmann et al., *Event generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034, arXiv: 1905.09127 [hep-ph].

[248] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, (2016), arXiv: 1610.07922 [hep-ph].

[249] M. Beneke, P. Falgari, S. Klein and C. Schwinn, *Hadronic top-quark pair production with NNLL threshold resummation*, Nucl. Phys. B **855** (2012) 695, arXiv: 1109.1536 [hep-ph].

[250] 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, arXiv: 1111.5869 [hep-ph].

[251] P. Bärnreuther, M. Czakon and A. Mitov, *Percent-Level-Precision Physics at the Tevatron: Next-to-Next-to-Leading Order QCD Corrections to $q\bar{q} \to t\bar{t} + X$*, Phys. Rev. Lett. **109** (2012) 132001, arXiv: 1204.5201 [hep-ph].

[252] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*, JHEP **12** (2012) 054, arXiv: 1207.0236 [hep-ph].

[253] M. Czakon and A. Mitov, *NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction*, JHEP **01** (2013) 080, arXiv: 1210.6832 [hep-ph].

[254] 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, arXiv: 1112.5675 [hep-ph].

[255] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-*, Phys. Rev. D **82** (2010) 054018, arXiv: 1005.4451 [hep-ph].

[256] N. Kidonakis, "Top Quark Production", *Proceedings, Helmholtz International Summer School on Physics of Heavy Quarks and Hadrons (HQ 2013)* (JINR, Dubna, Russia, 15th–28th July 2013) 139, arXiv: 1311.0283 [hep-ph].

[257] R. Frederix, E. Re and P. Torrielli, *Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO*, JHEP **09** (2012) 130, arXiv: 1207.5391 [hep-ph].

[258] 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, arXiv: 0907.4076 [hep-ph], Erratum: JHEP **02** (2010) 011.

[259] M. Aliev et al., *HATHOR – HAdronic Top and Heavy quarks crOss section calculatoR*, Comput. Phys. Commun. **182** (2011) 1034, arXiv: 1007.1327 [hep-ph].

[260] P. Kant et al., *HatHor for single top-quark production: Updated predictions and uncertainty estimates for single top-quark production in hadronic collisions*, Comput. Phys. Commun. **191** (2015) 74, arXiv: 1406.4403 [hep-ph].

[261] 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, arXiv: 1510.03823 [hep-ex].

[262] ATLAS Collaboration, *Recommendations of the Physics Objects and Analysis Harmonisation Study Groups 2014*, URL: https://cds.cern.ch/record/1743654.

[263] ATLAS Collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*, Eur. Phys. J. C **80** (2020) 47, arXiv: 1909.00761 [hep-ex].

[264] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*, JINST **15** (2020) P09015, arXiv: 2004.13447 [hep-ex].

[265] ATLAS Collaboration, *Athena*, version 22.0.1, URL: https://doi.org/10.5281/zenodo.2641997.

[266] ATLAS Collaboration, *TopCPToolkit*, URL: https://doi.org/10.5281/zenodo.19683083.

[267] ATLAS Collaboration, *FastFrames*, URL: https://doi.org/10.5281/zenodo.14773274.

[268] ATLAS Collaboration, *Rivet ATLAS_2018_I1677498*, https://rivet.hepforge.org/analyses/ATLAS_2018_I1677498.html, Accessed: October 2024.

[269] ATLAS Collaboration, *Measurement of $t\bar{t}$ production in association with additional b-jets in the $e\mu$ final state in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **01** (2025) 068, arXiv: 2407.13473 [hep-ex].

[270] ATLAS Collaboration, *Measurement of top-quark pair production in association with charm quarks in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. B **860** (2025) 139177, arXiv: 2409.11305 [hep-ex].

[271] 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** (2020) 689, arXiv: 2007.02645 [hep-ex].

[272] S. D. Ellis and D. E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. D **48** (1993) 3160, arXiv: hep-ph/9305266.

[273] ATLAS Collaboration, *Pile-up corrections for jets from proton–proton collisions at $\sqrt{s} = 7$ TeV in ATLAS in 2011*, ATLAS-CONF-2012-064, 2012, URL: https://cds.cern.ch/record/1459529.

[274] ATLAS Collaboration, *JetUncertaintiesRel22*, https://twiki.cern.ch/twiki/xbin/view/AtlasProtected/JetUncertaintiesRel22, Accessed: January, 2026.

[275] 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, arXiv: 1112.6426 [hep-ex].

[276] 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, 2022, URL: https://cds.cern.ch/record/2808016.

[277] ATLAS Collaboration, *Measurements of jet cross-section ratios in 13 TeV proton–proton collisions with ATLAS*, Phys. Rev. D **110** (2024) 072019, arXiv: 2405.20206 [hep-ex].

[278] 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, arXiv: 1607.08842 [hep-ex].

[279] ATLAS Collaboration, *Determination of jet calibration and energy resolution in proton–proton collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, Eur. Phys. J. C **80** (2020) 1104, arXiv: 1910.04482 [hep-ex].

[280] 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, arXiv: 1907.05120 [hep-ex].

[281] ATLAS Collaboration, *Measurement of the c-jet mistagging efficiency in $t\bar{t}$ events using pp collision data at $\sqrt{s} = 13$ TeV collected with the ATLAS detector*, Eur. Phys. J. C **82** (2021) 95, arXiv: 2109.10627 [hep-ex].

[282] ATLAS Collaboration, *Calibration of the light-flavour jet mistagging efficiency of the b-tagging algorithms with Z+jets events using 139 fb^{-1} of ATLAS proton–proton collision data at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **83** (2023) 728, arXiv: 2301.06319 [hep-ex].

[283] 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, arXiv: 2212.07338 [hep-ex].

[284] ATLAS Collaboration, *Inclusive and differential cross-sections for dilepton $t\bar{t}$ production measured in $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, JHEP **07** (2023) 141, arXiv: 2303.15340 [hep-ex].

[285] ATLAS Collaboration, *Measurement of $W^{\pm}Z$ production cross sections and gauge boson polarisation in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **79** (2019) 535, arXiv: 1902.05759 [hep-ex].

[286] G. Cowan, *Statistical data analysis*, 1998, ISBN: 978-0-19-850156-5.

[287] K. Cranmer, G. Lewis, L. Moneta, A. Shibata and W. Verkerke, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, tech. rep., New York U., 2012, URL: https://cds.cern.ch/record/1456844.

[288] M. Aly et al., *TRExFitter*, version 1.6.0, 2025, URL: https://doi.org/10.5281/zenodo.17790864.

[289] F. J. Tackmann, *Beyond Scale Variations: Perturbative Theory Uncertainties from Nuisance Parameters*, (2024), arXiv: 2411.18606 [hep-ph].

[290] R. J. Barlow and C. Beeston, *Fitting using finite Monte Carlo samples*, Comput. Phys. Commun. **77** (1993) 219.

[291] 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.

[292] 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, arXiv: 1007.1727 [physics.data-an],
Erratum: Eur. Phys. J. C **73** (2013) 2501.

[293] R. Brun and F. Rademakers, *ROOT – An object oriented data analysis framework*,
Nucl. Instrum. Meth. A **389** (1997) 81, ISSN: 0168-9002.

[294] A. Pinto et al., *Uncertainty components in profile likelihood fits*,
Eur. Phys. J. C **84** (2024) 593, arXiv: 2307.04007 [physics.data-an].

[295] L. Buonocore, M. Grazzini, S. Kallweit, J. M. Lindert and C. Savoini,
Towards NNLO QCD predictions for off-shell top-quark pair production and decays,
JHEP **10** (2025) 195, arXiv: 2507.11410 [hep-ph].

[296] R. D. Cousins, *Generalization of Chisquare Goodness-of-Fit Test for Binned Data Using Saturated Models, with Application to Histograms*, 2010,
URL: https://www.physics.ucla.edu/~cousins/stats/cousins_saturated.pdf.

[297] ATLAS Collaboration, *Precise measurement of the $t\bar{t}$ production cross-section and lepton differential distributions in e dilepton events from $\sqrt{s} = 13$ TeV pp collisions with the ATLAS detector*, Eur. Phys. J. C **86** (2026) 470,
arXiv: 2509.15066 [hep-ex].

[298] 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, arXiv: 1908.07305 [hep-ex],
Erratum: Eur. Phys. J. C **80** (2020) 1092.

[299] ATLAS Collaboration, *Differential $t\bar{t}$ cross-section measurements using boosted top quarks in the all-hadronic final state with 139 fb^{-1} of ATLAS data*,
JHEP **04** (2023) 080, arXiv: 2205.02817 [hep-ex].

[300] 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, arXiv: 2108.02803 [hep-ex].

[301] A. Höcker et al., *TMVA - Toolkit for Multivariate Data Analysis*, 2007,
arXiv: physics/0703039 [physics.data-an].

[302] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, ATL-PHYS-PUB-2017-013, 2017,
URL: https://cds.cern.ch/record/2273281.

[303] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*,
Comput. Phys. Commun. **191** (2015) 159, arXiv: 1410.3012 [hep-ph].

[304] ATLAS Collaboration,
Tagging and suppression of pileup jets with the ATLAS detector,
ATLAS-CONF-2014-018, 2014, URL: https://cds.cern.ch/record/1700870.

[305] ATLAS Collaboration,
Measuring the b-jet identification efficiency for high p_T jets using multijet events in proton−proton collisions at $\sqrt{s} = 13$ TeV recorded with the ATLAS detector,
ATL-PHYS-PUB-2022-010, 2022, URL: https://cds.cern.ch/record/2804062.

[306] ATLAS Collaboration,
Graph Neural Network Jet Flavour Tagging with the ATLAS Detector,
ATL-PHYS-PUB-2022-027, 2022, URL: https://cds.cern.ch/record/2811135.