Calibration of the ATLAS B-tagger and the search for the $t\bar{t}H(H \to b\bar{b})$ process at $\sqrt{s} = 13$ TeV with the ATLAS experiment at the LHC

Dissertation

zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades
„Doctor rerum naturalium“
der Georg-August-Universität Göttingen

im Promotionsprogramm ProPhys
der Georg-August University School of Science (GAUSS)

vorgelegt von

Jannik Geisen

aus Nordhorn

Göttingen, 2019

Betreuungsausschuss

Prof. Dr. Arnulf Quadt
Prof. Dr. Stan Lai

Mitglieder der Prüfungskommission:

Referent:	Prof. Dr. Arnulf Quadt II. Physikalisches Institut, Georg-August-Universität Göttingen
Korreferent:	Prof. Dr. Stan Lai II. Physikalisches Institut, Georg-August-Universität Göttingen

Weitere Mitglieder der Prüfungskommission:

Prof. Dr. Ariane Frey
II. Physikalisches Institut, Georg-August-Universität Göttingen

Prof. Dr. Wolfram Kollatschny
Institut für Astrophysik, Georg-August-Universität Göttingen

Prof. Dr. Karl-Henning Rehren
Institut für Theoretische Physik, Georg-August-Universität Göttingen

Prof. Dr. Steffen Schumann
Institut für Theoretische Physik, Georg-August-Universität Göttingen

Tag der mündlichen Prüfung: 8. März 2019

Referenz: II.Physik-UniGö-Diss-2019/01

Calibration of the ATLAS B-tagger and the search for the $t\bar{t}H(H \to b\bar{b})$ process at $\sqrt{s} = 13$ TeV with the ATLAS experiment at the LHC

Abstract

Top quarks and Higgs bosons are the heaviest particles in the Standard Model of particle physics and are the subject of many analyses performed with the ATLAS experiment at the LHC at CERN. The Higgs boson was discovered in 2012 and is expected to play a major role in the way fundamental particles acquire mass, but also in potential new physics beyond the Standard Model. However, many of its properties have not been measured yet. One such property is its interaction with the top quark, represented by the top Yukawa coupling. The best way to measure this coupling is by observing the associated production of a Higgs boson with a top-antitop quark pair ($t\bar{t}H$) at the LHC. Furthermore, investigating such cases in which the Higgs boson decays into a bottom-antibottom quark pair ($t\bar{t}H(H \to b\bar{b})$) opens a window to also measuring the Yukawa coupling to the bottom quark. As the top and antitop quarks are expected to decay via the charged-current weak interaction into bottom and antibottom quarks as well, this analysis is dependent on a very efficient and precise method to identify jets originating from bottom quarks. The calibration of these identification methods employed in the ATLAS experiment is presented. It is based on 80.5 fb^{-1} of data collected at a centre-of-mass energy of $\sqrt{s} = 13$ TeV in the years 2015, 2016, and 2017. This calibration produces scale factors which can be used to correct the predicted identification efficiency to the one measured in data. The relative uncertainties on these scale factors range from 8-9% for jets with a low transverse momentum (p_{T}) to 1% at a medium p_{T} and, finally, to 3-4% at high p_{T}.
The search for the $t\bar{t}H(H \to b\bar{b})$ process with 36.1 fb^{-1} of ATLAS data collected in 2015 and 2016 is presented thereafter. The cross-section of this production mode is measured by performing a profile likelihood fit over several analysis regions involving decays of the top-antitop quark pair that produce either one or two charged leptons in the final state. The most dominant sources of uncertainty originate from the modelling of physics processes involving a top-antitop quark pair in association with a bottom-antibottom quark pair which is the main background process of this search. The ratio of the measured cross-section with respect to the one expected in the Standard Model, μ, is found to be

$$\mu_{t\bar{t}H(H\to b\bar{b})} = 0.84 \pm 0.29(\mathrm{stat.})^{+0.57}_{-0.54}(\mathrm{syst.}) = 0.84^{+0.64}_{-0.61}.$$

This translates into an inclusive cross-section of $\sigma_{t\bar{t}H} = 426^{+326}_{-312}$ fb when neglecting correlations between related uncertainties. The result corresponds to an observed (expected) significance of 1.4 (1.6) standard deviations and thus is not sufficient to claim an observation, as it is well compatible with both hypotheses, namely the absence as well as the presence of the $t\bar{t}H$ signal.

Kalibration des ATLAS B-Taggers und die Suche nach dem $t\bar{t}H(H \to b\bar{b})$ Prozess bei $\sqrt{s} = 13$ TeV mit dem ATLAS-Experiment am LHC

Zusammenfassung

Top-Quarks und Higgs-Bosonen sind die schwersten Teilchen im Standardmodell der Teilchenphysik und stehen im Fokus zahlreicher Analysen, die mit dem ATLAS-Experiment am LHC am CERN durchgeführt werden. Das Higgs-Boson wurde im Jahr 2012 entdeckt; Physiker erwarten, dass es eine Hauptrolle im Mechanismus zur Erzeugung von Teilchenmassen annimmt, aber auch in Bezug auf mögliche neue Physik jenseits des Standardmodells. Viele Eigenschaften des Higgs-Bosons wurden jedoch noch nicht experimentell vermessen. Eine dieser Eigenschaften ist seine Wechselwirkung mit dem Top-Quark, gegeben durch die Top-Yukawa-Kopplung. Die sinnvollste Art diese Kopplung zu vermessen ist über die Beobachtung der Produktion eines Higgs-Bosons in Assoziation mit einem Top-Antitop-Quarkpaar ($t\bar{t}H$) am LHC. Wenn man insbesondere die Fälle studiert, in denen das Higgs-Boson weiter zu einem Bottom-Antibottom-Quarkpaar zerfällt ($t\bar{t}H(H \to b\bar{b})$), bietet sich die Möglichkeit, auch die Bottom-Yukawa-Kopplung zu vermessen. Die Top- und Antitop-Quarks zerfallen ihrerseits über die schwache Wechselwirkung fast ausschließlich in Bottom- und Antibottom-Quarks. Daher ist diese Analyse auf eine höchst effiziente und präzise Methode angewiesen, um solche Jets zu identifizieren, die von Bottom-Quarks gebildet werden.

Vorgestellt wird die Kalibration solcher Identifikationsmethoden, wie sie im ATLAS Experiment verwendet werden. Diese basiert auf 80.5 fb^{-1} Daten, welche bei einer Schwerpunktsenergie von $\sqrt{s} = 13$ TeV in den Jahren 2015, 2016 und 2017 gesammelt wurden. Die Kalibration liefert Skalenfaktoren, mit welchen man die simulierte Identifikationseffizienz zu derjenigen korrigieren kann, die in den gesammelten Daten beobachtet wird. Die relative Unsicherheit dieser Skalenfaktoren beginnt bei etwa 8-9% für Jets mit einem niedrigen Transversalimpuls (p_{T}), beträgt rund 1% bei mittlerem p_{T} und steigt schließlich auf 3-4% für einen hohen p_{T}.

Die Suche nach dem $t\bar{t}H(H \to b\bar{b})$-Prozess mit 36.1 fb^{-1} ATLAS-Daten aus den Jahren 2015 und 2016 wird im Anschluss präsentiert. Der Wirkungsquerschnitt dieses Produktionsprozesses wird vermessen, indem eine Likelihood-Funktion über mehrere Regionen gebildet und angepasst wird; und zwar solche Regionen, die Zerfälle des Top-Antitop-Quarkpaares beinhalten, in denen entweder ein oder zwei geladene Leptonen erzeugt werden. Die dominanten Quellen der Gesamtunsicherheit resultieren aus der Modellierung solcher physikalischen Prozesse, die ein Top-Antitop-Quarkpaar in Assoziation mit einem Bottom-Antibottom-Quarkpaar beinhalten, welche den Hauptuntergrund dieser Suche darstellen. Das Verhältnis des vermessenen Wirkungsquerschnittes zum im Standardmodell erwarteten Wert, μ, ergibt sich zu:

$$\mu_{t\bar{t}H(H\to b\bar{b})} = 0.84 \pm 0.29(\mathrm{stat.})^{+0.57}_{-0.54}(\mathrm{syst.}) = 0.84^{+0.64}_{-0.61}.$$

Daraus ergibt sich ein inklusiver Wirkungsquerschnitt von $\sigma_{t\bar{t}H} = 426^{+326}_{-312}$ fb bei Vernachlässigung von Korrelationen zwischen entsprechenden Unsicherheiten. Dieses Ergebnis entspricht einer beobachteten (erwarteten) Signifikanz von 1.4 (1.6) Standardabweichungen und ist somit nicht signifikant genug, um eine Beobachtung zu verkünden, denn es ist sowohl mit der Untergrund-Hypothese als auch mit der Signal-Plus-Untergrund-Hypothese kompatibel.

Contents

CHAPTER 1

Introduction

Thinking back in time, it is astonishing and fascinating how far and quickly humankind's knowledge advanced over the course of history. The never ending progress on technology and, consequentially, our increased possibilities to study our world's phenomena create a promising outlook for the future. It is especially motivating for those of us who go beyond and desire to find the fundamental laws of nature that describe how our cosmos works and what its constituents are made of. Is there a single trait that connects and explains everything we see in our Universe? And if yes, can and will we ever find it?
One of the most prominent fields of research that shares this idea is particle physics. It pictures the matter that we encounter in our everyday lives to be made up of the smallest, elementary particles such as electrons and quarks. It further describes the forces that we experience, such as the electromagnetic force, as matter particles exchanging force particles, in this case photons. The profound mechanisms behind this superficial statement are well understood by now and summarised as the Standard Model of particle physics (SM) which is discussed in more detail in the following chapter. What makes the SM such a powerful and promising theory is that it enables us to precisely predict and explain most of the data that have been collected in particle physics experiments over the last century. Probing the SM through all of these experiments and striving to push the limits on every possible measurement have extended our understanding of the Universe, in particular in the time interval directly after the big bang, substantially.
The discovery of a Higgs-like particle in 2012 by the ATLAS and CMS collaborations at CERN [1, 2] has been one of the greatest successes in the history of particle physics. While the existence of a Higgs boson allows us to answer many questions, it raises even more, and hence motivates a wide range of studies to understand and confirm all of its predicted properties. However, since there are data and observations that we currently cannot explain with the SM such as the existence of dark matter, many different extensions of the SM have been proposed, collectively referred to as physics beyond the

Standard Model (BSM). The various suggested extensions are intensively being studied by, among others, the experiments at CERN mentioned above as well as many theorists. One potential window to new physics is that the Higgs boson plays a key role in the breaking of the electroweak symmetry and that it prefers to interact with heavy rather than light particles. Here, its interaction with the heaviest SM elementary particle, the top quark [3], is of particular interest.

In this thesis, the measurement of the production of a Higgs boson in association with a top-antitop quark pair is presented. The focus of this analysis is on those cases where the Higgs boson decays into a bottom-antibottom quark pair, labelled $t\bar{t}H(H \to b\bar{b})$. The data have been collected with the ATLAS experiment at the Large Hadron Collider (LHC) at CERN with a centre-of-mass energy of $\sqrt{s} = 13$ TeV in the years 2015 and 2016, corresponding to an integrated luminosity of 36.1 fb^{-1} [4].

This dissertation is structured as follows: the SM and details relevant for the presented analysis are discussed in Chapter 2. Chapter 3 describes the experimental setup including the ATLAS detector and the LHC. Chapter 4 presents the theoretical methods to simulate the collected experimental data in order to compare and evaluate our physics models. The objects that are reconstructed in the ATLAS detector from the recorded particle collisions are explained in Chapter 5. The identification strategy of a special type of particles for LHC analyses, namely bottom quarks, is indispensable and therefore described in detail in Chapter 6. Chapter 7 gives then an overview of both the physical data as well as the simulated datasets necessary to perform the search for the $t\bar{t}H(H \to b\bar{b})$ process. Chapter 8 depicts the selections through which this analysis optimises its sensitivity to the measurement. The analysis strategy and corresponding methods of measurement are presented in Chapter 9. Chapter 10 discusses the various systematic uncertainties and their sources. The expected and observed results are presented in Chapter 11 in addition to consistency checks and validations of the fit method. Chapter 12 highlights the limiting factors of this search and presents studies that aim to reduce them in future analyses. Finally, Chapter 13 concludes the presented material and provides an outlook for possible studies in the future.

CHAPTER 2

The Standard Model of particle physics

Since the discovery of the electron in 1897 by J. J. Thomson [5] up to the discovery of the Higgs boson in 2012 by the ATLAS and CMS experiments at CERN [1, 2], the Standard Model of particle physics (SM) has evolved significantly. It comprises a large number of different elementary particles that are understood to constitute our Universe. Furthermore, it describes three of the four forces we observe in nature via matter particles, the so-called fermions, exchanging force-mediating particles, the so-called gauge bosons. These particle interactions are depicted by Feynman diagrams and calculated using the corresponding Feynman rules. A point that connects the interacting particles within these diagrams is called a vertex. In the following, a brief overview of the SM, the fundamental particles and their interactions are given. Afterwards, dedicated sections will succinctly describe the Higgs mechanism and the special properties of the top and bottom quarks which play a key role in the physics analyses presented in Chapters 6 and 7. The discussions in this chapter are based primarily on the basic theory of particle physics that can be found in all related standard textbooks such as in Ref. [6–9].

2.1 The fundamental particles

The fundamental, or elementary, particles of the SM can be divided into two categories. The first category consists of the fermions which comprise quarks as well as leptons and constitute the visible matter of our known Universe. The second category consists of the gauge bosons which, except for the Higgs boson, mediate the fundamental forces between particles. The former share the property of possessing a half-integer spin of 1/2, while the latter hold an integer spin of 1 or, in the special case of the Higgs boson, a spin of 0. The spin translates into an important behaviour of the particles in that fermions obey Fermi-Dirac statistics and bosons behave according to Bose-Einstein statistics. All elementary particles, along with their physical properties and the gauge bosons they

interact with, are shown in Figure 2.1.

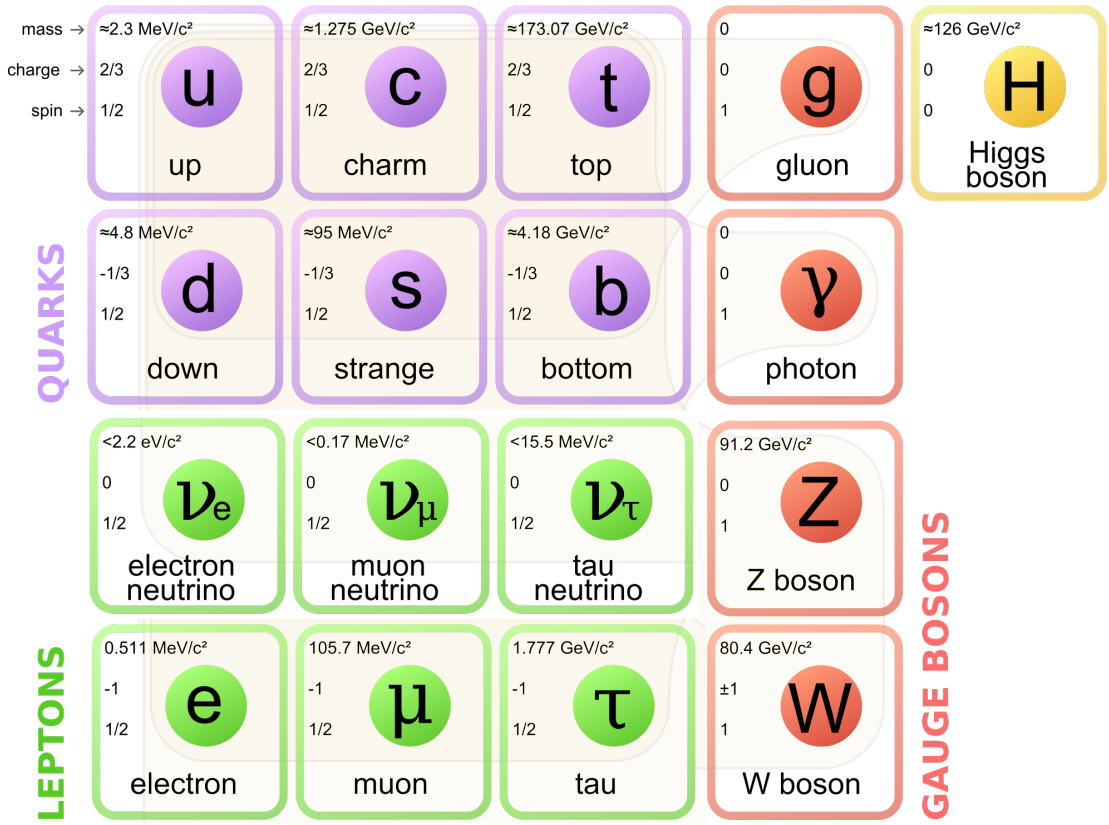

Figure 2.1: Overview of the known elementary particles of the Standard Model of particle physics and their physical properties.

The additional distinction between quarks and leptons within the fermions is attributable to the fact that quarks interact strongly via the exchange of gluons, while leptons do not. Furthermore, leptons and quarks exist in three families. The difference between the individual families is simply the mass of the fermions. In effect, the fermions in the second family can be thought of as copies of the fermions in the first family, but with a higher mass, while the fermions in the third family have an even higher mass. Current experimental data strongly suggests that our SM contains only three fermion generations [10], but a different scenario cannot be excluded entirely. Additionally, for each particle there exists an antiparticle which exhibits the same properties, but has all quantum numbers inverted.

It is useful to further divide the fermions according to their participation in the weak interaction which is given by their weak isospin I_{W}, more specifically the third component of its vector I_{W}^3. The charged-current weak interaction, further detailed in the next section, is mediated by the exchange of W^+ or W^- bosons between (anti)particles with

a (right-handed) left-handed chirality. (Anti)particles with a (left-handed) right-handed chirality do not participate in this interaction. Thus, the (right-handed) left-handed (anti)fermions form weak isospin doublets such that, for example, the electron forms a weak isospin doublet with the electron neutrino. In this doublet, the left-handed electron is the down-type partner with $I_{\mathrm{W}}^3 = -1/2$ and the left-handed electron neutrino is the up-type partner with $I_{\mathrm{W}}^3 = +1/2$. The right-handed counterparts form singlets instead, while right-handed neutrinos do not exist in the SM. Thus, a W^- boson may, for example, decay into a left-handed electron and, in order to conserve all relevant quantum numbers and charges, a right-handed electron antineutrino. In the same way, the muon and tau-lepton form doublets with their corresponding neutrinos as do the up, charm and top quarks ($I_{\mathrm{W}}^3 = +1/2$) with the down, strange and bottom quarks ($I_{\mathrm{W}}^3 = -1/2$), respectively. The arrangement into weak isospin doublets and singlets is depicted in the following:

$$\begin{pmatrix}\nu_e\\ e\end{pmatrix}_L, \begin{pmatrix}u\\ d\end{pmatrix}_L, e_R, u_R, d_R,$$
$$\begin{pmatrix}\nu_\mu\\ \mu\end{pmatrix}_L, \begin{pmatrix}c\\ s\end{pmatrix}_L, \mu_R, c_R, s_R,$$
$$\begin{pmatrix}\nu_\tau\\ \tau\end{pmatrix}_L, \begin{pmatrix}t\\ b\end{pmatrix}_L, \tau_R, t_R, b_R.$$

The key element of the charged-current weak interaction is that it is the only way that fermions of higher mass may decay into other fermions of lower mass, as long as all relevant quantum numbers, charges, energy, momenta and angular momenta are conserved. All other gauge bosons and the corresponding interactions they mediate conserve the so-called flavour of the particle, which means that, for example, a muon cannot turn into an electron by emitting a photon or a strange quark cannot turn into a down quark by emitting a Z^0 boson. Such a change of flavour is only possible via the emission of a charged W boson and a corresponding particle to ensure that all quantities mentioned above are conserved, as is illustrated in Figure 2.2.
Down-type quarks have an electric charge of -1/3, while up-type quarks carry an electric charge of +2/3. In addition to this, they also hold a colour charge and thus obey Quantum Chromodynamics (QCD) which describes the strong interaction via the exchange of gluons, further detailed in Section 2.4.
On the other hand, the down-type or charged leptons possess an electric charge of -1 and the up-type leptons, i.e. neutrinos, hold no electric charge. This electric charge as well as the weak isospin of the fermions play a key role in the electroweak interaction each fermion participates in.

2.2 The fundamental forces

The SM can be described succinctly as a renormalisable, locally gauge invariant quantum field theory that is based on the $\mathrm{SU}(3)_C \times \mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetry group. Symmetries

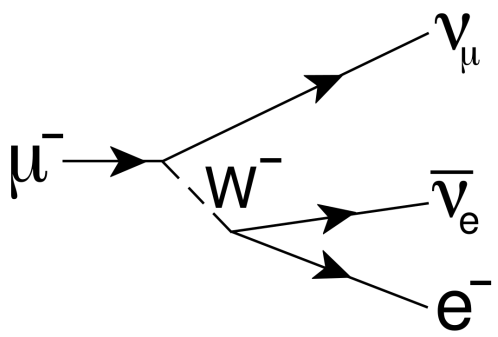

Figure 2.2: A Feynman diagram showing a muon decaying into a W^- boson and a muon neutrino. The W^- further decays into an electron and an electron antineutrino. The muon may only decay into an electron via the charged-current weak interaction mediated by an off-shell W boson and the two neutrinos are necessary to conserve all quantum numbers, flavours, energy, momentum and angular momentum at every vertex point during this decay.

are phenomena observed in nature that result from existing degrees of freedom within the considered system. They lead to conserved quantities that can be measured in experiments and can be described mathematically by the corresponding symmetry groups. In the case of the SM, the $\mathrm{U}(1)_{\mathrm{EM}}$ group represents the electromagnetic interaction with the electric charge acting as the conserved quantity. This symmetry group remains from the $\mathrm{U}(1)_Y$ symmetry which conserves the hypercharge Y after the spontaneous symmetry breaking. The concept of the latter will be discussed in Section 2.3.1. The $\mathrm{SU}(2)_L$ group depicts the weak interaction in which the weak isospin I_{W} is conserved. Finally, the strong interaction, which conserves the colour charge C, is defined by the $\mathrm{SU}(3)_C$ group. The force of gravity between individual particles at the energy scales available in current particle physics experiments is negligible compared to the other three forces and, therefore, can be completely neglected in the context of this thesis. Apart from this, gravity cannot be described as a renormalisable quantum field theory as the other three forces.
The aforementioned term "renormalisable" means that the physics we measure can be described independently of the distance, or correspondingly the energy scale, at which we observe an interaction between particles. Local gauge invariance is the second crucial principle which ensures that our laws of physics do not change if we perform a local gauge transformation according to the symmetry group under study, for example a simple phase shift in $\mathrm{U}(1)_Y$.

As indicated in Chapter 1, the SM is a powerful tool that allows to make remarkably precise predictions about the particle interactions known to us. However, in order to calculate such predictions, a mathematical description based on a locally gauge invariant quantum field theory is essential. The corresponding Lagrangians based on the

U(1), $SU(2)_L$ and $SU(3)_C$ symmetry groups can be constructed in a locally gauge invariant way by introducing the known gauge bosons with additional terms that describe their interactions with other particles as we observe them in nature. Additionally, this construction works similarly for the three symmetry groups. The difference between them is the structure of the Lagrangian, as the corresponding generators have a higher dimensionality for $SU(2)_L$ and $SU(3)_C$ compared to U(1). Because of the higher dimensionality, the weak and strong interactions are both non-Abelian theories which leads to self-interactions of the gauge bosons. In the case of $SU(2)_L$, it is reasonable to consider the two-dimensional Pauli matrices σ_i as generators of the symmetry group, while for $SU(3)_C$ the commonly used generators are the Gell-Mann matrices λ_k. However, this procedure only works if the interacting gauge bosons are massless. Since the masses of the $W^{\pm}$ and Z bosons of the weak interaction have been measured to be about 80 GeV and 91 GeV [10,11], respectively, another strategy has to be considered. This new strategy is the Higgs mechanism [12–14], which was proposed in 1964 as a solution to give masses to the three weak bosons and unify the electromagnetic and weak interaction, while keeping the Lagrangian invariant under local gauge transformations all at the same time. Since this is a critical aspect to understand the motivation behind the analysis presented in Chapter 7, the Higgs mechanism is the topic of the following section.

2.3 The Higgs boson

The Higgs mechanism was proposed to introduce mass terms for the gauge bosons of the weak interaction and elementary fermions to the SM Lagrangian without breaking its local gauge invariance. This theory was published by three independent groups in 1964 [12–14]. The idea is that there is a set of complex scalar fields throughout our Universe which our known massive particles interact with. The particles would initially be massless and thus travel at the speed of light. However, because the vacuum expectation value (VEV) v of this field, namely the Higgs field, is non-zero, the particles acquire mass and slow down. The Higgs boson itself is an excitation of this field. Since the Higgs boson holds neither electric charge nor colour charge [3], it does not interact directly with photons or gluons and hence, those gauge bosons are massless. Furthermore, as it is a scalar particle, the Higgs boson has no spin. Its discovery in 2012 by the CMS and ATLAS experiments at the LHC with a mass of about 125 GeV was a major success of the SM [1,2]. As a consequence, a new era of particle physics research has emerged, either probing the predicted properties of the Higgs boson or searching for new extensions of the SM compatible with available measurements of the Higgs boson.

In the following brief description of the Higgs mechanism, the unification of the electric and weak interaction as well as its spontaneous symmetry breaking arise naturally. This shows the elegance of the theory, because both interactions seem to manifest in fundamentally different ways at first.

2.3.1 The Higgs mechanism

The main ideas behind the Higgs mechanism are highlighted in the following based on the minimal Higgs model of the SM. Consider two complex scalar field, placed in a weak isospin doublet

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}, \tag{2.1}$$

where ϕ^+ is the charged scalar field which fulfils $(\phi^+)^* = \phi^-$. The two charged fields will yield the longitudinal degrees of freedom of the W^+ and W^- bosons, respectively. Correspondingly, the neutral scalar field ϕ^0 will yield the degrees of freedom of the Z^0 boson and the photon. This doublet has a Lagrangian

$$\mathcal{L} = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - V(\phi). \tag{2.2}$$

with the Higgs potential

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda (\phi^\dagger \phi)^2. \tag{2.3}$$

This potential only has a finite minimum if $\lambda > 0$, but μ^2 can be either greater or less than zero. In the case that $\mu^2 > 0$, the minimum of the potential is given by the trivial solution $\phi_1 = \phi_2 = \phi_3 = \phi_4 = 0$. More interestingly, however, is that for $\mu^2 < 0$, there is an infinite set of degenerate minima given by

$$\phi^\dagger \phi = \frac{1}{2} \left(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2 \right) = \frac{-\mu^2}{2\lambda} \equiv \frac{v^2}{2}, \tag{2.4}$$

where v is the VEV of the Higgs field. Writing out Equation 2.2 in terms of the real scalar fields ϕ_i, one finds

$$\mathcal{L} = \frac{1}{2} \sum_{i=1}^{4} (\partial_\mu \phi_i)(\partial^\mu \phi_i) - \frac{1}{2} \mu^2 \sum_{i=1}^{4} \phi_i^2 - \frac{1}{4} \lambda \left(\sum_{i=1}^{4} \phi_i^2 \right)^2. \tag{2.5}$$

Here, the first sum represents the kinetic energy of the scalar particle associated with the field, the second term can be interpreted as its mass and the third term describes its self-interactions.
Equation 2.5 is invariant under global transformations of the U(1) symmetry group, that is

$$\phi \to \phi' = e^{i\alpha} \phi, \tag{2.6}$$

because $\phi'^\dagger \phi' = \phi^\dagger \phi$. However, it is not invariant under local transformations, meaning $\alpha = \alpha(x)$, let alone local gauge transformations of the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetry.
Choosing a particular solution for the minimum, for example the simple case $\phi_1 = 0$, $\phi_2 = 0$, $\phi_3 = v$, $\phi_4 = 0$, spontaneously breaks the global gauge symmetry. This yields a massive scalar particle as well as three massless so-called Goldstone bosons in

the corresponding Lagrangian. These Goldstone bosons will yield the three longitudinal degrees of freedom for the three massive weak bosons. By expanding the field around the minimum chosen above with a new field such that $\phi_3(x) = v+\eta(x)$, it can be written as

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1(x) + i\phi_2(x) \\ v + \eta(x) + i\phi_4(x) \end{pmatrix}. \tag{2.7}$$

The photon has to remain massless after the spontaneous symmetry breaking, thus one can choose the so-called unitary gauge for the neutral field ϕ^0 such that

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \tag{2.8}$$

This gauge choice includes an entirely real scalar field, while $\eta(x)$ is written suggestively as the Higgs field $h(x)$.
The resulting Lagrangian is known as the Salam-Weinberg model. It can be written in a way such that it is invariant under local gauge transformations of the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ symmetry group, namely by introducing new gauge fields $\overrightarrow{W}_\mu$ and B_μ as well as replacing the derivatives with the following covariant derivatives:

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + ig_\mathrm{W} \frac{\vec{\sigma}}{2} \cdot \overrightarrow{W}_\mu + ig' \frac{Y}{2} B_\mu. \tag{2.9}$$

Here, $\vec{\sigma}$ are the Pauli matrices which are the three generators of the $\mathrm{SU}(2)_L$ symmetry group and Y is the weak hypercharge acting as the generator of the $\mathrm{U}(1)_Y$ symmetry group. It is related to the electric charge Q and the third component of the weak isospin I^3_W by the Gell-Mann-Nishijima formula:

$$Y = 2(Q - I^3_\mathrm{W}). \tag{2.10}$$

g' and g_W are the couplings of the $\mathrm{U}(1)_Y$ and $\mathrm{SU}(2)_L$ gauge symmetries, respectively. Their ratio can be expressed in terms of the so-called weak mixing angle θ_W:

$$\frac{g'}{g_\mathrm{W}} = \tan\theta_\mathrm{W}. \tag{2.11}$$

The new gauge fields in the now locally gauge invariant Lagrangian do not represent the physical massive gauge bosons. The latter can be understood as an interference or mixed states of the former. In particular, the fields of the physical charged-current bosons of the weak interactions are given by

$$W^\pm_\mu = \frac{1}{\sqrt{2}} \left(W^{(1)}_\mu \mp i W^{(2)}_\mu \right). \tag{2.12}$$

Additionally, by using Equation 2.11, the neutral photon and Z_0 boson fields can be written in the following way:

$$\begin{aligned} A_\mu &= \quad \cos\theta_\mathrm{W} B_\mu + \sin\theta_\mathrm{W} W^{(3)}_\mu, \\ Z_\mu &= -\sin\theta_\mathrm{W} B_\mu + \cos\theta_\mathrm{W} W^{(3)}_\mu. \end{aligned} \tag{2.13}$$

By rewriting the Lagrangian in terms of these mixed fields, one finds the mass of the $W^{\pm}$ bosons to be $m_W = \frac{1}{2} g_{\mathrm{W}} v$, the mass of the Z^0 boson to be $m_Z = g_{\mathrm{W}}^2 + g'^2 = \frac{1}{2} \frac{g_{\mathrm{W}}}{\cos\theta_{\mathrm{W}}} v = \frac{m_W}{\cos\theta_{\mathrm{W}}}$ and the mass of the photon to be $m_A = 0$. Furthermore, the mass of the new scalar particle H is found to be $m_H^2 = 2\lambda v^2$. Finally, while the VEV of the Higgs field v is a free parameter of the SM, the mass relations above are consistent with all available measurements of the SM and determine a value of $v = 246$ GeV [3].

The terms describing the interaction between the Higgs boson and the weak bosons can be extracted by writing out the corresponding products of the Higgs field and the physical gauge boson fields. Furthermore, it can be shown that the coupling strength of the Higgs boson to the weak bosons is proportional to their respective masses. Similarly, the Higgs mechanism can be used to introduce mass terms for the SM fermions. The coupling strength of the Higgs to these fermions, called the Yukawa coupling, is not predicted by theory, but likewise assumed to be proportional to the corresponding fermion masses. This will be further discussed in Section 2.7.1.

In the SM, the Higgs has neither electromagnetic nor colour charge and therefore does not directly couple to photons or gluons [3]. It may, however, decay into two photons or gluons via a triangular top quark or W boson loop, as shown in Figure 2.3. Another important decay channel of the Higgs boson is the four-lepton channel where the Higgs boson decays into one real and one virtual Z^0 boson which in turn decay into two leptons of the first and second family (ee or $\mu\mu$). The $\tau\tau$ channel is not considered due to the additional neutrinos from their subsequent decay. Even though the Higgs boson decay channels into two photons and four leptons have only small branching ratios [3] which is depicted in Figure 2.4, they have a very high signal to background ratio. It was in these two channels that the ATLAS and CMS experiments discovered the Higgs boson during Run 1 of the LHC in 2012 [1, 2].

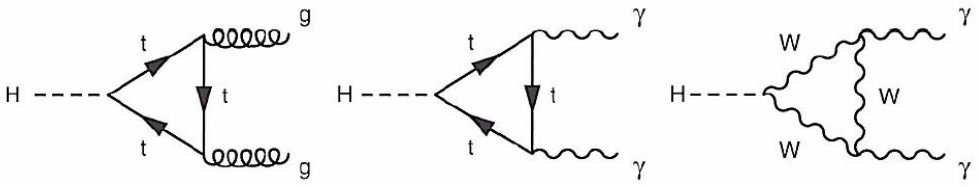

Figure 2.3: The Feynman diagrams show how the Higgs boson may decay indirectly into two gluons via a top quark loop (left) or into two photons via either a top quark loop (centre) or W boson loop (right) [7].

To understand the underlying principles of the SM as a renormalisable, locally gauge invariant theory, the strong interaction of the $\mathrm{SU}(3)_C$ has to be considered as well. The core principles of the strong interaction are important to understand the topology of the sought-after signal events introduced in Chapters 6 and 7 as well as their expected signature in the detector. These will be discussed briefly in the following section.

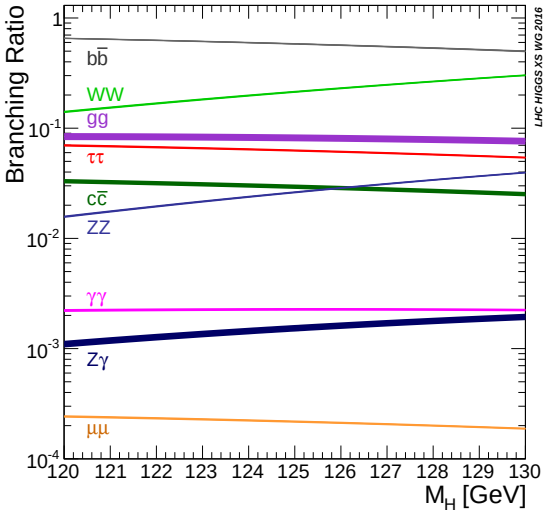

Figure 2.4: The decay branching ratios of the SM Higgs boson as a function of its mass [3, 15].

2.4 Quantum Chromodynamics

Quantum Chromodynamics (QCD) is the quantum field theory which is based on the $\mathrm{SU}(3)_C$ symmetry group to describe the strong interaction. It manifests itself via the exchange of gluons, i.e. the gauge bosons of the strong interaction, between particles that hold a so-called colour charge, which are quarks as well as gluons themselves. This colour charge represents an additional degree of freedom that can assume the three possible values 'red', 'green' and 'blue'. This idea was introduced to explain how spin-3/2 baryons in a symmetrical state of space, spin and flavour could still have a total antisymmetric wave function - namely by having different colour charges - and thus obey Fermi-Dirac statistics as all fermions do. While quarks possess one specific colour, there are eight different possible gluon states which comes from the structure of the $\mathrm{SU}(N)$ group, which for $N = 3$ has $N^2 - 1 = 8$ degrees of freedom.
This structure and number of degrees of freedom is reflected in the corresponding QCD Lagrangian below which describes quarks with colour a and b that interact with gluons. It contains kinetic terms describing the Dirac fermions as well as the gluons in addition to interaction terms introduced by the covariant derivative and is given by

$$\mathcal{L}_{\mathrm{QCD}} = \bar{q}_a \left(i\gamma^\mu D_\mu - m\right)_{ab} q_b - \frac{1}{4} F^A_{\mu\nu} F^{\mu\nu}_A. \tag{2.14}$$

Here, the covariant derivative

$$D_\mu = \partial_\mu + \frac{1}{2} i g_S G^A_\mu \lambda^A \tag{2.15}$$

introduces a new gauge field which is the gluon field G^A_μ where A runs over the eight colour degrees of freedom. The coupling strength of the strong interaction is given by g_S

and λ^A represent the Gell-Mann matrices that generate the $\mathrm{SU}(3)_C$ symmetry group. If the gluon field transforms as

$$G_\mu^A \rightarrow G_\mu^{A\prime} = G_\mu^A - \partial_\mu \alpha_A - g_S f^{ABC} \alpha_B G_\mu^C, \tag{2.16}$$

the Lagrangian is invariant under local $\mathrm{SU}(3)_C$ phase transformations of the form

$$q(x) \rightarrow q'(x) = \exp\left[\frac{1}{2} i g_S \vec{\alpha}(x) \cdot \vec{\lambda}\right] q(x). \tag{2.17}$$

In Equation 2.16, f^{ABC} are the structure constants of the $\mathrm{SU}(3)_C$ group, defined by the commutation relations $[\lambda^A, \lambda^B] = 2if^{ABC}\lambda^C$, and $\vec{\alpha}(x)$ is an arbitrary real function with eight components. Furthermore, Equation 2.14 contains the field strength tensor $F_{\mu\nu}^A$ derived from the gluon field. It is given by

$$F_{\mu\nu}^A = \left[\partial_\mu G_\nu^A - \partial_\nu G_\mu^A - g_S f^{ABC} G_\mu^B G_\nu^C\right] \tag{2.18}$$

Compared to the field strength tensor in QED, there is an additional third term on the right-hand side of Equation 2.18 which distinguishes QCD from QED. It exhibits the non-Abelian structure of the $\mathrm{SU}(3)_C$ symmetry group and gives rise to triple and quartic gluon self-interactions which ultimately yield the properties of asymptotic freedom and confinement.

These two properties can be explained by the so-called running of the strong coupling constant $\alpha_\mathrm{S} = g_S/4\pi$. First of all, the coupling "constants" of the SM interactions are not constant, but change with the energy scale q^2 at which the interaction takes place. However, in contrast to the electromagnetic and weak interactions, the strength of the strong interaction grows with increasing distance or decreasing energy. This follows from the fact that QCD is a renormalisable gauge theory based on the non-Abelian $\mathrm{SU}(3)_C$ symmetry group which involves triple and quartic self-interactions of gluons as mentioned above. The formula describing how α_S evolves with the energy scale q^2 is given by

$$\alpha_S(q^2) = \frac{\alpha_S(\mu^2)}{1 + \left(\frac{11N_C - 2N_f}{12\pi}\right) \alpha_S(\mu^2) \ln\left(\frac{q^2}{\mu^2}\right)}. \tag{2.19}$$

Here, $N_C = 3$ is the number of colours and N_f is the number of quark flavours which participate in the interaction at the energy scale μ^2. Since $N_f \leq 6$ in the SM, the term $11N_C - 2N_f$ is always positive and thus, α_S decreases with increasing q^2 and vice versa. Equation 2.19 can be used to determine α_S at any energy scale starting from the following value measured at the Z^0 boson mass at LEP [3, 10]:

$$\alpha_S(m_Z^2) = 0.1181 \pm 0.0011. \tag{2.20}$$

The first approach to calculate any particle physics interaction is to use the Feynman rules which rely on perturbation theory. Specifically for QCD, this means that the strong

interaction is approximated as a power series in α_S. And since α_S increases with decreasing energy, it attains a value of ≈ 1 around an energy of 1 GeV. This causes the perturbation theory to break down as the interaction probability diverges over all orders. This is the principle of confinement and it has another significant consequence for high energy particle physics experiments, namely infra-red divergence. The probability for coloured particles to radiate off soft (low energetic) or collinear (small angle) gluons increases over all boundaries and thus, quarks and gluons in collider experiments such as the LHC are not observed as individual particles, but rather as a bundle of hadronic bound states inside a cone, which is called a jet. The reconstruction of such jets is further discussed in Section 3.2.
With increasing travelling distance, the initial coloured particle has generated many additional quarks and gluons which eventually bind together and form hadrons. This process, called hadronisation, happens at such energy scales at which perturbation theory breaks down. It is part of the so-called parton shower (PS) evolution, which is simulated with Monte Carlo (MC) generators. These models are essential in order to make predictions that can be compared to the data measured in particle physics experiments involving such energy scales. Therefore, the method of modelling the PS and hadronisation is further described in Chapter 4.

The other property of the strong interaction is the asymptotic freedom and it describes the behaviour of coloured particles at very high energies, typically above 100 GeV. These energy scales are commonly observed during the primary interaction of high energy particle collisions as they occur, for example, at the LHC. Considering Equation 2.20 at the mass of the Z^0 boson which is around 91 GeV, the strong coupling strength is ≈ 0.1 at energies above 100 GeV. Very energetic quarks, therefore, behave as if they were free particles instead of being strongly confined into bound states. Thus, perturbation theory can be used to calculate the primary, hard interaction and allows physicists to predict them with high precision. The modelling of these hard interactions is the topic of Section 4.1. However, since α_S still has a value of ≈ 0.1, the higher order correction terms of the perturbation theory cannot be neglected which is in contrast to, for example, electromagnetic interactions with $\alpha_{EM}(m_Z^2) \approx 1/128$. These higher order corrections usually involve a large number of processes, especially quantum loops of virtual gluons, and thus prove to be a difficult challenge for modelling them with MC generators.

In order to explain the analyses presented in Chapters 6 and 7, two important particles have to be discussed, namely the bottom quark and the top quark. Their special properties as well as their relevance for these analyses will be described in the following sections.

2.5 The bottom quark

The bottom quark is the down-type quark of the third generation and thus closely related to the up-type top quark. Consequently, the bottom quark has a spin of 1/2 like

all fermions, an electric charge of -1/3, a colour charge, and the third component of its weak isospin is -1/2. It has, compared to most other elementary fermions, a relatively high mass of $m_b = 4.18^{+0.04}_{-0.03}$ GeV [3].
The existence of the bottom quark was proposed in 1973 to explain the observed CP violation in our Universe [16], where CP refers to the charge conjugation and parity symmetries. Applying the CP-symmetry to a particle state turns it into its corresponding antiparticle state and reflects the spatial coordinates. Since particles and antiparticles are commonly produced and annihilated together, one would expect the Universe to contain an equal amount of each. As this is not observed, the CP-symmetry must be violated in some form. The charged-current weak interaction allows for quarks of different generations to mix, a mechanism proposed for the first two quark families in 1963 [17]. It describes the interaction of up-type and down-type quarks via weak eigenstates instead of the physical mass eigenstates. However, this mechanism only leads to CP violation if there are at least three generations of quarks, hence the proposal of a third set of quarks [16]. The weak eigenstates are related to the mass eigenstates via the Cabibbo-Kobayashi-Maskawa (CKM) matrix, specifically in the following way:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} . \tag{2.21}$$

Here, V_{ij} indicates the amplitude of the interaction between up-type quark i and down-type quark j. Additionally, d, s and b represent the spinors of the mass eigenstates of the three down-type quarks. The CKM matrix is unitary and the absolute value of each of its elements has been measured in many different experiments [3]. They are approximately:

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} \approx \begin{pmatrix} 0.9742 \pm 0.0002 & 0.2243 \pm 0.0005 & 0.0039 \pm 0.0004 \\ 0.2180 \pm 0.0040 & 0.9970 \pm 0.0170 & 0.0042 \pm 0.0008 \\ 0.0081 \pm 0.0005 & 0.0039 \pm 0.0003 & 1.0190 \pm 0.0250 \end{pmatrix} . \tag{2.22}$$

As expected, quarks of the same generation have the highest interaction amplitude which is given by the diagonal elements. Furthermore, the off-diagonal elements are non-zero and therefore allow quarks from other generations to interact with each other. Three important examples for this are the top, bottom and charm quarks. Since these quarks belong to the second and third generation, they have a higher mass than those belonging to the first and thus may decay into lighter quarks via the charged-current weak interaction. The matrix element $|V_{tb}|$ is almost 1, which means that the top quark almost exclusively interacts with bottom quarks i.e. it decays weakly into bottom quarks and does so after an extremely short lifetime [3]. This is further outlined in Section 2.6. Since the bottom quark is lighter than the top quark, it can only decay into quarks of the second or first generation. These decay rates are proportional to $|V_{cb}|^2$ and $|V_{ub}|^2$, respectively. Since these are off-diagonal elements, their absolute value is extremely small which translates into a relatively long lifetime of bottom quark containing hadrons

(b hadrons) of $\tau_b \approx 1.5 \cdot 10^{-12}$ s [3]. The charm quark, on the other hand, may decay into a strange quark, which belongs to the same generation and thus, compared to b hadrons, hadrons containing a charm quark (c hadrons) possess a shorter lifetime of $\tau_c \approx 0.5-1.0 \cdot 10^{-12}$ s [3]. These lifetimes are in the range such that they can be exploited to identify bottom and charm quarks and distinguish them from lighter quarks in particle collisions at the LHC. This is a critical aspect of the multi-purpose detectors ATLAS and CMS for high energy particle physics analyses that is further described in Section 3.2.

The following section introduces the weak isospin partner of the bottom quark, namely the top quark. The focus of discussion lies on its special properties and its meaning for physics analyses at high energy hadron colliders such as the LHC which is crucial to motivate and understand the later chapters.

2.6 The top quark

According to the SM, the top quark is the weak isospin partner of the bottom (b) quark and belongs to the third generation of quarks. As all up-type quarks, it possesses an electric charge of $+2/3$, a colour charge, a spin of $1/2$ and the third component of its weak isospin is $+1/2$ [3]. Furthermore, as a fermion it has a unique trait which is its high mass of 173.0 ± 0.4 GeV [3] which makes it the heaviest particle in the SM.
After the discovery of the bottom quark in 1977, the top quark was expected to be discovered soon after. However, since it has such a high mass, an experimental setup with high energy was required which the LEP experiment did not offer. But in 1995, at the TEVATRON, a proton-antiproton collider with an energy of up to 1.96 TeV, the DØ and CDF experiments were both able to discover the top quark [18, 19].

The top quark exhibits special properties that distinguish it from all other quarks of the SM. Firstly, it is an unstable particle that may decay weakly into other down-type quarks, but it has a very short average lifetime. According to the CKM matrix elements, it decays into a b quark and a W boson in over 99.9% of all cases, while decays into strange or down quarks are difficult to observe in today's particle physics experiments. The decay width of the top quark, when neglecting higher order corrections, is predicted in the SM by [3]

$$\Gamma_t = \frac{G_F m_t^3}{8\pi\sqrt{2}} \left(1 - \frac{M_W^2}{m_t^2}\right)^2 \left(1 + 2\frac{M_W^2}{m_t^2}\right) \left[1 - \frac{2\alpha_S}{3\pi}\left(\frac{2\pi^2}{3} - \frac{5}{2}\right)\right]. \tag{2.23}$$

Here, G_F refers to the Fermi constant, m_t is the top quark pole mass, M_W is the W boson mass and α_S is the strong coupling constant. This gives a value of $\Gamma_t \approx 1.35$ GeV which translates into a very short average lifetime of $\tau_t = 1/\Gamma_t \approx 5 \cdot 10^{-25}$ s [3]. The hadronisation time scale is roughly two orders of magnitude slower, namely about 10^{-23} s, and thus, the top quark decays before it can form bound states or hadrons. Consequently, it is the only quark in the SM that can be measured experimentally as a bare quark, but only through its decay products. When studying top quarks at hadron colliders, the

primary production is the production of a top-antitop quark pair, called the $t\bar{t}$ process, via the strong interaction [7, 9]. The main $t\bar{t}$ production mechanisms are depicted in Figure 2.5. Then, as the top quarks decay before hadronising, their decay products can be analysed. Therefore, the $t\bar{t}$ final state is split into three categories as illustrated in Figure 2.6, depending on the decay modes of the two W bosons into either a charged lepton and neutrino or a quark-antiquark pair. A Feynman diagram of an example semileptonic decay ("lepton+jets") is shown in Figure 2.7. The difference in the identification of jets, hadrons, charged leptons, and neutrinos at hadron colliders as well as the production rates of those objects in background processes, which will be further described in Section 3.2 specifically for the ATLAS experiment at the LHC, are the reason that the all-hadronic decay channel ("alljets") has the smallest signal over background ratio (purity) compared to all channels. Furthermore, the dileptonic decay channel has an even higher purity compared to the semileptonic channel. However, it also has a smaller branching ratio, namely 9% compared to 45%. This circumstance must be taken into account when defining the strategy of the analyses presented in Chapters 6 and 7.

Another important property of the top quark is its high mass itself. Through quantum loops in the calculation of Feynman diagrams, the top quark and its mass contributes to various couplings and gauge boson decay widths. In particular, during their propagation, the Higgs boson, Z^0 boson and W boson may produce quantum loop diagrams involving virtual top quarks. The emphasis is on the word virtual, because the bosons are lighter than the top quark and thus cannot produce or decay into real top quarks. Furthermore, the Yukawa coupling of the Higgs boson to fermions, which is the topic of the following section, is expected to be proportional to the fermion mass. With the top quark being the heaviest particle in the SM, the top Yukawa coupling is expected to be the largest among all particles. Hence, the top quark mass significantly contributes to the Higgs boson decay width and it can play a key role in the electroweak symmetry breaking, and may be a potential window to physics beyond the SM.

The analysis in Chapter 6 will focus on $t\bar{t}$ events, specifically on the case when the $t\bar{t}$ pair decays into two b quarks and both W bosons decay into an electron or muon and corresponding neutrino. This channel has a branching ratio of about 2%, as can be seen in Figure 2.6, but it also has the highest purity which translates to a higher signal over background ratio than in the other decay channels.

The following section briefly introduces the production of a Higgs boson in association with a $t\bar{t}$ pair. It is a non-trivial final state involving many different types of particles and in particular b quarks. Thus, in order to study this production mechanism, it is necessary to exploit all previously discussed special properties of the Higgs boson, the bottom quark, and the top quark.

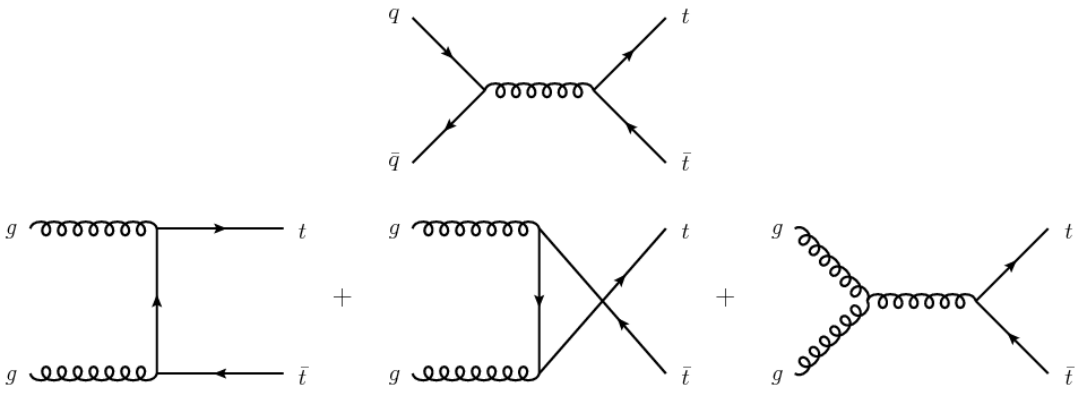

Figure 2.5: The leading order Feynman diagrams for $t\bar{t}$ production at a hadron collider. The production initiated by quarks was the most dominant process at the TEVATRON, while at the LHC it is the gluon-initiated production [3].

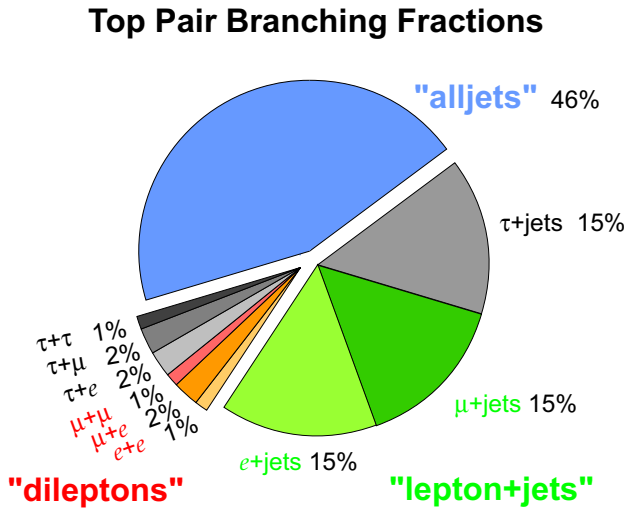

Figure 2.6: The different decay channels of a top-antitop quark pair and their respective branching ratios [3].

2.7 Higgs boson production in association with a $t\bar{t}$ pair

The production of a Higgs boson in association with a $t\bar{t}$ pair, referred to as $t\bar{t}H$ in the remainder of this thesis, is a physics process of great interest to the high energy particle physics community. It allows the direct measurement of the Yukawa coupling of the Higgs boson to the top quark, which represents another important property to verify whether the Higgs-like particle discovered in 2012 is actually the Higgs boson as it is predicted by the SM [20]. Any deviations from the expected coupling would hint at new physics phenomena that could challenge and eventually change our understanding

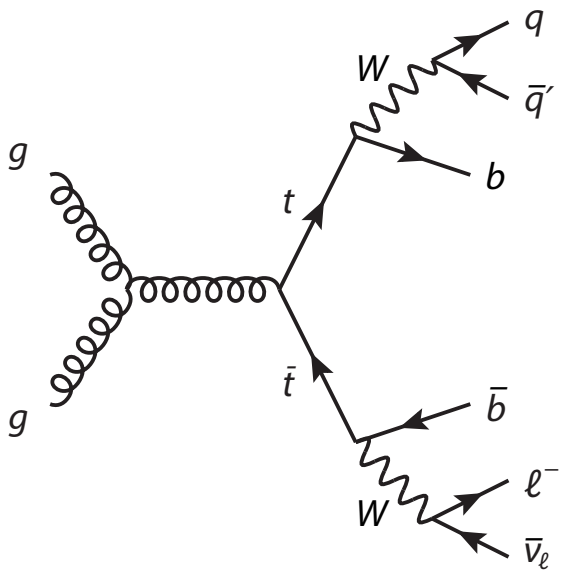

Figure 2.7: An example production of a semileptonic $t\bar{t}$ decay at the LHC. The top quarks are produced via the fusion of two gluons and they further decay weakly into b quarks and charged W bosons. The W^- boson decays into a negatively charged lepton and corresponding antineutrino, while the W^+ boson decays into a quark-antiquark pair.

of particle physics dramatically.

At the LHC, Higgs bosons are produced via four main production mechanisms, namely the fusion of two gluons, the fusion of two vector or weak bosons and the associated production modes together with either a vector boson or a $t\bar{t}$ pair. The Feynman diagrams of these four processes are depicted in Figure 2.8.
These production mechanisms have significantly different cross-sections at the LHC for different centre-of-mass energies. While the gluon-gluon fusion process has a cross-section of about 50 pb at $\sqrt{s} = 13$ TeV, the predicted cross-section of the $t\bar{t}H$ process is about 1% of that, namely 507^{+35}_{-50} fb [15]. The predicted production cross-sections are plotted in Figure 2.9 as a function of the centre-of-mass energy. The low cross-section of $t\bar{t}H$ compared to the gluon-gluon fusion process makes it a challenging process to observe and adds to the complexity of the analysis presented in Chapter 7.

2.7.1 The Yukawa Coupling

The Yukawa coupling of the Higgs boson to the top quark is an important parameter in the Higgs mechanism and the sector of electroweak symmetry breaking. It is essential as a test of the SM and a potential window to BSM physics. It can be measured directly in the $t\bar{t}H$ process. In general, this coupling strength of the Higgs boson to fermions can be used to introduce the mass terms for the Dirac fermions to the SM Lagrangian while

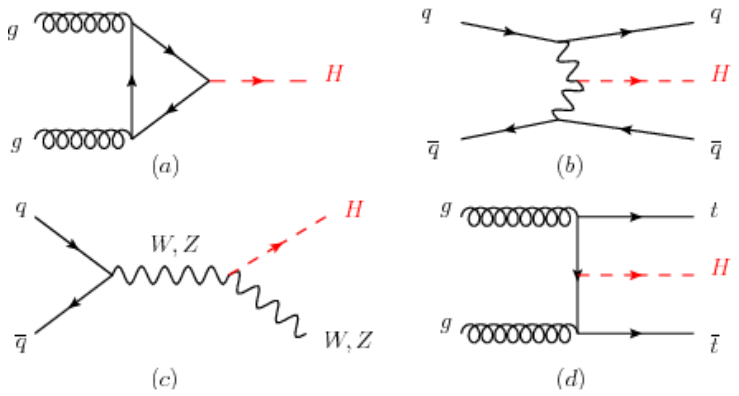

Figure 2.8: Feynman diagrams of the four main production mechanisms of the Higgs boson at the LHC. (a) shows the gluon-gluon fusion, (b) the vector boson fusion, (c) the associated production with a vector boson and (d) the associated production with a top-antitop quark pair.

keeping it invariant under local $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ gauge transformations. Such a strategy is necessary because of the different transformation properties of left- and right-handed chiral states, which was discussed in Section 2.1. Simply adding the Dirac fermion mass term to the Lagrangian, namely

$$-m\overline{\psi}\psi = -m\left(\overline{\psi}_R\psi_L + \overline{\psi}_L\psi_R\right), \tag{2.24}$$

where ψ_L and ψ_R refer to the left-handed and right-handed fermion fields ($\overline{\psi} \equiv \psi^\dagger\gamma^0$), respectively, would break the $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$ gauge symmetry. However, adding a term of the form

$$-g_f\left(\overline{L}\phi R + \overline{R}\phi^\dagger L\right) \tag{2.25}$$

to the Lagrangian preserves the local gauge invariance. In Equation 2.25, g_f is the Yukawa coupling to a fermion f, L is the left-handed SU(2) doublet of the fermion fields, R is the corresponding right-handed singlet and ϕ is the SU(2) doublet containing the two complex scalar fields of the Higgs mechanism. When choosing the unitary gauge and thus spontaneously breaking the symmetry, Equation 2.25 can be written as

$$-\frac{g_f}{\sqrt{2}}v\left(\overline{\psi}_L\psi_R + \overline{\psi}_R\psi_L\right) - \frac{g_f}{\sqrt{2}}h\left(\overline{\psi}_L\psi_R + \overline{\psi}_R\psi_L\right). \tag{2.26}$$

The Yukawa coupling is not predicted by the Higgs mechanism, but can be expressed by

$$g_f = \frac{\sqrt{2}\,m_f}{v} \equiv \frac{m_f}{\sqrt{2}\,m_W}g_W, \tag{2.27}$$

where m_f is the mass of the fermion, v is the vacuum expectation value of the Higgs field, m_W is the mass of the W boson and g_W is the strength of the $\mathrm{SU}(2)_L$ gauge coupling.

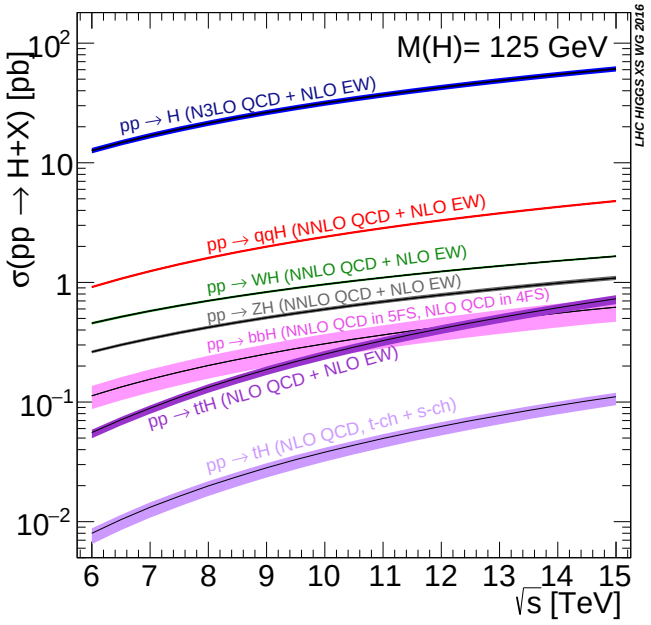

Figure 2.9: The inclusive cross-section of the different Higgs boson production mechanisms at proton-proton colliders as a function of the centre-of-mass energy in the range of $\sqrt{s} = 6 - 15$ TeV [15].

When inserting Equation 2.27 into Equation 2.26, one finds

$$-m_f\overline{\psi}\psi - \frac{m_f}{v}\overline{\psi}\psi h. \tag{2.28}$$

In Equation 2.28, the first term yields the mass of the fermion through its coupling to the Higgs field with a non-zero vacuum expectation value and the second term represents the coupling between the fermion and the Higgs boson. Therefore, Equation 2.27 is a reasonable choice that is consistent with the observed fermion masses in the SM. In addition to this, Equation 2.27 assumes the Yukawa coupling to increase linearly with the fermion mass. Combined measurements of the Higgs coupling to different particles by the ATLAS and CMS experiments are compared to the values expected in the SM in Figure 2.10. When inserting $m_t \approx 173$ GeV and $v \approx 246$ GeV into Equation 2.27, one finds the top Yukawa coupling to be $g_t \approx 1$. This is one of the reasons the top quark could play a special role in electroweak symmetry breaking.

Up until today, most of the previous measurements of the Yukawa couplings, such as those shown in Figure 2.10, were indirect measurements. For example, Figure 2.3 illustrates that the Higgs boson can decay into massless particles, i.e. gluons and photons, through virtual quantum loops involving top quarks and W bosons. By measuring these decay rates, it is possible to infer constraints on the top Yukawa coupling, but only when assuming no additional particles outside of the SM. Additionally, the gluon-gluon

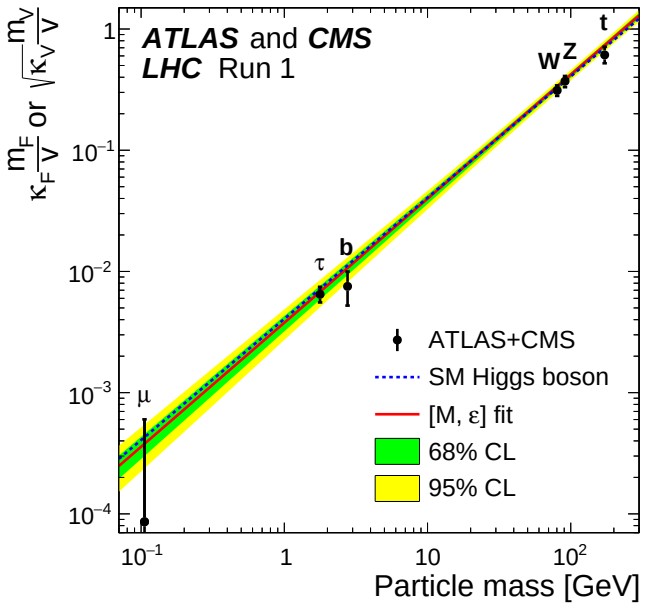

Figure 2.10: The combined measurement of the Higgs coupling to different particles by ATLAS and CMS as a function of the particle mass [21]. The measurements use data recorded at $\sqrt{s} = 7$ TeV and 8 TeV during Run 1 of the LHC physics programme. The dashed line is the prediction based on the SM Higgs boson with $m_H = 125$ GeV, the red line is the best fit with all data. The green and yellow uncertainty bands represent the 68% and 95% confidence level intervals, respectively.

fusion process (a) in Figure 2.8 gives an additional indirect measurement for the top Yukawa coupling, again assuming no additional BSM particles. The advantage of a direct measurement, such as measuring the $t\bar{t}H$ process, over an indirect measurement is its independence of such models and assumptions. Thus, any observed deviation from the SM expectation is a clear indication of new physics and results in stronger constraints on BSM models.

The $t\bar{t}H$ analysis presented in Chapter 7 and beyond is specific in the decay mode of the Higgs boson, however. This choice has several reasons. Firstly, as shown in Figure 2.4, the highest decay branching ratio of the Higgs boson at $m_H = 125$ GeV is into a pair of bottom quarks [3]. That means there are more $t\bar{t}H$ events in the recorded data where the Higgs boson decays into a $b\bar{b}$ pair, labelled $t\bar{t}H(H \to b\bar{b})$ in the remainder of this thesis, than any other decay channel. Secondly, by probing this decay the Yukawa coupling of the Higgs boson to the bottom quark can be measured simultaneously to the top Yukawa coupling, which represents an additional test of the SM.
At first, it seems reasonable to suggest measuring the bottom Yukawa coupling via an-

other Higgs production mechanism, for example the gluon-gluon fusion. However, this process simply involves a pair of b quarks in the final state at a very low cross-section. The simple production of a $b\bar{b}$ pair at the LHC via the strong interaction has a cross-section which is many orders of magnitude larger and thus, the $H \to b\bar{b}$ search requires a very sophisticated analysis strategy [3, 22]. On the other hand, when comparing the cross-section of $t\bar{t}H(H \to b\bar{b})$ to its main background process $t\bar{t} + b\bar{b}$, the difference in cross-section is significantly smaller, namely only roughly three orders of magnitude [22]. Hence, the $t\bar{t}H(H \to b\bar{b})$ measurement also provides a meaningful opportunity to directly measure the bottom Yukawa coupling at the LHC.

In this chapter, the SM was introduced, the special properties of top and bottom quarks and the Higgs boson were discussed and the measurement of the $t\bar{t}H(H \to b\bar{b})$ process was motivated. The following chapter describes the experimental setup used to collect the data for the analyses presented in Chapters 6 and 7.

CHAPTER 3

Experimental setup

This chapter depicts the experimental setup used to record the data that is analysed in the search for the $t\bar{t}H(H \to b\bar{b})$ process presented in Chapter 7. Since the analysis involves the production of three of the heaviest SM particles, namely a top-antitop quark pair as well as a Higgs boson, the energy necessary to produce such events is very high, namely at least the sum of their rest masses: $2 \cdot 173$ GeV $+ 125$ GeV $= 471$ GeV. With current technologies, such energies can only be achieved by a certain type of particle collider experiments. Lepton colliders, for example the Large Electron-Positron Collider (LEP) at CERN, could not achieve this energy [23].
Currently, such energies can only be achieved with synchrotron colliders, for which the collision of hadrons is the most reasonable. Synchrotron colliders accelerate particles on a circular orbit and thus can increase the kinetic energy of the particles with every orbit. The downside to this strategy is the loss of energy through radiation caused by the electrically charged particle on the circular orbit. This radiation is called synchrotron radiation and the loss of energy ΔE is proportional to

$$\Delta E \propto \frac{E^4}{m^4 R}, \tag{3.1}$$

where E and m are the energy and mass of the particle, respectively, and R is the radius of the circular orbit [7]. Since the electron is about 2000 times lighter than the proton, its energy loss per orbit is $\approx 2000^4$ times larger. For any given accelerator with fixed radius R, the energy loss of electrons or positrons thus quickly reaches a critical point that prevents such experiments achieving the energy levels mentioned above. Hadron colliders, on the other hand, are able to produce collisions with a centre-of-mass energy of several TeV.

The TEVATRON at FERMILAB was the first such hadron collider able to reach these energy levels by colliding protons with antiprotons at a centre-of-mass energy of $\sqrt{s} = 1.96$ TeV.

This energy was sufficient such that the top quark was discovered in 1995 by the DØ and CDF experiments [18, 19]. However, the energy as well as the rate at which the particle collisions were recorded did not suffice to discover the Higgs boson. Despite reaching a sufficiently large centre-of-mass energy to produce Higgs bosons, the DØ and CDF experiments did not have enough data to obtain a statistically significant result. Thus, any potential search for the $t\bar{t}H$ process was outside the realm of possibilities at the TEVATRON.
The LHC, on the other hand, was constructed in such a way that it provided both, a higher energy as well as a higher rate of collisions that could be recorded. This allowed the ATLAS and CMS experiments to discover the Higgs boson and focus on even more challenging searches such as the $t\bar{t}H$ process. The design of the LHC and the ATLAS experiment is the topic of the following sections.

3.1 The LHC

The Large Hadron Collider [24] (LHC) is a machine that accelerates hadrons on a circular orbit and eventually collides them at four specific interaction points of this orbit. This ring structure has a circumference of approximately 27 km and is located in depths between 50 and 175 m below the ground at CERN, which is at the France-Switzerland border near Geneva. The depth of the tunnel is necessary to shield the experiments from cosmic radiation. The four interaction points represent the centres of the detectors of the four main experiments at the LHC, namely ATLAS, CMS, ALICE and LHCB. An illustration of the ring structure, its dimensions and the four interaction points is shown in Figure 3.1. While ATLAS and CMS are multi-purpose detectors, ALICE specialises on heavy ion collisions and LHCB focuses on bottom quark physics. The LHC, therefore, mainly accelerates protons via cavities of 30 MV/m gradient, bends them onto circular trajectories with strong magnetic fields and eventually collides the protons after they have reached their target energy. The magnetic fields are generated through 1232 dipole magnets which create the circular trajectory; the proton beams are focused via 392 quadrupole magnets; and small imperfections in the field geometry are corrected by higher multipole order magnets, totalling over 10000 superconducting magnets. The centre-of-mass energy of the proton-proton collisions was $\sqrt{s} = 7$ TeV in 2010 and 2011, 8 TeV in 2012, comprising Run 1, and 13 TeV during 2015-2018, comprising Run 2. Additionally, in Run 1, the proton-proton collision rate was 50 ns [24], and it was 25 ns in Run 2.
The acceleration of protons at the LHC is performed in multiple steps by different systems [24]. Firstly, the protons originate from hydrogen tanks and are accelerated to 50 MeV by the linear particle accelerator LINAC 2. They are then fed into the Proton Synchrotron Booster, accelerated to 1.4 GeV and subsequently injected into the Proton Synchrotron. There, the protons are accelerated to 26 GeV and transferred to the Super Proton Synchrotron where they are accelerated to an energy of 450 GeV. It is only after these initial steps that the protons are finally injected into the LHC to be accelerated to their target energy mentioned above [24].

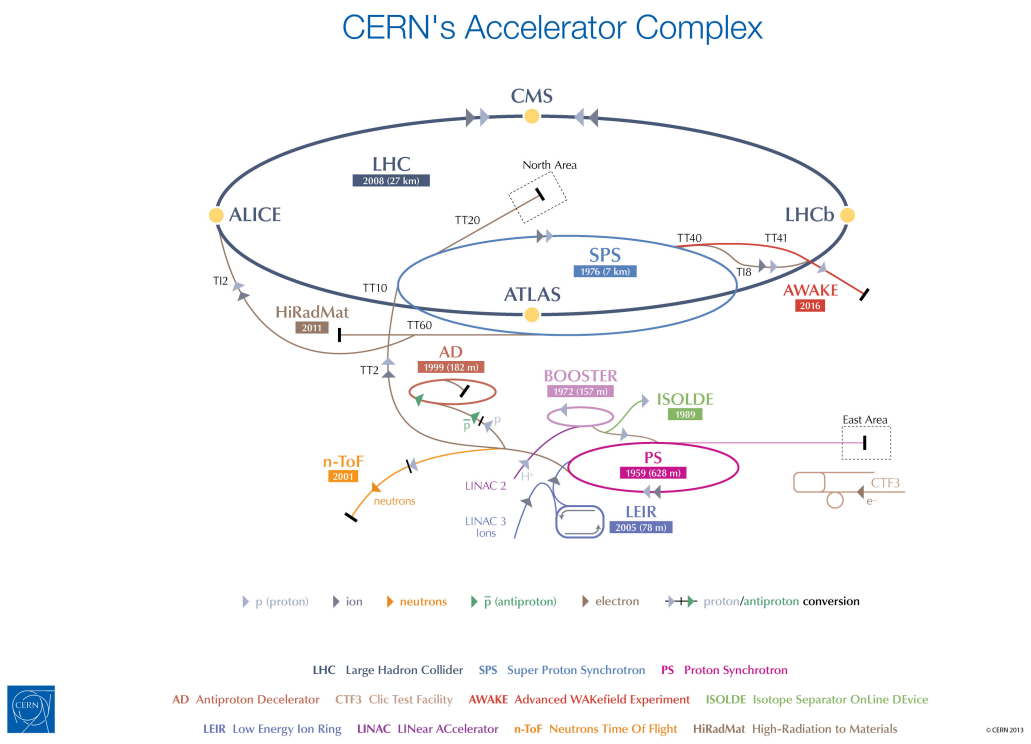

Figure 3.1: A sketch of the CERN accelerator complex at the France-Switzerland border near Geneva. Shown are the four interaction points of the LHC which mark the four main experiments at the LHC, namely ATLAS, CMS, ALICE and LHCB. Also shown are some of the other experiments at CERN as well as additional structures that pre-accelerate and feed the protons into the LHC tunnel. © 2013-2019 CERN.

Aside from this, the LHC also has a programme to accelerate and study heavy ions, mostly in lead-lead collisions. This topic, however, is outside the frame of this dissertation.

A crucial quantity to describe the amount of data i.e. particle collisions produced by the LHC is the so-called luminosity $\mathcal{L}$ [25]. The luminosity is the proportionality factor between the number of events per second $\mathrm{d}N/\mathrm{d}t$, also called event rate, and the cross-section σ. Thus, it can be written as

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \mathcal{L} \cdot \sigma. \tag{3.2}$$

In the case of the LHC, σ represents the total cross-section of proton-proton collisions, denoted as σ_{pp}. When considering the approximation of Gaussian beams colliding head-

on, one finds the following expression for the luminosity at the LHC [25]:

$$\mathcal{L} = \frac{f \cdot N_b \cdot N_1 N_2}{4\pi\sigma_x\sigma_y}. \tag{3.3}$$

Here, $\sigma_x = \sigma_y \approx 17\mu$m represent the width of the Gaussian beams in the x and y directions, respectively. $N_b \approx 2800$ is the number of proton bunches within the beams, $N_1 = N_2 \approx 1.1 \cdot 10^{11}$ the number of protons within the bunches of the respective beam and $f \approx 11245\ \mathrm{s}^{-1}$ describes the frequency with which the proton bunches travel through the LHC pipes. Inserting these values into Equation 3.3 gives a luminosity of $\approx 10^{34}\ \mathrm{cm}^{-2}\mathrm{s}^{-1}$ [24].

The luminosity introduced above is more specifically called the instantaneous luminosity. To determine the actual number of events that have been produced at the LHC, Equation 3.2 has to be integrated over the period of time during which the collisions occur, such that

$$N = \int \mathcal{L}\,\mathrm{d}t \cdot \sigma_{pp} \equiv \mathcal{L}_{\mathrm{int}} \cdot \sigma_{pp}. \tag{3.4}$$

$\mathcal{L}_{\mathrm{int}}$ is called the integrated luminosity and has to be measured precisely to estimate the amount of data. This is because the instantaneous luminosity decreases with passing time as the LHC is not refilled with new protons during a collision run.

The following section will lay out the design of the ATLAS detector at the LHC. Its individual detector parts are discussed as well as the signatures that different particles produce in each part which motivates their reconstruction strategy in data analysis.

3.2 The ATLAS experiment

The ATLAS experiment [26] is one of the two great multi-purpose experiments at the LHC, the other being CMS [27]. The idea behind these experiments, which operate at particle energies never achieved before, are diverse. The most important goal was to find the predicted Higgs boson, which both experiments did in 2012 [1,2]. Other goals include SM precision measurements, especially those involving top quarks and the electroweak symmetry breaking sector; finding or observing rare but expected SM processes such as $t\bar{t}H$; and, finally, searching for new physics beyond the SM such as Supersymmetry or dark matter.

The ATLAS detector consists of multiple different layers and each part serves a different purpose. Combining all parts of this onion-like structure allows different particles to be distinguished, because each expected elementary particle produces a different signature in the detector. The different layers, from the innermost to the outermost, are as follows [28, 29]:

- the inner detector (ID) which comprises the pixel detector (PD), the semi-conductor tracker (SCT), and the transition radiation tracker (TRT);
- the electromagnetic calorimeter (ECAL);
- the hadronic calorimeter (HCAL); and
- the muon spectrometers (MS).

In addition to this, the ID and MS are submerged in a strong magnetic field of about 4 T generated by a system of a central solenoid and three air-core toroids around the ID. A sketch of the ATLAS detector is shown in Figure 3.2.

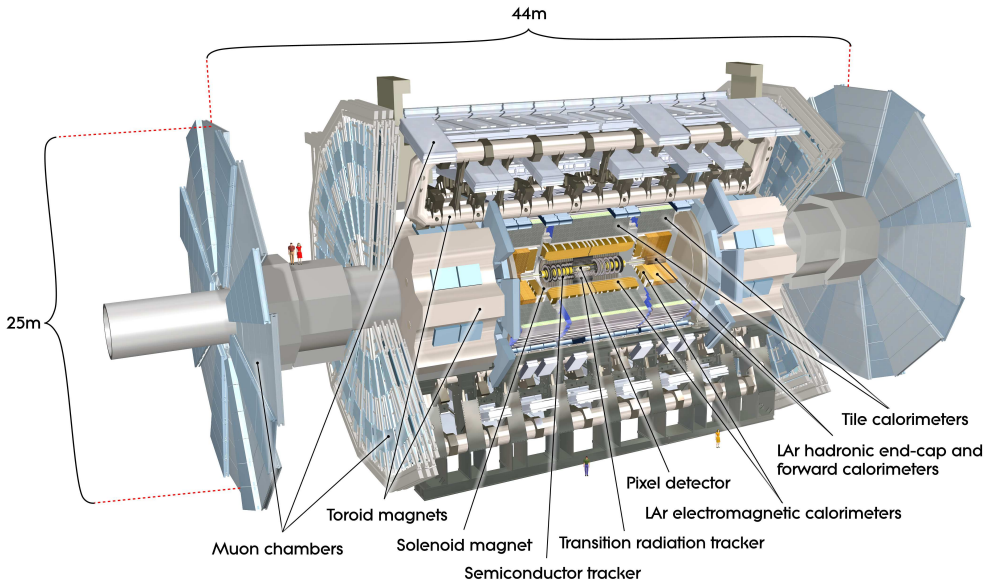

Figure 3.2: The different parts of the ATLAS detector at the LHC. © 2008-2019 CERN

The different sectors of the detector work in significantly different ways and complement each other. Before going into more details about each part, it is important to specify the coordinate system as well as important quantities used throughout the discussion [28].

First of all, the direction of the proton beams defines the z-axis in a right-handed coordinate system, while the $x-y$ plane is perpendicular to it. All transverse quantities considered in this thesis, such as the transverse momentum p_{T}, the transverse energy E_{T} as well as the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ or momentum $p_{\mathrm{T}}^{\mathrm{miss}}$ are defined in this $x-y$ plane. The positive x-axis points from the interaction point to the centre of the LHC ring and the positive y-axis points upwards. Additionally, the azimuthal angle ϕ and the polar angle θ are used to describe the position of particles in the detector. The angle around the beam axis is ϕ and θ is the angle from the beam axis. Given θ, one can define the so-called pseudorapidity $\eta = -\ln\tan(\theta/2)$. With this pseudorapidity, one

can define a very helpful variable, namely the distance in the pseudorapidity-azimuthal angle space $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.
Two more variables are crucial for the reconstruction of tracks in the ID, namely the transverse and the longitudinal impact parameters (IPs). The transverse IP d_0 of a track defines its transverse distance to the beam axis at the point of closest approach and the longitudinal IP z_0 defines its z position at this point.
All these quantities are useful to describe not only the area the detector parts are able to cover, but also the position of particles considered in the selection criteria of the analyses presented in Chapters 6 and 7.

The ID consists of the PD, the SCT, and the TRT. These parts are submerged in a magnetic field generated by the central solenoid which has a nominal strength of 2 T and a peak strength of 2.6 T at the superconductor [28]. The concept of the ID is to place semi-conducting material around the interaction point. Charged particles moving through the ID will then produce curved tracks by ionising the material around their trajectories. With this information, charged particles can be distinguished from each other in space and their momentum as well as the sign of their electric charge can be measured. Furthermore, the tracks allow the position of the main interaction point to be inferred, which is called primary vertex. The ID covers a pseudorapidity range of $|\eta| \leq 2.5$ and provides a track momentum resolution of $\sigma_{p_\mathrm{T}}/p_\mathrm{T} = 0.05\% \cdot p_\mathrm{T}[\mathrm{GeV}] \oplus 1\%$ [26].

The initial design of the PD comprises three barrel layers with 1456 pixel modules and three endcap disk layers with 288 modules to provide an environment of high granularity [28]. The insertable B-Layer (IBL) was installed during the first long shutdown (LS1) of the LHC between Run 1 and Run 2 [30, 31]. It was inserted as the innermost piece of the detector at a radius from the beam axis of 3.3 cm and contains 224 modules. It serves as an additional fourth layer in the PD to improve the reconstruction of particle tracks and withstand the increased radiation dose expected by the higher luminosity setup of Run 2 [30, 31]. Since each module holds $\approx 46,000$ pixels (27,000 for the IBL), the PD comprises over $80 \cdot 10^6$ semi-conducting pixel elements of size 50 μm $\times$ 400 μm (50 $\times$ 250 μm for the IBL) [28, 30, 31]. These consist of a silicon chip, which serves as the detecting material, bump-bonded to an electronic readout chip [28].

The SCT is made up of eight layers of silicon microstrips to give additional precision measurements of the $R\phi$ and z coordinates [28]. The layers are arranged into one set of radial strips and one set in which two stereo strips are glued together at an angle of 40 mrad. This improves the resolution of momentum, impact parameter and vertex position and reduces the hit ambiguity.

The TRT, finally, provides additional tracking points, typically 36 per track, to the ID using straw tubes [28]. The straws in the barrel are parallel to the beam direction, while the endcap straws are radially arranged into wheels which are perpendicular to the beam axis. With about 100,000 straws in the barrel and 320,000 straws in the endcaps,

the TRT allows for discrimination between tracking hits and transition radiation hits. In addition, a gas mixture containing Xenon, CO_2 and CF_4 helps to detect transition radiation photons which improves the identification of electrons. The overall setup is by itself radiation hard, but leads to a high counting rate when operating at the LHC luminosity quoted in the previous section.
The design of the ID without the IBL is depicted schematically in Figure 3.3.

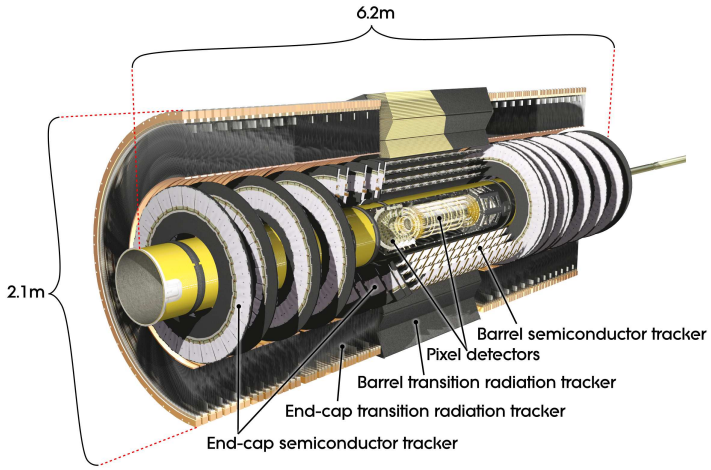

Figure 3.3: A sketch of the ATLAS ID. The PD, SCT and TRT are arranged into different layers of barrels and disks. © 2008-2019 CERN.

The calorimeters of the ATLAS detector are primarily split into two different parts, namely the ECAL and the HCAL [28]. As suggested by their names, they serve different purposes and help to identify different types of particles that are created in the collisions. Both are sampling calorimeters meaning they alternate between a material stimulating particles to deposit their energy into and another material which measures this deposited energy [28]. Both are cooled by cryostats surrounding the ID cavity [28]. The idea behind calorimeters is that particles such as electrons, photons and hadrons move through the active material and produce large showers of particles by interacting with the detector material. In particular, the showers are generated predominantly by pair production and bremsstrahlung of new particles until the energy of the initial particle becomes too small to initiate a new process and it is stopped by the detector material. Thus, the energy of particles can be inferred from the size of the showers i.e. the energy clusters deposited into the calorimeter cells.

The ECAL consists of modules made out of lead and liquid argon (LAr) and has the geometry of an accordion [28]. LAr is the chosen active medium throughout most of the ECAL and HCAL, because it exhibits a good level of intrinsic radiation hardness for its industrial cost compared to other materials. The ECAL is split into a barrel region which is contained by a barrel cryostat and two endcaps which are further divided into two coaxial wheels. The two endcaps as well as the hadronic calorimeters are contained

within two endcap cryostats. The total thickness of the ECAL is greater than 24 radiation lengths (X_0) in the barrel and greater than $26X_0$ in the endcap regions. In the $|\eta| < 1.8$ region, a presampler, built out of a layer of LAr, is used to correct for the energy lost by electrons and photons from unwanted interactions with the detector material. To stress this issue further, the region $1.37 < |\eta| < 1.52$ is not used for precision measurements involving photons in ATLAS, because the amount of material in front of the ECAL is too high. Apart from that, the total η coverage of the ECAL is $|\eta| < 3.2$ and its granularity ranges between $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ and 0.1×0.1, depending on the sampling segments [28]. The energy resolution in the ECAL is $\sigma_E/E = 10\%/\sqrt{E}[\text{GeV}] \oplus 0.7\%$ [26].

The HCAL is divided into three parts. The first two are, similarly to the other detector parts, a barrel and two endcap regions. The third one, however, is an additional forward calorimeter (FCAL), such that the HCAL covers $|\eta| < 4.9$ [28]. The barrel region contains plastic scintillator plates which are embedded in an iron absorber. The endcaps consist of alternating modules of LAr and copper. In the FCAL, very dense modules of LAr are embedded in a matrix of copper for the first segment and tungsten for the other two. This setup was chosen since the radiation dose to the material is higher for increasing η [28]. The total thickness of the HCAL is 11 interaction lengths at $|\eta| = 0$. This is sufficient to contain hadronic showers such that punch-through of particles into the MS is reduced to well below an irreducible level of expected prompt or decay muons [28]. At the same time, this setup provides a good resolution to measure highly energetic jets, as well as missing transverse energy $E_{\text{T}}^{\text{miss}}$. The energy resolution in the barrel and end-cap is $\sigma_E/E = 50\%/\sqrt{E}[\text{GeV}] \oplus 3\%$ and in the FCAL it is $\sigma_E/E = 100\%/\sqrt{E}[\text{GeV}] \oplus 10\%$ [26]. The granularity of the HCAL ranges between $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ and 0.2×0.2 [28].
In Figure 3.4, the ATLAS calorimeters are shown schematically. Altogether, they contain roughly 200,000 readout channels.

In the ECAL, all light particles that participate in the electromagnetic interaction are stimulated to produce particle showers. This means predominantly electrons and photons, because muons and hadrons are too massive to engage in significant pair production and bremsstrahlung processes [7]. On the other hand, all particles with a colour charge are stimulated to produce a shower via the strong interaction in the HCAL. This refers to the hadrons, such as neutrons and π mesons, which are formed by the quarks from the proton-proton collisions.

The MS is the outermost layer of the ATLAS detector and, as the name suggests, crucial for the identification of muons. The MS is submerged in a magnetic field generated through the barrel toroid and the two endcap toroids. The peak magnetic field strengths at their superconductors are 3.9 and 4.1 T, respectively [28]. This magnetic field is essential to bend the trajectories of high energetic muons and covers the range of $|\eta| \leq 2.7$. The MS is complementary to the ID based on the idea that all other SM particles are stopped before reaching the muon system and the muons produce up to three additional tracks which can be matched to the tracks they generated in the ID [28]. A system

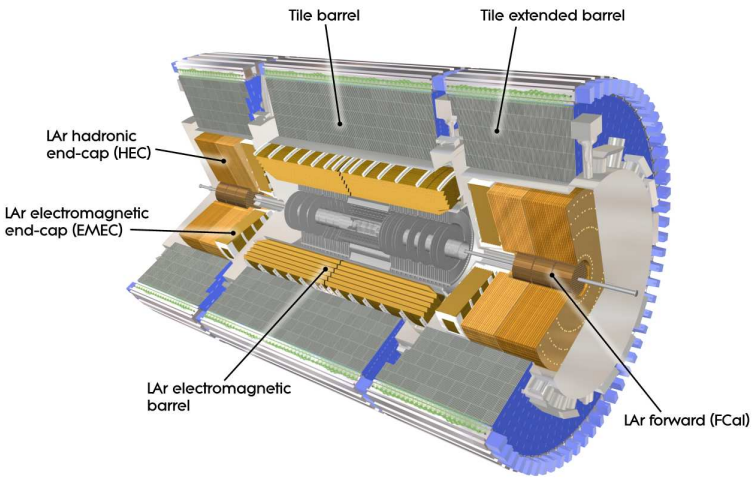

Figure 3.4: An illustration of the ATLAS calorimeters. The ECAL and HCAL are segmented into barrel and endcap regions, similarly to the ID. In addition, ATLAS uses a forward calorimeter to extend the η coverage of the HCAL. © 2008-2019 CERN

of monitored drift tubes and cathode strip chambers with higher granularity is used to perform precision measurements of the track coordinates. The track p_{T} resolution in the MS is $\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}} = 10\%$ for muons with $p_{\mathrm{T}} \approx 1$ TeV [26]. The drift tubes are made out of aluminium, while the cathode strip chambers are multiwire proportional chambers containing 30% Ar, 50% CO_2 and 20% CF_4. In addition to these measurement chambers, the muon system contains trigger chambers, namely resistive plate chambers and thin gap chambers, which are critical for the muon identification, but not described here [28]. A sketch of the ATLAS MS is shown in Figure 3.5.

Further specifications and details of the ATLAS detector, such as the cooling system and trigger chambers, can be found in the references [26, 28–31].

Taking the information of all these detector parts, the ATLAS experiment can reconstruct and identify electrons, muons, τ-leptons, photons and all SM hadrons. The only class of SM particles which cannot be detected are the neutrinos. However, their transverse momenta can be inferred from the total momentum that is missing in the transverse plane of an event. The different signatures of these particle classes are illustrated in Figure 3.6 to highlight that they are distinct from each other.

The ATLAS experiment operates additional forward detectors roughly 25 m away from the main detector. With these structures, the collaboration is able to perform a measurement of the total proton-proton collision cross-section and derive an estimate of the LHC luminosity [32, 33].

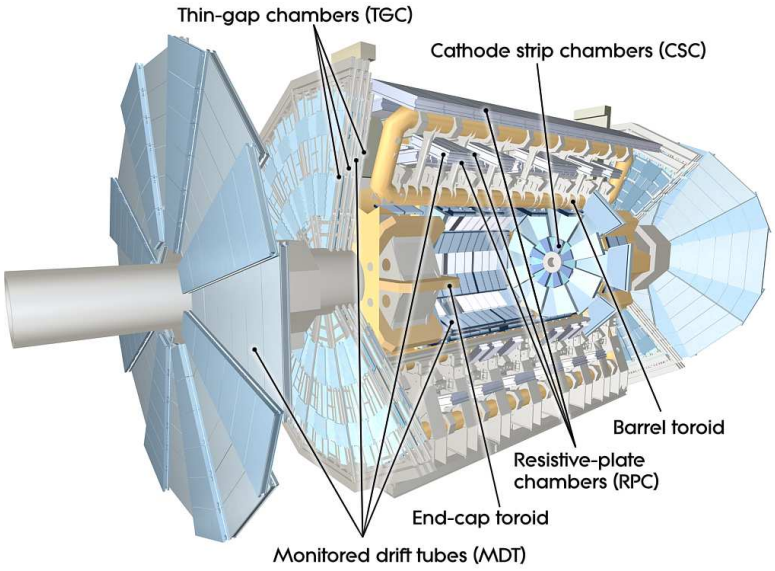

Figure 3.5: The ATLAS muon spectrometer is divided into a barrel and endcap regions, but contains various chambers for precision measurements of track coordinates as well as for the ATLAS trigger system. © 2008-2019 CERN.

As mentioned in the previous section, the crossing interval of proton bunches in the LHC ring during Run 2 is 25 ns. Consequently, when the LHC reaches the peak luminosity of 10^{34} cm^{-2}s^{-1}, the interaction rate approaches $\approx$ 1 GHz with a bunch crossing rate of 40 MHz. Currently, it would be impossible to store the raw data permanently at such a high event rate, because the amount of required disk space would reach 1 Petabyte per second [28]. Therefore, ATLAS applies a trigger system [34,35] to only accept events with particularly interesting features and, thus, decrease the event rate to $\approx$ 1 KHz, which results in a reasonable amount of data to handle and store.
In order to not reject any physics events that hold interesting signatures, such as decays of Higgs bosons or rare processes of the SM and beyond, the trigger system was designed in two stages with different strategies [35]. The first is the so-called level-1 trigger which is based on information from a subset of detectors using reduced granularity to determine Regions-of-Interest (RoIs) in the detector. The trigger chambers of the MS fire when they detect a muon with a p_{T} above a chosen threshold. In the same way, the calorimeter triggers fire when detecting high E_{T} electrons, photons, jets, τ-leptons that decayed into hadrons and/or a large total as well as missing transverse energy. With this, the level-1 trigger reduces the event rate to $\approx$ 100 KHz. The second trigger is the high-level trigger (HLT) which uses the RoIs as input and reduces the event rate to $\approx$ 1 KHz. For this, the HLT uses more sophisticated selection algorithms than a simple p_{T} cut and it has access to the full granularity detector information either in the RoI or the complete event. For example, the HLT can increase the p_{T} threshold of candidate

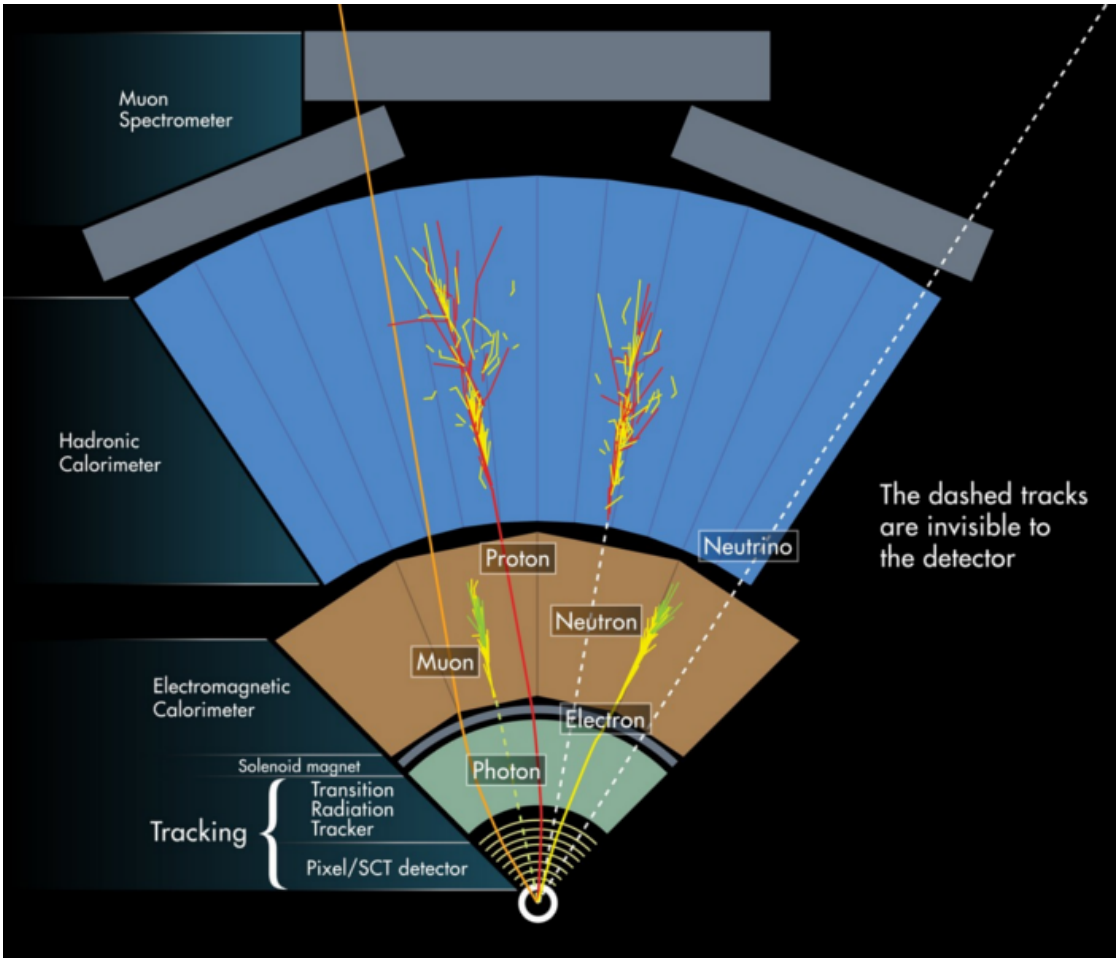

Figure 3.6: This figure shows that each elementary and composite particle, which travels a significant distance through the ATLAS detector, produces a different signature when combining the information from all detector parts. ATLAS Experiment © 2008 CERN.

objects, but also impose requirements on the isolation of muons, electrons and photons. It can further require high p_{T} tracks to match the other signatures of the charged leptons. On the other hand, the HLT can also require specific objects to be present in an event such as a pair of photons which is of high interest in the search for $H \to \gamma\gamma$ decays.

In this dissertation, two analyses will be presented, namely the calibration of jets originating from bottom quarks as well as the search for the Higgs boson in association with a $t\bar{t}$ pair where the Higgs boson decays into a $b\bar{b}$ pair. The final states of these processes may contain any leptons and quarks i.e. jets of the SM and, therefore, the information of all detector parts is required to analyse the data recorded with the ATLAS detector.

Two particularly important aspects are the calibration and reconstruction of objects as well as the Monte Carlo simulation of physics events that are expected to be recorded with the ATLAS detector. These topics will be discussed in the following two chapters, starting with the simulation using Monte Carlo generators.

CHAPTER 4

Physics modelling using the Monte Carlo method

In this chapter, the basic idea of the Monte Carlo (MC) method to simulate particle physics interactions is presented, especially for proton-proton collisions at the LHC and their possible outcomes that are of interest to the ATLAS experiment. The discussion is based primarily on the underlying concept of the MC method at collider experiments described in related standard textbooks such as [7, 9].

There are many mathematical problems which are very difficult if not impossible to solve analytically. For such cases, the MC method offers an alternative, namely numerical approach in which an efficient random number generator (RNG) is used to perform a large number of random experiments corresponding to the mathematical problem. In particle physics, these random experiments correspond to particle interactions which may happen at the boundary towards infra-red (low energy) or ultra-violet (high energy) divergence and thus require special treatment. According to the law of large numbers, if the number of random experiments is large enough, the average outcome over all events will approach the expected one. Furthermore, when computing the integral over a probability density function f, the central limit theorem of statistics states that the mean value of f over these random experiments is an unbiased estimator of the integral.

Since the LHC collides protons i.e. hadrons, the interacting particles are coloured particles that predominantly interact strongly. As discussed in Section 2.4, the Feynman calculus is based on perturbation theory to calculate the cross-section and most physical observables of the process under study. The resulting power series may have non-negligible terms of high order in the strong coupling strength α_S, depending on the energy scale at which the interaction takes place. However, the required CPU time to perform the perturbative computations increases roughly factorially with the order of α_S.

The main interaction is expected to happen at a very high energy scale, since the incoming protons possess an energy of several TeV, which is three orders of magnitude larger than their rest mass [3]. The strong coupling strength decreases with increasing energy and thus, the terms of higher order become less important. In consequence, this main (or hard) interaction can be calculated in a more CPU time efficient way with sufficient precision by only including terms up to a fixed order in α_S, usually the next-to-leading (NLO) or next-to-next-to-leading order (NNLO) with current state-of-the-art MC generators.
On the other hand, the incoming particles as well as the lower energetic coloured particles created by the hard interaction may produce more soft and/or collinear quarks and gluons. Since α_S increases at lower energies, the cross-section to produce these additional particles (or partons according to Feynman) increases over all boundaries and the power series breaks down. This process is called parton shower and is modelled with a different strategy, because higher order terms of the series cannot be neglected any longer. Thus, during the parton shower evolution, only the essential terms of the perturbation theory are taken into account, but to all orders of α_S. This is no longer an analytical treatment, but sufficient to achieve an approximate result and can be performed conveniently by MC generators. Figure 4.1 shows the simulation of an exemplary proton-proton collision including the hard interaction, parton shower and further parts of the process which will be described in the following sections.
It is critical to discuss the simulation of particle physics events with MC generators, because different generators exist that each have a different numerical implementation of simulating such events. Since these numerical models are based on parameters which cannot be derived from first principles or perturbation theory, as explained above, they all hold a degree of validity. Effectively, this means that the modelling of physical observables by different generators should be compared and evaluated. These comparisons may result in significantly different predictions which manifest as sources of large systematic uncertainties in physics analyses. This affects the total uncertainty and therefore the sensitivity of the studies presented in Chapters 6 and 7.

4.1 Matrix element generation

The matrix element generation is the first step during the MC simulation of physics events, in this case the collision of high energetic-coloured particles. It is referred to as the hard process or primary interaction. The goal of the simulation is to determine the cross-section of this particle interaction or scattering as well as the kinematic properties of the final state partons. This calculation is performed by integrating a probability density, which is represented by an element of the corresponding scattering matrix, over the physical phase space of the event according to the Feynman rules. During this calculation, a large number of events is generated and each event represents a sampling point in the multi-dimensional phase space spanned by the simulated final state particles.

Every strong interaction of such an event enters the cross-section formula as an addi-

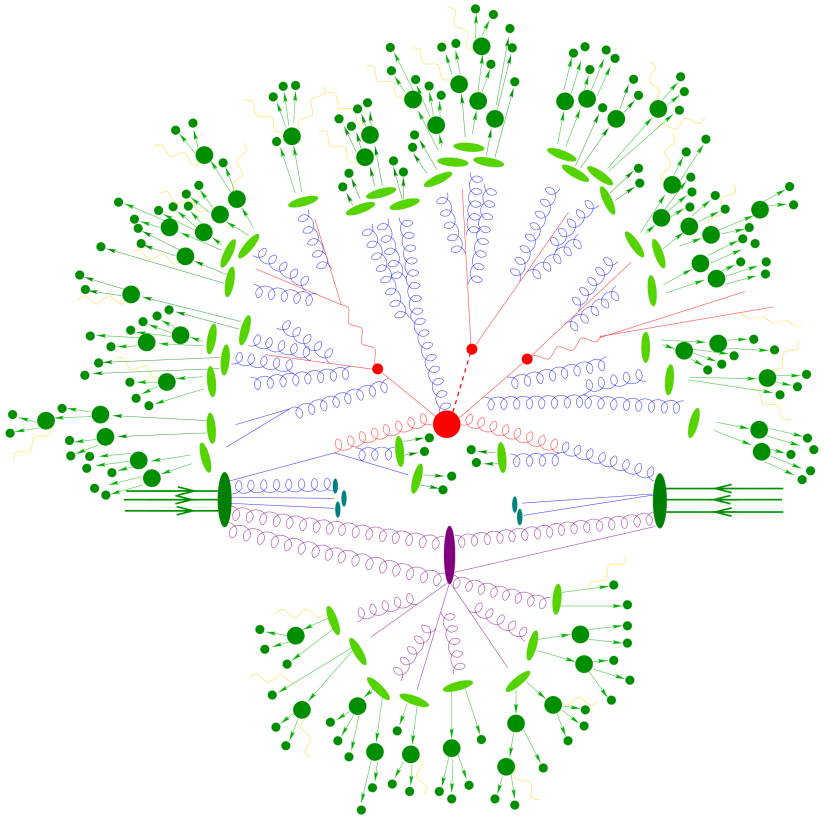

Figure 4.1: Sketch of a proton-proton collision at the LHC, with the hard process in red, underlying event in purple, radiation and splitting in blue and hadronisation in green [36].

tional factor of α_S. As discussed before, the hard process of proton-proton collisions at the LHC is usually generated at NLO; some complex processes are generated only at LO and some essential SM processes even at NNLO. The LO or tree level process refers to the primary interaction without any such additional strong interactions. Therefore, at LO, only the final state particles of the process are considered and they are decayed into the leptons and hadrons we expect to measure in the detector (neutrinos as $p_{\mathrm{T}}^{\mathrm{miss}}$), namely by using branching ratios, for example from the Particle Data Group [3]. Such an LO diagram is shown, for example, in Figure 2.7 for a semileptonic $t\bar{t}$ event. By implementing the various decay modes of the W bosons into the simulation, the branching fractions of the three $t\bar{t}$ decay channels shown in Figure 2.6 can be computed.

For any given order of α_S considered in the hard process, all possible Feynman diagrams up to that fixed order enter the cross-section calculation. More complex diagrams are neglected as an approximation of small values of α_S, typically around 0.11-0.12, as indicated in Equations 2.19 and 2.20. An NLO calculation, therefore, includes all LO

diagrams in addition to the LO diagrams with either a quantum loop of gluons or quarks, an additional emission of a gluon or a splitting of a gluon into a quark-antiquark pair. The NNLO calculation includes all NLO diagrams above plus LO diagrams with any two of the additional strong interactions combined, which shows the combinatorial problem of the simulation strategy leading to factorially increasing CPU times required.

To arrive at the total cross-section of the hard process, the MC generator determines the convolution of differential cross-sections and integrates them over the full phase space. While performing this integration, the four-momenta of all final state particles are calculated as well as the colour flow of coloured particles. The colour flow refers to the propagation of the initial colour that the incoming partons hold to the final state particles. Both the four-momenta and the colour flows are used as input to the parton shower evolution afterwards, because they are used to determine the starting energy scale Q_0^2 as well as possible parton branchings in the QCD shower.
In Equation 2.19, α_S is a function of the scale Q^2, which is called renormalisation scale μ_R and refers to the scale at which α_S is evaluated. In the hard process it may be, for example, given by

$$\mu_R = \sqrt[n]{\sum_i^n m_i^2}, \tag{4.1}$$

where m_i represents the invariant mass of particle i and the sum runs over all final state particles. From that point on, the scale Q^2 is then evolved for each parton branching using the so-called DGLAP equations of QCD [37], a process further described in the following section.

Another important aspect of the matrix element is the parton distribution function (PDF) of the proton. In the general case, that means when considering low energetic protons, they contain two up quarks and one down quark, the valence quarks. Thus, the hard interaction of two such protons would involve mostly these valence quarks. However, these quarks interact with each other and do so primarily via the strong interaction in which they exchange virtual gluons. In turn, these gluons may temporarily split into quark-antiquark pairs of the second or third generation such as strange, charm and bottom quarks, the so-called sea quarks. These processes are enhanced the more energy the proton holds. The consequence is that, even though two protons are collided at the LHC, the partons that initiate the primary interaction may be gluons or quarks of different flavours as well as antiquarks. At this point, the proton PDF f allows the probability to be determined that a parton with a certain flavour and momentum fraction x will participate in the interaction. In Figure 4.2, the momentum density $x \cdot f$ is shown as a function of x for the valence quarks, sea quarks and gluons. This product is also dependent on the scale Q^2 which is called factorisation scale μ_F. The different curves in Figure 4.2 are predictions using an example value of $Q^2 = 10$ GeV^2 and $Q^2 = 10^4$ GeV^2 at NLO precision [38].

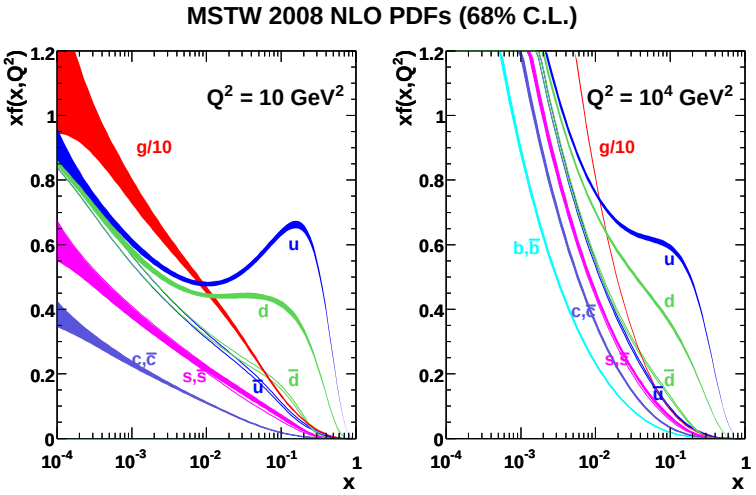

Figure 4.2: The momentum density xf of the parton is plotted as a function of its momentum fraction x. Here, the prediction of f is taken from [38] at a momentum transfer of $Q^2 = 10$ GeV^2 (left) and $Q^2 = 10^4$ GeV^2 (right) at NLO precision.

What can be seen is that the valence quarks typically carry large momentum fractions. On the other hand, sea quarks and especially gluons carry low momentum fractions and hence tend to dominate the interaction for smaller values of x. This is emphasised by the fact that about 50% of the total momentum of the proton is carried by gluons.

Two further aspects that arise at the start of proton-proton collisions are the multiple parton interactions (MPI) and the initial state radiation (ISR). MPI refers to the case that the beam remnants i.e. the partons of the protons that do not participate in the primary interaction, may still interact with each other and interfere with the hard process. On the other hand, ISR is simply the emission of additional gluons by the two incoming partons. Both topics are not part of the matrix element calculation, however, and are therefore discussed in later sections.

Finally, after the matrix element is computed, the hard process has to be matched to the parton shower evolution that follows. This matching is critical, because the assumptions about model parameters in the calculation, for example the choice of the functional form of μ_R and μ_F, propagate into the simulation of the shower evolution. There is a (limited) freedom of choice for these, since they cannot be derived from first principles and perturbation theory calculations are only available as approximations to NLO or NNLO in α_S. This freedom of choice leads to systematic uncertainties in the measurements, which will be described in more detail in Sections 6.5 and 10.3. However, the exact procedure of matching is out of scope of this dissertation and will not be

discussed.
In Chapter 6, the considered signal process is $pp \to t\bar{t} \to b\bar{b}e^{\pm}\mu^{\mp}\nu\bar{\nu}$, while in Chapter 7 it is $pp \to t\bar{t}H \to b\bar{b}b\bar{b}l^{\pm}l^{\mp}\nu\bar{\nu}$ or $b\bar{b}b\bar{b}q\bar{q}'l^{+}\nu(l^{-}\bar{\nu})$. The modelling specifications of the signal as well as background processes will be the topic in those respective chapters.

4.2 Parton shower evolution

In principle, the goal of the parton shower evolution is to add higher order corrections to the perturbative treatment of the hard process. During the hard process, the protons transfer a large momentum to the initial partons which in turn transfer their momentum to the final state partons. Therefore, the initial and final state partons possess a high kinetic energy and are strongly accelerated. And just as accelerated particles with an electric charge radiate off photons and produce Bremsstrahlung, accelerated coloured particles emit gluons. But in contrast to the QED radiation, the gluons themselves carry a colour charge and thus initiate new strong interactions and altogether produce a parton shower.

The modelling of the shower evolution begins at the energy scale Q_0^2, given by the hard process, which represents the upper limit of a possible squared momentum transfer t. The shower is then evolved downwards to an infra-red cut-off scale t_0, usually of the order of 1 GeV^2. This cut-off scale is necessary as the coupling strength α_S would grow too large and perturbation theory would break down. Thus, further emissions with $t < t_0$ involve partons which are too low energetic and thus declared unresolvable. The same holds for collinear emissions or splittings i.e. parton branchings under a very small angle which are also included by this cut-off scale. Both scenarios are crucial for the shower evolution and are the basis for the discussion below.
During the shower evolution, the MC generator iterates over each final state particle and computes possible parton branchings with an RNG until a step is reached where all momentum transfers are below the cut-off scale and no further branchings occur. At this point, the MC generator starts the hadronisation process for all coloured partons which is described in the next section.

The MC generator must determine the branching probabilities for parton emissions which can be derived as follows. Consider the differential cross-section for an n-parton final state

$$\mathrm{d}\sigma_n = \mathcal{F}|\mathcal{M}_n|^2\mathrm{d}\Phi_n, \tag{4.2}$$

where $\mathcal{F}$ is the initial-state flux factor, $\mathcal{M}$ is the matrix element and Φ_n is the final-state phase space for the n-parton process. Using the small-angle approximation, the infinitesimal $n+1$-parton phase space element can be expressed as

$$\mathrm{d}\Phi_{n+1} = \mathrm{d}\Phi_n \frac{1}{4(2\pi)^3}\mathrm{d}t\mathrm{d}z\mathrm{d}\phi, \tag{4.3}$$

where z is the momentum fraction of the emitted parton and ϕ is the azimuthal angle of the parton branching. The typical order of precision during parton shower evolutions with MC generators is LO. In this approximation, the differential $n+1$-parton cross section can be factorised such that, after integrating over ϕ, it is given by:

$$\mathrm{d}\sigma_{n+1} \approx \mathrm{d}\sigma_n \frac{\mathrm{d}t}{t} \mathrm{d}z \frac{\alpha_S}{2\pi} \hat{P}(z), \tag{4.4}$$

where $\hat{P}(z)$ is the so-called splitting function which represents the probability to emit another parton with momentum fraction z.
Equation 4.4 can be used to construct a more useful variable, namely the Sudakov form factor

$$\Delta(t) \equiv \exp\left[-\int_{t_0}^{t} \frac{\mathrm{d}t'}{t'} \int \mathrm{d}z \frac{\alpha_S}{\pi} \hat{P}(z) \right]. \tag{4.5}$$

The Sudakov form factor gives the probability to evolve from the cut-off scale t_0 upwards to t without a resolvable parton branching. More importantly, $\Delta(t_1)/\Delta(t_2)$ can be interpreted as the probability to evolve from t_1 downwards to t_2 without a resolvable parton branching. Based on this, the MC generator tries to solve

$$\Delta(t_1)/\Delta(t_2) = \mathcal{R}, \tag{4.6}$$

where $\mathcal{R}$ is a random variable uniformly distributed in the interval [0,1]. Thus, for a given t_1, the value of t_2 can be generated with the correct probability distribution. If $t_2 < t_0$, no further branching occurs. Otherwise, the momentum x_2 or rather the momentum fraction $z = x_2/x_1$ must be generated for the next parton branching. This is done by solving the equation

$$\int_{x_2/x_1}^{\epsilon} \mathrm{d}z \frac{\alpha_S}{2\pi} P(z) = \mathcal{R}' \int_{1-\epsilon}^{\epsilon} \mathrm{d}z \frac{\alpha_S}{2\pi} P(z), \tag{4.7}$$

where $\mathcal{R}'$ is another random number in [0,1] and ϵ is the infra-red cut-off for resolvable branchings in the z-space.
The above case describes the so-called forward evolution which is useful to simulate radiation from final state particles. This is because an upper limit on the energy scale is given by the final state particles of the hard process and during each step, the emitting partons evolve to a lower scale until all outgoing partons generate no further resolvable branching and the shower stops. On the other hand, radiation by the initial state partons is limited from below by the energy scale of the final state. For this type of radiation, another approach is commonly used in MC generators, namely the backward evolution. Here, the partons initiating the hard process evolve from the scale of the final state backwards to the hard interaction scale $s' = x_1 x_2 s$, where s is the centre-of-mass energy square of the collision and x_1 and x_2 are the respective momentum fractions of the partons. However, the backward evolution process cannot be performed by simply reversing the forward evolution strategy and using the same formulae. Since the forward

evolution approach is sufficient to understand the principles of the parton shower evolution, the backward evolution is not further discussed in this thesis.

Finally, during the parton shower evolution there is a freedom of choice of which parton branchings to generate first. Two showering strategies are commonly implemented in MC generators used for the analyses presented in Chapters 6 and 7, namely the coherent showering and the dipole showering. The coherent shower is implemented in the MC generator HERWIG [39, 40], while the dipole shower approach is used in PYTHIA [41, 42] and SHERPA [36]. In the coherent shower, parton branchings with a small opening angle are generated first, hence it is also called angular-ordered parton shower. In the dipole shower, the emission of gluons is not generated by parton splittings, but instead according to the dipole radiation pattern of a pair of partons. Discussing the details of these two showering algorithms is considered outside the scope of this thesis.

After the parton shower evolution stops, all generated final state partons are used as input to a hadronisation algorithm which aims at turning all coloured particles into observable hadrons. Two commonly used models are described in the next section.

4.3 Hadronisation

In this section, two different hadronisation models are presented. The first is the string model which is implemented in the MC generator PYTHIA and the second is the cluster model used in the generators HERWIG and SHERPA. At the end of the parton shower evolution, there are many partons with a virtual mass-squared in the order of the cut-off value t_0 introduced in the previous section. However, since this cut-off value is generally small, α_S becomes too large during the following processes and thus, the procedure of hadronisation cannot be performed in a perturbative way. Both hadronisation models aim at converting these partons into the hadrons that can be observed in physics experiments.

4.3.1 String model

To illustrate the QCD dynamics of the string model, the simple process $e^+e^- \rightarrow q\bar{q}$ is considered which is depicted in Figure 4.3. In the centre-of-mass frame, the quark and antiquark move apart in opposite directions. The principle of confinement in QCD was explained in Section 2.4 and applies in this scenario as well, since both partons interact strongly with each other over increasing distances. The increase of the strong coupling at growing distances and the gluon self-interactions lead to the property of confinement, but it can also be explained by an additional term in the strong potential $-kr$, where r is the distance and k is a constant, which results in a distance independent force of attraction. While the quark and antiquark move away from each other, their strong interaction is represented by a stretching gluon string with thickness $\approx$ 1 fm which connects the partons while respecting their colour charge. The potential energy

of the system increases while the kinetic energy decreases. At some point, the potential energy is sufficiently high that it is favourable for the gluon string to break and create a new quark-antiquark pair out of the vacuum with new corresponding gluon strings and appropriate colour connections to the initial antiquark and quark, respectively. As long as the kinetic energy is high enough, these new quark systems continue to move apart from each other and thus, the newly created strings stretch again until they break to create new quark-antiquark pairs and so on. This treatment continues until all the energy has been converted into quark-antiquark pairs which are connected by short string segments. These can be identified as colour neutral hadrons that are expected from the property of confinement.

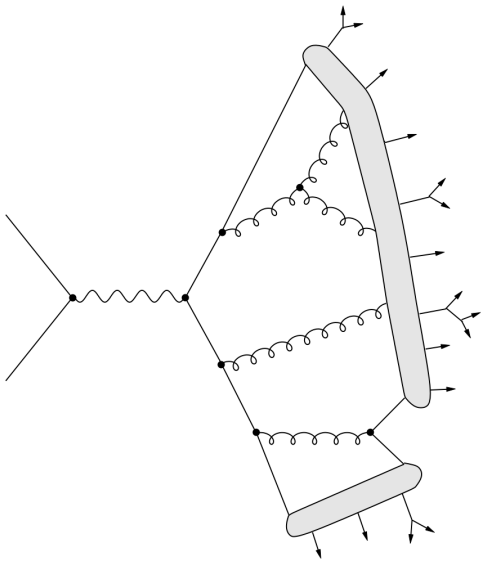

Figure 4.3: The string model illustrated in the process $e^+e^- \to q\bar{q}$ [9].

4.3.2 Cluster model

In contrast to the string model, which connects quarks and gluons via strings while respecting their colours, the cluster model groups the partons directly into colourless clusters. This strategy is motivated from the principle of preconfinement, which implies that pairs of colour-connected neighbouring partons have an asymptotic mass distribution which falls rapidly at high masses and is asymptotically Q^2-independent and universal [9]. Consider a non-perturbative splitting of gluons into $q\bar{q}$ pairs. In this case, the neighbouring quarks and antiquarks can be combined into colourless groups and their mass spectrum is universal, steeply falling at high masses and depends on the cut-off scale t_0. Given a typical value of $t_0 \approx 1$ GeV^2, the clusters have an average mass of the order of 3 GeV and can be decayed into hadrons by using an isotropic quasi-two-body phase space model. A graphical example of the cluster model is shown for the $e^+e^- \to q\bar{q}$ process in Figure 4.4.

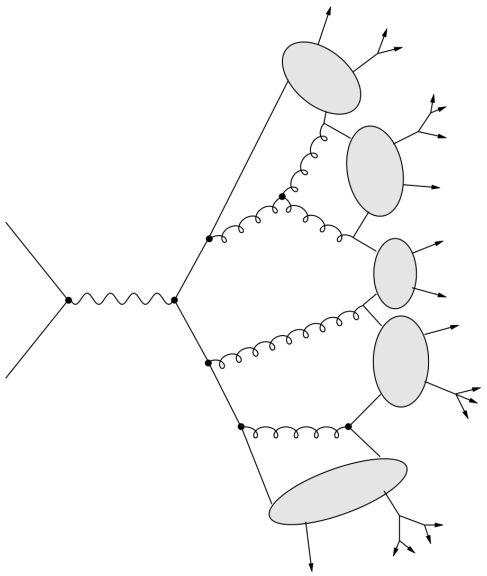

Figure 4.4: The cluster model shown with the process $e^+e^- \rightarrow q\bar{q}$ [9].

4.4 Underlying Event

As mentioned in Section 4.1, during a proton-proton collision, the protons contain more partons than the two which are chosen to initiate the hard process. These beam remnants i.e. remaining partons may also interact with each other and thus interfere with the hard process. This is referred to as the underlying event (UE) which results in the production of additional hadrons that enter the final state of hard processes. This additional activity in the hard process is also called multiple parton interactions (MPI).

The idea of the UE is that the elastic gluon-gluon scattering has the largest cross-section of all possible proton-proton interactions. It is even larger than the total proton-proton scattering cross-section itself which means the average number of gluon-gluon interactions per proton-proton collision is larger than one. This process diverges at small squared momentum transfers t and thus is only integrated up to a cut-off t_{min}. All possible combinations of elastic scatterings of quarks, antiquarks and gluons must be taken into account which all show the same divergence behaviour at LO. In addition, the probability of more and significant MPI correlates with the presence of a hard interaction and therefore should not be neglected.

The simulation of multiple parton interactions is built into all three previously mentioned MC event generators Herwig, Pythia and Sherpa. However, each generator uses a different approach which is not of significant importance to understand its phenomenological impact on the presented analyses as mentioned above.

Finally, after all hadrons are produced in the hadronisation process, meaning those in the parton shower and UE, any unstable hadrons are decayed according to the branching ratios given, for example, by the Particle Data Group [3]. This aspect cannot be neglected, particularly for cases of subsequent or multiple decays as with hadrons containing charm or bottom quarks.

4.5 Simulation of the ATLAS detector and pile-up

The simulation of proton-proton collisions from the matrix element to the final state objects is only the first step when performing an analysis based on actual physics data. The MC generators are able to predict the outcomes of proton-proton collisions in terms of physical particles, but in order to compare the generated events to experimental data, it is essential to simulate how the particles interact with the detector material.

The ATLAS experiment employs the GEANT 4 software package [43] in order to simulate the response of the ATLAS detector to the proton-proton collisions produced at the LHC [44]. This involves simulating how particles pass through the various parts of the detector and create hits (or not in case of a neutrino). The simulated hits are digitised to simulate the signals produced by the electronic readout system. The reconstruction algorithms turn these signals into tracks in the ID and MS as well as energy depositions in the calorimeter cells. These tracks and energy depositions are used to reconstruct electrons, muons, photons and hadrons i.e. jets.

Before the simulation of the detector response, however, one particular and very crucial aspect is taken into account within the MC simulated samples to reproduce the LHC conditions. This aspect is called pile-up and it refers to the concept that the mean number of proton-proton interactions per crossing of the proton bunches, $\langle\mu\rangle$, is significantly larger than one, with an average of $\langle\mu\rangle \approx 32$ interactions during the years 2015, 2016 and 2017 [45]. This is due to the high luminosity configuration at the LHC. The different pile-up profiles of the data recorded by ATLAS in the three years are shown in Figure 4.5. The large pile-up adds an additional challenge to physics analyses, because the neighbouring proton-proton interactions tend to produce many soft particles that are usually not of interest, but may still interfere with the reconstruction of the hard process. It is therefore essential to have a detector with high spatial resolution as well as good momentum resolution in order to identify the particles that originate from the primary vertex and distinguish them from those particles produced by neighbouring collisions. In simulation, pile-up is taken into account by overlaying the generated events with so-called minimum bias events assuming a value of $\langle\mu\rangle = 20$ [46]. Minimum bias events are inelastic collisions without any specific selections such as requiring the presence of a high p_{T} lepton or jet, hence the tendency to contain soft particles. In ATLAS, these events are generated using the PYTHIA 8 generator with the MSTW2008LO PDF set and the A2 tune [46–48]. After the full simulation of the detector response and reconstruction of objects, this simulated pile-up profile is corrected to the pile-up profile measured in data

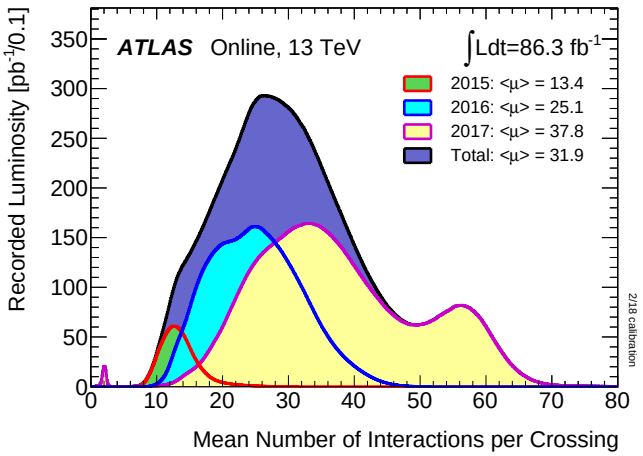

Figure 4.5: The pile-up profiles of the data recorded with the ATLAS experiment during the years 2015, 2016 and 2017 of Run 2 [45].

via certain weight factors which are applied to the generated events [49]. In addition to the pile-up modelling, the object reconstruction algorithms employed in ATLAS are designed in a sophisticated way to reject most particles that do not originate from the hard process, for example the jet vertex tagger algorithm outlined in Section 5.3.

The following chapter describes how or under which criteria the different relevant objects, namely leptons and jets, are reconstructed in the analyses presented in Chapters 6 and 7.

CHAPTER 5

Analysis objects

In the previous Section 4.5, the point was raised that analysers cannot simply compare an MC simulated sample of particles that are created during a proton-proton collision to actual physical events that are recorded with the ATLAS detector. First, the MC generator is required to model how the particles will interact with the detector, what signatures they produce and thus, which electric signals the readout system is ultimately able to record. Based on this information, the ATLAS collaboration applies algorithms that read in those particle signatures in the detector during a collision and reconstruct the particle candidate that produced the signature, along with its kinematic properties. Then, sophisticated trigger algorithms may fire a signal to store the event if certain criteria are met such as the presence of a high p_{T} lepton, as was discussed in Section 3.2.

This chapter describes how the individual physics objects, namely leptons and jets, are reconstructed based on the signatures that have been left in the detector. As depicted in Figure 3.6, each object produces a different signature which means that all the information from the detector has to be taken into account to maximise the reconstruction efficiency. Nonetheless, different particles may produce a similar signature inside the detector. Since the detector and readout systems are not perfect, the corresponding algorithms might reconstruct an object as a different one. This has to be considered especially for electrons, photons, jets and hadronically decaying τ-leptons.

5.1 Electrons

Electrons are the lightest charged fermions of the SM and thus do not decay [3,7]. They do, however, ionise the material of the ID and thus leave a track in it. Furthermore, they are expected to be stopped in the ECAL, predominantly by producing an electromagnetic shower through Bremsstrahlung and subsequent production of electron positron

pairs until all their energy is deposited in the ECAL [28].

The electron reconstruction algorithm [50–52] tries to find tracks in the ID with $p_{\mathrm{T}} > 0.5$ GeV that can be associated to energy depositions, called clusters, in the ECAL that fulfil $E_{\mathrm{cluster}}/\cosh(\eta_{\mathrm{track}}) \equiv E_{\mathrm{T}} > 2.5$ GeV. The energy of the cluster comprises different components besides the actual energy inside the ECAL, namely the energy deposited in front of the ECAL and the energy leakage outside of the cluster as well as after the ECAL. To reconstruct the four-momentum of the electron, the information from the ECAL and ID are combined. The energy of the electron is the cluster energy whereas its direction of movement is based on its track. The detector does not operate with a perfect efficiency, but instead suffers from a limited precision on the energy scale, energy resolution and other properties which are detailed further in Section 6.3.1 and Chapter 8. In order to match the reconstruction efficiency of simulated electrons (and other objects) to that of electrons in data, the above effects must be taken into account which is done via scale factors (SFs). These SFs are defined as $\epsilon_{\mathrm{data}}/\epsilon_{\mathrm{MC}}$, where ϵ represents the corresponding efficiency. The efficiency in data is measured experimentally and the SFs can be applied to simulated events in all analyses considering electrons.
As the ID is limited in η, only those candidates with $|\eta| < 2.47$ are considered. However, the region $1.37 < |\eta| < 1.52$ is excluded as well, because the reconstruction efficiency in this region suffers significantly from the large amount of passive material in front of the ECAL [28].

In addition to the reconstruction of the electron, the algorithms must consider multiple other mechanisms that could lead to another object being mis-reconstructed as an electron, thus called fake electron [28]. This involves primarily photons which undergo the same procedure in the ECAL, but leave no track in the ID unless they split into an electron positron pair at an early stage. Another example would be a photon that deposits energy into the ECAL and these clusters are matched wrongly to a track in the ID. On the other hand, jets also leave tracks in the ID and may deposit energy in the ECAL. If the jets are stopped already in the ECAL or do not deposit a lot of energy into the HCAL, they might be mis-identified as electrons due to the similar signature. These cases are discriminated by a cut-based classification which uses information from the ID as well as the calorimeters, which is described in more detail in Section 6.3.1.

5.2 Muons

Muons are the charged leptons of the second generation with roughly 200 times the mass of the electron [3]. Muons may decay into electrons and have an average lifetime of $\tau = 2.2 \cdot 10^{-6}$ s [3]. Furthermore, as they are typically produced with high energy during proton-proton collisions at the LHC, they possess a high $\beta\gamma$ value and thus decay outside of the detector. In addition to their larger mass, they are minimum ionising particles (MIPs) at LHC energies and do not participate in the strong interaction, hence they are neither stopped in the ECAL nor in the HCAL and tend to only deposit a small

fraction of their energy in both for p_{T} up to roughly 500 GeV. The resulting signature is a track in the ID and another track in the MS in addition to a negligible amount of energy deposits in the calorimeters in between the tracks [28].

The muon reconstruction algorithm [51,53,54] thus searches for tracks in the MS starting from the outermost layer and tries to match them to tracks in the ID while accounting for energy losses in the calorimeters. This matching procedure is critical, because tracks in the ID could potentially be associated with a jet or electron in the case of a significant energy deposit in the calorimeters. The ID and MS cover $|\eta| < 2.5$ for muon object candidates. As is the case for electrons, muons are not perfectly modelled as well and thus have to be corrected via similar SFs which is further detailed in Section 6.3.1.

Finally, since muons are MIPs and do not lose a lot of energy in the calorimeters, it is very rare for muons to fake or be faked by a jet or an electron. On the other hand, heavy flavour jets or hadrons, meaning those originating from b-quarks or c-quarks, may decay inside the detector into a muon (or electron). This additional source of muons not related to the primary interaction has to be taken into account in physics analyses and the strategy to reject them is detailed further in Section 6.3.1.

5.3 Jets

As discussed in Section 2.4, the principle of confinement implies that quarks are not observable as free particles. They will instead form colour neutral hadrons or decay weakly into lighter particles. In particular, the aspect of decaying is crucial for the reconstruction of top quarks and jets originating from b quarks, called b-jets. The hadrons or colour charged decay products deposit energy into the HCAL and, in case they are electrically charged, leave a track in the ID as well as deposit energy in the ECAL [28].

In the ATLAS experiment, jets are reconstructed from three-dimensional topological energy clusters in the calorimeter [55] with the anti-k_t algorithm [56] implemented in the FASTJET package [57] using a distance parameter of $R = 0.4$. This distance parameter is defined as $R = \sqrt{\eta^2 + \phi^2}$ and approximates a cone around the jet trajectory. As for the charged leptons, the energy of jets has to be corrected via calibration. The energy of the jet is initially based on the visible electromagnetic energy depositions and thus calibrated to the electromagnetic scale (EM scale). In the next step, the energy is corrected to the jet energy scale to bring it to the particle level. The jet energy scale calibration is derived from simulation and in situ calibration using 13 TeV data [58]. The four-momentum of the jet is then corrected to the position of the primary vertex. This jet energy scale calibration considers three dedicated systematic uncertainties with respect to the flavour of jets, namely the jet flavour composition as well as the jet response of gluon-initiated and b-quark-initiated jets [58]. This is important to note as the jet flavour plays a major role in the analyses presented in the following chapters.
The jets reconstructed with this procedure are called calorimeter jets. There are alter-

native approaches such as reconstructing jets only using tracks in the ID [59, 60]. As such approaches are not used in the two analyses presented in this thesis, they are not discussed further.

There are three additional aspects which affect the reconstruction of jets in the detector. The first is jets that originate from pile-up i.e. neighbouring proton-proton interactions in the same bunch crossing. To reduce the number of such jets contaminating the hard process, an algorithm called jet vertex tagger (JVT) is applied to jets in data as well as simulation [61]. For each jet within $|\eta| < 2.4$, the idea is to associate tracks from the ID to its axis and build a multivariate discriminant based on the detector information. On average, jets originating from the primary vertex achieve a higher JVT score and thus are more likely to pass a cut on this discriminant. In the following analyses, jets have to pass the condition of JVT > 0.59. This cut preserves 92% of jets originating from the primary vertex and approximately 2% of pile-up jets. For jets with $p_{\mathrm{T}} > 60$ GeV, the likelihood that it originates from a pile-up event is sufficiently small and so the JVT is not used in those cases.

The second aspect does not significantly contribute to the (mis-)reconstruction of objects, but should still be mentioned briefly. During collisions, the produced hadrons may travel through the detector with so much kinetic energy that they are not stopped in the HCAL and thus cause a punch-through. This means they may enter the MS and produce a signal there as well. However, the design of the HCAL is such that these cases are reduced well below the irreducible level of prompt or decay muons [28]. As it still affects the calibration of the jet energy scale, it is nonetheless taken into account as an uncertainty in the following analyses.

Finally, the third aspect is the removal of overlapping objects from events which aims to avoid double-counting. The metric commonly used to define overlapping objects is $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$.
During the reconstruction of jets, the considered energy deposits from jets are not discriminated against those from electrons. Therefore, jets within $\Delta R < 0.2$ of a reconstructed electron are removed. However, if the nearest jet surviving this cut is within $\Delta R < 0.4$ of that electron, the electron is discarded instead. This procedure is useful to only consider electrons which show a clear separation from close-by jets and hadrons.
Signatures from muons and jets can also overlap as muons may originate from the decay of heavy flavour jets or if highly energetic muons deposit a significant amount of energy in the calorimeter. This overlap is avoided by requiring the muon to be separated by $\Delta R > 0.4$ from the nearest jet. If this requirement fails, but the jet has at least three tracks in the ID associated to it, the muon is removed from the event, otherwise the jet is removed instead.

In this dissertation, the two main processes under study are $t\bar{t}$ and $t\bar{t}H(H \rightarrow b\bar{b})$. Therefore, the expected number of b-jets ranges between two and four at LO. A dedicated

algorithm to identify and tag these b-jets is applied in physics analyses within ATLAS and other high energy particle physics experiments. The underlying technique is called b-tagging and exploits the signature of b-jets which, on average, is significantly different from that of all other jets. How this algorithm works and performs and how it is calibrated in ATLAS is the topic of the next chapter.

5.4 τ-leptons

τ-leptons have a mass of $m_\tau \approx 1.777$ GeV and an average lifetime of $\tau_\tau \approx 2.9 \cdot 10^{-13}$ s [3]. Consequently, they travel a small distance in the order of a few millimetres through the ATLAS detector before decaying in $\approx 65\%$ of cases into either hadrons, mostly pions, or in $\approx 35\%$ of cases into either an electron or muon [3].

The τ-lepton itself is not reconstructed directly, but its production is inferred through its decay products [28]. In the $t\bar{t}$ and $t\bar{t}H$ events presented here, τ-leptons originate primarily from the decay of W bosons, similarly to electrons and muons and with roughly the same branching ratio of $\approx 11\%$ [3]. In contrast to a W boson directly decaying into an electron, muon or pair of quarks, the case in which the W boson first decays into a leptonically or hadronically decaying τ-lepton introduces one or two additional neutrinos. This increases, on average, the amount of missing transverse momentum in the event and, due to the different phase space in the τ-lepton decay, changes the kinematic properties of the decay products.

Both analyses in this dissertation are based on the $t\bar{t}$ process and select electrons, muons and jets in the final state. Thus, events where the W bosons first decay into τ-leptons which then subsequently decay into an electron, a muon or quarks are also considered. This is because such events are present in the recorded ATLAS data and it is difficult to efficiently discriminate a leptonically decaying τ-lepton from a W boson directly decaying into an electron or muon. Furthermore, τ-leptons may decay hadronically and, therefore, may fake the signature of jets in the calorimeter. Such a τ-lepton could be mis-identified as, for example, a b-jet. This latter circumstance contributes to the reconstruction efficiency and is considered as a small additional source of uncertainty in the search for the $t\bar{t}H(H \to b\bar{b})$ process, which is detailed further in Section 10.2.

5.5 Missing transverse momentum

The last object of importance in this dissertation is the missing transverse momentum. While the four-momenta of the initial protons are known, the actual colliding partons possess an unknown momentum in the plane perpendicular to the beam axis. Only their momenta in the beam direction is expected to be roughly zero before and after the collision due to momentum conservation [9].

Particles that were not reconstructed in the detector, such as neutrinos, contribute to a missing momentum in the transverse plane, labelled $p_{\mathrm{T}}^{\mathrm{miss}}$. It is defined as the negative vector sum of transverse momenta of all reconstructed and calibrated objects in the event plus a term representing all remaining energy depositions not associated to any objects. This latter term is calculated using tracks in the ID which are matched to the primary vertex which makes it more robust against tracks from pile-up [62, 63]. The missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, is the magnitude of the missing transverse momentum and may also be faked by noisy detector signals or even jets [28].

While the $p_{\mathrm{T}}^{\mathrm{miss}}$ consists of two unknowns, only the magnitude $E_{\mathrm{T}}^{\mathrm{miss}}$ can be inferred which at first suggests that the momenta of neutrinos can never be reconstructed in proton-proton collisions at the LHC. However, since the invariant masses of all SM particles are known, additional constraints can be applied to the neutrino momentum, for example in order to fulfil $p_{W^{\pm}}^2 = m_W^2$ and $p_{t/\bar{t}}^2 = m_t^2$ in $t\bar{t}$ decays. The missing transverse momentum is used during the reconstruction of $t\bar{t}H(H \to b\bar{b})$ events, as described in Chapter 8.

CHAPTER 6

Calibration of the heavy-flavour jet-tagging algorithm

This chapter discusses the underlying concepts of b-tagging in ATLAS. More specifically, it focuses on how the identification of b-jets in the detector is performed and how it influences physics analyses. In the following section, the three baseline algorithms in ATLAS are presented which exploit the distinct signatures of b-jets inside the detector and whose outputs are then combined to build the final b-tagging discriminant. Afterwards, the calibration strategy of this discriminant is motivated and the fit method described in detail. Section 6.3 discusses the setup of the calibration analysis including details about the studied dataset and simulated samples, the object reconstruction, event selection, and the differences with respect to previous calibrations. Following this, the robustness of the fit method is evaluated through several tests presented in Section 6.4. In Section 6.5, the considered statistical and systematic uncertainties are depicted as well as how they are estimated. Lastly, Section 6.6 summarises the results and evaluates their significance, followed by a short conclusion and outlook.

6.1 The b-tagging algorithm in Atlas

Algorithms aiming to identify b-jets try to exploit the distinct features of such jets. Firstly, b quarks are the lighter quarks of the third generation and thus may only decay weakly into quarks of the first or second generation. Such a decay is suppressed, since it involves off-diagonal CKM matrix elements, as was outlined in Section 2.5. Consequently, the average lifetime of b hadrons is long compared to other hadrons, namely in the order of $\tau_b \approx 1.5 \cdot 10^{-12}$ s, which corresponds to $c\tau_b \approx 450$ μm [3]. Thus, a typical b hadron produced at the LHC with a p_{T} in the order of 50 GeV travels, on average, about 3 mm in the transverse plane before it decays inside the detector [64]. This relatively long mean path of flight is visible in the ID as a secondary vertex which is displaced from the primary vertex, illustrated in Figure 6.1. Therefore, the most important input

for the b-tagging algorithm is the information about tracks from charged particles in the ID.

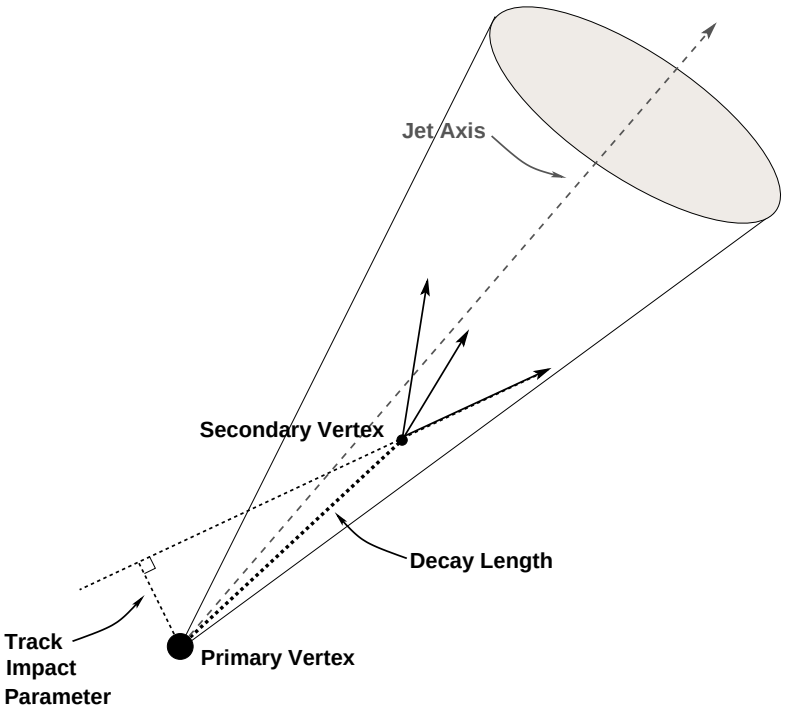

Figure 6.1: A scheme of the expected signature of a b-jet in the inner detector. There is a distinct track which connects the primary to the secondary vertex at which the b hadron decayed. The impact parameters of the tracks can be used to distinguish b-jets from other jets [65].

Three baseline algorithms are employed to optimise the b-tagging performance in the ATLAS experiment. Each of the algorithms exploits different properties of the expected signature of b-jets and thus provide complementary information. In the end, the outputs of these algorithms are combined into two multivariate discriminants, namely a boosted decision tree (BDT) as well as a deep learning neural network (DNN).

The following subsections describe the aforementioned sub-algorithms as well as the multivariate analysis techniques that are used to perform the b-tagging in ATLAS during Run 2 of the LHC physics programme. The multivariate discriminant based on the Run 2 BDT (DNN) training is called MV2 (DL1). They are constructed and trained in a different way with respect to the corresponding b-tagging algorithms employed by ATLAS during Run 1 which are discussed in Ref. [64]. These differences are a result of the different detector layout during Run 1, particularly in regards to the ID and the IBL. The expected performance of the new b-tagging algorithms due to the detector upgrades

from Run 1 to Run 2 is presented in Ref. [66]. Further optimisation studies of the Run 2 algorithms are given in Refs. [67, 68].

6.1.1 Impact parameter based algorithms: IP2D and IP3D

The expected signature of a b hadron in the ID is a secondary vertex which is significantly displaced from the primary vertex, meaning the hard collision vertex, and additional tracks originating from that secondary vertex. Two quantities can be constructed for each such track with respect to the r-ϕ plane. The transverse impact parameter (IP), d_0, is defined as the distance of closest approach in the r-ϕ plane of the track to the primary vertex [28]. This is shown in Figure 6.1. The longitudinal IP is defined based on this point of closest approach in the r-ϕ plane, namely as the distance between the track and the primary vertex in the longitudinal plane and thus is given by $z_0 \sin\theta$ where z_0 is the position along the beam axis and θ is the polar angle of the track [28]. The corresponding significances of these quantities are defined as the IP divided by its estimated uncertainty, namely $S_{r\phi} := d_0/\sigma_{d_0}$ and $S_z := z_0 \sin\theta/\sigma_{z_0 \sin\theta}$, respectively.
One might argue that tracks from decay products of b hadrons may interfere with the association of tracks to the primary vertex. However, as the b hadron possesses a long lifetime, the secondary vertex is displaced and the decay products will, on average, have large IP values [67]. Furthermore, the significance of the IPs of such tracks is significantly different from zero whereas the significance of the IPs of tracks from light-jets tends to be consistent with zero [67].
Additionally, one can assign a sign to the IP, depending on whether the secondary vertex is behind or in front of the primary vertex with respect to the jet axis which represents its direction of motion. If the secondary vertex is behind the primary vertex, the track is likely not to come from b-jets and thus its IP is negative. This is useful to estimate the b-tagging rate of light-jets, also called mis-tag rate [64].

Tracks that enter the IP algorithms must fulfil the following criteria [67]:

- Track $p_{\mathrm{T}} > 1$ GeV;
- $|d_0| < 1$ mm and $|z_0 \sin\theta| < 1.5$ mm;
- ≥ 7 hits in PD and SCT combined (silicon hits) and ≤ 2 silicon holes of which at most one may be in the PD. A hole refers to a hit that is expected to be associated with the track, but is not present [51].

The difference in the IP algorithms is suggested by their names and lies in the information they use. The IP2D algorithm uses only the transverse IP significance $S_{r\phi}$ as a discriminating variable, while IP3D uses S_z as well, namely in a two-dimensional template which takes their correlation into account [67]. Figure 6.2 shows the respective significances for the different jet categories which are b-jets, c-jets and light-jets. These templates are given in the form of probability density functions (PDFs) and derived from simulated $t\bar{t}$ events. The templates are different depending on the hit pattern of

the track and the jet flavour hypothesis, which could be either b, c or light [67].

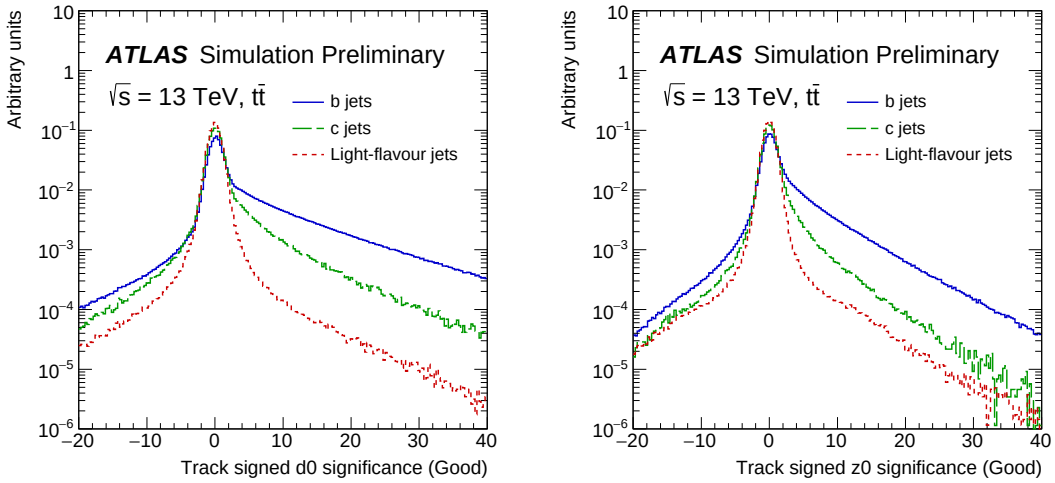

Figure 6.2: The significances of the transverse IP (a) and longitudinal IP (b) are shown in simulated $t\bar{t}$ events for b-jets (solid blue), c-jets (dashed green) and light-jets (dotted red). The jets fulfil the 'Good' category defined in Ref. [67].

The final discriminant of the IP algorithms is given by a log-likelihood ratio (LLR) in the form of a sum of per-track contributions, namely $\sum_{i=1}^{N} \log\left(\frac{p_b}{p_u}\right)$. In this sum, N is the number of tracks associated to the jet and p_b and p_u are the template PDFs for the b-jet and light-jet hypotheses of each track, respectively, while the tracks in the sum are assumed to be uncorrelated [67]. Thus, this LLR is specialised to discriminate between b-jets and light-jets. Correspondingly, there are two additional LLR using p_c to separate b-jets from c-jets or c-jets from light-jets [67]. Both algorithms, IP2D as well as IP3D, are included in the training of the multivariate b-tagging discriminants [67].

6.1.2 Secondary Vertex Finder: SV1

The Secondary Vertex Finder algorithm (SV1) [69] aims to reconstruct all secondary i.e. displaced vertices produced by jets originating from b hadrons in an event [67]. This algorithm is based on the tracking information from the ID, similar to the IP based algorithms and also the JetFitter algorithm described in the next subsection. The selection of tracks for the SV1 and JF relies on d_0, z_0 and the number of hits in the ID associated to the jet. This selection strategy has been optimised separately for the two algorithms [67].

The SV1 algorithm considers all tracks inside a jet and converts them into a set of all possible two-track vertices. SV1 then iterates over all track pairs and tests the hypothesis that they come from a vertex with two tracks. Vertices are rejected that are compatible

with the decay of long lived particles such as Λ or K_S, photon conversions or hadronic interaction with the detector material [67]. The tracks from all vertices remaining after this process are used to fit a new secondary vertex. Outlier tracks are removed iteratively from the fit, while tracks from b or c hadron decays are allowed to be associated to a single common secondary vertex [70].

In order to improve the performance of the fit, several additional requirements and strategies are applied, namely [67]:

- The tracks are required to have at least 7 silicon hits, while at most 1 hit may be shared in the two detectors. Additionally, the χ^2 value of the vertex fit divided by the numbers of degrees of freedom (dof) is required to be less than 3. At $|\eta| > 1.5$, tracks are required to have at least 8 silicon hits instead of 7 to mitigate the effects of the larger amount of detector material which causes a worse track resolution and higher occurrence of hadronic interactions.
- Tracks with $S_{d_0} < 2$ and $S_{z_0} > 6$ are likely to originate from fake vertices due to high pile-up and are therefore rejected.
- The significance of the distance of each track to the primary vertex should be above 2, the sum of the two IP significances of both tracks should be greater than 2 and the χ^2 of the fitted tracks should be less than 4.5.
- Jets with energies above 300 GeV tend to have a high number of tracks from the jet fragmentation as shown in Figure 6.3 on the left. This increases the likelihood of reconstructing fake vertices inside the jet. Therefore, at most those 25 tracks with the highest p_{T} are considered in the reconstruction. Since b hadron decays have, on average, 5 tracks from charged particles, this selection is sufficiently inclusive to keep a high reconstruction efficiency [67].
- The invariant mass of the vertex can be further exploited to reject tracks from unwanted sources. The vertices from b or c hadron decays tend to have masses between $1 - 5$ GeV, as is depicted in Figure 6.3 on the right. Thus, the invariant mass of candidate vertices is required to be less than 6 GeV. In addition, the mass spectra of $\pi^+\pi^-$ and $p\pi$ systems are taken into account to reject two track vertices coming from K_S or Λ decays. All tracks associated to such a vertex are rejected.

The SV1 algorithm uses eight discriminating variables during the training, including the number of tracks associated to the vertex, its invariant mass, its energy fraction which is defined as the total energy of tracks associated to the vertex divided by the total energy of tracks associated to the jet, and the three-dimensional decay length significance which is defined as the distance between the primary and secondary vertices divided by its uncertainty [67]. A similar set of variables is used for the training of the JetFitter algorithm described in the next subsection [67].
The most significant reason for the degradation of performance of the SV1 algorithm at high p_{T} is the increase in the number of tracks which, as explained above, increases the chance to reconstruct fake vertices [67].

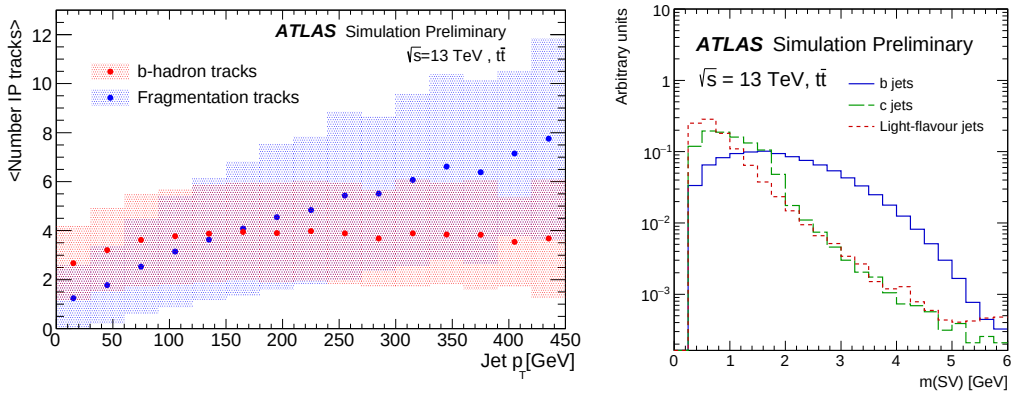

Figure 6.3: Left: Average number of tracks selected for the IP algorithms as a function of the jet p_{T} originating from b hadrons (red) and jet fragmentation (blue). The shaded areas represent the RMS for each p_{T} bin [67]. Right: the fraction of events that a b-jet (solid green line), c-jet (dashed blue line) or light-jet (dotted red line) has a secondary vertex with a certain invariant mass associated to it [67].

6.1.3 Decay Chain Multi-Vertex Finder: JetFitter

The JetFitter algorithm (JF) [71] tries to reconstruct the decay chain of b hadrons in a topological way along the jet axis. The underlying idea is that b hadrons are most likely to decay into a c hadron which then decays further and thus produces a tertiary vertex [67].
The JF is based on a modified Kalman filter [72] and aims to find a single common line connecting the primary vertex, the b hadron decay vertex and the c hadron decay vertex [67]. This approach allows to reconstruct the vertices even with only a single track associated to them, as long as the track resolution is high enough. As mentioned in the previous subsection, a similar set of variables is used during the training of both the SV1 and JF algorithms, including the number of tracks, invariant mass, energy fraction and three-dimensional decay length significance [70].

6.1.4 Multivariate discriminant for b-tagging: MV2

In order to maximise the b-tagging performance, the outputs of all three sub-algorithms and the jet kinematics, meaning the p_{T} and η, are combined to train a multivariate discriminant, namely a BDT which is called MV2 [51,67,68]. The BDT is trained using the ROOT Toolkit for multivariate analyses, called TMVA [73], on two simulated samples. The first contains 5 million $t\bar{t}$ events and the other 3 million Z' bosons decaying into two jets [68], where Z' refers to a commonly proposed new heavy vector boson which is not part of the SM, described for example in Ref. [74]. It behaves similarly to the Z boson but has a significantly higher rest mass of the order of $m_{Z'} \approx 4$ TeV. Here, the physics

details of this boson are not as important as the fact that it allows physicists to simulate the production of highly energetic $b\bar{b}$ pairs at the LHC. This is because, while the b-jets produced in $t\bar{t}$ decays have, on average, a p_{T} significantly less than 250 GeV, the b-jets from Z' decays sufficiently often reach a p_{T} of up to ≈ 2 TeV. This is critical for the training of the algorithm on b-jets with high p_{T} and its subsequent calibration. The contribution of both simulated samples to the b-jet p_{T} spectrum as well as the light-jet p_{T} spectrum is shown in Figure 6.4.

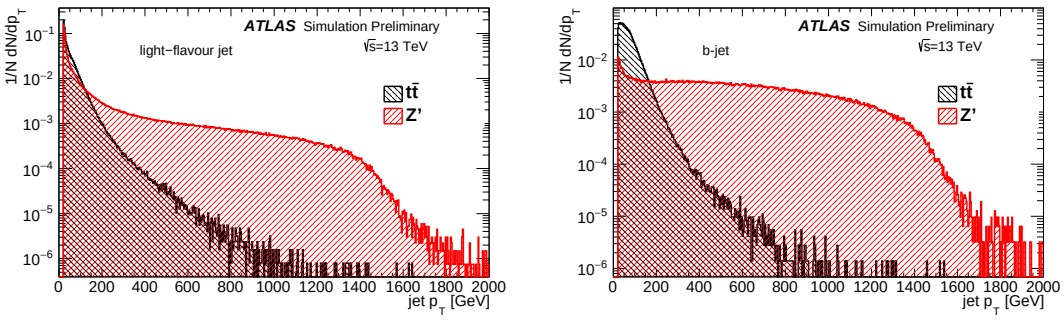

Figure 6.4: The normalised differential cross-section of the light-jet p_{T} (left) and b-jet p_{T} (right) spectra [68]. The simulated $t\bar{t}$ sample contributes mostly to jets with $p_{\mathrm{T}} \leq 250$ GeV (black). Therefore, an additional sample of simulated Z' decays is added to increase the number of jets with a higher p_{T} (red).

During the training of MV2, b-jets are set as the signal, while the background contains a mixture of 7% c-jets and 93% light-jets. Other compositions of backgrounds have been tested as well, but this particular mixture has been found to give the best separation between the different flavours. It has, therefore, been chosen as the default in ATLAS during Run 2 [67].
The p_{T} and $|\eta|$ of jets are included in the training to exploit correlations with other discriminating variables from the input sub-algorithms [67]. However, these kinematic distributions are different between signal and background jets and the BDT could interpret these differences as discriminating by the training. In order to avoid this, the p_{T} and $|\eta|$ distributions of b-jets and c-jets are reweighted to match those of light-jets [67]. A penalty weight of 10^{-6} is given to a jet during the training if and only if it fails to produce a result in all three sub-algorithms [67]. In the case of the SV1 algorithm, for example, this happens if it is not able to reconstruct a secondary vertex that can be associated to the jet. The fraction of jets affected by this is 0.6% for light-jets and 0.1% for c-jets and b-jets, respectively.

The discriminant based on this training is called MV2c10 to reflect the chosen background mixture in which 'c10' represents the roughly 7% c-jets [67]. The performance of the MV2c10 algorithm can be evaluated by its ability to separate between the different

jet flavours. The BDT output for the different jet flavours is presented in Figure 6.5. Further details about the training and performance can be found in [67].

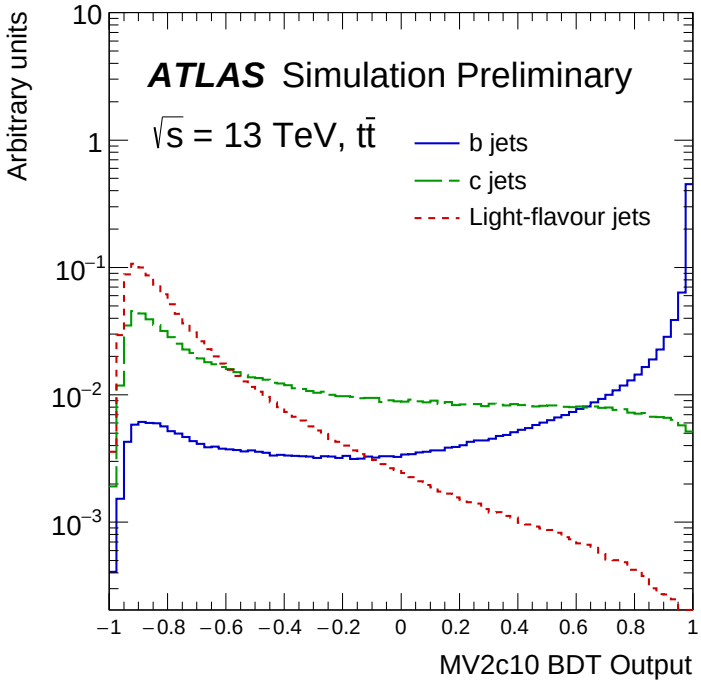

Figure 6.5: The MV2c10 output is shown for b-jets (solid blue), c-jets (dashed green) and light-jets (dotted red), evaluated using $t\bar{t}$ events [67].

Finally, the MV2c10 algorithm can be used at different working points (WPs). Analysers in ATLAS may choose to apply the b-tagging of jets at four different efficiencies, which are 60%, 70%, 77% and 85%. These percentages represent the fraction of b-jets which remain in simulated $t\bar{t}$ events after the corresponding cut on the MV2c10 output value is applied. The lower the fraction, the fewer tagged b-jets remain, but at the same time far more c-jets, light-jets or τ-jets are rejected. This is illustrated in the so-called receiver operator curve in Figure 6.6. Table 6.1 summarises the BDT cut values corresponding to the WPs as well as the respective rejection factors for c-jets, light-jets and τ-jets. If analysers have a sufficiently large number of events containing b-jets, they may choose to apply the tightest working point of 60% to reject as many unwanted background jets as possible.

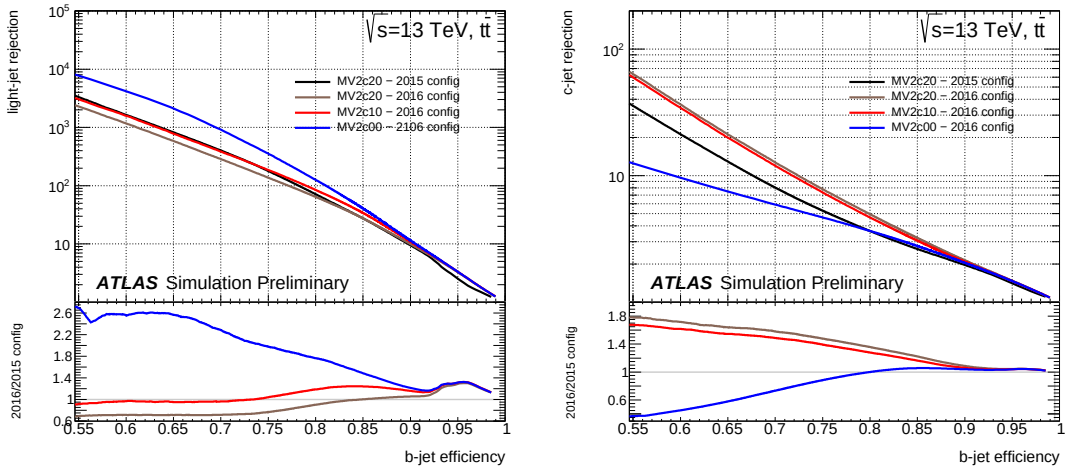

Figure 6.6: Light-jet rejection (left) and c-jet rejection (right) as a function of the b-jet tagging efficiency for an old MV2c20 configuration based on 2015 data (black) compared to several improved configurations based on 2015 and 2016 data, namely MV2c20 (brown), MV2c10 (red) and MV2c00 (blue), described in Ref. [67]. The distributions are evaluated on simulated $t\bar{t}$ events. MV2c00 (20) corresponds to the discriminant trained on a sample with 0% (15%) c-jets out of all background jets. The bottom panels show the ratio between the improved configurations with respect to the old 2015 configuration.

b-tagging WP	MV2c10 cut value	c-jet rejection	light-jet rejection	τ-jet rejection
85%	0.1758	3.1	33	8.2
77%	0.6459	6	134	22
70%	0.8244	12	381	55
60%	0.9349	34	1538	184

Table 6.1: Shown are the four b-tagging WPs and the corresponding MV2c10 BDT cut values along with the rejection factors for c-jets, light-jets and τ-jets.

6.1.5 Deep Learning algorithm: DL1

In addition to the MV2 discriminants, another set of b-jet tagging algorithms have been developed, trained and calibrated. For these algorithms, the same input variables and training samples are used as for MV2. But, instead of a BDT, a deep learning neural network is used which results in a slightly different performance. This is reflected in, on average, higher light-jet rejection factors and smaller c-jet rejection factors for a given b-jet tagging efficiency with respect to the MV2 discriminants. This set of algorithms is called DL1. A detailed discussion on their concept and training can be found in Refs. [68, 70].

6.2 The strategy to measure ϵ_b

In ATLAS, the b-tagging algorithms can be applied to jets at different WPs by cutting on specific output values of the discriminant. However, the corresponding b-tagging efficiencies are derived from simulation. The performance of the b-tagging algorithms in actual data is not necessarily equal to their predicted performance, for example if the jet properties are mis-modelled. In this chapter, the calibration of the MV2c10 and DL1 algorithms is presented and the goal of this calibration is to measure the b-tagging efficiency in data (ϵ_b^{data}). This is done by selecting events which are enriched in b-jets and counting the number of jets which are tagged by the algorithm. The ratio of the measured efficiency with respect to the predicted efficiency (ϵ_b^{MC}) defines a scale factor (SF) that analysers in ATLAS apply in their analyses to correct the simulated b-tagging efficiency to the one measured in data.

Different methods to measure ϵ_b^{data} exist such as the Tag & Probe method and the Combinatorial Likelihood (LH) method which are described in Ref [75]. The latter is based on probability density functions (PDFs) and is thus called PDF method in the remainder of this chapter. The calibration analysis presented in this thesis is based on such a PDF method, but employs an improved strategy [76] with respect to the previous calibration in ATLAS that is detailed in Ref. [75]. The essential differences between the two calibrations are depicted in the following sections.
First, the basic concept of the old PDF method and how it can be used to measure the b-tagging efficiency in data is described in the following. This is important, because the $t\bar{t}H(H \to b\bar{b})$ analysis presented in Chapter 7 applies this older version of the b-tagging calibration. After this, the new improved approach is presented which allows to extract more information from data compared to the old method and is thus less dependent on the modelling of the selected physics processes.

To start, consider a sample of proton-proton collision events which contain reconstructed jets in the final state. Naively, the fraction of jets which are tagged by the b-tagging algorithm can be written as follows [75]:

$$f_{\text{tagged}} = f_b \epsilon_b + (1 - f_b)\, \epsilon_l. \tag{6.1}$$

Here and in the following, f_b is the fraction of b-jets in that sample and ϵ_b and ϵ_l are the tagging efficiencies for b-jets and non-b-jets, respectively. Non-b-jets represent all c-jets, light-jets and τ-jets in the following. This approach assumes that all b-jets are tagged with efficiency ϵ_b and all non-b-jets are tagged with efficiency ϵ_l. The b-tagging efficiency can thus be extracted by measuring f_{tagged} in data and deriving f_b and ϵ_l from simulation [75].
This approach, however, treats jets individually on a case-by-case basis and does not account for possible correlations between flavours of multiple jets within a single event. Consider, for example, an event that contains two reconstructed, highly energetic jets which originate from the same vertex and move in opposite directions. If one of the jets

is tagged as a b-jet, the likelihood that the other one is a b-jet increases, since they are expected to have been produced as a pair, thus increasing the chance to also tag the second jet. This is relevant to this calibration analysis, because it selects events with exactly two reconstructed jets in addition to other selections which are further detailed in the next section. Thus, in order to exploit such per-event jet flavour correlations in events with exactly two jets, one can define a system of two equations describing the fraction of events with 1 and 2 b-tags, respectively in the following way [75]:

$$\begin{aligned} f_{1\ b-\text{tag}} &= 2f_{bb}\epsilon_b\left(1-\epsilon_b\right) + f_{bl}\left(\epsilon_b\left(1-\epsilon_l\right) + \epsilon_l\left(1-\epsilon_b\right)\right) + 2\left(1-f_{bb}-f_{bl}\right)\epsilon_l\left(1-\epsilon_l\right), \\ f_{2\ b-\text{tags}} &= f_{bb}\epsilon_b^2 + f_{bl}\epsilon_b\epsilon_l + \left(1-f_{bb}-f_{bl}\right)\epsilon_l^2. \end{aligned} \tag{6.2}$$

f_{bb} and f_{bl} represent the fraction of events containing exactly two b-jets or exactly one b-jet and one non-b-jet, respectively. Both, as well as ϵ_l, can be derived from simulation and the two fractions of events $f_{1\ b-\text{tag}}$, $f_{2\ b-\text{tags}}$ can be measured in data to determine ϵ_b, similarly to the previous strategy. In this case, however, the jet flavour correlations among the two jets are accounted for, since ϵ_b is extracted in the two distinct tag regions simultaneously which improves the precision of the measurement [75].

The old measurement is performed in N different kinematic bins [75]. That means it is split up by different ranges of the two jet p_{T}, labelled $p_{\text{T},1}, p_{\text{T},2}$ from now on. This leads to an over-constrained system of $2 \cdot N^2$ coupled non-linear equations for N free parameters of interest (POIs). This requires a strategy to determine the most likely values of the POIs. Therefore, this calibration employs a LH function $\mathcal{L}$ [77] and performs a likelihood fit to data which is explained in the following.
In general, LH functions are products of distinct PDFs that depend on the POIs. Thus, for a given observed data, the LH function is a measure for the probability or compatibility of the POI values to describe the observed data. The goal is to find the POI values that maximise this LH function over all observed data. With this LH approach, the system of equations above is replaced by appropriate PDFs $\mathcal{P}$ and the values of the POIs are estimated by numerically maximising the LH function or, equivalently, minimising the negative logarithm of it (LLH function). Another advantage of the PDF method is its flexibility as it allows to include PDFs that are a function of the b-tagging weights w_1, w_2 which represent the outputs of the tagging discriminant for the two jets. The LH function for each event can thus be expressed as follows [75]:

$$\begin{aligned} \mathcal{L}_{\text{event}}\left(w_1, w_2, p_{\text{T},1}, p_{\text{T},2}\right) = [&f_{bb}\mathcal{P}_{bb}(p_{\text{T},1}, p_{\text{T},2})\mathcal{P}_b(w_1|p_{\text{T},1})\mathcal{P}_b(w_2|p_{\text{T},2}) \\ &+ f_{bl}\mathcal{P}_{bl}(p_{\text{T},1}, p_{\text{T},2})\mathcal{P}_b(w_1|p_{\text{T},1})\mathcal{P}_l(w_2|p_{\text{T},2}) \\ &+ f_{ll}\mathcal{P}_{ll}(p_{\text{T},1}, p_{\text{T},2})\mathcal{P}_l(w_1|p_{\text{T},1})\mathcal{P}_l(w_2|p_{\text{T},2}) \\ &+ \text{same terms, but replacing } 1 \leftrightarrow 2]/2. \end{aligned} \tag{6.3}$$

$\mathcal{P}_{f_1 f_2}(p_{\text{T},1}, p_{\text{T},2})$ are the PDFs for an event to contain two jets with flavours f_1, f_2 that have $p_{\text{T},1}$ and $p_{\text{T},2}$. $\mathcal{P}_f(w_i|p_{\text{T},i})$ is the PDF for a jet with flavour f and $p_{\text{T},i}$ to be assigned a b-tagging weight w_i by the b-tagging algorithm. f_{bb} and f_{bl} are the same parameters as in Equation 6.2 and $f_{ll} = (1 - f_{bb} - f_{bl})$. $\mathcal{P}_b(w|p_{\text{T}})$ can be extracted from data after

minimising the negative logarithm of the above LH function in which all other PDFs are derived from simulation. For a given jet p_{T}, $\mathcal{P}_b(w)$ is defined by a histogram with two bins, where the bin for a weight w above the cut value of the chosen WP w_{cut} represents the b-tagging efficiency. This can be translated into:

$$\epsilon_b = \int_{w_{\mathrm{cut}}}^{\infty} \mathrm{d}w' \mathcal{P}_b(w'|p_{\mathrm{T}}). \tag{6.4}$$

With this, the minimisation of the negative LLH function over all data events allows to determine ϵ_b from the observed data [75].
The new PDF method has several improvements compared to the above approach which is outlined in the following [76]. Firstly, it employs an extended binned LLH function which treats the total number of observed events in the analysed dataset as a parameter that is constrained by a Poissonian distribution [77]. Secondly, the new method measures the b-tagging efficiency in a pseudo-continuous way with five instead of two output bins corresponding to the WPs which includes a fifth WP for jets not passing the 85% WP. This means that if a WP of, for example, 70% is chosen for the b-tagging algorithm, then discriminant output weights that would pass the 60% WP fall into that '60%' bin. This is in contrast to the measurement based on Equation 6.4 which would keep all weights passing the 70% WP in one bin. On the other hand, the b-tagging efficiency in data for a given WP can be derived easily from the pseudo-continuous calibration through a cumulative sum.
In the new LLH definition, the pseudo-continuous WP definition is taken into account by considering the PDF $\mathcal{P}_f\left(w \in W|p_{\mathrm{T}}\right)$ which represents the PDF for a jet with flavour f and given p_{T} to have a weight w which falls into the WP bin W [76]. Thus, in the pseudo-continuous calibration, $\mathcal{P}_b(w|p_{\mathrm{T}})$ is given by a histogram with five bins for each p_{T} bin [76].
By writing an extended LLH function for each $(p_{\mathrm{T},1}, p_{\mathrm{T},2})$ bin, which are labelled as k from now on, one finds [76]:

$$\log \mathcal{L}_k\left(\nu_{\mathrm{tot}}, \mathcal{P}_b\left(w \in W|p_{\mathrm{T}}\right)\right) = -\nu_{\mathrm{tot}}^k + \sum_{w_1,w_2} n_{w_1,w_2}^k \log\left(\nu_{\mathrm{tot}}^k \cdot \sum_{f_1,f_2} \mathcal{P}_{f_1,f_2|k} \mathcal{P}_{f_1}\left(w_1 \in W_1|p_{\mathrm{T},1}\right) \mathcal{P}_{f_2}\left(w_2 \in W_2|p_{\mathrm{T},2}\right)\right). \tag{6.5}$$

In this equation, ν_{tot}^k is the expected total number of events and n_{w_1,w_2}^k is the observed number of events in bin k. The form is written in a more compact way than Equation 6.3 by taking the logarithm and including a sum which goes over all jet flavour combinations f_1, f_2. Here, ν_{tot} is considered a nuisance parameter, while $\mathcal{P}_b\left(w_i \in W_i|p_{\mathrm{T},i}\right)$ are the POIs, which will be estimated from the fit to data.

The precision of the previous measurement is limited strongly by the MC modelling uncertainty of the $t\bar{t}$ process [75]. More specifically, our understanding of the two jet flavour compositions in simulation is one of the dominant systematic uncertainties of the measurement. To account for this, correction factors c_{f_1,f_2}^k are introduced as nuisance

parameters to the LLH functions. These allow the two jet flavour compositions to be extracted from the fit for each bin k. In practice, the following change is made to Equation 6.5:

$$\nu^k_{\text{tot}} = \sum_{f_1,f_2} \nu^k_{f_1,f_2} \rightarrow \sum_{f_1,f_2} c^k_{f_1,f_2} \cdot \nu^k_{f_1,f_2}. \tag{6.6}$$

This calibration analysis aims to select as many events as possible that contain exactly two reconstructed b-jets. Such events are considered signal events and, consequently, this selected region of phase space is defined as the signal region (SR) and labelled 'bb'. In addition to this, other regions of phase space are selected in order to produce samples enriched in background events where either the first or second jet, ordered by p_{T}, or both are non-b-jets. These are considered control regions (CRs) and labelled 'lb', 'bl' and 'll', respectively. The exact definition of the SR and CRs is given in the next section which is different with respect to the old calibration described in Ref. [75].
Several tests have been performed to determine which regions should be used to extract the b-tagging efficiency with highest precision [76]. As a result, only the new SR is considered to measure ϵ_b^{data} which reduces the impact of non-b-jets. However, the CRs can still be exploited to estimate certain background components in the analysis such as the number of background events contaminating the SR [75, 76].
The old calibration only included events from their corresponding SR in the LH fit [75]. The new PDF method introduces additional parameters to the fit such as the two jet flavour correction factors described above. Thus, in order to have meaningful constraints on all fit parameters, the three CRs are included in the fit as well, but not used to extract ϵ_b^{data} [76]. This is achieved by creating dedicated LLH functions for each region according to Equation 6.5 and removing the dependence on the POIs in the LLH functions for the CRs.
Taking Equation 6.6 into account, the LLH function for the SR is given by [76]:

$$\log \mathcal{L}_{k,\text{SR}} \left(c^k_{f_1,f_2}, \mathcal{P}_b \left(w \in W | p_{\text{T}} \right) \right) = - \sum_{f_1,f_2} c^k_{f_1,f_2} \nu^{k,\text{SR}}_{f_1,f_2}$$
$$+ \sum_{w_1,w_2} n^{k,\text{SR}}_{w_1,w_2} \log \left(\sum_{f_1,f_2} c^k_{f_1,f_2} \nu^{k,\text{SR}}_{f_1,f_2} \mathcal{P}_{f_1,f_2|k,\text{SR}} \mathcal{P}_{f_1} \left(w_1 \in W_1 | p_{\text{T},1} \right) \mathcal{P}_{f_2} \left(w_2 \in W_2 | p_{\text{T},2} \right) \right). \tag{6.7}$$

The LLH functions for the CRs are written similarly except that all w_1, w_2 bins are merged together. This way, their dependence on the POIs are removed and they are only sensitive to the two jet flavour correction factors. They can be expressed as follows [76]:

$$\log \mathcal{L}_{k,\text{CR}} \left(c^k_{f_1,f_2} \right) = - \sum_{f_1,f_2} c^k_{f_1,f_2} \nu^{k,\text{CR}}_{f_1,f_2} + n^{k,\text{CR}} \log \left(\sum_{f_1,f_2} c^k_{f_1,f_2} \nu^{k,\text{CR}}_{f_1,f_2} \mathcal{P}_{f_1,f_2|k,\text{CR}} \right). \tag{6.8}$$

Estimates for $\nu^{k,\text{SR}}_{f_1,f_2}$ and $\nu^{k,\text{CR}}_{f_1,f_2}$ are derived from simulation, but they are corrected along k, f_1, f_2 by the two jet flavour correction factors $c^k_{f_1,f_2}$. Thus, only the shape of the

estimated ν distribution along the SR and CRs is taken from MC.

The differences of the new PDF method with respect to the old method are [76]:

- A reduced dependence on MC simulation, because the two jet flavour compositions are extracted from the fit to data;
- The fraction of signal i.e. bb events is much higher in the SR of the new calibration; thus, the extraction of the b-tagging efficiency in this region decreases the impact from non-b-jets;
- Since the w_1, w_2 bins are merged together in the LLH functions in the CRs, those functions are insensitive to ϵ_b which minimises the correlation between the two jet flavour correction factors and the POI $\mathcal{P}_b\left(w|p_{\mathrm{T}}\right)$.

Finally, the robustness of the new fit method is evaluated through a sanity test, a closure test and a stress test. These are presented in Section 6.4.

6.3 Analysis setup

In this section, the setup of the MV2c10 and DL1 calibrations will be detailed. This includes the chosen processes and corresponding final states, the selection strategy to enrich the data sample with these final states, the simulated signal and background samples considered as well as the dataset under study.

As outlined in the previous section, the calibration of the b-tagging algorithm is performed in events with exactly two reconstructed jets. Additional selections are performed to increase the fraction of events that contain exactly two b-jets. This is done by selecting a reconstructed electron (positron) and antimuon (muon) in addition to the two reconstructed jets. Thus, the selected final state is labelled as '$e\mu$+2j' in the remainder of this chapter. Many processes may contribute to it, but the most dominant one at the LHC is the $t\bar{t}$ process. This process is able to produce the selected final state, because each top quark is assumed to decay into a bottom quark and a W boson and the two W bosons may subsequently decay into the selected leptons. The most significant non-$t\bar{t}$ processes at the LHC that contribute to this final state are:

- the production of a single top quark in association with a W boson (Wt);
- the production of two vector bosons with additional jets, also called 'diboson' production ($WW/WZ/ZZ$+jets including off-shell Z boson contributions); and
- the production of a Z boson decaying into τ-leptons with additional jets ($Z \rightarrow \tau\tau$+jets).

The $e\mu$+2j final state bears the advantage that exactly two b-jets are expected from the $t\bar{t}$ decay at LO; any additional jets must come from initial state or final state radiation.

These may originate, for example, from the hard process at NLO, but are primarily generated during the PS and the production cross-section of these jets decreases exponentially with their p_{T}. This results in a high purity of bb events compared to background events. The disadvantage is the significantly smaller production cross-section, since the chosen decay channel accounts only for about 2% of all $t\bar{t}$ events with exactly two jets [3]. Despite this, the signal purity is increased even more compared to, for example, a full dileptonic decay channel selection in [75], because the contribution from background processes such as $Z \to l^+l^-$ is reduced dramatically. In addition to this, the datasets under study are collected during the years 2015, 2016 and 2017 with the ATLAS experiment, which correspond to an integrated luminosity of $\mathcal{L}_{\mathrm{int}} = 80.5 \pm 1.6$ fb^{-1} [33]. This combined data sample is sufficiently large to contain a large enough amount of $e\mu$+2j events. Therefore, the measurement of $\epsilon_b^{\mathrm{data}}$ is not limited by the statistical uncertainty associated to the dataset. This motivates the chosen event selection and will be discussed further in Section 6.5.

Numerous changes have been made with respect to the previous iteration of this analysis, namely [75, 76]:

- Larger datasets and MC simulated samples which decreases their associated statistical uncertainties. Moreover, the simulated samples have been regenerated with an improved setup. Additionally, new samples of physics processes have become available that contribute to the target final state.

- The increased data and MC statistics available allow to increase the jet p_{T} range in this calibration from 300 GeV to 600 GeV. The current measurement is performed in the following nine p_{T} bins: 20−30 GeV; 30−40 GeV; 40−60 GeV; 60−85 GeV; 85−110 GeV; 110 − 140 GeV; 140 − 175 GeV; 175 − 250 GeV; 250 − 600 GeV.

- Instead of a BDT to select the signal events, this version only applies a certain selection cut discussed below [75, 76]. In general, employing a BDT can benefit the selection efficiency. However, the training of a BDT introduces additional sources of systematic uncertainty to the measurement. The event selection is potentially biased from the choice of MC generator used to generate the training and testing samples. Furthermore, when estimating PS and hadronisation model uncertainties by using a different MC generator, one would either have to retrain the BDT with that generator or apply MC-to-MC SFs to match the predicted performance of the nominal MC generator with the alternative generator of choice. This requires another strategy to avoid double counting of the corresponding systematic uncertainties. Moreover, the output distribution of a complex BDT is less likely to be well modelled compared to basic kinematic variables. The gain from using a BDT and dealing with the corresponding complications compared to using a simple cut on kinematic variables has been estimated to not be significant enough.
  First, the basic selection cuts on reconstructed objects in this analysis are:
  - Exactly two jets, each with $p_{\mathrm{T}} > 20$ GeV and $|\eta| < 2.5$;

  - An electron (positron) and an antimuon (muon), each with $p_{\mathrm{T}} > 28$ GeV. Because of the ID and ECAL conditions, electrons are required to have $|\eta| < 1.37$ or $1.52 < |\eta| < 2.47$. Muons must fulfil $|\eta| < 2.5$.

  The SR and CRs are defined by cutting on an additional variable that has been employed in this calibration. It is the invariant mass of the charged lepton-jet pairs $m_{j,l}$. From energy and momentum conservation, the sum of the four-momenta of a charged lepton, neutrino and b-jet from a top quark decay would give its initial four-momentum. Thus, the square of the summed four-momenta equals m^2_{top}. Since the neutrino cannot be detected, the invariant mass of the charged lepton and jet can be considered instead. Its distribution should have an endpoint at the top quark mass and reaches it in case the four-momentum of the neutrino is negligible. However, there are two possible pairings of charged leptons and jets per event. The chosen pairing strategy is to pair the charged leptons with jets such that the sum of the squared invariant masses is minimal. This choice is motivated by tests showing this strategy to create a SR with the highest purity in bb events [76]. By requiring this invariant mass to be below 175 GeV, which is roughly the top quark mass, the jet from this pairing is more likely to be a b-jet from a top quark decay. If $m_{j,l} > 175$ GeV, the jet is more likely to be a light-jet. Figure 6.7 shows the $m_{j,l}$ distribution for the leading jet in $e\mu$+2j events split either by the considered simulated processes or by the two jet flavour fractions representing the SR and CRs. The SR is defined by requiring $m_{j1,l}$ and $m_{j2,l}$ to be below 175 GeV, hence the selected phase-space is enriched in events with two b-jets, labelled bb. The three CR are defined by reversing either one of the two or both $m_{j_i,l}$ requirements and are thus labelled as bl, lb and ll, respectively [76]. This selection is illustrated in Figure 6.8 and predictions of the two jet flavour fractions are discussed in more detail in Section 6.6.1.

- The previous analysis considers the other dileptonic decay channels as well as events with exactly three jets [75]. The idea is to constrain the background, for example from $Z \to ll$ events, as well as radiation of additional jets and to increase the available statistics. But each of those channels has a smaller signal purity and thus is impacted more strongly by background events containing non-b-jets. Thus, the gain from this approach has been estimated to be insignificant compared to the additional amount of time and work required to properly account for all uncertainties introduced by these channels, one example being the $Z \to ll$ background.

- The estimation of the background from non-prompt leptons is performed following a data-driven strategy in which the charged leptons are required to have the same electric charge [76]. It is described in more detail in Appendix A.

- Updated recommendations by theorists on how to estimate uncertainties resulting from the imperfect modelling of physics processes [76]. This primarily involves upward and downward variations of scale choices in the ME and PS calculation

and uncertainties related to the parton distribution function of the proton.

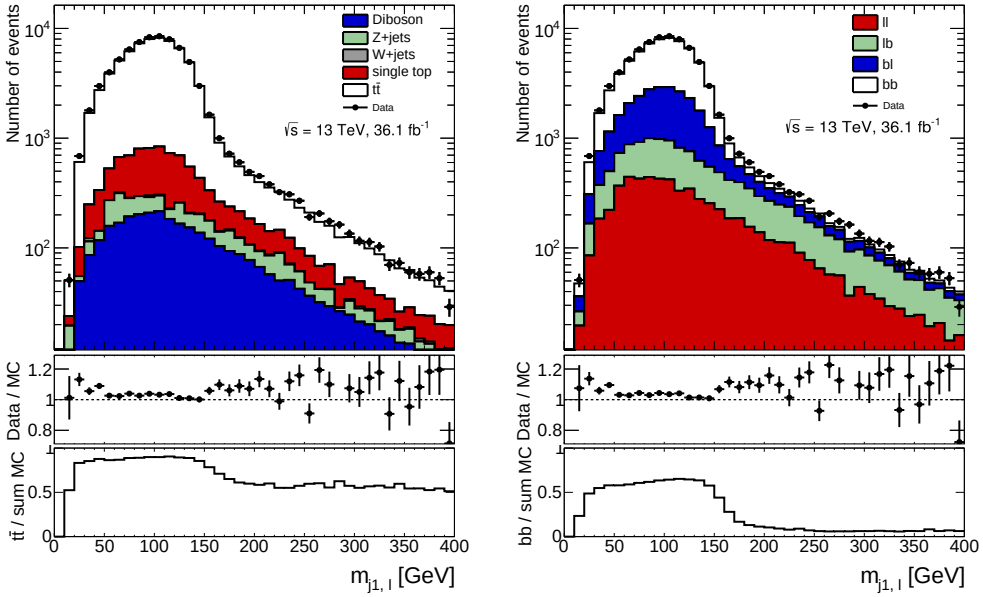

Figure 6.7: The invariant mass distribution of the leading jet and the charged lepton pair combined such that the sum of the squared invariant masses of both jet-lepton pairs in the event is minimal. $e\mu$+2j events are selected and the data corresponds to 36.1 fb^{-1} collected with the ATLAS detector in 2015 and 2016. Left: the distribution is shown split by the contributions from the individual simulated signal and background processes considered in this analysis. The bottom panel shows the data/MC ratio as well as the fraction of $t\bar{t}$ events out of all simulated MC events. Right: the distribution is shown split by the SR and three CRs. The bottom panel shows the data/MC ratio as well as the fraction of events containing two b-jets out of all simulated MC events.

6.3.1 Object reconstruction

In this measurement, electrons, muons and jets are reconstructed on the basis of the details given in Chapter 5. τ-leptons are not considered or reconstructed as a final state objects, but contribute to the selected events by decaying into electrons or muons or by being mis-identified as another object. In addition to those basic reconstruction algorithms, further quality criteria are required to optimise the object reconstruction efficiency for this analysis [76].
Electrons and muons must both pass certain, but different quality criteria, each summarised under the label 'TIGHT', which are defined and described by the corresponding working groups in ATLAS [52, 54]. The term TIGHT is suggestive in that electrons and muons passing these quality criteria are very likely to be prompt leptons which are those

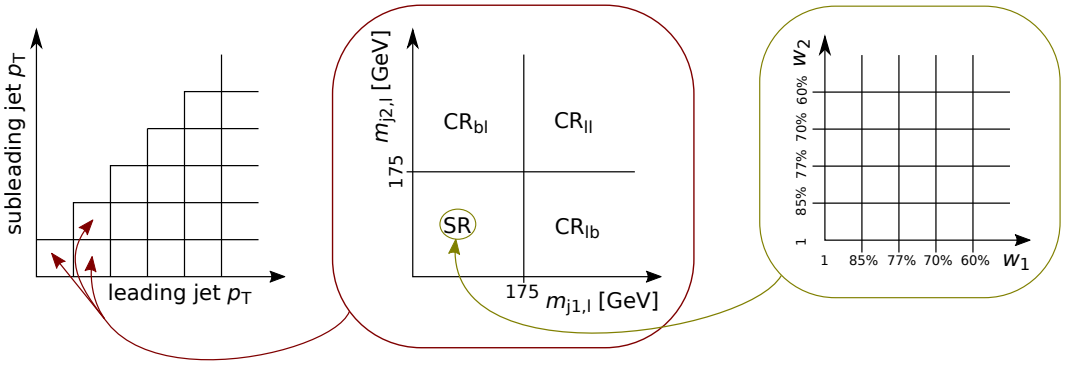

Figure 6.8: A schematic of how the various signal and control regions of each event are selected according to the leading and subleading jet and the corresponding $m_{j,l}$ quantities. The b-tagging efficiencies are then extracted for leading and subleading jet from the SR for each WP. This schematic is included with the courtesy of Julian Schmoeckel (DESY).

originating from W and Z boson decays. Furthermore, both electrons and muons must pass certain isolation criteria which are defined as 'GRADIENT' in Refs. [52,54], ensuring that both leptons are sufficiently separated from other objects and do not overlap with significant energy deposits in the calorimeters or tracks with a high p_T. The $p_T > 28$ GeV cut mentioned above serves the purpose of avoiding significant trigger efficiency uncertainties in the turn-on region of those un-prescaled single lepton triggers having the lowest p_T requirement [78], which is 26 GeV in the data samples 2016 and 2017. For both leptons, the trigger with the highest p_T threshold has no isolation requirement. If a so-called 'bad' muon is found, the whole event is discarded. A bad muon is a muon reconstructed either from highly energetic jets producing sufficiently high hit multiplicities in the MS from punch-through or from badly measured tracks in the ID associated to a jet and wrongly matched to segments of the MS [54].

Considering the trigger system, at least one of the charged leptons is required to cause at least one of the single lepton triggers [78] to fire for a candidate event to be stored. Furthermore, the leptons passing the trigger algorithms are then matched to the reconstructed leptons that pass the event selection. More details about the trigger selection is given in [76].

Similar to the b-tagging efficiency SFs this analysis aims to measure, several SFs exist for electrons and muons, namely to correct the simulated efficiencies to those in data for the reconstruction, trigger, identification and isolation of both leptons. These SFs are derived in $Z \to ll$ events [52,54]. Additionally, the calorimeter energy of electrons is calibrated to the true electron energy in simulation and remaining differences between simulation and data in the electron energy scale and resolution are corrected using the $Z \to ll$ mass peak of $m_Z \approx 91$ GeV [52].

Jets reconstructed from the detector information based on Section 5.3 also must fulfil the 'LooseBad' cleaning criteria defined in Ref. [79]. In addition to this, the difference between the simulated JVT efficiency [61] and the efficiency in data is corrected, similar to previously discussed object algorithm efficiencies.

Finally, as shown in Section 4.5, the effect of pile-up has to be taken into account when analysing LHC data. During the data recording periods of 2015 to 2017, an average of $\langle\mu\rangle = 32$ proton-proton collisions occur simultaneously per bunch crossing, which adds more background jets that could interfere with the primary interaction of interest. The basic simulated pile-up profile is corrected to the actual collision data by applying appropriate weight factors [49]. The pile-up conditions of the LHC are different for each year of data taking and thus need to be accounted for via different profiles as depicted in Figure 4.5.

The number of $e\mu$+2j events are 76204 in the combined years of 2015 and 2016, while there are 83718 events in the 2017 data collection, totalling 159922 events for all three years [76]. From these, 68441 and 74693 events, respectively, summing up to 143134 events, fall into the SR and are used to extract the b-tagging efficiency [76].

6.3.2 MC generator setups and simulation specifications

Several MC generator setups are used to simulate the various signal and background processes that could contribute significantly to the $e\mu$+2j final state. As indicated in Section 4.5, all nominal MC samples were processed through the complete ATLAS detector simulation infrastructure which is based on GEANT 4 [43, 44]. In addition to this, alternative setups are used to estimate modelling uncertainties; most of these alternative samples were processed through a faster simulation procedure called AtlFast2 [80]. For each simulated sample and MC generator setup, the corresponding production cross-section is taken from the setup itself at the order of perturbation theory that the sample is generated with.

The nominal dileptonic $t\bar{t}$ sample is generated with POWHEG-BOX v2 (POWHEG) [81–84] for the matrix element (ME) using the PDF4LHC parton distribution functions [85]. This setup as well as all other setups employed to model processes involving top quarks in the remainder of this chapter assume a top quark mass of $m_{\text{top}} = 172.5$ GeV. The hard process is then interfaced to PYTHIA 8 which generates the parton shower (PS) and hadronisation using the A14 tune [48]. The decay of heavy flavour hadrons is performed by the EVTGEN package [86]. An important parameter, h_{damp}, sets the cut-off scale for the first gluon emission in the $t\bar{t}$ simulation. The generator setup used to produce this sample fixed the parameter to $1.5 \cdot m_{\text{top}}$ which was found to give the best modelling of the top and $t\bar{t}$ system p_{T} [87].

To estimate the uncertainty from the choice of the generator setup, alternative $t\bar{t}$ samples were generated using either POWHEG interfaced to HERWIG 7 [40] for the PS and hadronisation; or MADGRAPH5_aMC@NLO [88] for the ME interfaced to PYTHIA 8 with the A14 tune. Furthermore, the effects of the scale choices for ISR are estimated in two ways: firstly, the same nominal sample is used, while the renormalisation and factorisation scales are multiplied by 2 and a parameter variation of the A14 tune is performed ('Var3c') which reduces the radiation during the PS [48,87]. Secondly, a new $t\bar{t}$ sample is used where h_{damp} is set to $3 \cdot m_{\text{top}}$, the two scales are multiplied by 0.5 and another parameter variation of the A14 tune is performed to increase PS radiation. Both simulations are generated with the fast simulation setup and compared with a fast simulation sample of the nominal setup to assess this ISR uncertainty.

The $t\bar{t}$ samples contribute, as expected, the majority of events to the $e\mu$+2j final state. However, as mentioned at the beginning of this section, other processes contribute as well. The three most significant ones considered are the single top production in association with a W boson (Wt), diboson production ($WW/WZ/ZZ$+jets and off-shell Z contributions) and a Z boson decaying into a pair of τ-leptons in association with jets ($Z \to \tau\tau$+jets). All these samples, $t\bar{t}$ and non-$t\bar{t}$, contain $e\mu$+2j events with two prompt leptons, thus considered signal events, but they also contain background events in which at least one of the leptons is non-prompt. The estimation of this non-prompt lepton background is detailed in Appendix A.

The Wt production is simulated using the diagram removal scheme [89] and is the most dominant non-$t\bar{t}$ sample that contributes to the signal process. The Wt sample is generated using the same generator setup as the nominal $t\bar{t}$ setup above. The generator and modelling uncertainties for this process are estimated in an almost coherent way to the $t\bar{t}$ case, except that the ISR uncertainty is estimated without an alternative h_{damp} sample and that an additional Wt sample is generated using the diagram subtraction scheme [90] to compare it with the nominal scheme. The difference between the two schemes is in the way they account for interference between the single top and $t\bar{t}$ Feynman diagrams.

The inclusive $Z \to ll$+jets process can contribute to the signal via the Z boson decaying into two τ-leptons that further decay into the $e\mu$ final state. The nominal sample is modelled using the SHERPA v2.2.1 [36] setup with the NNPDF30NNLO [91] PDF set. The diboson process with additional jets covers the $4l, lll\nu, ll\nu\nu$ and $l\nu\nu\nu$ final state by decays of weak boson pairs. The baseline is modelled using SHERPA v2.2.1 and SHERPA v2.2.2, depending on the specific process, and using the same PDF set as the Z+jets setup.

The dominant production of $e\mu$+2j with a non-prompt lepton comes from a semileptonic $t\bar{t}$ decay, which is modelled in the same way as the nominal dileptonic sample. Another contribution comes from single top production in the t-channel and s-channel, which

are simulated similarly to the Wt process. Finally, the W+jets process contributes to this background as well if the W decays leptonically including a leptonically decaying τ-lepton. The nominal W+jets process is modelled in the same way as $Z \to ll$+jets.

Alternative samples for W+jets, Z+jets and diboson are generated with POWHEG +PYTHIA 8 using the AZNLO tune [92] and the CTEQ6L1 PDF set [93]. Finally, another alternative sample for Z+jets is considered which is modelled by MADGRAPH 5 [88] for the ME interfaced to PYTHIA 8 with the A14 tune.

All these MC generator setups, those producing the nominal samples as well as the alternative ones, are critical and contribute a significant amount to the total systematic uncertainty which dominates this measurement. These uncertainties are further discussed in Section 6.5.3.

6.4 Sanity, closure and stress tests of the PDF method

Different tests have been performed to evaluate the robustness and precision of the new PDF likelihood method presented in Section 6.2. The tests work as follows:

1. Pseudo-data with a known b-tagging efficiency ($\epsilon_b^{\text{truth}}$) is created by MC generators;
2. The LLH functions from Equations 6.7 and 6.8 are fitted to the pseudo-data;
3. The POI $\epsilon_b^{\text{pseudo-data}}$ is extracted from the fit and compared to $\epsilon_b^{\text{truth}}$, similarly to the measurement based on real data.

6.4.1 Sanity tests

At first, a very simple sanity check is performed. The LLH functions are used directly to build the pseudo-data from simulation. Thus, the PDF method should extract $\epsilon_b^{\text{truth}}$ from the fit or, correspondingly, the SF $\epsilon_b^{\text{measured}}/\epsilon_b^{\text{truth}}$ should be unity. This test should find its input and thus a closure of the distributions before and after the fit is expected. As shown in Figure 6.9, this is exactly the case.

After this, a second sanity test is employed. In this check, the pseudo-data is fluctuated according to a Poisson distribution before the PDF method is applied. One would expect the SFs from the fit to still be compatible with unity within the statistical uncertainty of the pseudo-data. And indeed, this is observed in Figure 6.10. Moreover, this serves as a validation of the method with which the statistical uncertainty in data is estimated, described in Section 6.5.1.

6.4.2 Closure tests

In Section 6.2, the LLH functions were derived with the assumption that the b-tagging weight of a b-jet depends only on the p_{T} bin they fall into. However, this is only an

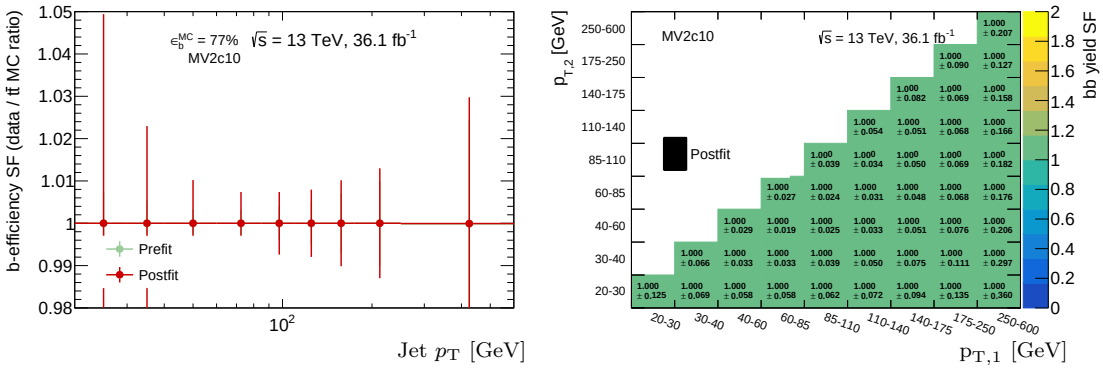

Figure 6.9: The *b*-tagging efficiency SFs (left) and the *bb* yield correction factor (right) for the MV2c10 algorithm at the 77% WP. This test is performed using pseudo-data directly created from the LLH functions in Equations 6.7 and 6.8. Thus, a perfect closure is expected and observed.

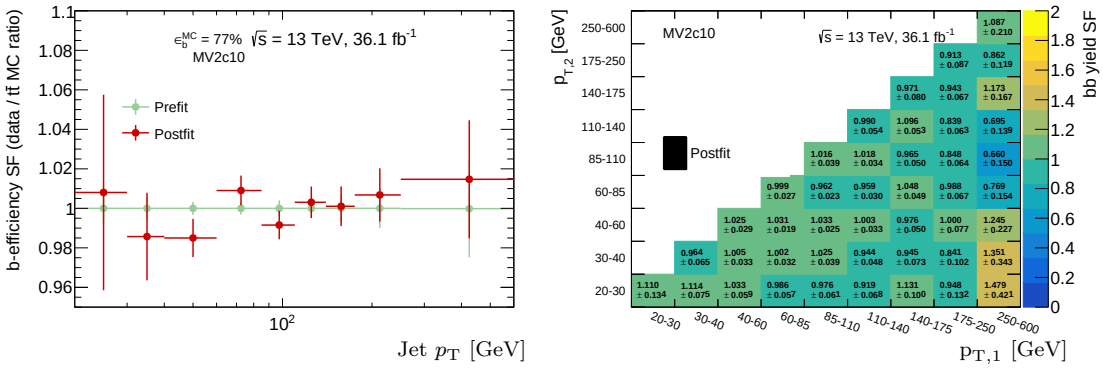

Figure 6.10: The *b*-tagging efficiency SFs (left) and the *bb* yield correction factor (right) for the MV2c10 algorithm at the 77% WP. This test is performed using pseudo-data directly created from the LLH functions in Equations 6.7 and 6.8, but the pseudo-data is fluctuated using a Poisson distribution before the fit. Thus, a perfect closure within statistical uncertainty of the pseudo-data is expected and observed.

approximation for actual collision data, since the MV2c10 algorithm is trained on numerous variables that may be correlated to, but not entirely dependent on, the p_T of the jet such as the pseudo-rapidity η. For example, both the p_T and η distributions of the leading and subleading jets may be different in any given p_T bin of the calibration [76]. Furthermore, there may be additional or hidden variables that have different distributions for the leading and subleading jets and thus affect the *b*-tagging algorithm, but are not considered at all [76].

To assess the impact of this assumption, a closure test is performed as follows: the nominal simulated samples of the considered physics processes are used both as the pseudo-data as well as the MC input to the LLH function. In this test, however, the LLH function uses a significantly restricted fraction of information compared to the total MC. This test is depicted in Figure 6.11 which indeed shows a small non-closure effect. Its highest value is 0.3% in the lowest p_{T} bin which is expected, because the b-tagging efficiency variation is the highest from the lower to the upper bin threshold, i.e. 20 GeV to 30 GeV, with respect to other bins. Despite this small observed non-closure effect, it is always significantly smaller than the Poisson uncertainty that results from the limited pseudo-data statistics. Therefore, this effect is considered to be negligible.

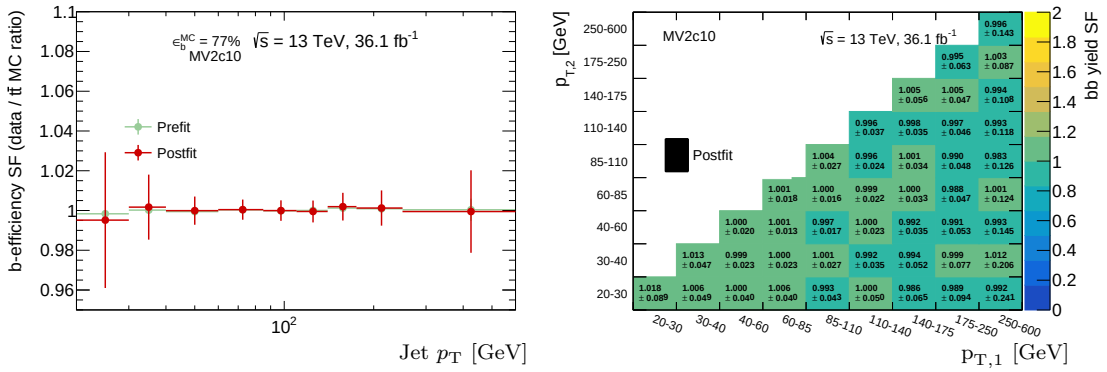

Figure 6.11: The b-tagging efficiency SFs (left) and the bb yield correction factor (right) for the MV2c10 algorithm at the 77% WP for a closure test. In this test, the same MC generated samples are used to create the pseudo-data and MC inputs to the LLH functions in the fit. Here, the LLH functions only include a significantly smaller fraction of information compared to the total MC. A small non-closure effect is expected, because of the physical differences between leading and subleading jets within any given p_{T} bin, which is also observed.

6.4.3 Stress tests of the simulation

This final test measures the impact that the MC generator choice has on the PDF fit method. The idea is to use an alternative $t\bar{t}$ sample combined with the nominal non-$t\bar{t}$ samples as pseudo-data. The ϵ_b value extracted from the fit is then compared to $\epsilon_b^{\mathrm{truth}}$ to assess the dependence of the method on the MC generated input samples. The alternative $t\bar{t}$ sample chosen for this stress test was generated using Sherpa v2.2.1, because of three reasons:

1. Compared to the previous test, the alternative $t\bar{t}$ sample leads to a controlled difference between pseudo-data and simulated inputs for the LLH function.

2. The alternative $t\bar{t}$ sample generated with SHERPA is not used for the estimation of the modelling uncertainties in Section 6.5.3.

3. The MC generator setup of SHERPA is able to simulate the $t\bar{t}$ process beyond NLO precision in the ME. Because of this, the modelling of additional jets, especially light-jets, is improved and thus the setup is believed to predict data more precisely.

This procedure is performed twice, namely for the old PDF approach and the new method. The mis-identification efficiency of light-jets as b-jets, called mis-tag rate, is corrected in all MC samples by applying appropriate SFs derived from an individual calibration in ATLAS, detailed in Ref. [94], as well as MC-to-MC SFs that bridge the difference between POWHEG +PYTHIA 8 and SHERPA v2.2.1. Therefore, the mis-tag rate for both generator setups should correspond to the one found in data. The comparison of the two PDF fit methods is shown in Figure 6.12.

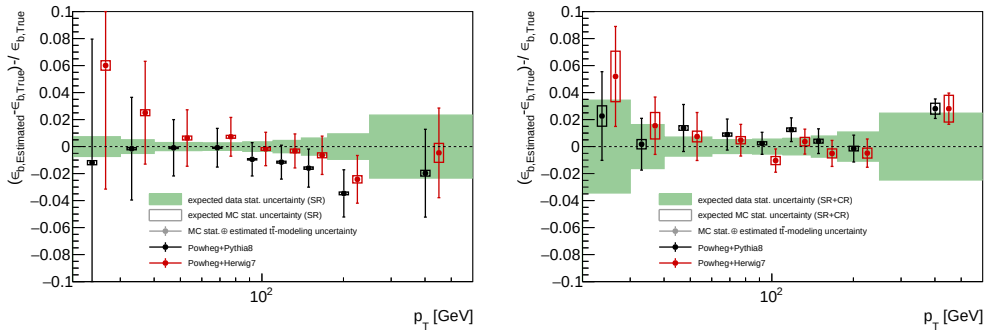

Figure 6.12: The PDF method is performed twice using an alternative $t\bar{t}$ sample generated with SHERPA v2.2.1 combined with the remaining nominal non-$t\bar{t}$ samples as pseudo-data. The extracted b-tagging efficiency is compared to the true one from the MC input. The left plot shows the result with the old LLH functions, while the right plot shows the results with the new LLH functions. The green band represents the statistical uncertainty of the pseudo-data, taken from the MINUIT package [95] while assuming Poissonian errors. The statistical uncertainty of the $t\bar{t}$ MC input is represented by an open rectangle and estimated using the bootstrap resampling technique [96] described in Section 6.5.2, which means performing the fit for each bootstrap replica and feeding it into the LLH function while leaving the pseudo-data unchanged. The vertical error bars are given by the sum of this squared statistical MC uncertainty and the squared $t\bar{t}$ modelling uncertainty that is taken from the actual measurement based on real data. This procedure is done for the nominal $t\bar{t}$ sample generated with POWHEG interfaced to PYTHIA 8 and repeated for an alternative setup, namely POWHEG +HERWIG 7.

Here, the statistical uncertainty of the pseudo-data is estimated in the same way as the actual data in the calibration, that is by using MINUIT [95] and assuming Poissonian errors, which is given by the green band. The statistical uncertainty of the $t\bar{t}$ MC sample is represented by the open rectangle and estimated using the same bootstrap resampling technique [96] described in Section 6.5.2. This means the measurement is repeated and fed into the LLH functions for each bootstrap replica while leaving the pseudo-data and non-$t\bar{t}$ samples untouched. The vertical error bars are the sum of the squared $t\bar{t}$ MC statistical uncertainty and the squared $t\bar{t}$ modelling uncertainty that is quoted in the actual calibration based on real data. Apart from these, no other uncertainties are taken into account, because only those related to the $t\bar{t}$ inputs are considered relevant for this comparison.

The comparison of the two fit methods exhibits the following features:

- The statistical uncertainty of the pseudo-data is slightly higher with the new fit method. This is expected, because more information is extracted from data via the fit, namely the two jet flavour correction factors.

- The statistical uncertainty of the MC input is higher as well. But this is also expected, since more information is extracted from the MC samples in the new approach. In particular, the two jet flavour fraction templates i.e. bb, bl, lb, and ll, are also needed for the LLH of the CRs to constrain them further in the likelihood fit, as opposed to the previous method which only used the SR to do so.

- The new method highlights one of its key motivations, namely that the dependence on the modelling of the $t\bar{t}$ inputs is significantly smaller compared to the old fit method. The old configuration shows a negative slope as a function of the jet p_{T}, while the new one does not. The reason for this is the different jet flavour compositions predicted by SHERPA compared to the prediction by POWHEG interfaced to PYTHIA 8. This could hint at a problem, namely that the central values of the b-tagging efficiencies in the calibration are biased if the jet flavour composition is different between actual data and simulation. With the new approach, however, the CRs are used to additionally constrain the jet flavour compositions in the SR and thus, the new extraction method produces less biased b-tagging efficiency SFs.

- The new fit shows a peculiar behaviour in the last p_{T} bin, but this has been found to originate from the statistical uncertainty of the $t\bar{t}$ sample simulated with SHERPA. The problem is that its statistical uncertainty from the simulation is larger than the Poissonian uncertainty that is assumed in the determination of the pseudo-data statistical uncertainty. In other words: $\sqrt{N} < \sqrt{\sum w^2}$, where w refers to the MC event weights of the SHERPA $t\bar{t}$ sample.

Finally, the significantly decreased dependence on the modelling of the $t\bar{t}$ inputs outweighs the slightly increased statistical uncertainties of the data and MC inputs to the fit. Consequently, the total combined uncertainty is reduced with the new fit approach.

To conclude, the new PDF method behaves as expected. Both sanity and closure tests are passed without any significant issues and thus, no additional uncertainty is added to the measurement [76].

6.5 Uncertainties

This section presents which uncertainties of both statistic or systematic origin are considered to have an impact on the calibration and how these uncertainties are estimated.

6.5.1 Statistical uncertainties in data

The minimisation of the LLH function is performed using MINUIT [95]. In order to estimate the statistical uncertainties of the fitted parameters arising from data, the error matrix of the fit is used which is given by the two processors MIGRAD and HESSE [76]. However, these uncertainties are correlated between the fitted parameters. On one hand, this is because the fit is performed in two dimensions according to the p_{T} bins of the two jets, meaning $(p_{\mathrm{T},1}, p_{\mathrm{T},2})$. On the other hand, the b-tagging efficiency is extracted for a pseudo-continuous weight discriminant distribution in five bins which must sum up to unity.

In order to work with uncorrelated statistical uncertainties, a principal component analysis method is used [97]. For this, the error matrix is diagonalised and the statistical uncertainty from data is split into 36 uncorrelated components, which corresponds to the product of the nine jet p_{T} bins and the four independent b-tagging discriminant bins. The 36 components are then summed in quadrature to obtain the total statistical uncertainty from data. Further details of this procedure are given in Ref. [76].

This uncertainty ranges from $< 1\%$ around a jet p_{T} of 100 GeV up to about 4% at very low and very high jet p_{T} values. The impact on the measurement is only significant at high jet p_{T}, however, because the systematic uncertainty dominates at low jet p_{T}.

6.5.2 Statistical uncertainties in the simulation

The statistical uncertainties in the MC generated distributions are not directly taken from the simulation itself via the sum of weights. Instead, an approach based on the bootstrap resampling method [96] is used. The idea is to create an ensemble of, in this case, one hundred statistically equivalent measurements. This is performed by randomly sampling a Poisson distribution with a mean of 1 for each nominal MC event and then multiplying this number to the weight of that event. The statistical uncertainty of the simulation is then understood as the standard deviation of the resulting distribution of all these measurements.

This uncertainty is of the order of a few % at low jet p_{T} and consistently below 1% for jet p_{T} above 40 GeV. With respect to the total uncertainty, this statistical uncertainty

is negligible over the full p_{T} range. The ensemble of equivalent measurements, however, is used to smooth certain systematic uncertainties, which is detailed in the following subsection. Finally, the total statistical uncertainty is given by summing the statistical uncertainties from data and simulation in quadrature.

6.5.3 Systematic uncertainties

The full procedure to measure $\epsilon_b^{\mathrm{data}}$ is repeated for each source of systematic uncertainty considered in this analysis while applying the corresponding systematic variation. In case there are two systematic variations of parameters, understood as 'up' and 'down' such as increased or decreased ISR during the PS, the following strategies are applied:

- If the two variations up and down affect the distributions in opposite directions, a symmetric uncertainty corresponding to their average is taken which is (up-down)/2;
- If both variations impact the distributions in the same direction or one variation is compatible with zero, the maximum of both i.e. $\max(\mathrm{up}, \mathrm{down})$ is taken as symmetric uncertainty instead;
- If there is only one variation, meaning it is one-sided, a symmetric uncertainty is constructed as the full difference between this variation and the nominal measurement.

Systematic uncertainties that are affected significantly by the limited available statistics of the simulated samples are smoothed using the large ensemble of statistically equivalent measurements created by the bootstrap resampling method introduced in the previous subsection [76]. This is important to avoid the statistical uncertainty of the simulated samples to enter the total uncertainty twice. This could happen, for example, when considering alternative samples that have smaller available statistics than the nominal samples, or when evaluating a systematic variation which causes a significant amount of events to migrate from one bin to another in a nominal sample [76].

The smoothing of systematic uncertainties as a function jet p_{T} is performed for the calibration of cumulative WPs in the following way [76, 98–100]:

1. For each systematic uncertainty, the statistical uncertainty in the simulated sample is estimated as explained in Section 6.5.2;
2. the affected histogram of the parameter as a function of jet p_{T} is rebinned, while the bins are assumed to be independent, until a certain level of significance is achieved. The chosen level of significance is $\geq 2\sigma$ from zero except in the two highest p_{T} bins where the uncertainty is kept even if it is insignificant. This allows to account for the behaviour at high jet p_{T} of the respective systematic uncertainty regardless of the statistical fluctuations in simulation;

3. the histogram is then rebinned back to its original binning, but the bin contents and errors are replaced by the significant ones determined in the previous step;
4. lastly, a smoothing based on a Gaussian kernel is applied to the histogram.

Physics modelling uncertainties

The dominant systematic uncertainties of this analysis stem from our limited understanding and thus modelling the involved physics processes. The simulation of collision events is subject to a certain degree of freedom, as explained in Chapter 4. This is because perturbation theory cannot be used to calculate the full process due to divergences arising from QCD principles. The choice of which proton PDF to use, the matching of the ME to the PS, the PS itself, the hadronisation process as well as the UE all involve strategies and parameters that are not given by first principles of nature.
The modelling uncertainties for the different physics processes considered are summarised in the following [76].

1. $t\bar{t}$ modelling uncertainties
   - In the previous iteration, the modelling of the ME and its matching to the PS was estimated by comparing the nominal sample generated with Powheg +Pythia 8 to an alternative sample generated with MadGraph5_aMC@NLO +Pythia 8. However, the latter setup exhibits a significant disagreement between the prediction and existing data to which the improved PDF method in this calibration is very sensitive. Consequently, the comparison of the two setups leads to an unreasonably large systematic uncertainty that completely dominates the measurement at low jet p_{T}. Therefore, this systematic variation is dropped. Instead, FSR uncertainties which the new PDF method is also more sensitive to are estimated as described below. On the other hand, cross-checks with an alternative sample generated with Sherpa indicate that the systematic uncertainty from the ME model only has a small impact on this measurement.
   - The uncertainty from choosing a certain PS and hadronisation strategy is estimated by comparing the nominal sample to an alternative sample generated from interfacing Powheg to Herwig 7 instead of Pythia 8.
   - The PDF used to assess the interacting particles within the protons is determined by following the recommendations for the PDF4LHC set [85].
   - The ISR uncertainty is estimated as described in Section 6.3.2.
   - The impact of the FSR scale choices is estimated similarly to the ISR case, namely via certain MC weights that multiply the corresponding renormalisation and hadronisation scales by either a factor of 2 (up) or 0.5 (down)
2. Single top modelling uncertainties
   - The uncertainties related to the ME, PS, hadronisation, FSR and proton PDF are estimated in the same way as for the $t\bar{t}$ process.

- The ISR uncertainty is estimated similar to the $t\bar{t}$ case which is described in Section 6.3.2. The estimation for the single top process is slightly different, because there is no additional sample with a different h_{damp} parameter. Therefore, only scale and tune parameter variations are considered.
- The uncertainty from choosing the diagram removal scheme is estimated by comparing it to another sample simulating the Wt process using the same generator setup, but employing the diagram subtraction scheme [90]. The difference between both schemes lies in the treatment of interference between single top and $t\bar{t}$ diagrams. This uncertainty is consistently below 1% and negligible for the total uncertainty.

3. Diboson, Z+jets and W+jets modelling uncertainties
   - Even though these backgrounds are sub-dominant, their modelling uncertainty is estimated by comparing the likelihood fit results based on the nominal SHERPA setup with results using the alternative generator setups specified in Section 6.3.2.

The modelling uncertainties from $t\bar{t}$ and single top range from $\approx 3-4\%$ at low jet p_{T} to 1% at high p_{T}. This is a significant improvement of the new PDF method that extracts the two jet flavour compositions from the fit in contrast to the old configuration [75,76]. The dominant components at low p_{T} originate from ISR and FSR as well as the PDF variations. At high p_{T}, the uncertainty from the choice of the PS and hadronisation models becomes significant as well. The modelling uncertainties from the other physics processes is only significant at low p_{T}, where they reach $\approx 2-3\%$.

Detector modelling uncertainties

Even though the reconstruction, identification and calibration of objects can never be perfect under the experimental conditions that exist in the physical world, the corresponding systematic uncertainties are small as long as the detector performance is modelled well. Therefore, the following uncertainties are considered due to the imperfect simulation of the ATLAS detector [76]:

- For electrons, the considered detector related uncertainties are: the electron energy scale and energy resolution; the reconstruction efficiency, identification efficiency, isolation efficiency and trigger efficiency [52];
- For muons, the following uncertainties are considered: the muon momentum scale and resolution; the identification efficiency, isolation efficiency, trigger efficiency as well as the track-to-vertex association efficiency [54];
- For jets, they are: the jet energy scale and resolution and the JVT efficiency [58,61];
- The light-jet and c-jet tagging efficiencies [64, 101];
- The modelling of the pile-up profiles [49]; and

- The background arising from mis-identified i.e. fake leptons or from non-prompt leptons. The strategy to estimate this background is described in Appendix A.

From these uncertainties, the ones related to the charged leptons and the JVT are negligible with respect to the total uncertainty. The light-jet and c-jet tagging efficiencies and the pile-up modelling uncertainties are only significant at low jet p_{T}, where they contribute $\approx 1-2\%$. The most dominant uncertainties are the jet energy scale and resolution, which amount to $\approx 6\%$ in the lowest p_{T} bin. The uncertainty from non-prompt lepton background contribution is only significant in the lowest p_{T} bin, where it reaches $\approx 2\%$.

Finally, since the PDF method has performed well throughout and passed all three tests described in the previous section, no additional systematic uncertainties are considered related to the extraction of the b-tagging efficiency [76].

6.6 Fit results

Before presenting the calibration results, the two jet flavour fractions are discussed in more detail below, as their definition and content are essential to the final result.

6.6.1 Two jet flavour fractions

The four two jet flavour fractions are based on whether the leading and/or subleading jet of an event is a b-jet or a non-b-jet and are labelled bb, bl, lb and ll, accordingly. Their definition was previously described in Section 6.3. The four fractions are shown for the SR of this measurement in Figure 6.13.

The bb fraction is $\approx 50\%$ or higher for $p_{\mathrm{T},2} > 30$ GeV and $\geq 75\%$ for $p_{\mathrm{T},2} > 60$ GeV. Lower bb purities are only found for the lowest $p_{\mathrm{T},2}$ bin, namely $20 < p_{\mathrm{T},2} < 30$ GeV. In those bins, the bb fraction is similar to the bl fraction, namely about 40% for $p_{\mathrm{T},1} > 40$ GeV, while the lb and ll fractions amount to roughly 10% and 5%, respectively. The bb fraction decreases as $p_{\mathrm{T},1}$ becomes smaller, while the ll fraction increases and even becomes the dominant component in the very first i.e. lowest p_{T} bin. The high bb fractions at high leading and subleading jet p_{T} is explained by the increasing performance of the two $m_{j,l}$ variables used to discriminate between the SR and CRs.

In contrast to this, the bl, lb and ll fractions in their respective CR are depicted in Figure 6.14. As expected from the $m_{j,l}$ selection, the CRs are enriched with their corresponding target events.
In CR_BL, the bl fraction is always above 30% and most often above 45%. The fraction decreases slightly if both jet p_{T} are below 40 GeV or above 85 GeV, because the ll and bb fractions increase in those areas, respectively. In CR_LB, the purity of lb events is above 40% for $p_{\mathrm{T},2} > 40$ GeV, but is decreased dramatically if $p_{\mathrm{T},2} < 40$ GeV, because the fraction of ll events increases at low jet p_{T}, as is the case in all regions. On the

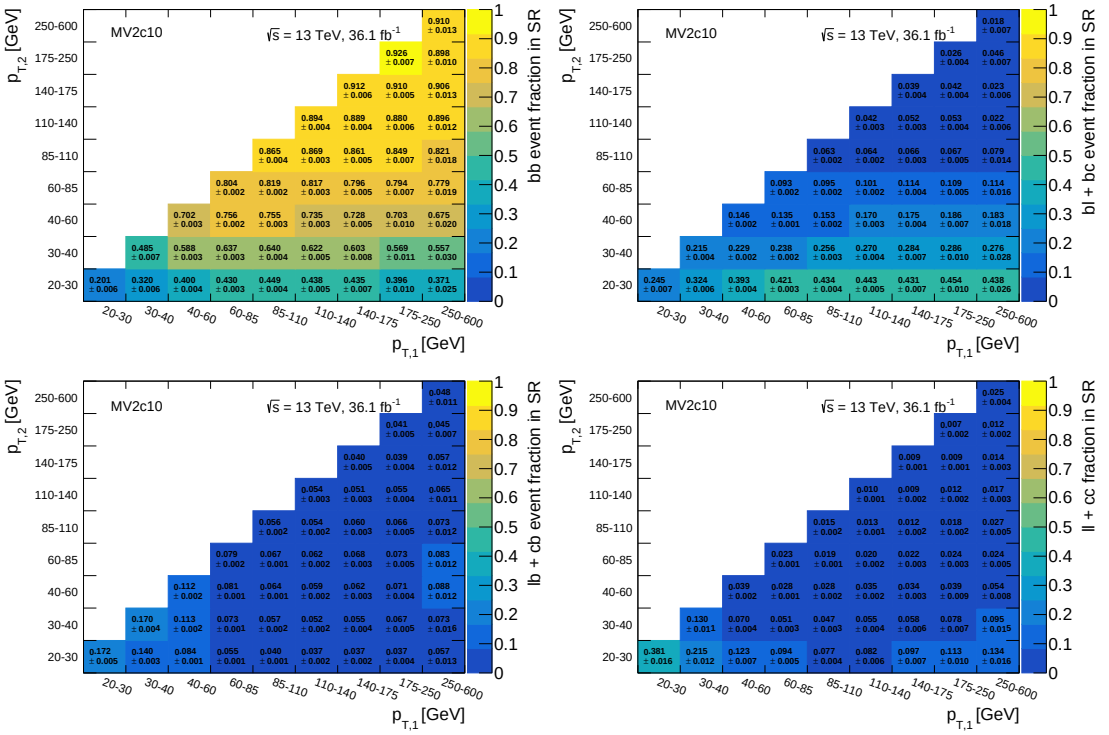

Figure 6.13: The fraction of bb, bl, lb and ll events in the SR as a function of leading and subleading jet p_{T} using MC predictions before the fit to data taken during the years 2015 and 2016 with the ATLAS detector [76].

other hand, the ll fraction in CR_LL is mostly above 50%, but decreases to $\approx 30\%$ if $60 < p_{\mathrm{T},1}, p_{\mathrm{T},2} < 110$ GeV, because the contamination from lb events is high in those bins [76].

In conclusion, the CRs are always enriched with the events that they are supposed to contain and overall behave as expected. For each two dimensional p_{T} bin, the estimation of the dominant background in the SR is improved by the corresponding CR which has a high purity. The least performing CR, which is CR_LB and thus is responsible for contributing to the lb estimation in the SR, especially for $p_{\mathrm{T},1} < 40$ GeV, suffers from a significant ll contribution in these bins. Despite that, the purity of lb events in this p_{T} range is still above 30% [76].

6.6.2 b-tagging efficiency and data/MC SF

In this analysis, a global χ^2 variable is constructed in order to evaluate the agreement of data and simulation before the fit. Furthermore, it is used to define a measure for the goodness of the fit. This is detailed in Appendix B. To summarise this study, the fit behaves as expected and the defined goodness-of-fit is a good estimator of the fit

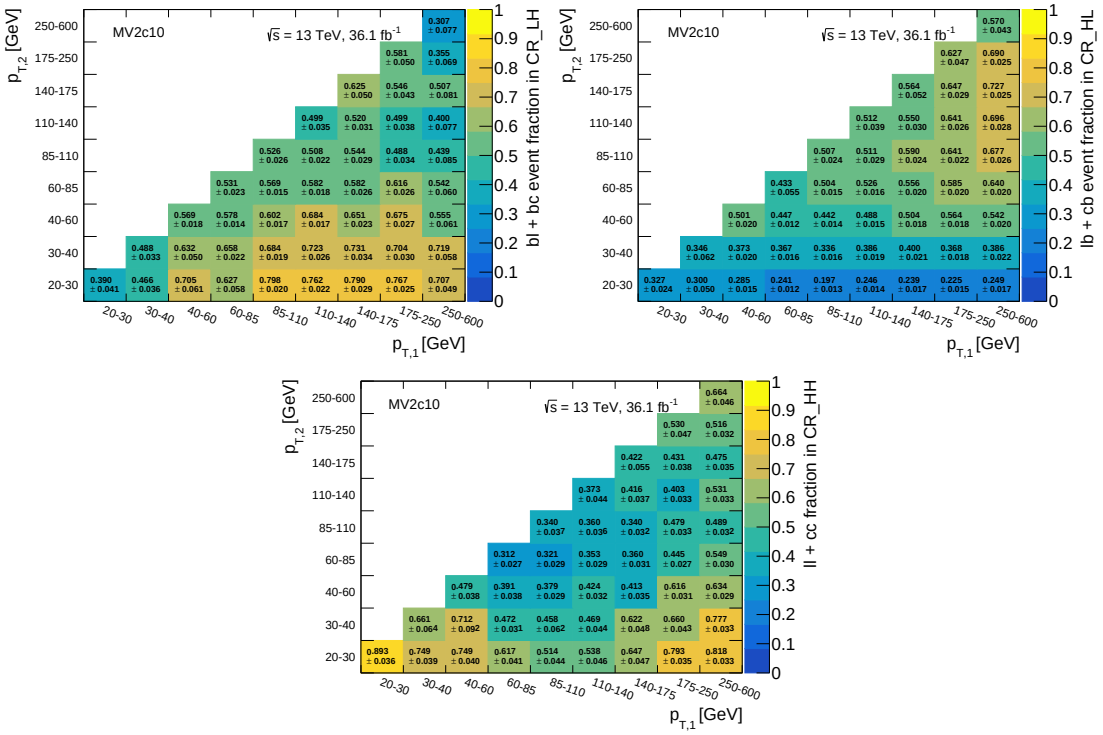

Figure 6.14: The fraction of bl, lb and ll events in their respective CR as a function of leading and subleading jet p_{T} using MC predictions before the fit to data taken during the years 2015 and 2016 with the ATLAS detector.

performance.

The following plots show the calibration results of the MV2c10 algorithm at the 77% WP. Figure 6.15 shows the b-tagging efficiency measurements based on the LLH fit to different datasets collected with the ATLAS detector at $\sqrt{s} = 13$ TeV. Figure 6.16 depicts the corresponding SF which is the ratio of the b-tagging efficiency measured in data with respect to the one found in the nominal $t\bar{t}$ MC sample. These MC efficiencies are derived using all $t\bar{t}$ events generated by POWHEG +PYTHIA 8 that pass the full $e\mu$+2j selection from Section 6.3. In Figure 6.17, the statistical, total systematic and total combined uncertainties are presented for the measurements. The most dominant systematic uncertainties for the measurement based on 2015, 2016 and 2017 data are plotted in Figure 6.18. The corresponding calibration results for the DL1 b-tagging algorithm are shown in Appendix C. Results for the other WPs as well as the pseudo-continuous calibration can be found in Ref. [76].

It is important to note that the b-tagging efficiency as a function of the jet p_{T} is affected by the WP definition, the training of the tagging algorithm and, of course, the perfor-

mance of the ATLAS tracking detector and corresponding algorithms. In the presented case of the MV2c10 algorithm at the 77% WP, the b-tagging efficiency starts at 60% for the lowest jet p_{T} and then smoothly increases to reach a plateau at 85% for a jet p_{T} $\approx$ 100 GeV. Starting from a jet p_{T} of about 200 GeV, the efficiency slightly decreases again. Despite this, the efficiency SFs are compatible with unity within one standard deviation over the full jet p_{T} range considered in this measurement. Additionally, the measurements based on 2015 and 2016 data are compatible with those based on 2017 data, again within one standard deviation.

The total uncertainty of the efficiency measurement starts at a maximum of 8-9% in the lowest jet p_{T} bin, decreases to $\approx$ 1% at a jet p_{T} of $\approx$ 100 GeV and then starts to increase again, ending up at 3-4% for high jet p_{T}. At jet p_{T} values between $20-100$ GeV, the total uncertainty is dominated by the systematic uncertainty, namely by the jet energy scale and resolution as well as the $t\bar{t}$ and Z+jets modelling uncertainties. At higher jet p_{T} values, the statistical uncertainty from data starts to become significant and eventually the limiting factor of this analysis. This is in contrast to the previous calibration which was limited by the systematic uncertainty, in particular the modelling of the $t\bar{t}$ process, over the full jet p_{T} range [75]. This aspect motivates the combination of the datasets from the three years to help reduce the total uncertainty at high p_{T} which is highlighted in the comparison plot in Figure 6.17.

6.6.3 Conclusion

This chapter presented the new calibration strategy of the ATLAS b-tagging algorithm MV2c10 and its latest results based on data recorded in the years 2015, 2016 and 2017. The SFs for the DL1 algorithm are shown in Appendix C for the various datasets. The differences with respect to the previous PDF method [75] were described in detail throughout the sections. This new method holds a significant advantage over the old method in that it significantly reduces the dependence of the measurement on the MC modelling of the $t\bar{t}$ process. This is achieved by adding a new variable, namely the invariant mass of pairs of charged leptons and jets, to split the events passing the selection into a signal region and three control regions for additional constraining power; and by adding new terms to the likelihood function that allow to extract more information from data, namely the two jet flavour compositions. This improvement is reflected in the fact that the total uncertainty of this measurement is decreased by a factor of up to two compared with previous publications. To summarise the uncertainties of the presented results: the total uncertainty starts at $\approx$ 8-9% at low jet p_{T}, drops to a minimum of $\approx$ 1% for jet p_{T} values between $100 - 200$ GeV and then increases again to $\approx$ 3-4% at high p_{T}. However, the most important aspect is that the total uncertainty is not completely dominated by the imperfect $t\bar{t}$ modelling anymore which is in contrast to previous calibration methods. Instead, detector uncertainties and limited data statistics start to play a significant role as well. This means that this new calibration approach will gain a significant improvement in precision in the coming years of the LHC physics programme.

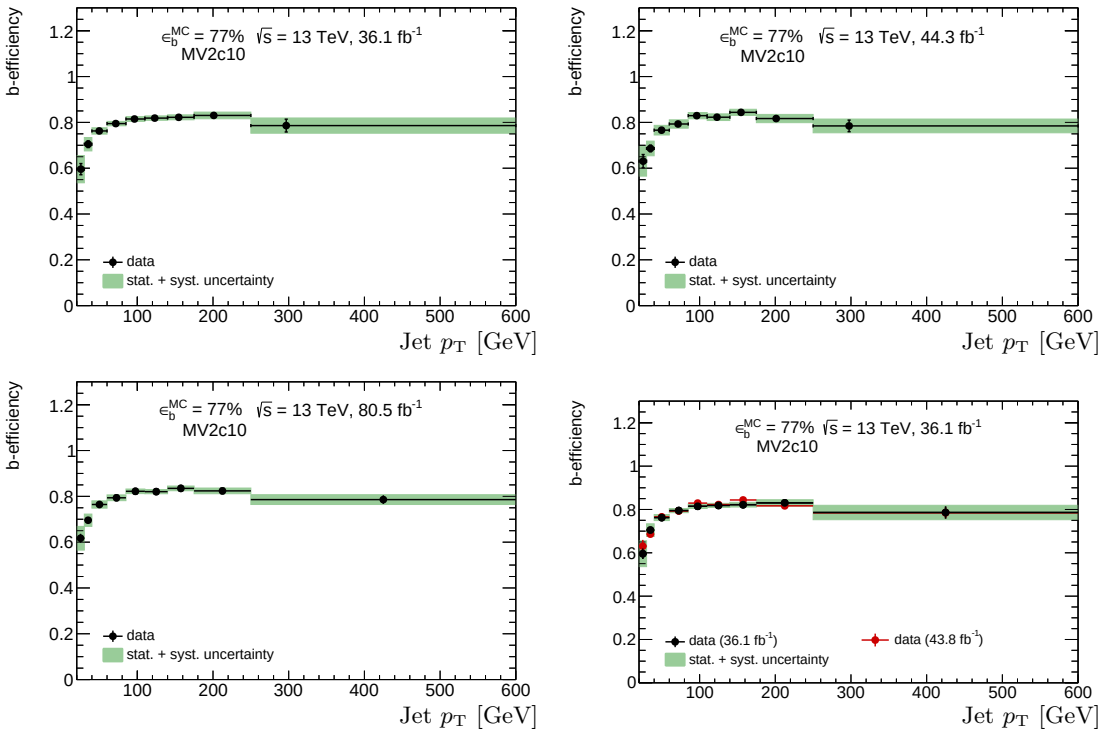

Figure 6.15: The b-tagging efficiency as a function of the jet p_T extracted from the LH fit method to data for the MV2c10 algorithm at the 77% WP. The fit result is shown based on ATLAS data from the years 2015 and 2016 (top left), 2017 (top right) and all three years combined (bottom left). In the bottom right plot, a comparison of the 2015 and 2016 results (black) to the 2017 result (red) is depicted. The individual fit results (top left and top right) show the bin centres at the mean jet p_T value of each bin. The vertical error bars include only the statistical uncertainties from data, while the green error band represents the sum in quadrature of the statistical and systematic uncertainties.

Another important aspect to consider for the remainder of this thesis is that the $t\bar{t}H(H \to b\bar{b})$ analysis presented in the following chapters employs the older calibration of the MV2c10 algorithm which is documented in Ref. [75]. This primarily affects the uncertainty on the b-tagging efficiency SFs.

As an outlook, efforts to improve current b-tagging algorithms and develop new strategies are ongoing constantly within the ATLAS collaboration. To give one example, the IP based algorithms introduced in Section 6.1.1 treat the tracks within a jet as uncorrelated. This is suboptimal, because b hadrons may decay into several charged particles that each possess a large IP. These IPs are correlated, since finding one track within a b-jet with

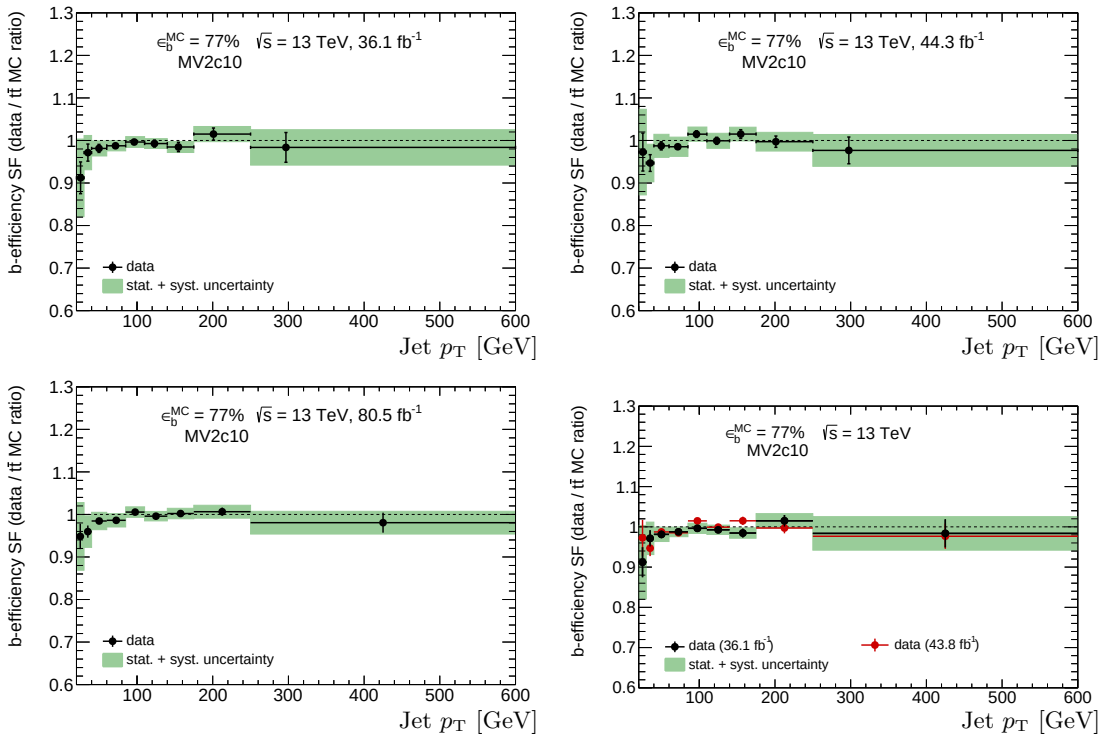

Figure 6.16: The b-tagging efficiency SF as a function of the jet p_T extracted from the LH fit method to data for the MV2c10 algorithm at the 77% WP. The fit result is shown based on ATLAS data from the years 2015 and 2016 (top left), 2017 (top right) and all three years combined (bottom left). In the bottom right plot, a comparison of the 2015 and 2016 results (black) to the 2017 result (red) is depicted. The individual fit results (top left and top right) show the bin centres at the mean jet p_T value of each bin. The vertical error bars include only the statistical uncertainties from data, while the green error band represents the sum in quadrature of the statistical and systematic uncertainties.

a large IP makes it more likely to find a second track with a large IP. For light-jets, on the other hand, such a correlation is not expected, because the tracks naturally tend to have smaller IPs. On the other hand, the IP3D algorithm uses likelihood templates to compute per-flavour conditional likelihoods and the computation of these templates require large sample statistics. Therefore, the IP3D algorithm assumes that the properties of each track are independent of all other tracks. To solve this problem, a Recurrent Neural Network (RNN) is applied to the IP algorithm. The resulting algorithm is called RNNIP [68, 70].

It has the same approach as the IP algorithms described in Section 6.1.1, but differs in the training. For each track, in addition to S_{z_0} and $S_{r\phi}$, the RNN takes the p_T fraction of

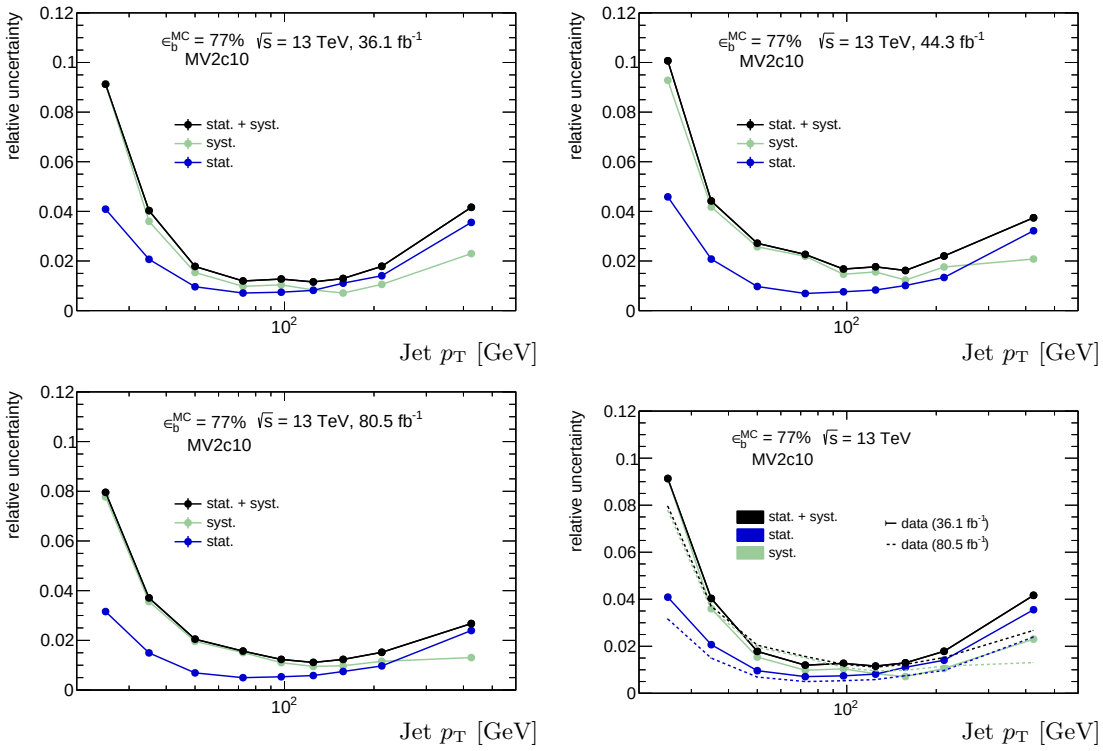

Figure 6.17: The uncertainty components relative to the extracted b-tagging efficiency as a function of the jet p_T from the LH fit method to data for the MV2c10 algorithm at the 77% WP. Shown are the statistical uncertainty from data (blue), the total systematic uncertainty (green) and their sum in quadrature (black). The uncertainty estimations are based on the fit to ATLAS data from the years 2015 and 2016 (top left), 2017 (top right) and all three years combined (bottom left). In the bottom right plot, a comparison of the 2015 and 2016 results (solid lines) to the combined result (dashed lines) is depicted.

the track with respect to the jet and the angular distance between them, $\Delta R(\text{track}, \text{jet})$. The output of the RNN is the different jet-flavour probabilities. With these, the final discriminant for the b-jet hypothesis is given by [68, 70]:

$$D_{\text{RNN}}(b) = \ln\left(\frac{p_b}{f_c p_c + (1 - f_c)\, p_u}\right). \tag{6.9}$$

In Equation 6.9, f_c represents the c-jet fraction during the training which was fixed to 7%. While jets originating from τ decays could in principle be considered, they have not been included in the RNNIP training [68, 70]. A new set of b-tagging algorithms, called MV2r and DL1r, have been developed and trained recently. They are based on the combination of IP3D and the RNNIP discriminant, instead of using the IP2D, and are

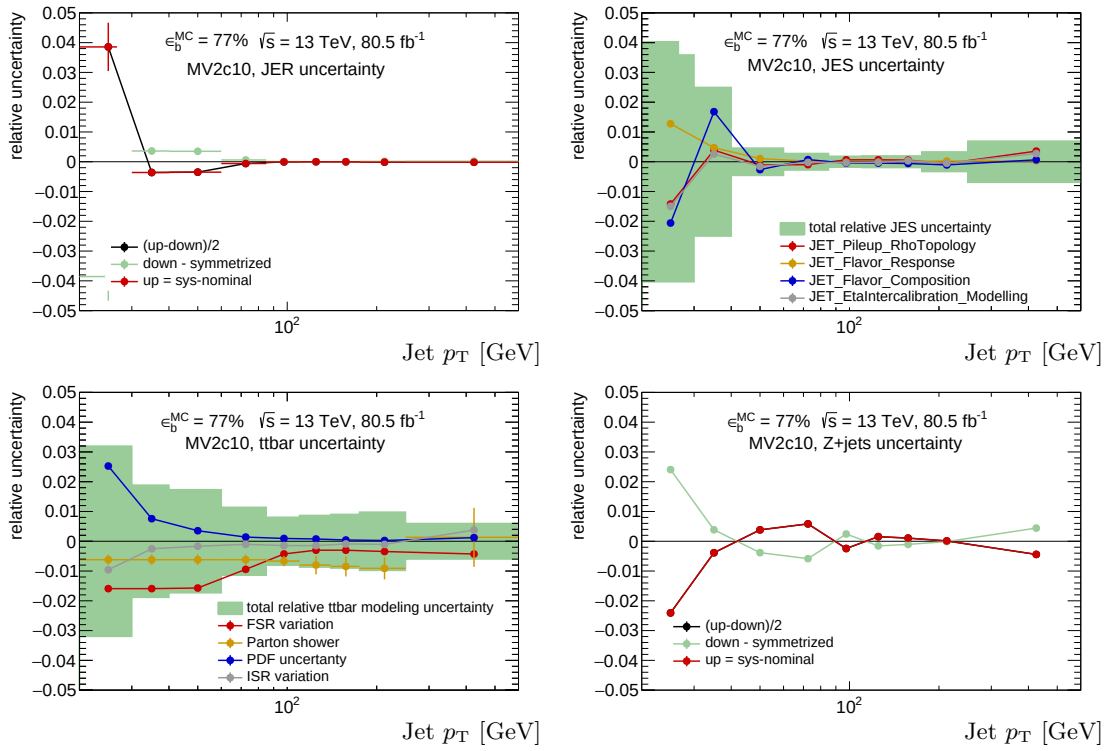

Figure 6.18: The dominant systematic uncertainties as a function of the jet p_T during the b-tagging efficiency measurement for the MV2c10 algorithm at the 77% WP. The uncertainties shown are from the fit results based on ATLAS data from the years 2015, 2016 and 2017. They represent the jet energy resolution (top left), jet energy scale (top right), $t\bar{t}$ modelling (bottom left) and Z+jets modelling (bottom right) as a function of jet p_T.

thus expected to perform better than the corresponding default b-tagging algorithms [68, 70]. Therefore, a dedicated calibration of these new RNNIP based tagging algorithms is a highly motivated effort in order to reduce b-tagging related systematic uncertainties in future ATLAS analyses. The statistical uncertainty component from data will be reduced by recording more collisions with the ATLAS experiment in future runs and using them in the calibration.

CHAPTER 7

The search for the $t\bar{t}H(H \to b\bar{b})$ process

In the remaining chapters of this dissertation, the ATLAS search for a SM Higgs boson produced in association with a $t\bar{t}$ pair at $\sqrt{s} = 13$ TeV is presented in which the Higgs boson decays into a $b\bar{b}$ pair [4]. The final state of this process, labelled $t\bar{t}H(H \to b\bar{b})$, is expected to consist of at least four b-jets at LO. Thus, a high performance of the b-tagging algorithm as well as small uncertainties on the b-tagging efficiency are crucial to the precision of this complex search.

The goal of this analysis is to measure the production cross-section of the $t\bar{t}H(H \to b\bar{b})$ process. In principle, this is done by counting the amount of such signal events in the recorded dataset. The predicted cross-section of the inclusive $t\bar{t}H$ process at $\sqrt{s} = 13$ TeV is $\sigma_{t\bar{t}H}^{\text{SM}} = 507^{+35}_{-50}$ fb at NLO accuracy in QCD including NLO electroweak corrections [15, 102–106]. The main challenge is to discriminate the signal events from the $t\bar{t}$ background which has a predicted inclusive production cross-section that is roughly three orders of magnitude higher, namely $\sigma_{t\bar{t}}^{\text{SM}} = 832^{+46}_{-51}$ pb at NNLO in QCD including the resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms [107–110]. This motivates an optimised event selection that includes as many signal events while rejecting as many background events as possible.

This $t\bar{t}H(H \to b\bar{b})$ search is performed in the semileptonic ('single lepton') and dileptonic decay channels of the associated $t\bar{t}$ pair. Furthermore, the analysis considers a so-called boosted decay channel as well in which the decay products of at least one top quark and those of the Higgs boson are boosted such that they cannot be properly separated in the detector and are, therefore, reconstructed as single large radius jets [4, 111]. This is a sub-channel of the semileptonic decay with only very few events passing the selection. However, those few events are removed from the resolved semileptonic decay channel to keep both measurements orthogonal to each other. Discussing the boosted channel

further, however, is outside the scope of this dissertation, therefore it is left out.

After the basic reconstruction and pre-selection, events are split into several analysis regions. These are multiple signal-enriched regions (SRs) in which the cross-section is measured as well as dedicated control regions (CRs) which are enriched in specific backgrounds and used to constrain associated uncertainties. These aspects are explained in more detail in Chapter 8. Since the expected signal-to-background ratio in this search is extremely small, this search employs several multivariate analysis techniques (MVAs) to discriminate signal from background events as well as possible. A combined profile likelihood method is used to simultaneously fit the MVA distribution in all SRs and other quantities in the CRs to extract $\sigma_{t\bar{t}H}$ from the data. The MVAs and fit method are described further in Chapter 9. Chapter 10 discusses the various sources of systematic uncertainties in this analysis and how they are estimated. The fit results are presented in Chapter 11 and compared to other $t\bar{t}H$ searches in ATLAS and CMS in Appendix F. Finally, various studies are presented in Chapter 12 that aim to reduce the most significant systematic uncertainties limiting future $t\bar{t}H(H \to b\bar{b})$ analyses.

The following sections give details about the recorded dataset and the modelling of the $t\bar{t}H$ signal as well as the $t\bar{t}$ and other background processes considered in this search. In general, the simulated samples are very similar to the ones in the presented b-tagging calibration analysis. That means most of the physics processes that contribute significantly to the selected final state are the same with a few exceptions such as the $t\bar{t}H$ signal process itself. In addition to this, most simulated samples are processed through the full ATLAS detector simulation [44] which uses the GEANT 4 toolkit [43], while the remaining samples are processed with a faster detector simulation strategy [80]. The baseline pile-up profile in each sample is corrected to the pile-up profile measured in data by applying specific weights to the MC events [49]. Lastly, the decay of heavy flavour hadrons in simulation is done by the EVTGEN package [86], except for samples generated with SHERPA [36].

7.1 Dataset

The presented search for the $t\bar{t}H(H \to b\bar{b})$ process is performed on proton-proton collision data recorded with the ATLAS experiment at $\sqrt{s} = 13$ TeV during the years 2015 and 2016 [4]. Only such events are considered where all relevant parts of the ATLAS detector were fully operational and where at least one vertex exists that has two or more tracks passing $p_{\mathrm{T}} > 0.4$ GeV. The primary vertex is then defined as the vertex with the largest sum of squared p_{T} of associated tracks. This helps to reduce the impact of pile-up, a concept that was introduced and discussed in Section 4.5. During the years 2015 and 2016, the LHC conditions were such that each collision involved roughly between 8 and 45 interactions per bunch crossing with an average of $\langle\mu\rangle = 24$ depicted in Figure 7.1. The integrated luminosity of this combined dataset is the same as in the b-tagging calibration, namely 36.1 ± 0.8 fb^{-1} [33].

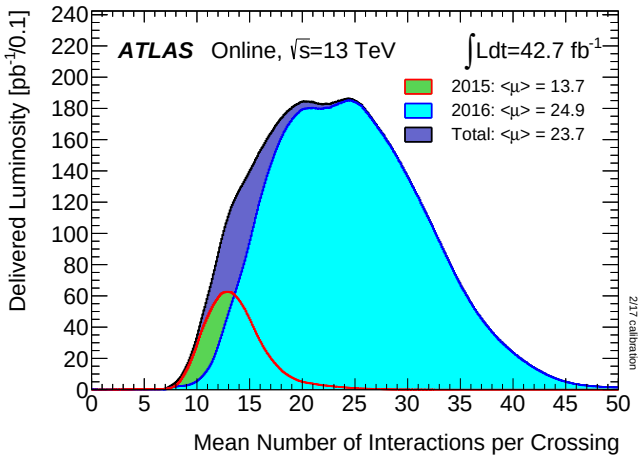

Figure 7.1: The pile-up profiles of the data recorded during the years 2015 and 2016 of Run 2 with the ATLAS experiment [45].

7.2 $t\bar{t}H$ signal

In this analysis, the $t\bar{t}H$ signal is simulated using MADGRAPH5_aMC@NLO v2.3.2 for the ME at NLO with the NNPDF3.0NLO PDF set [91]. This setup as well as all other setups employed to model processes involving top quarks in the remainder of this thesis assume a top quark mass of $m_{\text{top}} = 172.5$ GeV. This sample is then interfaced to PYTHIA 8 [41,42] with the A14 tune [48] which models the PS and hadronisation. The factorisation and renormalisation scales were both set to $\mu_F = \mu_R = H_{\text{T}}/2$, where H_{T} represents the scalar sum of the transverse masses of all final state particles, i.e. $\sum_i \sqrt{p_{\text{T},i}^2 + m_i^2}$. In order to preserve the spin correlations among the top quarks and their decay products, the top quarks in the ME are decayed with the MADSPIN programme [112]. In this sample, the mass of the Higgs boson is assumed to be $m_H = 125$ GeV and all of its decay modes are included while the corresponding branching ratios are computed by HDECAY [15,113]. The predicted cross-section of the inclusive $t\bar{t}H$ process is the same as the value given in the beginning of this chapter, namely $\sigma_{t\bar{t}H}^{\text{SM}} = 507^{+35}_{-50}$ fb at NLO accuracy in QCD including NLO electroweak corrections [15, 102–106].

7.3 MC driven backgrounds

The final state of the $t\bar{t}H(H \to b\bar{b})$ process involves a $t\bar{t}$ pair and an additional $b\bar{b}$ pair, the latter coming from the Higgs boson decay. This final state is identical to the $t\bar{t} + b\bar{b}$ process, i.e. the production of a $t\bar{t}$ and an additional $b\bar{b}$ pair via the strong interaction, making it an irreducible background. This is illustrated in Figure 7.2.

The fundamental difference between the $t\bar{t}H(H \to b\bar{b})$ and $t\bar{t} + b\bar{b}$ processes lies in

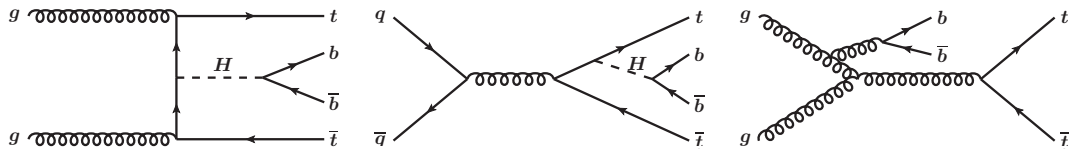

Figure 7.2: Feynman diagrams of the $t\bar{t}H(H \to b\bar{b})$ process initiated by gluons (left), quarks (centre) and the main background, which is $t\bar{t}+b\bar{b}$ production (right).

the event kinematics characterised by numerous different variables. This search tries to exploit these variables to train an MVA, namely a BDT which will be described in more detail in Section 9.1. Both processes involve eight final state particles at LO, some of which are heavy such as the four b-jets. However, in the particular case of the $t\bar{t}+b\bar{b}$ process, it is challenging for current state-of-the-art MC generators to simulate all of the kinematic properties with a high precision in perturbative QCD due to the additional gluon splitting. The CPU time required to simulate one such event with, for example, Sherpa is in the order of minutes. Considering that millions of events are necessary to perform a statistically significant and meaningful measurement, various analysis strategies are employed to optimise the modelling of the signal and background. The MC generation of this background is detailed in the following section.

7.3.1 $t\bar{t}$ +jets background

As the title suggests, the main considered background does not only involve $t\bar{t}+b\bar{b}$ production, but it includes the full $t\bar{t}$ +jets process. In this inclusive process, the additional jets in the event can also be c-jets or light-jets. The main reason to use this inclusive sample is to create dedicated CRs that have a lower jet and b-jet multiplicity. These CRs are depleted in signal events, but enriched in specific background events and are used to constrain the uncertainties associated to these backgrounds. One of these uncertainties is, for example, the mis-tag probability of c-jets and light-jets as b-jets.

The nominal $t\bar{t}$ +jets sample is generated using Powheg [81–84] with the NNPDF3.0NLO PDF set and $h_{\text{damp}} = 1.5 \cdot m_t$ [87] for the ME interfaced to Pythia 8 with the A14 tune during the PS and hadronisation process. This $t\bar{t}$ sample is inclusive, meaning that it contains the $t\bar{t}$ decay generated at NLO in the ME, but only includes additional jets beyond the first parton via radiation during the PS. At the time of this analysis, there was no specific mode of this generator setup to specifically model the $t\bar{t}+b\bar{b}$ process in the ME by itself and thus, this inclusive sample was studied instead. The fact that the additional jets in this sample only originate from gluon splitting during the PS, which is simulated at LO+LL precision and assumes b-quarks to be massless, means that the predicted cross-sections of different $t\bar{t}$ +jets categories as well as the kinematics of the expected additional jets are generated at a sub-optimal level of precision.

In this nominal $t\bar{t}$ +jets sample, the renormalisation and factorisation scales are both

set to an event dependent variable, namely the transverse mass of the top quark, i.e. $\mu_R = \mu_F = m_{\mathrm{T},t} = \sqrt{m_t^2 + p_{\mathrm{T},t}^2}$, with $p_{\mathrm{T},t}$ being the p_{T} of the top quark in the centre-of-mass frame of the $t\bar{t}$ pair. The predicted cross-section of the inclusive $t\bar{t}$ process is taken from the Top++2.0 programme [114]. For $\sqrt{s} = 13$ TeV, it is the same value quoted in the beginning of this chapter, namely $\sigma_{t\bar{t}}^{\mathrm{SM}} = 832^{+46}_{-51}$ pb at NNLO in QCD which includes the resummation of NNLL soft gluon terms [107–110].

Alternative generator setups are employed to produce more $t\bar{t}$ samples that are used for the evaluation of systematic uncertainties. They are described in more detail in Section 10.3.
This $t\bar{t}$ +jets background sample can be split into different categories of events. It is useful to choose a categorisation based on the flavour of the additional jets in order to estimate the corresponding uncertainties associated to the flavour-tagging of jets and to study the kinematic properties of individual event categories. This categorisation scheme was already employed in the $t\bar{t}H(H \to b\bar{b})$ search at $\sqrt{s} = 8$ TeV during Run 1 [115] and has been labelled Heavy Flavour Classification (HFC). This classification exploits the truth information within the simulated samples and is based on the number of jets with certain flavours where the flavour labels are applied as follows:

- A jet at particle level with $p_{\mathrm{T}} > 15$ GeV not coming from the top quark or W boson decay, i.e. an additional jet, is labelled a b-jet if one b hadron with $p_{\mathrm{T}} > 5$ GeV is found within $\Delta R = 0.4$ of the jet.
- If a second b hadron is found close to the jet, the jet is labelled as a B-jet instead. No p_{T} requirement is applied to the second hadron.
- If none such hadrons are found, this procedure is repeated in an analogous way for c hadrons.
- If no heavy flavour hadrons are found, the jet is labelled as a light-jet.

The first HFC is performed based on these jets, namely in the following way:

- An event with at least one additional b- or B-jet is labelled $t\bar{t}+ \geq 1b$. This is the dominant and irreducible background category in this analysis.
- If there are no additional b- or B-jets, but at least one additional c- or C-jet, the event is flagged as $t\bar{t}+ \geq 1c$.
- In all other cases, the event is labelled as $t\bar{t} +$ light.

As a second step of the HFC, the $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ events can be split up further, namely as follows:

- Events with exactly one additional b-jet (B-jet) are labelled $t\bar{t} + b$ ($t\bar{t} + B$).
- Events with exactly two additional b-jets are labelled $t\bar{t} + b\bar{b}$

- All other events in the $t\bar{t}+ \geq 1b$ category are flagged as $t\bar{t}+ \geq 3b$.
- Any events that only contain additional b-jets originating from MPI or FSR are placed into a separate category that is $t\bar{t} + b$(MPI/FSR).
- The $t\bar{t}+ \geq 1c$ category is further split up in an analogous way.

In this analysis, a crucial strategy is employed to improve the modelling of the $t\bar{t}+ \geq 1b$ background. The idea is to use the highest available precision of simulation and apply it to the nominal sample. While the Powheg +Pythia 8 setup is not able to generate the ME of the $t\bar{t} + b\bar{b}$ process, another setup is able to do so. This setup consists of Sherpa v2.1.1 interfaced to OpenLoops [116, 117] using the CT10 four flavour (4F) PDF set [118,119]. It simulates the full $t\bar{t} + b\bar{b}$ ME at NLO precision in QCD as well as the PS and hadronisation process. In this PDF set, the b-quarks are considered massive particles, which is referred to as the 4F scheme and accounts for effects in the simulation from the high b-quark mass. Thus, the predicted kinematics of the additional b-quarks within the $t\bar{t}+ \geq 1b$ category are believed to be more precise than those of the nominal sample.
In this sample, the renormalisation scale is set to the CMMPS value, meaning $\mu_R = \mu_{\text{CMMPS}} = \Pi_{i=t,\bar{t},b,\bar{b}} E_{\text{T},i}^{1/4}$ [116]. Both the factorisation scale as well as the resummation scale are set to $H_{\text{T}}/2 = \frac{1}{2}\sum_{i=t,\bar{t},b,\bar{b}} E_{\text{T},i}$. The resummation scale μ_Q sets an upper limit for the momentum transfer of gluon radiation and splittings during the PS. This sample is labelled Sherpa 4F from now on.
The currently available Sherpa setups do not offer the option to merge this 4F sample with an inclusive $t\bar{t}$ +jets sample and remove overlapping events. Therefore, the event fractions of the four sub-categories of $t\bar{t}+ \geq 1b$ in the Powheg +Pythia 8 sample are scaled to those found in this Sherpa 4F sample, instead. Since the $t\bar{t}+b$(MPI/FSR) sub-category is not modelled with NLO precision in the Sherpa 4F sample, it is excluded from this scaling procedure. The event fractions of the $t\bar{t}+ \geq 1b$ sub-categories for the two setups are compared in Figure 7.3, again excluding the $t\bar{t} + b$(MPI/FSR) sub-category. The $t\bar{t} + b$(MPI/FSR) events amount to $\approx 10\%$ of the $t\bar{t}+ \geq 1b$ events in the nominal sample. Additional studies regarding the modelling of the $t\bar{t} + b\bar{b}$ process are presented in Chapter 12.

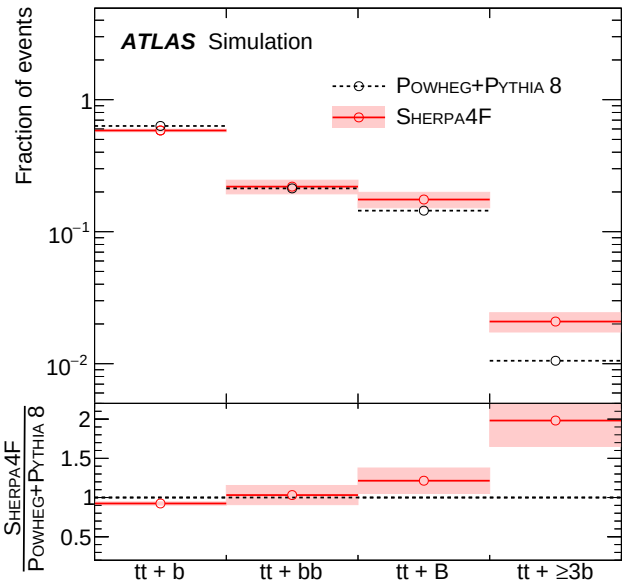

Figure 7.3: Shown are the relative event fractions of the four $t\bar{t}+ \geq 1b$ sub-categories, namely $t\bar{t}+b$, $t\bar{t}+B$, $t\bar{t}+b\bar{b}$ and $t\bar{t}+ \geq 3b$, before any event selection. The fractions are normalised to sum up to unity without considering the $t\bar{t}+b$(MPI/FSR) category. Here, the predicted fractions in the inclusive $t\bar{t}$ +jets sample from Powheg +Pythia 8 are compared to those in the dedicated $t\bar{t}+b\bar{b}$ sample generated by Sherpa 4F. The uncertainty band of the Sherpa 4F sample considers several sources which is discussed in Section 10.3.

7.3.2 Other backgrounds

Various additional background processes are considered, some of which are similar to those in the b-tagging calibration analysis. These involve the production of W+jets and Z+jets, which are simulated by the same setups given in Section 6.3. The Z+jets events falling into the $t\bar{t}+ \geq 1b$ or $t\bar{t}+ \geq 1c$ category are scaled by 1.3, a factor determined from dedicated CRs in data, namely e^+e^- and $\mu^+\mu^-$ events with 83 GeV $< m_{ll} <$ 99 GeV, with m_{ll} being the invariant mass of the charged lepton pair [4]. Corresponding W+jets events have been studied and found to not require additional scaling.
The Diboson background is generated using Sherpa v2.1.1, which is a different setup compared to the b-tagging analysis, described further in [120].
The single top process is generated with Powheg-Box v1 at NLO. While the Wt and s-channel are simulated using the CT10 PDF set during the ME, the t-channel employs the CT10 4F set and uses MadSpin to decay the top quarks. All three ME setups are interfaced to Pythia 6.428 [41, 42] with the Perugia tune [121] for the PS and hadro-

nisation. The cross-sections of the single top samples are normalised to approximate NNLO predictions given in [122–124].

In addition to the above processes, this analysis considers background processes with significantly smaller cross-sections, but similar final states. Samples of $t\bar{t}W$ and $t\bar{t}Z$, summarised as $t\bar{t}V$, are generated at NLO with MadGraph5_aMC@NLO +Pythia 8 with the NNPDF3.0NLO PDF set and A14 tune. The rare production of tWH is modelled with MadGraph5_aMC@NLO +Herwig++ [39] with the CTEQ6L1 PDF set. The production of tH in association with jets ($tHqb$), is simulated at LO with MadGraph5_aMC@NLO +Pythia 8 with the CT10 4F PDF set. Other backgrounds involving the simulation of Higgs bosons were found to be negligible and are thus not included [4]. The production of four top quarks ($t\bar{t}t\bar{t}$) and of $t\bar{t}WW$ is simulated with MadGraph5_aMC@NLO +Pythia 8 at LO in the ME. The tZ background is generated with the same ME setup, but interfaced to Pythia 6. Finally, the tZW process is simulated using MadGraph5_aMC@NLO +Pythia 8, but with a NLO ME.
Table 7.1 summarises the employed MC samples as well as the generator setups including PDF sets used to simulate them.

Sample	ME generator	PDF set	PS generator
$t\bar{t}H$	MadGraph5_aMC@NLO	NNPDF3.0NLO	Pythia 8
$t\bar{t}$ +jets	Powheg	NNPDF3.0NLO	Pythia 8
W+jets/Z+jets	Sherpa v2.2.1	NNPDF3.0NNLO	Sherpa v2.2.1
Diboson	Sherpa v2.1.1	CT10	Sherpa v2.1.1
Single top (Wt and s-channel)	Powheg	CT10	Pythia 6
Single top (t-channel)	Powheg	CT10 4F	Pythia 6
$t\bar{t}$ +V	MadGraph5_aMC@NLO	NNPDF3.0NLO	Pythia 8
tWH	MadGraph5_aMC@NLO	CTEQ6L1	Herwig++
$tHqb$	MadGraph5_aMC@NLO	CT10 4F	Pythia 8
$t\bar{t}$ $t\bar{t}$, $t\bar{t}WW$ and tZW	MadGraph5_aMC@NLO	CT10	Pythia 8
tZ	MadGraph5_aMC@NLO	CT10	Pythia 6

Table 7.1: The MC generator setups and PDF sets used to simulate the signal and background MC samples used in the $t\bar{t}H(H \to b\bar{b})$ analysis.

All mentioned samples in this subsection, except for the $t\bar{t}V$ samples, are labelled as non-$t\bar{t}$ in the plots and tables in the following chapters. Their contribution to the total background ranges from 4% to 15%, depending on the considered SR or CR.
Aside from these MC driven backgrounds, additional backgrounds are considered, but estimated using data driven techniques instead. These are discussed in the following section.

7.4 Data driven backgrounds

In this search, dileptonic and semileptonic $t\bar{t}$ decays are considered and thus, the charged leptons in the final state may be faked by other objects such as jets or photons or they may originate from other sources than W or Z boson decays. The estimation of this fake and non-prompt lepton background is performed in the single lepton channel using a data driven (DD) approach, namely a matrix method (MM) [125], explained in the following.

The goal is to estimate of the number of fake or non-prompt leptons that pass the identification and isolation criteria of this analysis. The nominal object reconstruction criteria will be described in more detail in the next chapter. For this estimation, a data sample is selected where the isolation criteria is removed and the identification criteria for electrons are loosened (LOOSE) [4]. This new sample should, therefore, contain a significant amount of fake and non-prompt leptons. The efficiency of leptons in this sample to pass the nominal (TIGHT) criteria is measured in data, but separately for prompt leptons (ϵ_r) and for fake and non-prompt ones (ϵ_f). ϵ_r is estimated using Z+jets events and ϵ_f is extracted from events with low values of $E_{\mathrm{T}}^{\mathrm{miss}}$ and reconstructed transverse mass of the leptonically decaying W boson. The latter quantity is defined as $m_{W,\mathrm{T}} = \sqrt{2p_{\mathrm{T}}^{\mathrm{lepton}} E_{\mathrm{T}}^{\mathrm{miss}} (1 - \cos\Delta\phi)}$.

With the two efficiencies, a matrix can be defined by the following system of equations:

$$N^l = N_r^l + N_f^l; \quad N^t = \epsilon_r N_r^l + \epsilon_f N_f^l, \tag{7.1}$$

where $N^l(N^t)$ is the total number of events with leptons passing the LOOSE (TIGHT) selection criteria and $N_r^l(N_f^l)$ is the number of events with a prompt (fake or non-prompt) lepton passing the loose criteria. This matrix can be inverted to extract the number of fake and non-prompt leptons passing the TIGHT criteria, given by $\epsilon_f \cdot N_f^l$. This information is applied in the form of a weight factor to all events with a lepton passing the LOOSE selection. This allows to extract the shape of kinematic distributions as well as the overall normalisation of this background. In the most signal-enriched regions in the single lepton channel, this background is found to be consistent with zero within uncertainties and is, therefore, neglected [4].

In the dilepton channel, on the other hand, the fake and non-prompt lepton contribution is instead taken from MC simulation and normalised to data using a CR with two same-sign (SS) leptons [4]. This background is found to be a very small in the dilepton channel as well.

CHAPTER 8

Event reconstruction, selection, and categorisation

In Chapter 5, the physical objects that are reconstructed from the detector signals are presented, namely the charged leptons, jets and neutrinos in the form of missing transverse momentum. In addition to those employed reconstruction algorithms, certain criteria are applied to them in order to further reduce the probability of mis-reconstruction or mis-identification. This increases the purity of selected events in the dedicated regions which is critical in this statistically and systematically limited search. The chosen additional quality criteria are given in the following.

- **Electrons** must pass additional quality criteria labelled as TIGHT [52] to reduce the contribution from fake objects. They must also fulfil $|z_0 \sin\theta| < 0.5$ mm and $|S_{d_0}| < 5$ [4]. The first electron in an event must pass $p_{\mathrm{T}} > 27$ GeV and the second must have at least $p_{\mathrm{T}} > 10$ GeV. In addition to this, the GRADIENT isolation criterion [52] is required to reduce background contamination from non-prompt electrons.

- **Muons** must pass MEDIUM [126] and GRADIENT isolation criteria [54]. Additionally, they must fulfil $|z_0 \sin\theta| < 0.5$ mm and $|S_{d_0}| < 3$ [4]. The first muon in an event must have $p_{\mathrm{T}} > 27$ GeV and any second muon must pass $p_{\mathrm{T}} > 10$ GeV, in agreement with the electron requirement. The value of 27 GeV stands in contrast to the 28 GeV applied in the b-tagging calibration analysis. This is because the search for $t\bar{t}H(H \to b\bar{b})$ was performed using data from the years 2015 and 2016, while the b-tagging calibration also measures the b-tagging efficiency in data from 2017. In 2017, however, the trigger threshold for charged leptons was increased slightly and, to avoid any reconstruction inefficiencies around this threshold, the lepton p_{T} cut was increased to 28 GeV accordingly. Then, for consistency reasons, the same p_{T} cut is used for all three years.

- **Jets** in this search are required to have a p_{T} above 25 GeV in contrast to the 20 GeV in the b-tagging calibration analysis, but are otherwise treated in an identical way. The jet p_{T} requirement is motivated from the expected jet kinematics in signal events and a reduced impact from the uncertainty on the jet energy scale. Additionally, it reduces the contribution of mis-reconstructed jets or other objects mimicking the signature of jets in the detector. The overlap between jets and electrons or jets and muons in an event is handled according to the procedure described in Section 5.3.

- In this analysis, ***b*-jets** are selected using the pseudo-continuous (PC) b-tagging strategy with the MV2c10 algorithm. Thus, instead of being labelled either as a b-jet or non-b-jet, jets can be put into one of five bins according to which WP they pass and thus which PC bin of the MV2c10 distribution they fall into. A jet not passing the lowest WP (85%) is assigned a label '1', while a jet passing the tightest WP (60%) is assigned a label '5' and so on.

- The **missing transverse momentum**, $p_{\mathrm{T}}^{\mathrm{miss}}$, or its magnitude i.e. the missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, is defined in the same way as presented in Section 5.5 [62, 63]. It is not used for the event selection, but for the reconstruction of the $t\bar{t}H$ final state.

- In general, ***τ*-leptons** are reconstructed via their decay products which are either charged leptons or hadrons, as described in Section 5.4. This analysis reconstructs hadronically decaying τ-leptons in order to discriminate them from jets. For this, a strategy is followed that is based on the track multiplicity and a multivariate discriminant based on the collimation of the tracks, jet substructure variables and kinematic variables [4]. A candidate object, labelled as τ_{had} from now on, must pass the MEDIUM τ-identification WP [127] and fulfil $p_{\mathrm{T}} > 25$ GeV as well as $|\eta| < 2.5$. If the τ_{had} candidate is within $\Delta R = 0.2$ of any selected electron or muon, it is discarded.

In the following sections, the event reconstruction and selection are presented and the various analysis regions are defined.

8.1 Event Selection

For both semileptonic and dileptonic decay channels of the $t\bar{t}$ pair, events are recorded as soon as the trigger system identified one charged lepton fulfilling the following criteria. In 2015 (2016) data, electrons must have at least a $p_{\mathrm{T}} > 24$ (26) GeV, while muons are required to have a $p_{\mathrm{T}} > 20$ (26) GeV. These leptons also must pass dedicated isolation criteria [35, 78] not to be confused with those mentioned in the previous section. Additional trigger algorithms are employed in a logical "or" that require a higher p_{T} threshold of the leptons, but with looser identification and no isolation requirement. They are 60 and 120 GeV (60 and 140 GeV) for electrons and 50 GeV for muons in 2015

(2016), respectively.

Events must contain at least one reconstructed charged lepton with $p_{\mathrm{T}} > 27$ GeV that has the same lepton flavour (e or μ) as the one reconstructed by the trigger algorithm that fired and also be within $\Delta R < 0.1$ of it. Specifically for the dilepton channel, exactly two leptons are required that are oppositely charged. The second lepton, ordered by p_{T}, must have $p_{\mathrm{T}} > 15$ (10) GeV in the ee ($e\mu$ and $\mu\mu$) channel(s). Furthermore, in the ee and $\mu\mu$ channels, the invariant mass of the leptons m_{ll} is required to be above 15 GeV in order to not include hadronic resonance decays with low mass. Additionally, m_{ll} must be sufficiently far outside the window of the Z boson mass, here chosen to be $83 - 99$ GeV, which rejects unwanted leptonic Z boson decays to a large extent.

Selected events are also required to contain at least three jets with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$ of which at least two must pass the b-tagging WP of 77%. In the single lepton channel specifically, at least five jets with $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$ are required. In case there are exactly five such jets, at least three of them have to be b-tagged at the 77% WP. On the other hand, if there are at least six such jets, either at least three of them must pass the 77% b-tagging WP or at least two of them must be b-tagged at the 60% WP. If an event passes the dilepton event selection, it is removed from the single lepton channel, similarly to the treatment of the boosted decay channel.

As there are other searches for $t\bar{t}H$ production, in particular measurements targeting hadronically decaying τ-leptons, certain events are excluded from the search for $t\bar{t}H(H \to b\bar{b})$. In the dilepton channel, these are events with at least one τ_{had} candidate and in the single lepton channel these are events containing at least two such candidates.

Finally, about 2.5% of all simulated $t\bar{t}H(H \to b\bar{b})$ events pass the dilepton event selection and roughly 8.7% pass the single lepton event selection [4].

8.2 Analysis regions

For each of the two leptonic $t\bar{t}$ decay channels, several regions are defined that aim to split the analysed samples into categories enriched either with signal events or with specific background components. These involve a SR which contains mostly $t\bar{t}H(H \to b\bar{b})$ and $t\bar{t}+b\bar{b}$ events, referred to as $t\bar{t}+ \geq 2b$; and three signal-depleted CRs $t\bar{t}+b$, $t\bar{t}+ \geq 1c$ and $t\bar{t}+$ light. The definition for a region to be signal-enriched is given by the expected number of signal (S) and background (B) events. If the signal-to-background ratio S/B is larger than 1% and its significance with respect to Poissonian background fluctuations $S/\sqrt{B}$ is above 0.3 at the same time, the region is considered signal-enriched.

For both the single lepton and dilepton channel, multiple SRs are defined that all fulfil the above criteria of being signal-enriched. One essential feature of the $t\bar{t}H(H \to b\bar{b})$ final state is its high total number of jets and particularly its number of b-jets. Therefore,

events are split up according to the number of jets and according to the PC b-tagging bins that the four jets (or three for corresponding dileptonic events) with the highest MV2c10 score, ordered by this score in descending order, fall into. These four (three) jets build the basis for the categorisation of events and ultimately the definition of the different analysis regions which is presented in the following subsections.

8.2.1 Regions in the single lepton channel

In the single lepton channel, the hadronically decaying W boson from the $t\bar{t}$ pair contributes c-jets to the final state. This worsens the selection efficiency and makes it difficult to define a pure '$t\bar{t}$ + heavy flavour' control region that is enriched in $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ background events. While an incorrect selection of such jets could be avoided by applying a tight b-tagging requirement, this would lead to a selection of signal events. This problem does not exist in the dilepton channel. However, there are far more semileptonic than dileptonic $t\bar{t}$ decays due to the larger decay branching ratio. Therefore, the single lepton SRs contain considerably more signal events compared to the dilepton SRs, which ultimately dominates the sensitivity in this search.

Below are the definitions of the regions for the single lepton channel with at least six reconstructed jets. The content of the b-tagging categories is estimated via MC:

- $\mathrm{SR}_1^{\geq 6j}$ contains all events where the four jets with highest MV2c10 score pass the 60% WP, labelled as the '(5,5,5,5)' category, which has a contribution from $t\bar{t}+ \geq 2b$ of at least 60%. This is the most signal-enriched region and has the highest purity;
- $\mathrm{SR}_2^{\geq 6j}$ includes all remaining b-tagging categories with a $t\bar{t}+ \geq 2b$ contribution of at least 45%;
- $\mathrm{SR}_3^{\geq 6j}$ includes all remaining b-tagging categories with a $t\bar{t}+ \geq 2b$ contribution of at least 30%;
- $\mathrm{CR}_{t\bar{t}+b}^{\geq 6j}$ includes all remaining b-tagging categories with a $t\bar{t}+b$ contribution of at least 30%;
- $\mathrm{CR}_{t\bar{t}+\geq 1c}^{\geq 6j}$ includes all remaining b-tagging categories with a $t\bar{t}+ \geq 1c$ contribution of at least 30%;
- $\mathrm{CR}_{t\bar{t}+\mathrm{light}}^{\geq 6j}$ includes all remaining b-tagging categories.

The regions for single lepton events with exactly five jets are defined as follows:

- SR_1^{5j} is defined in the same way as $\mathrm{SR}_1^{\geq 6j}$;
- $\mathrm{CR}_{t\bar{t}+b}^{5j}$ includes all remaining b-tagging categories with a $t\bar{t}+b$ component of at least 20%;

- SR_2^{5j} includes all remaining b-tagging categories with a $t\bar{t}+ \geq 2b$ component of at least 20%;
- $\mathrm{CR}_{t\bar{t}+\geq 1c}^{5j}$ includes all remaining b-tagging categories with a $t\bar{t}+ \geq 1c$ component of at least 20%;
- $\mathrm{CR}_{t\bar{t}+\mathrm{light}}^{5j}$ includes all remaining b-tagging categories.

These region definitions are illustrated in Figure 8.1. For each of the two iterative region selections, a few b-tagging categories are moved between regions by hand to avoid disconnecting areas between them in Figure 8.1 [4].
As an example, the SRs with at least six reconstructed jets are constructed as follows: $\mathrm{SR}_1^{\geq 6j}$ contains the (5,5,5,5) category. $\mathrm{SR}_2^{\geq 6j}$ is composed of all other categories that have a $t\bar{t}+ \geq 2b$ contribution above 45% which are (5,5,5,4) and (5,5,5,3), representing events in which the fourth ranked jet only passes the 70% and 77% b-tagging WP, respectively. Finally, $\mathrm{SR}_3^{\geq 6j}$ contains all other categories with a $t\bar{t}+ \geq 2b$ component above 30% which are (5,5,5,2), (5,5,4,4) and (5,5,4,3).

8.2.2 Regions in the dileptonic channel

Since the dileptonic decay involves two fewer jets than the semileptonic one, the regions are now split as either having at least four or exactly three jets. The category with at least four jets does not contain a $t\bar{t}+b$ CR, because of the small number of corresponding events in the dilepton sample. It is split as follows:

- First, remove all b-tagging categories that have fewer than eight background events ($B < 8$) or a signal-to-background ratio of $S/\sqrt{B} < 0.08$ to have a more clean region selection that is not strongly affected by limited statistics;
- $\mathrm{SR}_1^{\geq 4j}$ includes all remaining b-tagging categories with a $t\bar{t}+ \geq 2b$ component of at least 70%;
- $\mathrm{SR}_3^{\geq 4j}$ includes all remaining b-tagging categories with a $t\bar{t} + b$ component of at least 30%;
- $\mathrm{SR}_2^{\geq 4j}$ includes all remaining b-tagging categories with a $t\bar{t}H$ component of at least 1.5%;
- $\mathrm{CR}_{t\bar{t}+\geq 1c}^{\geq 4j}$ includes all remaining b-tagging categories with a $t\bar{t}+ \geq 1c$ component of at least 25%;
- $\mathrm{CR}_{t\bar{t}+\mathrm{light}}^{\geq 4j}$ includes all remaining b-tagging categories.

The dileptonic events with exactly three jets do not contain any signal-enriched regions, again due to the low amount of signal events. They are defined in the following way:

- First, remove all b-tagging categories that have fewer than eight background events ($B < 8$) or a signal-to-background ratio of $S/\sqrt{B} < 0.08$;

- $\mathrm{CR}^{3j}_{t\bar{t}+b}$ includes all remaining b-tagging categories with a $t\bar{t}+b$ component of at least 30%.;
- $\mathrm{CR}^{3j}_{t\bar{t}+\mathrm{light}}$ includes all remaining b-tagging categories.

The definitions for the dilepton regions are summarised in Figure 8.2. Also here, a few b-tagging categories are moved between the regions by hand. Furthermore, the categories with limited statistics taken out before defining the regions are recovered in order to help create connected regions in Figure 8.2 [4].

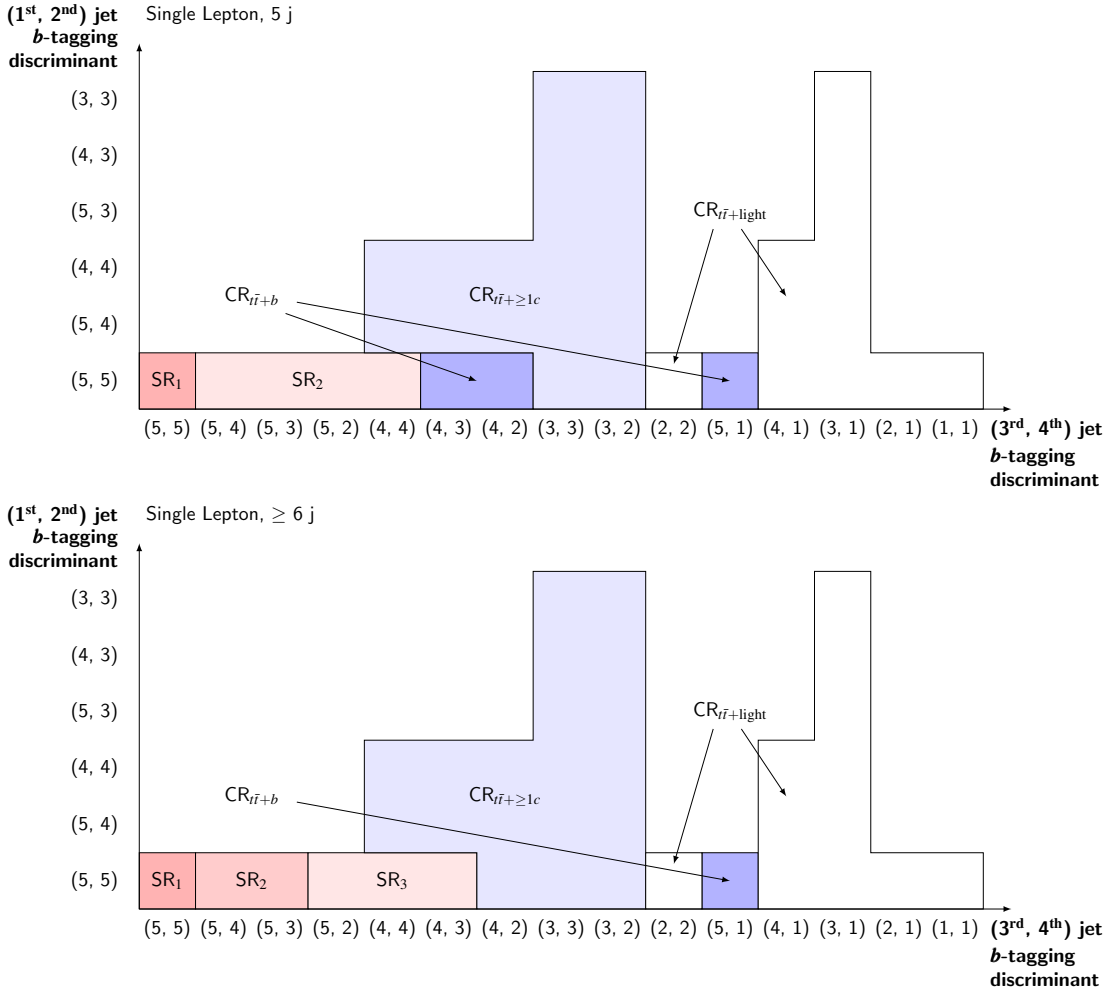

Figure 8.1: Schematic of the definition of the SRs and CRs in the single lepton channel for the exactly five (top) and at least six (bottom) jets category. The regions are shown as a function of the PC b-tagging bins that the four jets with the highest MV2c10 score, ordered by this score in descending order, fall into.

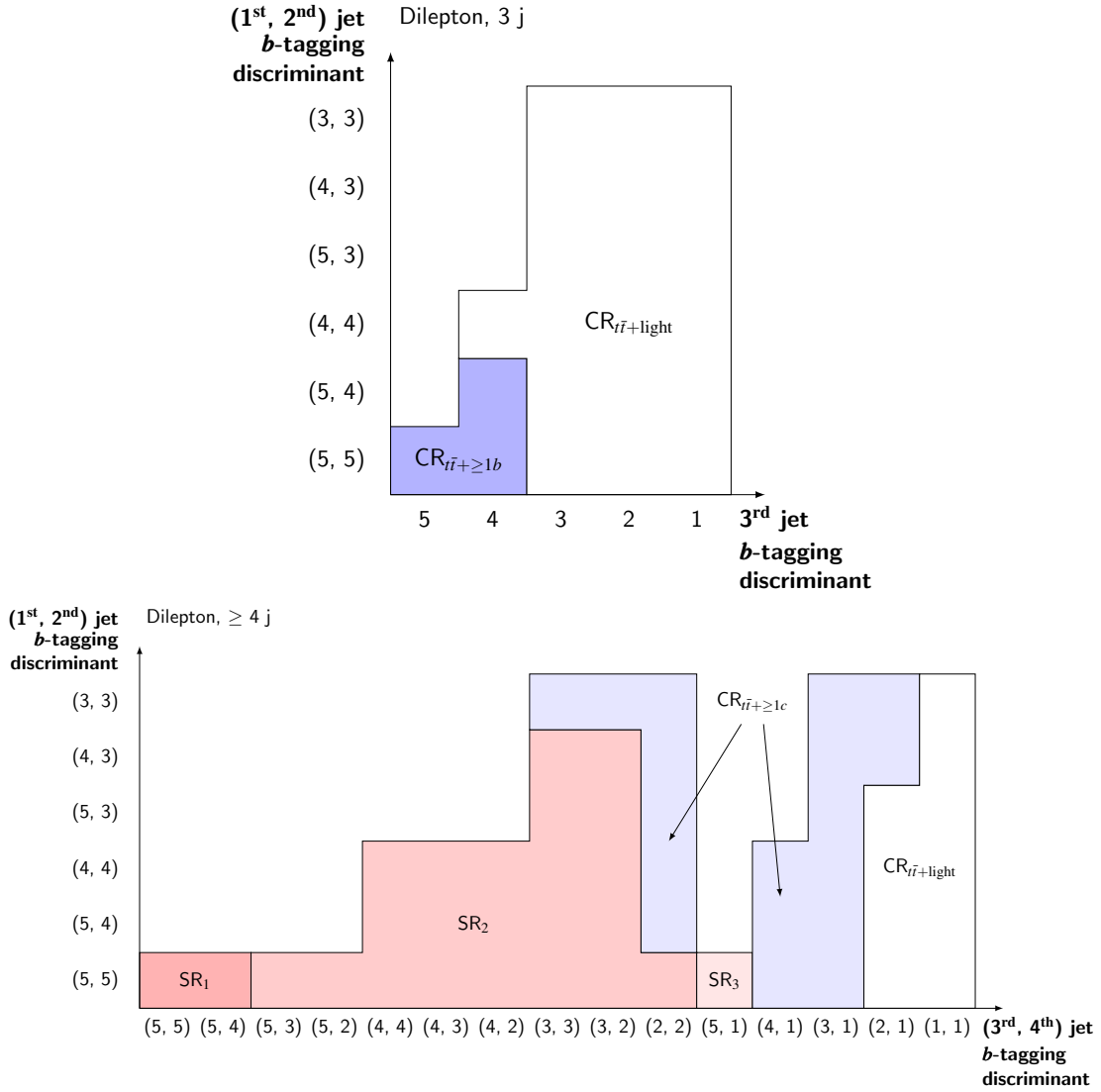

Figure 8.2: Schematic of the definition of the SRs and CRs in the dilepton channel for the exactly three (top) and at least four (bottom) jets category. The regions are shown as a function of the PC b-tagging bins that the three (top) and four (bottom) jets with the highest MV2c10 score, ordered by this score in descending order, fall into.

The predictions of the different signal and background components of each region are shown in Figure 8.3 for both channels. The predicted signal-to-background ratio S/B as well as its significance $S/\sqrt{B}$ is depicted in Figure 8.4 for each region of the two channels. The $H \to b\bar{b}$ decay is present in 89% (96%) of selected $t\bar{t}H$ signal events in the dilepton (single lepton) SRs. Further details about the analysis regions are given in Ref. [4].

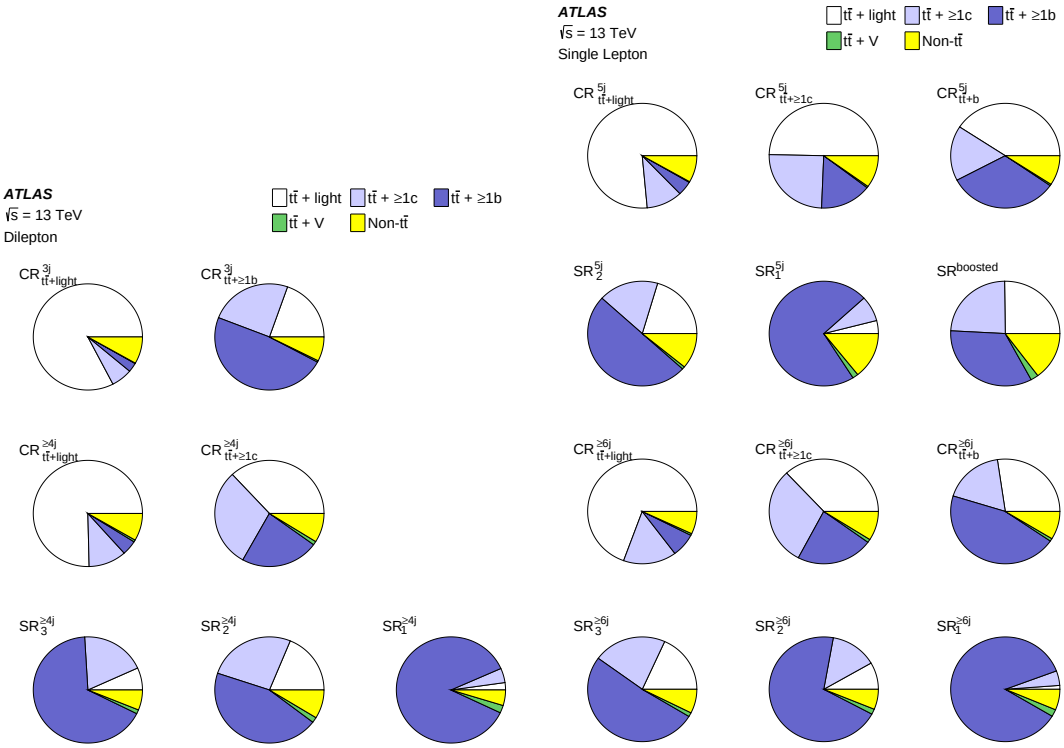

Figure 8.3: Pie charts showing the fraction of events belonging to the different categories within each of the dilepton (left) and single lepton (right) regions. The categories are defined in Section 7.3.1.

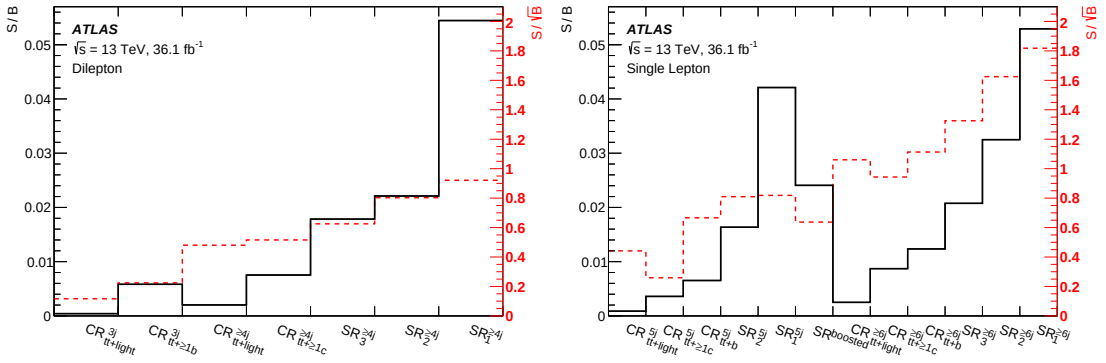

Figure 8.4: The signal-to-background ratio (solid black) and its significance (dotted red) are shown for each of the dilepton (left) and single lepton (right) regions.

CHAPTER 9

Extraction of the $t\bar{t}H(H \to b\bar{b})$ cross-section

The previous chapter introduced and defined the various SRs and CRs in the two channels. These not only help to separate signal from background events, but also to have dedicated samples enriched in a specific type of background in order to study and control it, for example its contribution to the shape of an observable in a signal-enriched region. SRs have been defined for each channel, but even the purest among them are still dominated by $> 94\%$ background events, as can be seen in Figure 8.4. This is mainly attributable to the $t\bar{t}$ +jets background and its cross-section which is roughly three orders of magnitude larger than the $t\bar{t}H$ signal cross-section. Therefore, a boosted decision tree (BDT) is trained for each SR in order to further discriminate $t\bar{t}H(H \to b\bar{b})$ signal events from background events. This BDT is labelled 'Classification BDT' and it uses numerous input variables for its training, depending on the specific SR it is applied to [4]. Three essential inputs are

1. the so-called 'Reconstruction BDT'. The leptonic $t\bar{t}H(H \to b\bar{b})$ final states at LO consist of four to six jets of which four are b-jets. As a consequence, there are many possible combinations of jets that could be used to reconstruct the top quarks and the Higgs boson of an event, while only one is technically correct disregarding degenerate assignments. The reconstruction BDT is trained to select the best possible (i.e. the correct) matching of final state partons to jets and use those jets to build the top quark and Higgs boson candidate objects;

2. a likelihood discriminant (LHD). A LHD can be built which, for each possible combination of matching final state partons to physics objects in an event, contains corresponding terms that represent the probability of the event to be compatible with the signal or background hypotheses; and

3. a discriminant based on the matrix element method (MEM). As the name suggests, it is based on the calculation of the ME of the $t\bar{t}H(H \to b\bar{b})$ signal compared to

the ME of the background processes. Its idea is similar to the LHD, but is based on CPU intensive calculations of ME of Feynman diagrams, instead of using the fully simulated nominal MC samples.

The combined profile likelihood fit, further detailed in Section 9.2, uses the classification BDT in all SRs. In the CRs, the fit is performed using the number of events in that region, except for the two single lepton CRs dedicated to the $t\bar{t}+\geq 1c$ component, which are $\mathrm{CR}^{5j}_{t\bar{t}+\geq 1c}$ and $\mathrm{CR}^{\geq 6j}_{t\bar{t}+\geq 1c}$. In these regions, $H_{\mathrm{T}}^{\mathrm{had}}$ is used instead to have a stronger constraint on the $t\bar{t}+\geq 1c$ background. Table 9.1 summarises which discriminant, along with its binning, is applied to which region.

Single lepton channel					
≥ 6 jets			exactly 5 jets		
Region	Discriminant	Bins	Region	Discriminant	Bins
SR1	BDT	8	SR1	BDT	8
SR2	BDT	8	SR2	BDT	8
SR3	BDT	8			
$\mathrm{CR}_{t\bar{t}+b}$	N_{events}	1	$\mathrm{CR}_{t\bar{t}+b}$	N_{events}	1
$\mathrm{CR}_{t\bar{t}+\geq 1c}$	$H_{\mathrm{T}}^{\mathrm{had}}$	8	$\mathrm{CR}_{t\bar{t}+\geq 1c}$	$H_{\mathrm{T}}^{\mathrm{had}}$	6
$\mathrm{CR}_{t\bar{t}+\mathrm{light}}$	N_{events}	1	$\mathrm{CR}_{t\bar{t}+\mathrm{light}}$	N_{events}	1
Dilepton channel					
≥ 4 jets			exactly 3 jets		
Region	Discriminant	Bins	Region	Discriminant	Bins
SR1	BDT	6			
SR2	BDT	8			
SR3	BDT	8			
$\mathrm{CR}_{t\bar{t}+\geq 1c}$	N_{events}	1	$\mathrm{CR}_{t\bar{t}+b}$	N_{events}	1
$\mathrm{CR}_{t\bar{t}+\mathrm{light}}$	N_{events}	1	$\mathrm{CR}_{t\bar{t}+\mathrm{light}}$	N_{events}	1

Table 9.1: The discriminants and the respective number of bins used in each region during the profile likelihood fit.

Before constructing the MVAs and performing the fit, the modelling of various observables has been checked in the dedicated CRs of the two channels. Such a test is necessary to ensure that those variables are simulated and understood properly before using them to train an MVA. The physics motivation behind the selected observables is discussed in Appendix D. If some variables are found to be described poorly by the simulation, the analysis strategy as well as the final result can be biased. In this test, basic event-based and object-based variables have been studied such as the p_{T} and η of reconstructed objects, H_{T} and $H_{\mathrm{T}}^{\mathrm{had}}$, the number of jets as well as the number of b-tagged jets and many more. Example distributions of $H_{\mathrm{T}}^{\mathrm{had}}$ in the single lepton CRs are presented in Figures 9.1 and 9.2. Showing further relevant plots would go beyond the scope of this thesis and hence they are left out. In general, the observables which are required to

describe and reconstruct the $t\bar{t}H(H \to b\bar{b})$ final state have been found to be modelled well. Furthermore, those variables which show a significant mis-modelling have not been included in the training of the MVAs.

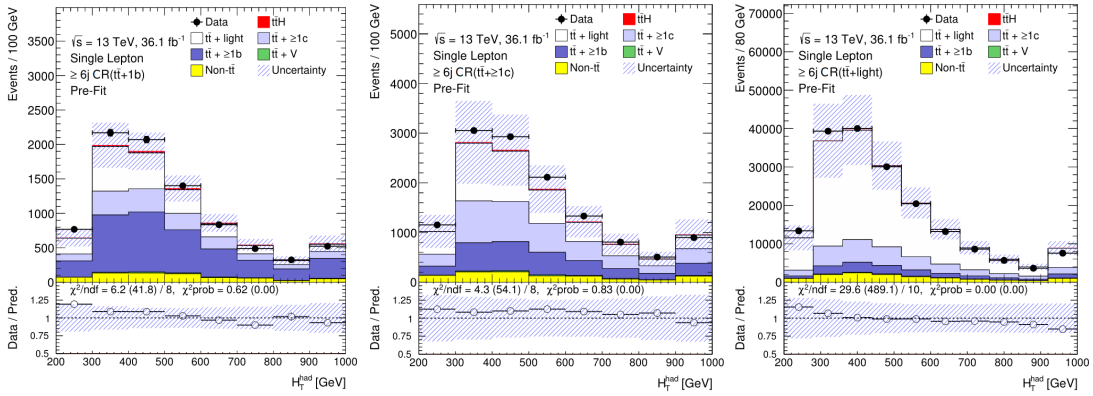

Figure 9.1: Comparison of the predicted event yields to the observed events in data as a function of $H_{\mathrm{T}}^{\mathrm{had}}$ in the single lepton CRs selecting at least six jets before the fit. The distribution is shown in the $\mathrm{CR}^{\geq 6j}_{t\bar{t}+b}$ (left), $\mathrm{CR}^{\geq 6j}_{t\bar{t}+\geq 1c}$ (centre) and $\mathrm{CR}^{\geq 6j}_{t\bar{t}+\mathrm{light}}$ (right). The filled red area represents the $t\bar{t}H$ signal stacked on top of the background normalised to the SM cross-section before the fit. The total uncertainty in the simulated yields is represented by the hatched area, while an uncertainty in the $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ normalisations is not considered before the fit.

As an additional check, the predicted event yields for each channel and each region are compared to the observed data yields before the fit. A statistically significant mis-match could hint at several possible problems such as missing contributions from other sources or a false prediction of a cross-section. The yields before the fit ('pre-fit') are summarised in Appendix E where the yields after the fit ('post-fit') are included as well. Tables E.1–E.6 show that the total predicted yields are well in agreement with the observed data within the considered uncertainties pre-fit as well as post-fit.

9.1 Reconstruction of the $t\bar{t}H$ signal

This section describes how the reconstruction and classification of the $t\bar{t}H(H \to b\bar{b})$ final state is performed in the two channels by the different MVAs. First, the reconstruction BDT is detailed, followed by the LHD and MEM which are ultimately fed into the classification BDT to select the signal events that enter the profile likelihood fit.

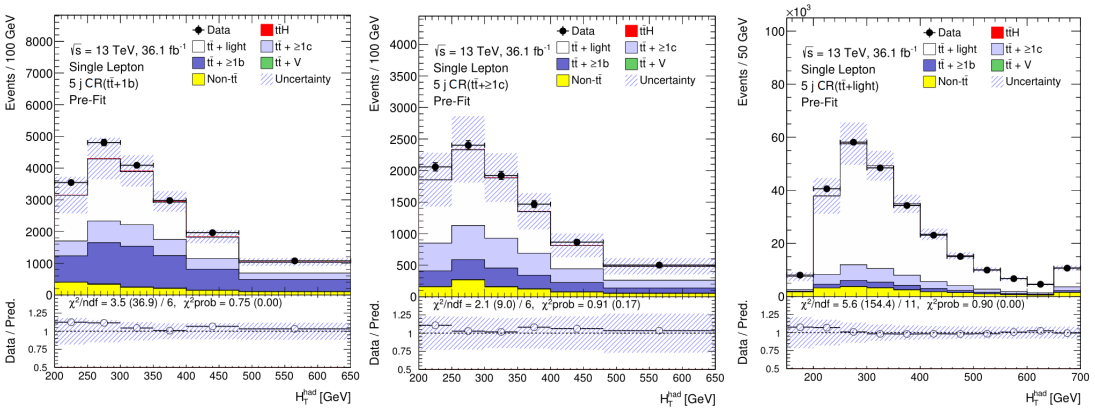

Figure 9.2: Comparison of the predicted event yields to the observed events in data as a function of $H_{\mathrm{T}}^{\mathrm{had}}$ in the single lepton CRs selecting exactly five jets before the fit. The distribution is shown in the $\mathrm{CR}^{5j}_{t\bar{t}+b}$ (left), $\mathrm{CR}^{5j}_{t\bar{t}+\geq 1c}$ (centre) and $\mathrm{CR}^{5j}_{t\bar{t}+\mathrm{light}}$ (right). The filled red area represents the $t\bar{t}H$ signal stacked on top of the background normalised to the SM cross-section before the fit. The total uncertainty in the simulated yields is represented by the hatched area, while an uncertainty in the $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ normalisations is not considered before the fit.

9.1.1 Reconstruction BDT

The reconstruction BDT tries to correctly assign each final state parton of an event to a jet and then, using those jets as well as the lepton(s), build the candidate particles for the Higgs boson as well as the W bosons and from those the top quarks. The training is performed with the TMVA Toolkit [73] using exclusively simulated $t\bar{t}H(H \to b\bar{b})$ events. The signal is the one configuration where all partons are matched to the correct jets in an event, all other configurations are considered background. The BDT is trained separately for the single lepton (dilepton) regions with exactly five (three) and at least six (four) jet regions. However, the training is performed inclusively on events with at least four jets b-tagged at the 85% WP (exactly three for the respective dilepton regions), instead of having one dedicated training per signal region. Both strategies have been found to perform similarly well, while the former is much simpler to control and implement technically [4]. Additionally, the four (three) jets that fall into the highest PC bins of the MV2c10 algorithm are considered b-jets, while all others are considered light-jets. If two or more jets fall into the same bin they are ordered by their p_{T} in descending order. This allows to exploit the finer differentiation between b-jets through PC b-tagging.

In the single lepton channel, the leptonic W boson is reconstructed from the momenta of the lepton and neutrino, i.e. p_l and p_ν. The momentum of the neutrino is derived from the $p_{\mathrm{T}}^{\mathrm{miss}}$, while the z component is calculated by solving the following equation

given by the invariant mass of the W:

$$m_W^2 = (p_l + p_\nu)^2 \,. \tag{9.1}$$

This is a quadratic equation and thus holds two solutions. Both of them are used and treated as individual configurations. However, if no real solution exists, the discriminant of the quadratic solution is set to zero such that there is one unique solution.
The hadronic W boson (Higgs boson) is reconstructed from a pair of light (b-tagged) jets. After this, the top quark candidates are each reconstructed from one W boson and one jet. In the following, the top quark reconstructed from the hadronically (leptonically) decaying W boson is referred to as hadronically (leptonically) decaying top quark candidate.
It is important to keep in mind that the single lepton $t\bar{t}H(H \to b\bar{b})$ final state expects six jets at LO. Thus, in events with exactly five jets, one jet could not be reconstructed. In over 70% of those cases, this jet comes from the hadronic W boson. In order to account for this, the hadronic top quark in events with exactly five jets is reconstructed from two jets of which exactly one must be b-tagged.

In contrast to this, the BDT in the dilepton channel does not try to reconstruct the W bosons, which would prove difficult given two undetected neutrinos. Therefore, it instead reconstructs the top quark candidates directly from one lepton and one jet.
For both channels, two BDT configurations are trained for each of the two different jet regions: one that exploits the full information from the $t\bar{t}H$ final state, and one that does not use any information from the Higgs boson decay. The configuration without the Higgs boson information is less efficient in reconstructing the final state correctly, but its output is less biased, because it does not shape the result to resemble the signal [4]. Thus, this second configuration is assumed to better discriminate between signal and other background processes. Nonetheless, both BDT configurations as well as variables built from jet-to-parton assignments from the former, the latter or both are used as input for the classification BDT. The BDT output and variables are still slightly correlated among the different configurations, however they provide complementary information which outweighs the correlation.

The variables used for the BDT training are invariant masses of various final state objects and angular separations between them in addition to more kinematic variables. These variables are chosen to maximise the efficiency of correctly matching jets to the Higgs boson, but at the same time as few variables as possible are included to avoid strong correlations in the training.

In each event, the BDT assigns a score or weight to each configuration and the one with the highest score is assumed to be the correct one. It is taken as input to the classification BDT, but it also serves as the basis to construct additional kinematic variables from the jets of this configuration. In the most sensitive single lepton SR ($\mathrm{SR}_1^{\geq 6j}$), the Higgs boson is correctly reconstructed in 48% (32%) of all selected $t\bar{t}H(H \to b\bar{b})$ events when

including (leaving out) the Higgs boson information in the training. The corresponding efficiency in the most sensitive dileptonic SR ($\mathrm{SR}_1^{\geq 4j}$) is 49% (32%).

The input variables for the training of the single lepton reconstruction BDT is given in Table 9.2. The input variables for the dileptonic reconstruction BDT is listed in Table 9.3.

Variable	Region	
	$\geq 6j$	$5j$
Topological information from the $t\bar{t}$ pair		
$m_{t_{\mathrm{lep}}}$	✓	✓
$m_{t_{\mathrm{had}}}$	✓	-
$m_{t_{\mathrm{had}}}$ using only one jet and one b-jet	-	✓
$m_{W_{\mathrm{had}}}$	✓	-
$m_{W_{\mathrm{had}},b_{t_{\mathrm{lep}}}}$	✓	✓
$m_{W_{\mathrm{lep}},b_{t_{\mathrm{had}}}}$	✓	✓
$\Delta R(W_{\mathrm{had}}, b_{t_{\mathrm{had}}})$	✓	✓
$\Delta R(W_{\mathrm{had}}, b_{t_{\mathrm{lep}}})$	✓	✓
$\Delta R(l, b_{t_{\mathrm{had}}})$	✓	✓
$\Delta R(l, b_{t_{\mathrm{lep}}})$	✓	✓
$\Delta R(b_{t_{\mathrm{had}}}, b_{t_{\mathrm{lep}}})$	✓	✓
$\Delta R(q_{1,W_{\mathrm{had}}}, q_{2,W_{\mathrm{had}}})$	✓	-
$\Delta R(b_{t_{\mathrm{had}}}, q_{1,W_{\mathrm{had}}})$	✓	-
$\Delta R(b_{t_{\mathrm{had}}}, q_{2,W_{\mathrm{had}}})$	✓	-
$\min_i (b_{t_{\mathrm{had}}}, q_{i,W_{\mathrm{had}}}))$	✓	-
$\min_i (b_{t_{\mathrm{had}}}, q_{i,W_{\mathrm{had}}})) - \Delta R(l, b_{t_{\mathrm{lep}}})$	✓	✓
Topological information from the Higgs boson		
m_H	✓	✓
$m_{H,q_{1,W_{\mathrm{had}}}}$	✓	✓
$\Delta R(b_{1,H}, b_{2,H})$	✓	✓
$\Delta R(b_{1,H}, l)$	✓	✓
$\Delta R(b_{1,H}, b_{t_{\mathrm{had}}})$	-	✓
$\Delta R(b_{1,H}, b_{t_{\mathrm{lep}}})$	-	✓

Table 9.2: Input variables for the reconstruction BDT used in the single lepton SRs [4]. $t_{\mathrm{lep}}(t_{\mathrm{had}})$ refers to the leptonically (hadronically) decaying top quark, $W_{\mathrm{lep}}(W_{\mathrm{had}})$ refers to the leptonically (hadronically) decaying W boson, $b_{t_{\mathrm{lep}}}(b_{t_{\mathrm{had}}})$ is the b-jet from the leptonically (hadronically) decaying top quark, l is the charged lepton, $q_{1,W_{\mathrm{had}}}(q_{2,W_{\mathrm{had}}})$ is the leading (subleading) jet from the hadronically decaying W boson, H refers to the Higgs boson and, finally, $b_{1,H}(b_{2,H})$ is the leading (subleading) b-jet from the Higgs decay.

Variable	Region	
	$\geq 4b$	$3b$
Topological information from the $t\bar{t}$ pair		
$m_{l^+,b} - m_{l^-,\bar{b}}$	✓	✓
$m_{l^+,b}$	✓	✓
$m_{l^-,\bar{b}}$	✓	✓
$\Delta R\left(l^+, b\right)$	✓	✓
$\Delta R\left(l^-, \bar{b}\right)$	✓	✓
$\Delta\phi\left(b\bar{b}\right)$	✓*	✓
$\Delta R\left(b\bar{b}\right)$	✓*	-
p_{T}^{b}	✓	✓
$p_{\mathrm{T}}^{\bar{b}}$	✓	✓
$\Delta R\left(l^+, b\right) - \Delta R\left(l^-, \bar{b}\right)$	✓	✓
$\min \Delta\eta\left(b, l\right)$	✓	✓
Topological information from the Higgs boson		
m_H	✓	✓
$\Delta\phi(H, t\bar{t})$	-	✓
$\Delta R(H, t\bar{t})$	✓	-
$\Delta R(b_{1,H}, b_{2,H})$	✓	✓
$p_{\mathrm{T}}^{b_{2,H}}$	-	✓
$\min \Delta R\left(b_H, l\right)$	-	✓
$\max \Delta R\left(H, b\right)$	✓	-

Table 9.3: Input variables for the reconstruction BDT used in the dileptonic SRs [4]. The description of variables using the information from the $t\bar{t}$ pair imply the charged lepton ($l^{\pm}$) and b-jets ($b, \bar{b}$) to originate from the $t\bar{t}$ pair decay. Among the variables using information from the Higgs boson, H refers to the Higgs boson, $b_{1,H}, b_{2,H}$ are the two b-jets from the Higgs boson decay, ordered by their PC b-tagging bin, and b in the last variable means any b-jet in the event. The superscript * indicates two special variables: $\Delta\phi\left(b\bar{b}\right)$ is considered for all BDT configurations except the one including the Higgs boson information for ≥ 4 b-jets, in which $\Delta R\left(b\bar{b}\right)$ is used for the training instead.

9.1.2 Likelihood Discriminant

The LHD is based on a combinatorial likelihood approach and defined in the same way as in Ref. [120]. The underlying idea of a LHD is to construct likelihood functions $\mathcal{P}^{\mathrm{sig}}(x)$ and $\mathcal{P}^{\mathrm{bkg}}(x)$, which represent the probability that the event under study is compatible with the signal or background hypothesis, respectively. Here, those hypotheses are that the event originates from a $t\bar{t}H(H \to b\bar{b})$ or $t\bar{t} + b\bar{b}$ process, respectively, and x refers to the kinematics of the reconstructed final state objects in the event. The LHD D is

defined as

$$D = \frac{\mathcal{P}^{\text{sig}}}{\mathcal{P}^{\text{sig}} + \mathcal{P}^{\text{bkg}}}. \tag{9.2}$$

Thus, the event is more compatible with the signal hypothesis as D approaches a value of one.

These likelihood functions are the product of multiple one-dimensional probability density functions (PDFs), where each PDF represents the probability for a given value of one observable in the event to be compatible with the signal or background hypothesis. Consider, for example, an additional $b\bar{b}$ pair, separated from the $t\bar{t}$ decay, with an invariant mass close to the Higgs boson mass. When neglecting all other information of the event for this example, the signal PDF for reconstructing an additional $b\bar{b}$ pair with such an invariant mass is expected to be higher than the corresponding background PDF. Consequently, $\mathcal{P}^{\text{sig}}(x)$ is higher than if the invariant mass were significantly smaller or larger than the Higgs boson mass and thus the event is expected to be more compatible with the signal hypothesis.

The individual PDFs are based on invariant mass and angular separation variables built from the reconstructed objects in the event, similarly to those used for the reconstruction BDT. As discussed in Appendix D, such observables are useful and expected to discriminate the $t\bar{t}H(H \to b\bar{b})$ signal from the dominant background $t\bar{t} + b\bar{b}$. For this method, the latter is represented by two background hypotheses, namely $t\bar{t}+ \geq 2b$ and $t\bar{t} + b$. The likelihood functions of both categories are merged to a weighted average, where the weights are given by their relative fraction of events taken from $t\bar{t}$ +jets simulation.

$\mathcal{P}^{\text{sig}}(x)$ is the product of the aforementioned PDFs which are primarily those of the invariant masses of the Higgs boson, the leptonically and hadronically decaying top quarks and the hadronically decaying W boson. In a significant amount of semileptonic signal and background events with at least six jets, only one jet from the hadronically decaying W boson is reconstructed and selected. This is taken into account by adding another hypothesis for both signal and background scenarios. Furthermore, in events with exactly five jets, variables based on the hadronic top quark are built in a similar way to those for the reconstruction BDT. Before computing the product, the PDFs are averaged over all possible jet-to-parton matching configurations. These configurations are weighted using b-tagging information to reduce the contribution from parton-to-jet matchings with an inconsistent flavour pattern, because four b-jets are expected in the final state and assigned to the Higgs boson and the two top quarks, and one or two light-jets are assigned to the hadronically decaying W boson.

$\mathcal{P}^{\text{bkg}}(x)$ is constructed similarly to $\mathcal{P}^{\text{sig}}(x)$, but instead of the Higgs boson mass it includes the invariant mass of the additional b-jet pair. More variables are included such as the angular separation between final state objects, but the invariant masses above are the most significant contributions to the two likelihood functions. The distributions are

derived from simulation, but only their shape information is used. In the high mass tails of the PDFs, the number of contributing events may become very small and, correspondingly, statistical fluctuations may occur. These cases are taken care of by a smoothing procedure [4].

The most important difference of the LHD with respect to the reconstruction BDT is that the former takes advantage of all possible configurations in the event. However, because the LHD contains products of one-dimensional PDFs, it cannot properly take correlations between the variables within one configuration into account.

9.1.3 Matrix Element Method

The general idea of the MEM is to construct a discriminant, MEM_{D1}, based on ME calculations similarly to what has been done in the corresponding analysis at $\sqrt{s} = 8$ TeV [115]. At first, the MEM resembles the LHD as one builds two likelihoods, namely L_S and L_B, that express how compatible an event is with the signal and background hypotheses, respectively. More specifically, they represent the probabilities that a final state is produced either via a $t\bar{t}H(H \to b\bar{b})$ signal Feynman diagram or a $t\bar{t}$+$b\bar{b}$ background Feynman diagram, which are depicted in Figure 7.2. But in contrast to the fully simulated MC samples used for the LHD, these likelihoods are built from ME calculations at the parton level. These ME calculations cost a significant amount of CPU time, since the involved processes comprise very complex final states. Therefore, the MEM is only applied in the most sensitive single lepton region which is $\mathrm{SR}_1^{\geq 6j}$.

The two likelihoods are computed as follows:

$$L_i = \sum \int \frac{f_1\left(x_1, Q^2\right) f_2\left(x_2, Q^2\right)}{|\vec{q}_1||\vec{q}_2|} |\mathcal{M}_i\left(\mathbf{Y}\right)|^2 T\left(\mathbf{X}; \mathbf{Y}\right) \mathrm{d}\Phi_n\left(\mathbf{Y}\right). \tag{9.3}$$

In this definition, f_j are the PDFs for a certain parton of the proton with momentum $\vec{q}_j$ to carry the momentum fraction x_j of the proton in a collision at energy scale Q^2. $\mathcal{M}_i$ is the ME for either the signal or background Feynman diagram and is calculated for a phase space configuration $\mathbf{Y}$ at the parton level. T is the transfer function that represents the probability that a jet at the reconstruction level $\mathbf{X}$ originates from the parton level configuration $\mathbf{Y}$. Since only the information about $\mathbf{X}$ is available, integration over all other parameters is carried out via the infinitesimal phase space element $\mathrm{d}\Phi_n$. This integration considers the undetected neutrino in the final state. Furthermore, the transfer function T helps to constrain the size of the phase space significantly. The sum goes over all different possible initial states, but, specifically in this analysis, only gluon-induced Feynman diagrams are considered to reduce the required CPU time.

The ME calculations are performed with MadGraph5_aMC@NLO at LO with the CT10 PDF set which is interfaced via the LHAPDF package [128]. The transfer functions that map the detector signals to parton level quantities are derived from a $t\bar{t}$ sample simulated

with POWHEG +PYTHIA 6 and are validated with the nominal $t\bar{t}$ generated by POWHEG +PYTHIA 8. Moreover, the direction in η and ϕ of all visible final state objects is expected to be well measured and thus, the corresponding transfer functions are given by δ functions. The neutrino momentum in the calculation is constrained as the total p_{T} is conserved in each event, while the z-coordinate of its momentum is integrated over using VEGAS [129] and following the principles in Ref. [130]. As for the other MVAs, the MEM exploits the information from b-tagging to reduce the number of possible jet-to-parton matching configurations. Finally, the MEM discriminant is defined as the difference of the logarithms of the signal and background likelihood:

$$\mathrm{MEM}_{D1} = \log\left(L_S\right) - \log\left(L_B\right). \tag{9.4}$$

9.1.4 Classification BDT

In this search, the classification BDT is used to perform the final discrimination of the $t\bar{t}H(H \to b\bar{b})$ signal from the $t\bar{t} + b\bar{b}$ background after the event has been fully reconstructed. Similarly to the reconstruction BDT, it is trained using the TMVA Toolkit. The training and evaluation is done on two statistically independent samples, while the $t\bar{t}H(H \to b\bar{b})$ process is set as the signal and $t\bar{t} + b\bar{b}$ as the background.
The inputs to the BDT are:

1. general kinematic variables such as invariant masses and angular separations of final state objects;
2. topological variables such as event shapes or the number of jets with a certain minimum p_{T};
3. additional kinematic variables based on objects reconstructed using the reconstruction BDT;
4. the reconstruction BDT score itself;
5. the LHD, but only in the single lepton SRs;
6. the MEM discriminant, but only in $\mathrm{SR}_1^{\geq 6j}$; and
7. variables built from the b-tagging weights of selected jets.

Only the single lepton BDT includes all three MVAs from the previous sections, namely the reconstruction BDT, the LHD and the MEM discriminant. All three appear to exploit similar variables and thus be strongly correlated. However, the variables are always used from different perspectives and therefore provide not entirely uncorrelated, but still complementary information that ultimately increases the discrimination power of this BDT. In addition to this, the most powerful variables in the training of the single lepton BDT are specifically the respective outputs of these three intermediate MVAs. Despite this, the dileptonic BDT does not include the LH and MEM discriminants, because of the

associated uncertainties that would be introduced and the additional CPU time required.

During the training, one variable is added at a time and the best BDT performance as a function of the number of input variables is determined. If the performance reaches a plateau, no more variables are added. Therefore, the number of variables may differ between the regions.

For the single lepton classification BDT, the training is done separately for regions with exactly five and at least six jets, but inclusively on events with at least four b-jets tagged at the 85% WP, which is the same strategy as for the reconstruction BDT training. However, a dedicated training on the $\mathrm{SR}_1^{\geq 6j}$ is pursued, because it includes the MEM discriminant. Furthermore, variables based on b-tagged jets are built from the four jets falling into the highest PC bins of the MV2c10 algorithm, where jets falling into the same bin are ordered by their p_{T}.

On the other hand, each dileptonic SR follows its own dedicated training to further optimise the variables used for the different background components. While different b-tagging WPs are considered for the jets in these regions, only those WPs are chosen for which those regions always contain b-jets [4]. Ultimately, only $\mathrm{SR}_2^{\geq 4j}$ employs variables relying on the b-tagging weights of jets. Further details about the various b-tagging WPs used to construct the input variables can be found in Ref. [4].

The input variables to the classification BDT for the single lepton SRs are listed in Table 9.4 and the ones for the dileptonic SRs are given in Table 9.5.

9.2 Profile likelihood fit method

The search for the $t\bar{t}H(H \to b\bar{b})$ process is performed with the help of the classification BDT described in the previous section. For this, the distributions of the chosen discriminants from each analysis channel and region, specified in Table 9.1, are combined. The search is executed by measuring the signal strength $\mu = \sigma^{\mathrm{Data}}_{t\bar{t}H(H\to b\bar{b})}/\sigma^{\mathrm{SM}}_{t\bar{t}H(H\to b\bar{b})}$, which is the ratio of the observed $t\bar{t}H(H \to b\bar{b})$ cross-section with respect to the value expected in the SM. The measurement of μ is done in the form of a statistical analysis, namely by constructing a likelihood function that depends on μ and maximise this function in a profile likelihood fit to data. In addition to this, the fit can be used to determine the normalisation of the most important background components and constrain the systematic uncertainties associated to them which improves the precision of the measurement as indicated below [4].

This statistical analysis is based on a binned likelihood function (LHF), $\mathcal{L}$, which is the product of Poisson probabilities over all bins of all distributions considered in the analysis. The regions themselves are not distinguished in the fit, it is only the distribution entering the fit that changes, as listed in Table 9.1. The binning of the classification

Variable	Definition	Region	
		$\geq 6j$	$5j$
General kinematic variables			
$\Delta R_{bb}^{\text{avg}}$	Average ΔR for all b-jet pairs	✓	✓
$\Delta R_{bb}^{\max p_{\text{T}}}$	ΔR between the b-jet pair with largest vector sum p_{T}	✓	-
$\Delta\eta_{jj}^{\max \Delta\eta}$	Maximum $\Delta\eta$ between any two jets	✓	✓
$m_{bb}^{\min \Delta R}$	Mass of the b-jet pair with the smallest ΔR	✓	-
$m_{jj}^{\min \Delta R}$	Mass of the jet pair with the smallest ΔR	-	✓
N_{30}^{H}	Number of b-jet pairs with 95 GeV $< m_{bb} <$ 155 GeV	✓	✓
$H_{\text{T}}^{\text{had}}$	Scalar sum p_{T} of all jets	-	✓
$\Delta R_{\text{lep},b\bar{b}}^{\min \Delta R}$	ΔR between lepton and the b-jet pair with smallest ΔR	-	✓
Aplanarity	$1.5\lambda_2$, where λ_2 is the second eigenvalue of the momentum tensor [131] built from all jets	✓	✓
H_1	The second Fox-Wolfram moment [132] built from all jets and the lepton	✓	✓
Variables from the reconstruction BDT output			
BDT score	BDT output	✓*	✓*
m_H	Higgs boson mass	✓	✓
$m_{H,b_{t_{\text{lep}}}}$	Mass of the Higgs boson and b-jet from the leptonic top	✓	-
$\Delta R_{b\bar{b}}^{H}$	ΔR between the b-jets from the Higgs boson	✓	✓
$\Delta R_{H,t\bar{t}}$	ΔR between the Higgs boson and $t\bar{t}$ pair	✓*	✓*
$\Delta R_{H,t_{\text{lep}}}$	ΔR between the Higgs boson and the leptonic top	✓	-
$\Delta R_{H,b_{t_{\text{had}}}}$	ΔR between the Higgs boson and the b-jet from the hadronic top	-	✓*
Variable from the LHD calculation			
D	Likelihood discriminant	✓	✓
Variable from the MEM calculation			
MEM_{D1}	Matrix element method	✓	-
Variables from b-tagging (not used in $\text{SR}_1^{\geq 6j}$)			
w_b^H	Sum of binned b-tagging weights of jets from best Higgs candidate	✓	✓
B_{j^3}	b-tagging weight of the third b-jet, sorted by weight	✓	✓
B_{j^4}	b-tagging weight of the fourth b-jet, sorted by weight	✓	✓
B_{j^5}	b-tagging weight of the fifth b-jet, sorted by weight	✓	✓

Table 9.4: Input variables for the classification BDT used in the single lepton SRs. The variables with (without) a * superscript are built from objects reconstructed by the reconstruction BDT with (without) using Higgs boson information

BDT is optimised to maximise the sensitivity of the search and at the same time keep the total MC statistical uncertainty in each bin such that the predicted number of events is not biased from statistical fluctuations.
The LHF depends on the signal strength μ as well as θ which is the set of nuisance parameters that represent the impact of the considered systematic uncertainties on the expected number of signal and background events. The nuisance parameters are implemented as either Gaussian, log-normal or Poissonian prior probabilities in the LHF [77]. θ also contains the normalisation factors (k-factors) for the $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ background components. However, no prior terms are assumed for these factors and thus,

Variable	Definition	Region		
		$SR_1^{\geq 4j}$	$SR_2^{\geq 4j}$	$SR_3^{\geq 4j}$
General kinematic variables				
m_{bb}^{min}	Minimum invariant mass of b-jet pairs	✓	✓	-
m_{bb}^{max}	Maximum invariant mass of b-jet pairs	-	-	✓
$m_{bb}^{\text{min}\,\Delta R}$	Mass of the b-jet pair with the smallest ΔR	✓	-	✓
$m_{jj}^{\text{max}\,p_{\text{T}}}$	Mass of the jet pair with the largest p_{T}	✓	-	-
$m_{bb}^{\text{max}\,p_{\text{T}}}$	Mass of the b-jet pair with the largest p_{T}	✓	-	✓
$\Delta\eta_{bb}^{\text{avg}}$	Average $\Delta\eta$ of b-jet pairs	✓	✓	✓
$\Delta\eta_{l,j}^{\text{max}}$	Maximum $\Delta\eta$ between a jet and a lepton	-	✓	✓
$\Delta R_{bb}^{\text{max}\,p_{\text{T}}}$	ΔR between the b-jet pair with maximum p_{T}	-	✓	✓
N_{30}^{H}	Number of b-jet pairs with 95 GeV $< m_{bb} <$ 155 GeV	✓	✓	-
$N_{\text{jets}}^{p_{\text{T}}>40}$	Number of jets with $p_{\text{T}} > 40$ GeV	-	✓	✓
$\Delta R_{\text{lep},b\bar{b}}^{\text{min}\,\Delta R}$	ΔR between lepton and the b-jet pair with smallest ΔR	-	✓	-
Aplanarity	$1.5\lambda_2$, where λ_2 is the second eigenvalue of the momentum tensor [131] built from all jets	-	✓	-
H_{T}	Scalar sum p_{T} of all jets and leptons	-	-	✓
Variables from the reconstruction BDT output				
BDT score	BDT output	✓**	✓**	✓
m_H	Higgs boson mass	✓	-	✓
$\Delta R_{H,t\bar{t}}$	ΔR between the Higgs boson and $t\bar{t}$ pair	✓*	-	-
$\Delta R_{H,l}^{\text{min}}$	Minimum ΔR between the Higgs boson and the lepton	✓	✓	✓
$\Delta R_{H,b}^{\text{min}}$	Minimum ΔR between the Higgs boson and b-jet from top	✓	✓	-
$\Delta R_{H,b}^{\text{max}}$	Maximum ΔR between the Higgs boson and b-jet from top	-	✓	-
ΔR_{bb}^{H}	ΔR between the b-jets from the Higgs boson	-	✓	-
Variables from b-tagging				
w_b^H	Sum of binned b-tagging weights of jets from best Higgs candidate	-	✓	-

Table 9.5: Input variables for the classification BDT used in the dileptonic SRs. The variables with (without) a * superscript are built from objects reconstructed by the reconstruction BDT with (without) using Higgs boson information, while those variables with ** are based on reconstructed objects from both BDT configurations.

they are only constrained by the fit to data. Before the fit, the two k-factors are set to unity, corresponding to the prediction from Powheg +Pythia 8 for the fraction of both components relative to the inclusive $t\bar{t}$ +jets prediction.
With this, the total number of expected events in each bin depends on both μ and θ. The CRs are strongly dominated by background and thus contain a large number of events. By fitting the nuisance parameters not only in the SRs, but also in those CRs, they are much more constrained and, therefore, their impact on the sensitivity of this search is reduced significantly which improves the precision of this analysis.

From the LHF $\mathcal{L}(\mu, \theta)$, one can construct a test statistic for hypothesis testing, which

in this case is the profile likelihood ratio (PLR):

$$t_\mu = -2 \ln \left(\mathcal{L}\left(\mu, \hat{\hat{\theta}}_\mu\right) / \mathcal{L}\left(\hat{\mu}, \hat{\theta}\right) \right), \tag{9.5}$$

where $\hat{\mu}$ and $\hat{\theta}$ represent the respective parameter values that maximise the LHF under the condition that $0 \leq \hat{\mu} \leq \mu$; and $\hat{\hat{\theta}}_\mu$ are the nuisance parameter values that maximise the LHF for a given value of μ. Given that this search is performed as a statistical analysis with two hypotheses, namely either the presence or the absence of a $t\bar{t}H(H \to b\bar{b})$ signal, the PLR is the most powerful test statistic according to the Neyman-Pearson lemma [77]. The PLR is not only used to measure the compatibility of the observed data with the hypothesis of an absent signal (background-only hypothesis, given by $\mu = 0$), but also to make statistical inferences about μ, for example through upper limits on μ with the CL_S method [133–135].

By using the LHF, μ is extracted from simultaneously fitting the distributions of the discriminants in all fifteen considered regions to data. In this profile likelihood fit, μ is allowed to float freely, but it must have the same value in all regions. The normalisation of each background component is taken from the fit as well. Regarding these, the contamination of the signal from the $t\bar{t}$, V+jets, single top, Diboson and $t\bar{t} + V$ processes is constrained by the uncertainties of the corresponding theory calculations, the uncertainty on the luminosity, and the data. The statistical uncertainty of the predictions is given by the statistical uncertainty in the MC events and in the data-driven fake and non-prompt lepton estimate. This uncertainty is included as additional nuisance parameters, namely one dedicated parameter for each of the included bins. Finally, the uncertainty of the best-fit value of the signal strength $\hat{\mu}$ can be determined by varying t_μ by one unit.

The validation and performance of this fit procedure is discussed in Chapter 11. Before showing those results, the following chapter introduces the various systematic uncertainties considered in this analysis and describes them in detail.

CHAPTER 10

Systematic uncertainties

In this chapter, each individual source of systematic uncertainty considered in the $t\bar{t}H(H \to b\bar{b})$ search is discussed in detail. Except for a few instances of cross-section or normalisation uncertainties, the impact of each systematic uncertainty on the signal strength μ is established for ranking purposes by repeating the profile likelihood fit while only applying the associated systematic variation and leaving all other parameters untouched. Outside of that, all uncertainties are included and fit simultaneously. The considered uncertainties may either affect only the normalisation of simulated samples or change both the normalisation and shape of the distributions of the discriminants that enter the fit.
The uncertainties considered in this analysis are related to the estimation of the total integrated luminosity; the reconstruction and identification of leptons and jets; and the modelling of signal and background processes using MC generators. All of the experimental uncertainties considered affect both the normalisation and shape of distributions in all simulated samples, except for the luminosity uncertainty which only changes the overall normalisation. Uncertainties related to the physics modelling of signal and background events impact either both the normalisation and shape of distributions of their corresponding processes, or only the normalisation such as cross-section uncertainties. Because the normalisation uncertainties change the relative fractions of different backgrounds, they have an effect on the shape of the discriminant distributions in the different analysis regions [4].

Each individual source of a systematic uncertainty is considered uncorrelated, but correlations among the systematic uncertainties themselves are maintained over all regions and channels. Each source of a systematic uncertainty is represented by an individual independent nuisance parameter. These are described in more detail in the following sections. Some systematic uncertainties, especially most of the experimental ones de-

scribed in the next section, comprise multiple independent sources. Similarly, some of the modelling uncertainties, in particular those related to the $t\bar{t}$ +jets modelling, are split up into multiple sources of uncertainties that impact different sub-components of a certain process in an independent way. These are then accounted for by adding corresponding nuisance parameters to the fit.

The dominant sources of systematic uncertainties are related to our limited understanding i.e. modelling of the $t\bar{t}$ +jets process, in particular the $t\bar{t} + b\bar{b}$ and $t\bar{t} + c\bar{c}$ processes; the uncertainties associated to the b-tagging efficiencies which are based on the old calibration analysis; and the jet energy scale uncertainty. On the other hand, uncertainties related to leptons, $E_{\mathrm{T}}^{\mathrm{miss}}$ and the trigger system have a sub-dominant impact on this measurement.

10.1 Luminosity

In this analysis, the integrated luminosity, derived from the instantaneous luminosity, is a measure for the amount of data events recorded with the ATLAS detector at the LHC in the years 2015 and 2016. However, the instantaneous luminosity is estimated via dedicated measurements, the discussion of which are outside the scope of this thesis.

The uncertainty on the integrated luminosity of the analysed data is 2.1%. This value has been derived following a similar method to the one in Ref. [33]. The idea is to perform a preliminary calibration of the luminosity scale using $x - y$ beam separation scans. These have been performed in August 2015 and May 2016 and their uncertainties are partially correlated.
In addition to this, a variation in the pile-up reweighting of MC events is considered which covers the uncertainty in the ratio of the simulated to observed inelastic cross-sections in the fiducial volume given by $M_X > 13$ GeV, where M_X represents the mass of the hadronic system [136].

10.2 Reconstructed objects

This section summarises the systematic uncertainties related to the trigger system as well as the reconstruction and identification of physics objects and corresponding momentum or energy scale and resolution effects.

1. Jets
    - The uncertainty on the jet energy scale is estimated from test-beam data, LHC collision data and simulation [58] which amount to eight independent sources. However, additional uncertainties are taken into account which are related to pile-up corrections as well as the flavour, η and p_{T} of jets, summing up to twenty independent sources. This uncertainty is in the order of 1% to 6% per jet, depending on its p_{T}. But due to the large number of jets expected

in the final state of the $t\bar{t}H(H \to b\bar{b})$ and $t\bar{t}+b\bar{b}$ processes, the impact of this uncertainty on the measurement becomes one of the dominant factors.

- The uncertainty on the jet energy resolution is considered as well and consists of two independent sources.
- An uncertainty is added related to the efficiency of the JVT which is used to suppress the contamination from pile-up jets.

2. Flavour-tagging of jets

   - The b-tagging efficiency calibration is based on the measurement of dileptonic $t\bar{t}$ events in $\sqrt{s} = 13$ TeV data [75]. This measurement is employing the old PDF method and in consequence, the uncertainties on the efficiency SFs are still largely dominated by the modelling of the $t\bar{t}$ process. They lie between 2% and 10% depending on the WP and p_{T} of the jet. Because the PC b-tagging strategy is used in this search, thirty independent sources are considered for these uncertainties.
   - Other jets can be falsely b-tagged as well which is represented by the mis-tag rate. The mis-tag rate of c-jets is measured in semileptonic $t\bar{t}$ events where one of the W bosons decays hadronically into a cs quark pair [101]. The uncertainties on the c-jet mis-tag rate ranges between 5% to 20%, depending on the WP and jet p_{T} and consists of fifteen independent sources.
   - The light-jet mis-tag rate is measured in QCD-induced multi-jet events with secondary vertices and tracks that have an IP which suggests a negative lifetime [64], see Section 6.1.1. The uncertainties on this mis-tag rate are between 10% and 50%, depending on the WP, jet p_{T} and jet η and comprise eighty independent sources.
   - These uncertainties are extracted for each WP as a function of the jet kinematics and then combined into a calibration of the MV2c10 distribution and corresponding uncertainties that correctly take the correlations between the WPs into account.
   - Jets from τ_{had} candidates are treated like c-jets in terms of the mis-tag rate corrections and systematic uncertainties. An additional nuisance parameter is included to account for the extrapolation between c-jets and jets from τ_{had} candidates.

3. Charged leptons

   - The charged lepton objects are subject to uncertainties of the trigger system [78] as well as the reconstruction, identification and isolation algorithms. Moreover, uncertainties are associated to the energy (momentum) scale and resolution of the electron (muon) candidate. These are measured in data and are based on the leptons in $Z \to ll$, $J/\Psi \to ll$ as well as $W \to e\nu$ events [52, 54]. These uncertainties are split up into twenty-four independent sources, but overall have only a small impact on the end result.

4. Missing transverse momentum
    - All uncertainties in the energy scales and resolutions are propagated to the $p_{\mathrm{T}}^{\mathrm{miss}}$ that represents the undetected neutrinos.
    - A total of three additional uncertainties are added to account for the scale and resolution of the soft term which is used to reconstruct the $E_{\mathrm{T}}^{\mathrm{miss}}$ [63].

10.3 Signal and background modelling

Below is a summary of the physics modelling i.e. MC simulation related uncertainties that are taken into account during the fit. The uncertainties are mostly represented by individual nuisance parameters that enter the fit, but some sources are combined into one single parameter.

1. Modelling of the $t\bar{t}H$ signal process:
    - The inclusive production cross-section uncertainty from simulation is $^{+5.8\%}_{-9.2\%}$(scale) $\pm 3.6\%$(PDF), where the first term originates from the QCD scale uncertainty and the second from the PDF+α_S uncertainty [15, 102–106]. Both components are treated as uncorrelated in the fit. Their contribution to the shape of the distributions of the discriminants is found to be insignificant and thus neglected [4].
    - The branching fractions of the Higgs boson also carry uncertainties, which for the $b\bar{b}$ decay are 2.2% [15].
    - The nominal sample was simulated with MadGraph5_aMC@NLO +Pythia 8. To estimate the parton shower and hadronisation model uncertainty, the fit is repeated with a sample modelled by MadGraph5_aMC@NLO interfaced to Herwig++.
2. Modelling of the $t\bar{t}$ +jets background process:
    - The $t\bar{t}$ +jets sample is normalised to the NNLO+NNLL cross-section prediction [114]. During the calculation, the factorisation and renormalisation scale, the PDF, α_S and the top quark mass are systematically varied. To account for these effects, an uncertainty of $\pm 6\%$ is considered.
    - The primary HFC categories i.e. $t\bar{t}+ \geq 1b$, $t\bar{t}+ \geq 1c$ and $t\bar{t}+$ light all are subject to different systematic uncertainties:
        - The $t\bar{t}+$ light simulation has contributions from additional diagrams and thus profits from precise measurements based on data [4].
        - The modelling of $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ depends on the flavour scheme used for the PDF. Therefore, these categories can have similar or different diagrams. The mass difference between c- and b-quarks further contributes to the difference between both processes.
        - Because of this, every $t\bar{t}$ +jets modelling related uncertainty is represented by an independent nuisance parameter for the three categories, except for the inclusive cross-section.

- The normalisations of the $t\bar{t}$ + heavy flavour components, $k\,(t\bar{t}+\geq 1b)$ and $k\,(t\bar{t}+\geq 1c)$, are free parameters of the fit. The uncertainties that affect their shape contributions are determined by comparing the nominal sample to multiple alternative setups described further below. The events from these alternative samples are reweighted such that the predicted $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ fractions match those in the nominal sample. Specifically for the $t\bar{t}+\geq 1b$ category, the resulting uncertainties associated with the relative normalisation of each of its four sub-components $t\bar{t}+b, t\bar{t}+B, t\bar{t}+b\bar{b}$ and $t\bar{t}+\geq 3b$, are each represented by an individual nuisance parameter. All other alternative samples, i.e. those not used to derive uncertainties related to these fractions, are rescaled such that the relative fractions of the $t\bar{t}+\geq 1b$ sub-components match those in the Sherpa 4F setup, excluding the $t\bar{t}+b$(MPI/FSR) contribution.
- The $t\bar{t}$ nominal MC generator Powheg +Pythia 8 is compared to Powheg +Herwig 7 to estimate the uncertainty associated to the choice of the PS and hadronisation model. Here, the two Powheg setups are identical, while the PS and hadronisation of the alternative sample are modelled by Herwig v7.0.1 with the H7-UE-MMHT tune for the underlying event [40].
- Then, by also comparing the nominal generator to another generator, specifically Sherpa v2.2.1 with the ME+PS@NLO setup interfaced to OpenLoops, which changes the NLO ME as well as the PS and hadronisation model, the impact of the NLO ME generator choice can be inferred. This alternative setup is able to model one additional parton in the ME at NLO and up to four additional partons at LO. It employs the NNPDF3.0NNLO PDF set, while the renormalisation and factorisation scales are both set to $\sqrt{\left(m^2_{\mathrm{T},t}+m^2_{\mathrm{T},\bar{t}}\right)/2}$. This setup is labelled Sherpa 5F, because it uses the five flavour scheme in the PDF during the ME calculation which treats the b-quark as a massless parton, in contrast to the four flavour PDF set used in the Sherpa 4F setup which takes the b-quark mass into account and is described in Section 7.3.1.
- Two additional samples are generated with Powheg +Pythia 8 but slightly different settings each time in order to estimate the impact of the ISR and FSR modelling [137], exactly in the same way as in the b-tagging calibration analysis, see Section 6.3.2. That means, one sample increases the contribution from radiation by decreasing the renormalisation and factorisation scales by a factor of two, setting $h_{\mathrm{damp}} = 3 \cdot m_{\mathrm{top}}$ and employing the upwards A14 tune parameter variation 'Var3c' to increase PS radiation. The second sample decreases the radiation by multiplying both scales by a factor of two and using the 'Var3c' downward variation, while keeping the nominal h_{damp} value of $1.5 \cdot m_{\mathrm{top}}$. These radiation uncertainties are summarised as three independent sources for each of the $t\bar{t}+\geq 1b$, $t\bar{t}+\geq 1c$ and $t\bar{t}+$ light predictions.
- The problem that the nominal $t\bar{t}$ +jets sample simulated by Powheg +Pythia 8 only includes the additional $b\bar{b}$ pair in the PS affects a possible additional $c\bar{c}$

pair as well. The modelling of the $t\bar{t}+ \geq 1c$ component is thus believed to have suboptimal precision and hence, a dedicated setup to model the $t\bar{t}+c\bar{c}$ process in the ME is employed [4]. This sample is generated with MadGraph5_aMC@NLO to model the $t\bar{t}+c\bar{c}$ ME at NLO while using a three flavour scheme PDF to account for the non-zero mass of the c-quarks in the proton PDF. The sample is then interfaced to Herwig++ for the PS and hadronisation. The full setup is described in Ref. [138]. These same generators, but with a five flavour scheme PDF in the ME, are then used to produce an inclusive $t\bar{t}$ sample where the additional c-jets are simulated in the PS. The difference in the predictions between these two samples is taken as another independent uncertainty for the $t\bar{t}+ \geq 1c$ background.

- Following up with the previous point, the difference between the $t\bar{t}+ \geq 1b$ predictions from Powheg +Pythia 8 and the $t\bar{t}+b\bar{b}$ prediction from the Sherpa 4F setup are taken as another independent uncertainty, but for the $t\bar{t}+ \geq 1b$ component. In this uncertainty, the modelling of the $t\bar{t}+b$(MPI/FSR) sub-category is not included, as it is not simulated at NLO precision by the Sherpa 4F setup.
- The relative fractions of the $t\bar{t}+b$, $t\bar{t}+B$, $t\bar{t}+b\bar{b}$ and $t\bar{t}+ \geq 3b$ sub-components in the Powheg +Pythia 8 sample are fixed to the Sherpa 4F prediction and thus, related uncertainties are not accounted for by the above comparison. They are split into seven individual components and are estimated separately. Three of the seven uncertainties are estimated by a) multiplying the renormalisation scale by a factor of two or one half; b) changing the functional form of the resummation scale to μ_{CMMPS}, see Section 7.3.1; and c) setting all relevant scales to the CMMPS scale, i.e. $\mu_Q = \mu_R = \mu_F = \mu_{\text{CMMPS}}$. The remaining four are accounted for by d) using alternative PDF sets, namely MSTW2008NLO [139] and NNPDF2.3NLO; e) employing an alternative shower recoil scheme; and f) using an alternative tune for the UE simulation. In addition to this, a 50% normalisation uncertainty is assigned to the $t\bar{t}+ \geq 3b$ sub-component, because its prediction by the Sherpa 4F setup shows significant differences to the considered five flavour scheme setups which are not covered by the variations above [4].
- A 50% uncertainty is included for the normalisation of the $t\bar{t}+b$(MPI/FSR) sub-component of the $t\bar{t}+ \geq 1b$ category which is based on studies of different UE sets of tuned parameters. Its shape uncertainty is already accounted for by the comparison of the nominal sample to the alternative samples described above [4].
- All of the $t\bar{t}$ +jets components detailed above are summarised in Table 10.1. A total of thirteen independent systematic uncertainties are associated to the $t\bar{t}+ \geq 1b$ component, four to the $t\bar{t}+ \geq 1c$ component, three to the $t\bar{t}+$ light component and, finally, one for the inclusive $t\bar{t}$ +jets production cross-section.

3. V+jets modelling:

- The W+jets and Z+jets processes are simulated by the SHERPA generator as mentioned in Section 7.3.2. The normalisation uncertainties for these processes described below are estimated from varying the factorisation and renormalisation scales as well as the matching parameters in the simulation.
- For the W+jets normalisation, an uncertainty of 40% is taken. Additionally, specifically for the heavy flavour component, another 30% is added which is treated as uncorrelated between events with exactly two and with at least three heavy flavour jets.
- In case of the Z+jets normalisation, an uncertainty of 35% is assumed which also accounts for the correction factor of 1.3 for the heavy flavour component derived from a CR in data [4]. This uncertainty is considered uncorrelated across all jet bins.

4. Single top modelling:
   - The production cross-sections of the three considered single top processes which are Wt, t-channel and s-channel production, are each assigned an uncertainty of $^{+5\%}_{-4\%}$ [122–124].
   - The uncertainties related to the choice of PS and hadronisation model as well as ISR and FSR are estimated for the Wt and t-channel samples in a similar way to the $t\bar{t}$ +jets process, namely via alternative generator setups. Thus, for the former uncertainty, the nominal setup is compared to POWHEG interfaced to Herwig++. For the ISR and FSR uncertainties, the nominal setup is compared to alternative POWHEG +PYTHIA 6 samples where the factorisation and renormalisation scales as well as the Perugia 2012 tune parameters are varied appropriately [121].
   - An additional uncertainty is considered for the estimation strategy of the interference between Wt and $t\bar{t}$ diagrams at NLO. The nominal sample is generated with the diagram removal scheme and compared to an alternative sample which uses the diagram subtraction scheme instead [89, 90].

5. Diboson modelling:
   - For the Diboson background, a total normalisation uncertainty of 50% is employed to account for uncertainties in the inclusive cross-section and additional jet production [140].

6. $t\bar{t}+V$ modelling:
   - An uncertainty of 15% is assumed for the $t\bar{t}+V$ production cross-section at NLO, which is split into PDF and scale uncertainties [141].
   - The nominal $t\bar{t}+V$ setup is compared with alternative samples simulated by SHERPA to estimate ME, PS and hadronisation model uncertainties.
   - Uncertainties from $t\bar{t}+W$ are considered uncorrelated to the ones from $t\bar{t}+Z$.

7. Modelling of rare processes:

- The $t\bar{t}t\bar{t}$ background is assigned a 50% normalisation uncertainty
- Two uncertainties related to the cross-section are assigned to each of the small background contributions from tZ, $t\bar{t}WW$, $tHjb$ and WtH, namely the PDF and scale uncertainties. The tWZ background only has one such uncertainty which accounts for both aspects.

8. Estimate on background with non-prompt leptons:
   - A 50% uncertainty is considered for the overall estimated number of non-prompt leptons in the single lepton channel. This uncertainty is considered uncorrelated between events containing an electron or events containing a muon as well as between events with exactly five or at least six jets.
   - In the dileptonic channel, a 25% uncertainty is considered instead which is correlated across the lepton flavours and all analysis regions.

Systematic uncertainty	Description	$t\bar{t}$ categories
$t\bar{t}$ cross-section	Up or down by 6%	All, correlated
$k\,(t\bar{t}+\geq 1c)$	Free-floating $t\bar{t}+\geq 1c$ normalisation	$t\bar{t}+\geq 1c$
$k\,(t\bar{t}+\geq 1b)$	Free-floating $t\bar{t}+\geq 1b$ normalisation	$t\bar{t}+\geq 1b$
Sherpa 5F vs. nominal	Related to the choice of the ME generator	All, uncorrelated
PS and hadronisation	Powheg +Herwig 7 vs. Powheg +Pythia 8	All, uncorrelated
ISR/FSR	Variations of $\mu_R, \mu_F, h_{\text{damp}}$ and A14 Var3c parameters	All, uncorrelated
$t\bar{t}+\geq 1c$ ME vs. inclusive	MadGraph5_aMC@NLO +Herwig++: $t\bar{t}+c\bar{c}$ with 3F vs. $t\bar{t}+\geq 1c$ with 5F	$t\bar{t}+\geq 1c$
$t\bar{t}+\geq 1b$ Sherpa 4F vs. nominal	Comparison of $t\bar{t}+b\bar{b}$ with 4F vs. $t\bar{t}+\geq 1b$ with 5F	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ renormalisation scale	Multiply by a factor of 2 or 1/2	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ resummation scale	Change μ_Q from $H_{\text{T}}/2$ to μ_{CMMPS}	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ global scales	Set $\mu_Q = \mu_R = \mu_F = \mu_{\text{CMMPS}}$	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ shower recoil scheme	Alternative model scheme	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ PDF choice (MSTW)	MSTW vs. CT10	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ PDF choice (NNPDF)	NNPDF vs. CT10	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ UE	Alternative tune choice for the UE	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 1b$ MPI	Up or down by 50%	$t\bar{t}+\geq 1b$
$t\bar{t}+\geq 3b$ normalisation	Up or down by 50%	$t\bar{t}+\geq 1b$

Table 10.1: The sources of systematic uncertainties related to the physics modelling of the $t\bar{t}$ +jets background processes [4].

10.4 Pruning and smoothing of systematic uncertainties

Similarly to the strategy depicted in Section 6.5.3, this analysis employs the pruning and smoothing of certain systematic uncertainties. This is because some of the nuisance parameters would otherwise behave in an uncontrolled way during the profile likelihood fit and lead to an unstable result or even preventing the fit from converging. Furthermore, statistical fluctuations in the parameters can lead to non-physical constraints on the uncertainties. Another motivation is that the pruning and smoothing of the systematic uncertainties proves essential in order to reduce the CPU time required to perform the fit.

The normalisation and shape effects of systematic uncertainties are pruned separately through bin-by-bin variations relative to each sample in each region. The smoothing is then done in an analogous way to the b-tagging calibration analysis in that the histograms are rebinned meaning the bins are merged and then smoothed, constrained by the maximum number of shape variations allowed which is configurable for each systematic uncertainty. Here, the pruning thresholds are set to 1%. This means that a systematic uncertainty is pruned if its impact on the normalisation of a sample or if the difference of each individual bin variation with respect to the average bin variation in the discriminant distributions is less than 1%. The pruning is done separately for each systematic uncertainty and each region. This threshold value is a result found to significantly speed up the fit procedure without any observable changes to the uncertainty on μ or to constraints or pulls on any of the nuisance parameters.

CHAPTER 11

Results of the $t\bar{t}H(H \rightarrow b\bar{b})$ analysis

In this chapter, the results of the search for the presence of a $t\bar{t}H(H \rightarrow b\bar{b})$ signal are presented. They are based on the combined profile likelihood fit described in Section 9.2. The fit method as well as the results presented throughout this chapter are subject to consistency checks in Section 11.1. The plots presented in Section 11.2 show the pre-fit and post-fit modelling of the discriminant distributions and the event yields in all considered analysis regions, split up into the different signal and background categories introduced in Section 7.3. The signal strength parameter is extracted twice, once for the single lepton and dilepton channels individually and once for the combination of both channels. Along with the signal strength parameter, the most dominant uncertainties limiting the measurement are listed and the behaviour of the nuisance parameters during the fit is highlighted. Finally, upper limits on the expected and observed $t\bar{t}H(H \rightarrow b\bar{b})$ cross-sections are set. The analysis results are combined with those of other $t\bar{t}H$ search channels in the ATLAS collaboration and compared to those of the CMS experiment in Appendix F. Studies that aim to reduce the most dominant systematic uncertainties i.e. those associated to the signal and main background modelling are presented in the next chapter.

11.1 Consistency checks of the fit result

As will be highlighted in the following section, the simulation of the $t\bar{t}+ \geq 1b$ process is subject to large uncertainties. One aspect of this evaluation involves comparing $t\bar{t}+ \geq 1b$ events modelled by different MC generators and this comparison reveals significant differences in the normalisation and shape of relevant distributions. With the generators that employ the five flavour scheme PDF in the ME, the $t\bar{t}+b\bar{b}$ process, which is the actual background of interest, is only accessible via the PS simulation, not the ME generation. Therefore, the nominal generator setups currently available in ATLAS

are only able to model the $t\bar{t} + b\bar{b}$ process with a sub-optimal precision in QCD. As a consequence, the systematic uncertainties associated to the modelling of this process are dominant and limit the sensitivity of this search. Thus, it is reasonable to explore possibly false evaluations of these systematic uncertainties and their sources.
Firstly, the choice of nuisance parameters which encode the systematic uncertainties associated to the $t\bar{t}+ \geq 1b$ background is studied. For this background, thirteen independent nuisance parameters are included in the fit. The profile likelihood fit should have the capability to correct any potential mis-modellings of this background without introducing any bias in the fitted signal strength parameter. This capability to correct for mis-modellings, aside from those shown in the distributions in the following section, is tested by evaluating the agreement between data and the simulation of all input variables to the classification BDT post-fit. The result is positive as no significant disagreement between data and prediction is observed [4]. On the contrary, the agreement is improved post-fit. Alternative strategies have been considered and tested for this background, namely regarding its modelling or the definition of corresponding uncertainties and their correlations. The respective results are found to be compatible with the nominal result quoted in the following section [4].

Some of the employed nuisance parameters are shifted away from their nominal values by the fit which can be seen in Figure 11.10 in the following section. It shows the twenty nuisance parameters from independent sources of systematic uncertainty that have the largest impact on the total uncertainty on μ, ranked by decreasing impact. In order to find the cause of these shifts, those nuisance parameters are set as uncorrelated between all analysis categories and samples and the fit procedure is repeated. This reveals that the fit uses these nuisance parameters mostly to correct the $t\bar{t}$ +jets background to match the observed data in multiple regions [4]. When the fit employs the background-only hypothesis and the most signal-enriched bins are taken out of the distributions, a similar trend can be observed when repeating the fit. The impact of these shifts on the signal strength can be tested and quantified by fixing the nuisance parameters individually to their pre-fit values throughout the fit and then comparing the new result with the nominal fit result. Using this approach, the contribution of these parameter shifts are found to be smaller than the uncertainty on μ [4]. In addition to this, a strategy is followed in which independent signal strength parameters are fitted using different sets of analysis categories in the dilepton and single lepton channels. The parameter values of this test are also compatible with the nominal results [4].

It should be noted that some of the systematic uncertainties are decreased by the fit. This is because the fit strongly constrains the corresponding nuisance parameters if their variation would affect the distributions of the discriminants in such a way that they would lead to a significant discrepancy between prediction and data. To test this hypothesis, the capability of the profile likelihood fit to constrain systematic uncertainties is validated through a fit to a so-called Asimov dataset [133] which is a pseudo-dataset based on all nominal simulated samples [4].

In order to further validate the robustness of the fit method, another pseudo-dataset is constructed from all nominal simulated background samples except for the $t\bar{t}$ +jets background which is replaced by an alternative sample generated with POWHEG +PYTHIA 6. This sample is not used in the definition of any uncertainty and is similar to the sample used in the search for $t\bar{t}H(H \to b\bar{b})$ based on $\sqrt{s} = 8$ TeV ATLAS data collected during Run 1 of the LHC [115]. If there is a bias on the extraction of the signal strength, the fit to this pseudo-dataset should yield a value of μ that is incompatible with zero under the signal-plus-background hypothesis which assumes the presence of a $t\bar{t}H$ signal. After performing this fit, no such bias can be observed [4]. Finally, the ability of the fit to constrain the systematic uncertainties is validated in this fit as well.

11.2 Results before and after the fit to data

Depicted in this section are the pre-fit and post-fit distributions of the discriminants used in the combined profile likelihood fit. Figure 11.1 summarises the event yields in the various considered analysis regions of the single lepton and dilepton channels before and after the fit assuming the presence of a $t\bar{t}H$ signal. Similarly, the scalar sum p_{T} of selected jets, $H_{\mathrm{T}}^{\mathrm{had}}$, is shown in the two $t\bar{t}+ \geq 1c$ CRs of the single lepton channel in Figure 11.2. The Figures 11.3–11.7 show the output distributions of the Classification BDT in the SRs of the dilepton and single lepton channels. As can be seen, all these distributions are modelled well pre-fit within the considered uncertainties assigned to the respective predictions. In the post-fit distributions, this level of agreement between data and simulation improves, because the profile likelihood fit is able to adjust the considered nuisance parameters accordingly [4]. This is particularly important for the two k-factors of the $t\bar{t}$ + heavy flavour background, whose post-fit values correspond to $k(t\bar{t}+ \geq 1b) = 1.24 \pm 0.10$ and $k(t\bar{t}+ \geq 1c) = 1.63 \pm 0.23$, respectively [4]. The quoted uncertainties do not consider any theory related uncertainties of the $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ cross-sections. Their post-fit uncertainty is significantly reduced as well, because the fit constrains and creates correlations between the nuisance parameters [4]. Aside from this, the input variables to the classification BDT are checked post-fit in the respective SR and no significant disagreement between data and simulation is found [4]. As an example, Figure 11.8 illustrates the post-fit predictions of the Higgs boson candidate mass distribution in the single lepton and dilepton channels.

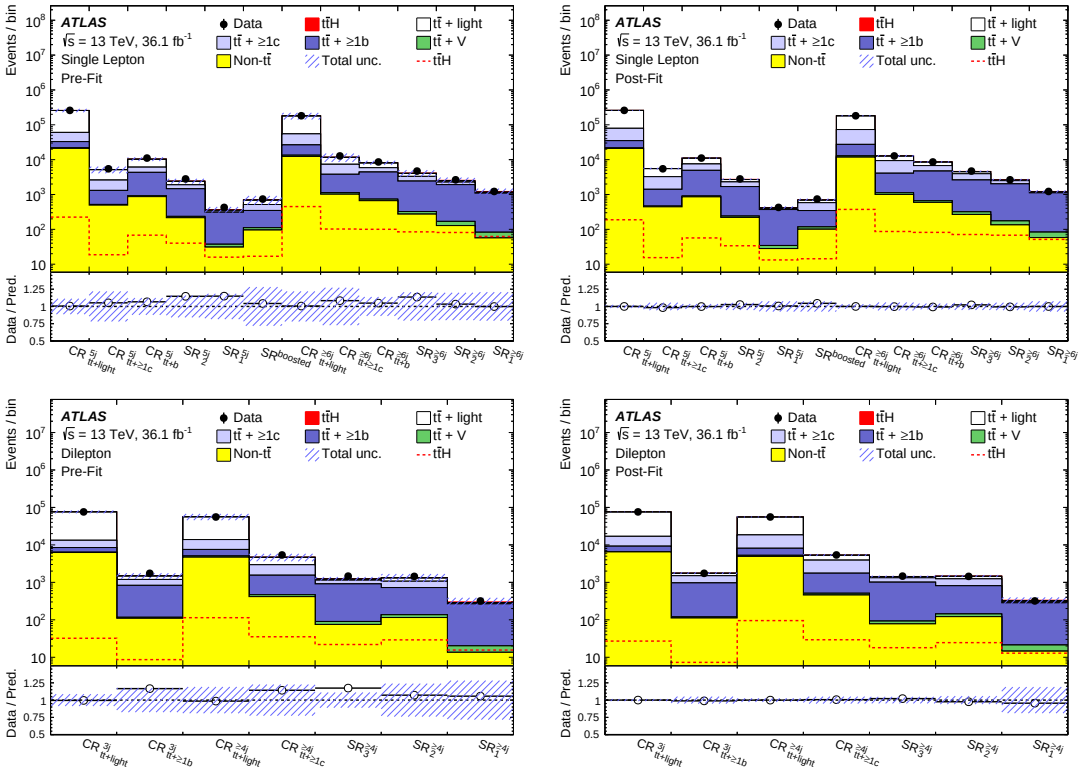

Figure 11.1: Comparison of the predicted event yields to the observed events in data in all considered analysis regions in the single lepton channel before (top left) and after the combined fit (top right) and in the dileptonic channel before (bottom left) and after the combined fit (bottom right). The filled red area (dashed red line) represents the $t\bar{t}H$ signal stacked on top of the background (shown separately) normalised to the SM cross-section before the fit and to the extracted μ value after the fit. The total uncertainty in the simulated yields is represented by the hatched area, while the histograms before the fit do not consider an uncertainty in the $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ normalisations.

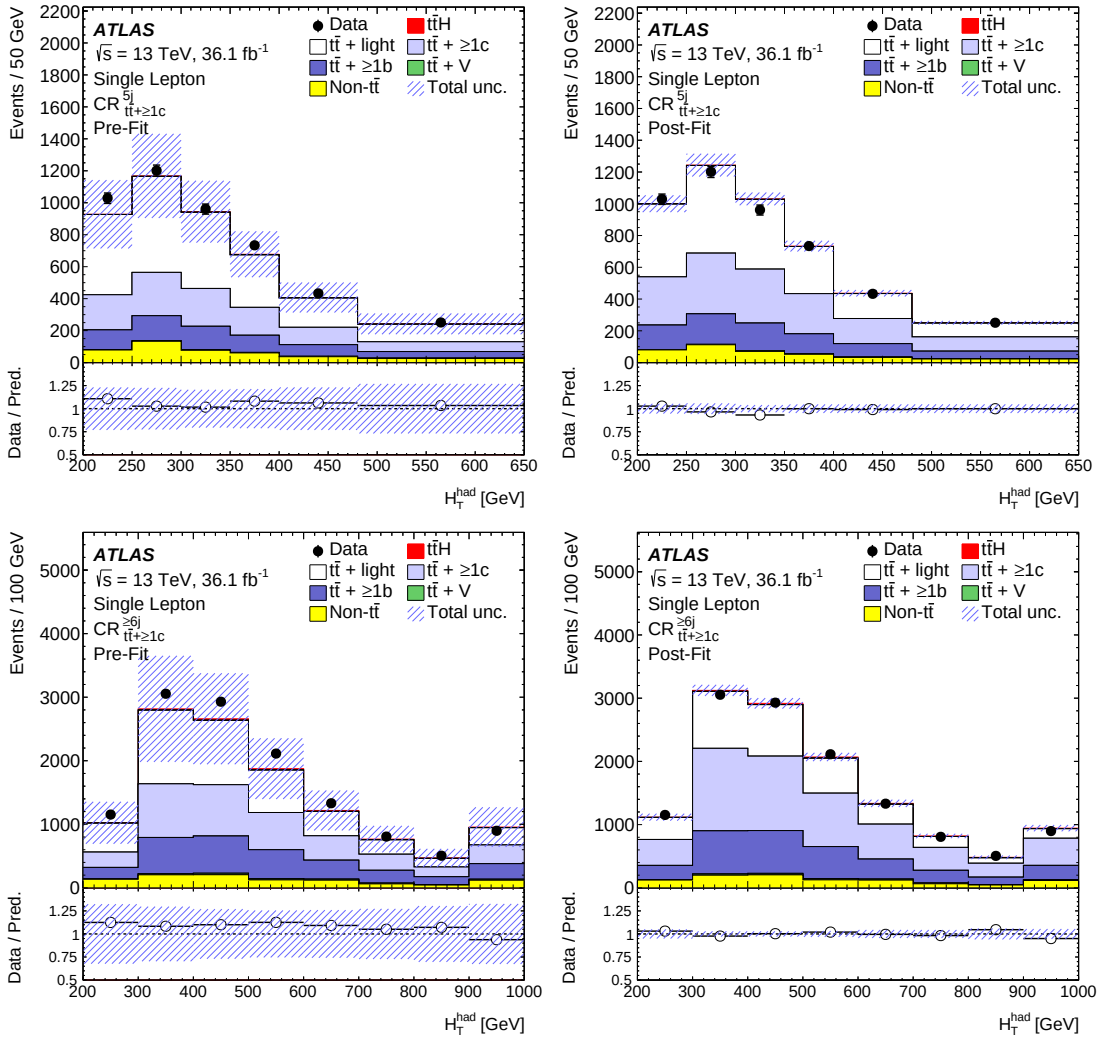

Figure 11.2: Comparison of the predicted event yields to the observed events in data as a function of $H_{\mathrm{T}}^{\mathrm{had}}$ in the single lepton $t\bar{t}+ \geq 1c$ CR with exactly five jets before (top left) and after the combined fit (top right) and in the corresponding CR with at least six jets before (bottom left) and after the combined fit (bottom right). The filled red area represents the $t\bar{t}H$ signal stacked on top of the background normalised to the SM cross-section before the fit and to the extracted μ value after the fit. The total uncertainty in the simulated yields is represented by the hatched area, while the histograms before the fit do not consider an uncertainty in the $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ normalisations.

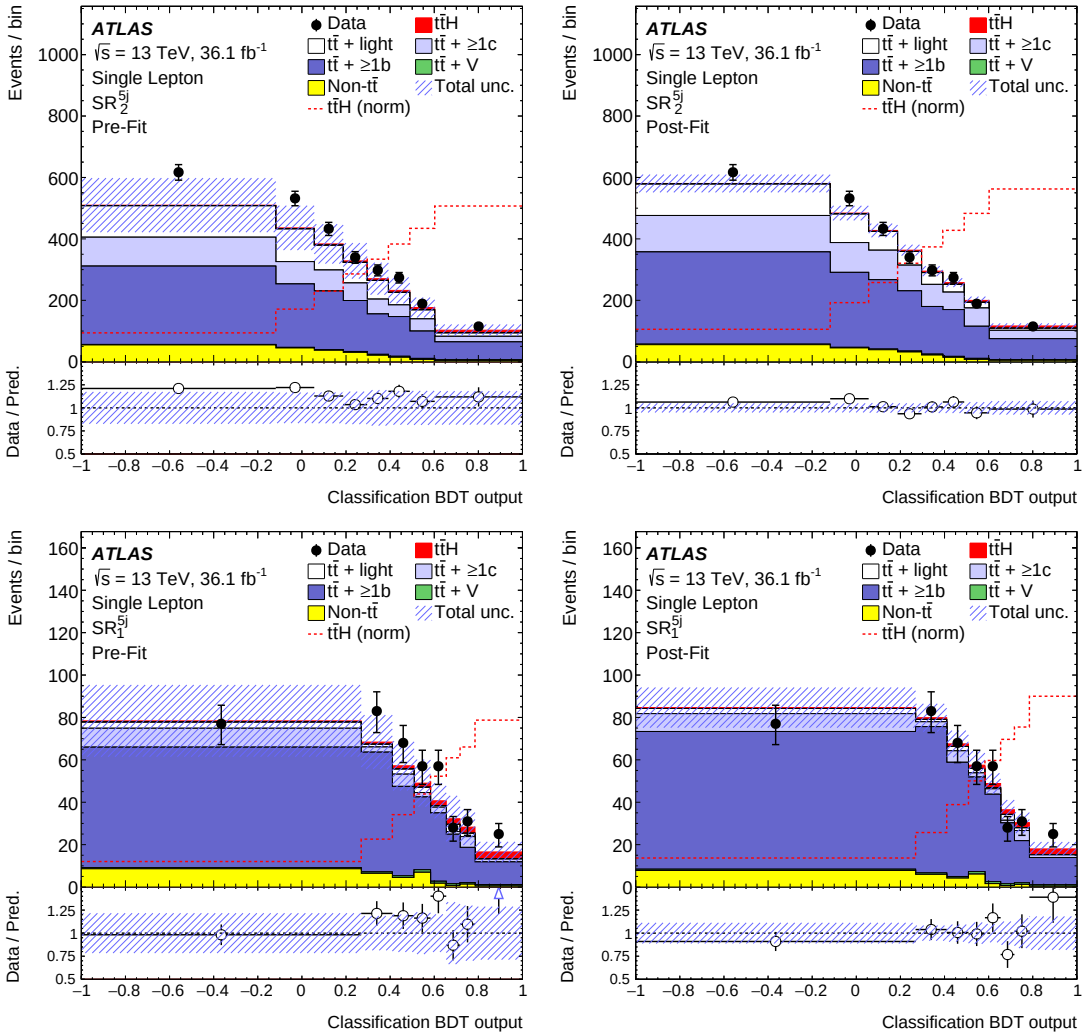

Figure 11.3: Comparison of the predicted event yields to the observed events in data as a function of the Classification BDT output in the single lepton SR_2^{5j} before (top left) and after the combined fit (top right) and in the single lepton SR_1^{5j} before (bottom left) and after the combined fit (bottom right). The filled red area represents the $t\bar{t}H$ signal stacked on top of the background normalised to the SM cross-section before the fit and to the extracted μ value after the fit. The dashed red line shows the $t\bar{t}H$ signal separately, normalised to the total background prediction. The total uncertainty in the simulated yields is represented by the hatched area, while the histograms before the fit do not consider an uncertainty in the $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ normalisations.

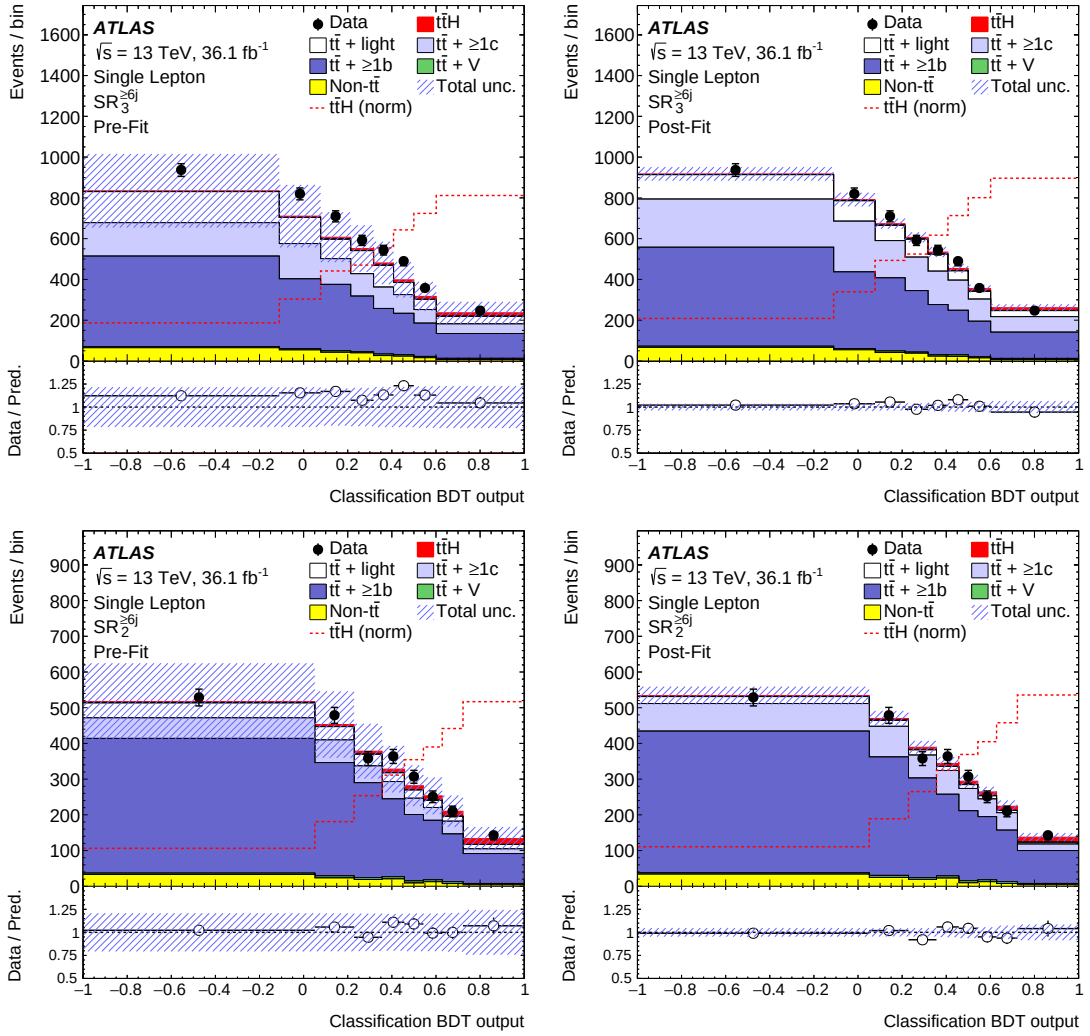

Figure 11.4: Comparison of the predicted event yields to the observed events in data as a function of the Classification BDT output in the single lepton $\mathrm{SR}_3^{\geq 6j}$ before (top left) and after the combined fit (top right) and in the single lepton $\mathrm{SR}_2^{\geq 6j}$ before (bottom left) and after the combined fit (bottom right). The filled red area represents the $t\bar{t}H$ signal stacked on top of the background normalised to the SM cross-section before the fit and to the extracted μ value after the fit. The dashed red line shows the $t\bar{t}H$ signal separately, normalised to the total background prediction. The total uncertainty in the simulated yields is represented by the hatched area, while the histograms before the fit do not consider an uncertainty in the $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$ normalisations.

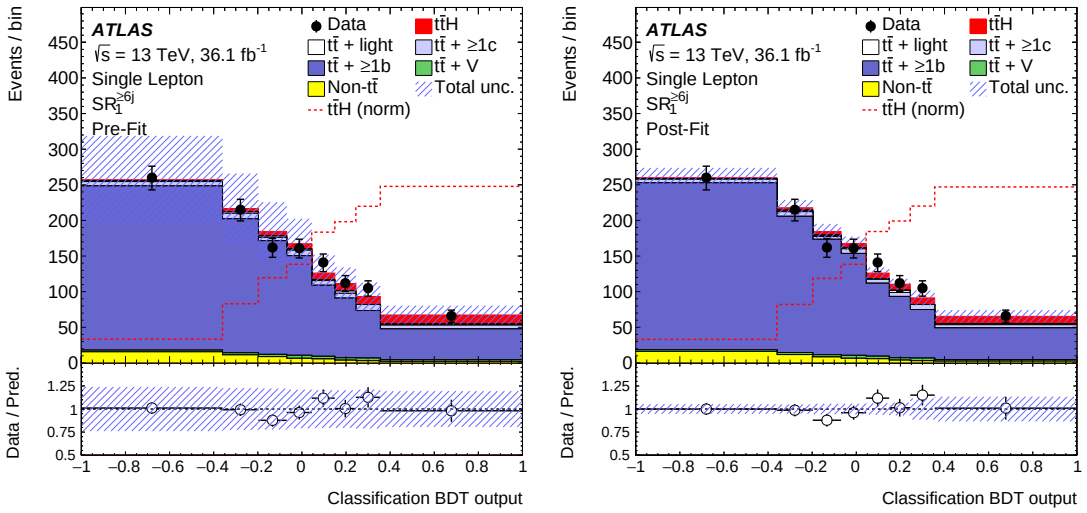

Figure 11.5: Comparison of the predicted event yields to the observed events in data as a function of the Classification BDT output in the single lepton $\mathrm{SR}_1^{\geq 6j}$ before (left) and after the combined fit (right). The filled red area represents the $t\bar{t}H$ signal stacked on top of the background normalised to the SM cross-section before the fit and to the extracted μ value after the fit. The dashed red line shows the $t\bar{t}H$ signal separately, normalised to the total background prediction. The total uncertainty in the simulated yields is represented by the hatched area, while the histograms before the fit do not consider an uncertainty in the $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ normalisations.

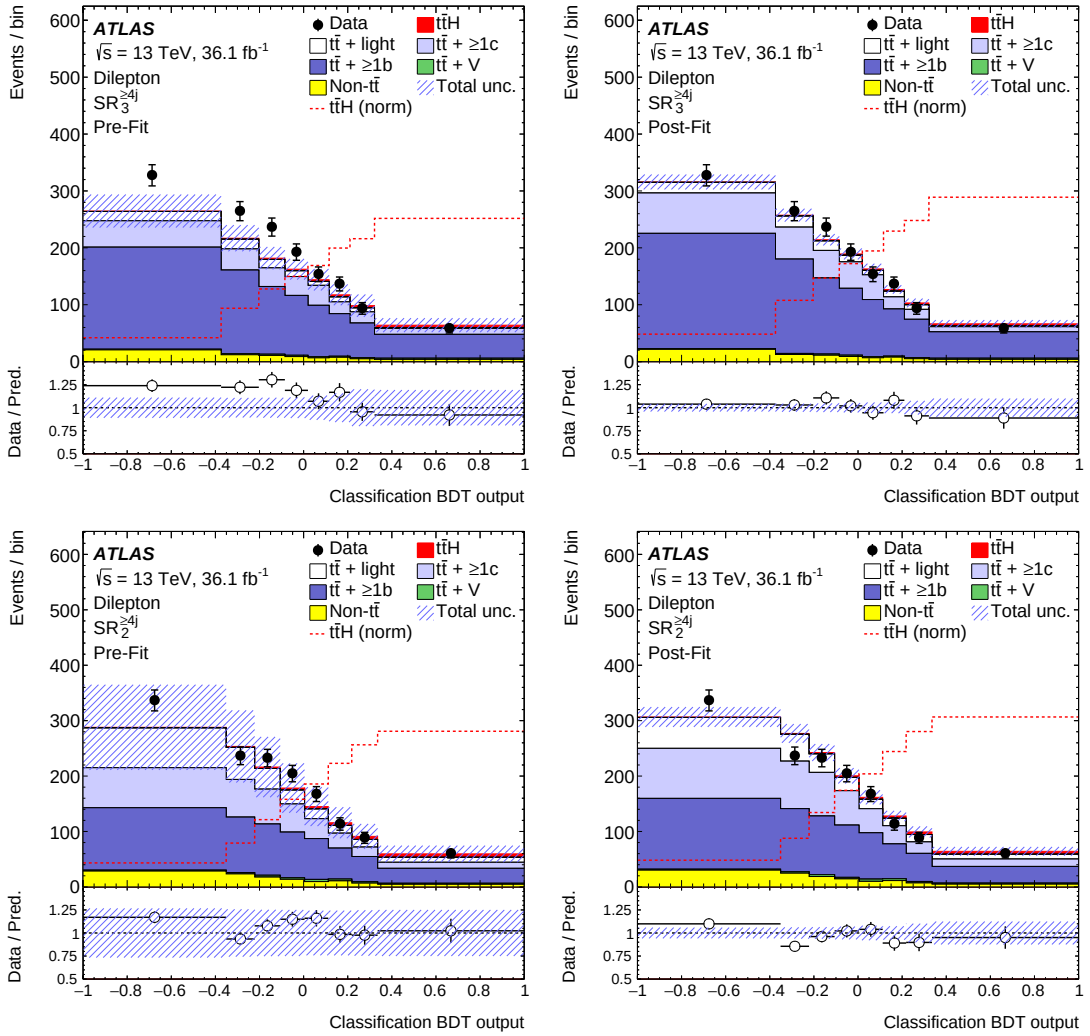

Figure 11.6: Comparison of the predicted event yields to the observed events in data as a function of the Classification BDT output in the dilepton $\mathrm{SR}_{3}^{\geq 4j}$ before (top left) and after the combined fit (top right) and in the dilepton $\mathrm{SR}_{2}^{\geq 4j}$ before (bottom left) and after the combined fit (bottom right). The filled red area represents the $t\bar{t}H$ signal stacked on top of the background normalised to the SM cross-section before the fit and to the extracted μ value after the fit. The dashed red line shows the $t\bar{t}H$ signal separately, normalised to the total background prediction. The total uncertainty in the simulated yields is represented by the hatched area, while the histograms before the fit do not consider an uncertainty in the $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ normalisations.

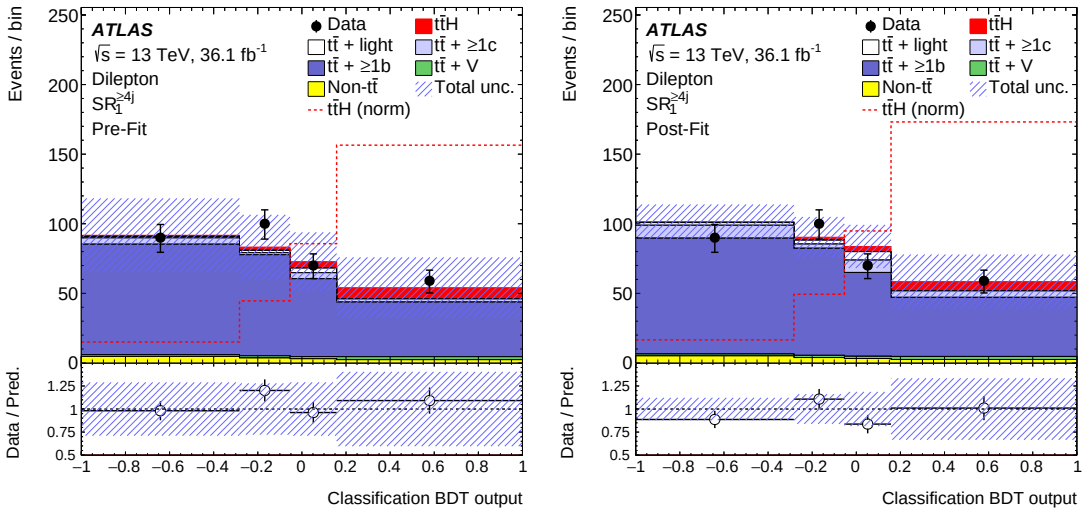

Figure 11.7: Comparison of the predicted event yields to the observed events in data as a function of the Classification BDT output in the dilepton $\mathrm{SR}_1^{\geq 4j}$ before (left) and after the combined fit (right). The filled red area represents the $t\bar{t}H$ signal stacked on top of the background normalised to the SM cross-section before the fit and to the extracted μ value after the fit. The dashed red line shows the $t\bar{t}H$ signal separately, normalised to the total background prediction. The total uncertainty in the simulated yields is represented by the hatched area, while the histograms before the fit do not consider an uncertainty in the $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ normalisations.

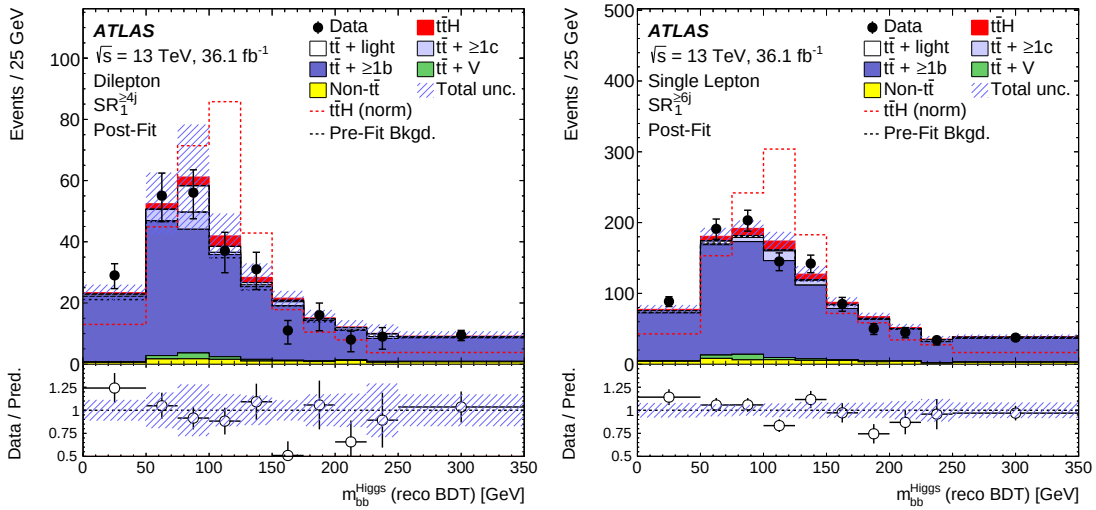

Figure 11.8: Comparison of the predicted event yields to the observed events in data as a function of the Higgs boson candidate mass from the reconstructed BDT trained without using Higgs boson information in the dilepton $\mathrm{SR}_1^{\geq 4j}$ (left) and the single lepton $\mathrm{SR}_1^{\geq 6j}$ (right) after the combined fit. The filled red area represents the $t\bar{t}H$ signal stacked on top of the background normalised to the SM cross-section before the fit and to the extracted μ value after the fit. The dashed red line shows the $t\bar{t}H$ signal separately, normalised to the total background prediction. The total uncertainty in the simulated yields is represented by the hatched area.

The signal strength is extracted from the combined fit to data, namely the simultaneous fit of all fifteen single lepton and dilepton regions. The best-fit value is [4]:

$$\mu = 0.84 \pm 0.29(\text{stat.})^{+0.57}_{-0.54}(\text{syst.}) = 0.84^{+0.64}_{-0.61}. \tag{11.1}$$

The observed uncertainty is identical to the one expected from the fit to the Asimov dataset. A separate fit is also performed in which both the single lepton and dilepton channels are included in the combined fit, but allowed to have independent signal strength parameters. The corresponding results are:

$$\begin{aligned} \mu_{\text{dilepton}} &= -0.24^{+1.02}_{-1.05}, \\ \mu_{\text{single-lepton}} &= 0.95^{+0.65}_{-0.62}. \end{aligned} \tag{11.2}$$

The probability to observe a discrepancy between these two parameters that is equal to or larger than the quoted values is 19% [4]. The three signal strength parameters and their respective uncertainties are depicted in Figure 11.9.

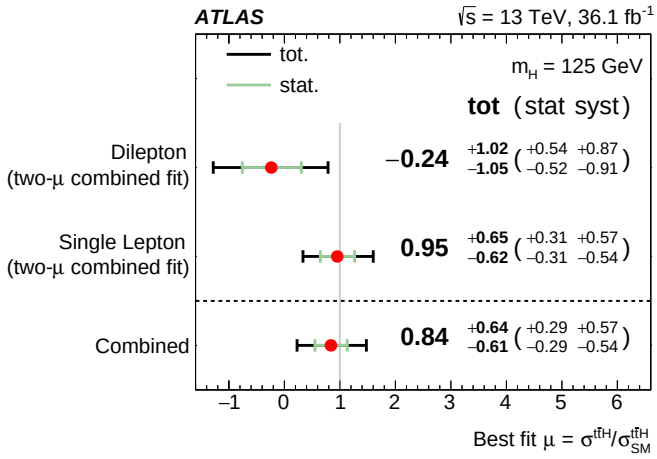

Figure 11.9: The signal strength parameter values μ of the simultaneous profile likelihood fit to ATLAS data in all analysis regions. Shown is the combined result as well as the individual results of the single lepton and dilepton channels, where the signal strengths are treated uncorrelated while keeping all nuisance parameters correlated across the channels.

The statistical uncertainty on μ is determined by repeating the fit to data after fixing all nuisance parameters to their corresponding post-fit values, except for the normalisation factors allowed to float freely, namely $k(t\bar{t}+\geq 1b), k(t\bar{t}+\geq 1c)$ and μ. The total systematic uncertainty is then computed by subtracting this statistical uncertainty in quadrature from the total uncertainty. It is important to note that the statistical uncertainty is only a minor component of the total uncertainty, the systematic component is significantly more dominant.

In addition to the combined fit above, another fit configuration is studied in which the single lepton and dilepton regions are fitted completely independently of each other. The corresponding signal strength parameters are:

$$\begin{aligned} \mu_{\text{dilepton}} &= 0.11^{+1.36}_{-1.41}; \\ \mu_{\text{single-lepton}} &= 0.67^{+0.71}_{-0.69}. \end{aligned} \tag{11.3}$$

As is expected and can be seen, the total uncertainties of both results are significantly larger than those of the combined fit. It is also important to note that the best-fit value of both parameters is smaller than the observed combined result of μ quoted in Equation 11.1. This has been found to originate from the large correlations between the systematic uncertainties associated to the simulated background contributions in both channels [4].

The impact of each independent source of uncertainty on μ in the combined fit is listed in Table 11.1. The total systematic uncertainty is clearly dominated by the modelling uncertainty on the $t\bar{t}+\geq 1b$ background prediction. The second source in the ranking is the limited number of events in the MC samples which leads to significant statistical fluctuations in certain distributions, but it also includes the uncertainty component in the data-driven estimation of the non-prompt and fake lepton background in the single lepton channel. After these, the most highly ranked sources are the uncertainties in the b-tagging efficiency SFs, the jet energy scale and resolution, and the MC simulation of the $t\bar{t}H$ signal.
Figure 11.10 shows the twenty nuisance parameters from independent sources of systematic uncertainty that have the largest impact on the total uncertainty on μ and they are ranked by decreasing impact. Here, the best-fit value for each nuisance parameter and its corresponding uncertainty post-fit is depicted. The nuisance parameters with the largest impact on the signal strength parameter uncertainty, $\Delta\mu$, are all related to the imperfect modelling of the $t\bar{t}+\geq 1b$ background: the highest contribution to the total uncertainty originates from the comparison between the nominal prediction of $t\bar{t}+\geq 1b$ and the one simulated by the Sherpa 5F setup which represents the uncertainty in the choice of the ME modelling. The three highest contributing nuisance parameters after this are also related to the $t\bar{t}+\geq 1b$ modelling. In addition to this, sources of systematic uncertainty are listed that are associated to the modelling of the $t\bar{t}H$ signal, the $t\bar{t}+\geq 1c$ and $t\bar{t}+$ light background processes and to experimental sources of uncertainties, for example the b-tagging efficiency as well as the jet energy scale and resolution.

However, their impact on $\Delta\mu$ is significantly lower than those related to the background modelling. The twenty nuisance parameters listed here contribute to 95% of the total uncertainty on the best-fit value of μ [4].
At this point, a trend appears indicating that the physics modelling is the critical limiting factor of this search. As more data is recorded and the experimental conditions improve over the next years, the systematic uncertainties originating from our imperfect understanding of these physics processes remains the biggest challenge. In order to improve the modelling of the $t\bar{t}H$ signal and the $t\bar{t} + b\bar{b}$ background processes, various studies have been performed up until now and they show promising results. Some of these studies are presented in Chapter 12.

Uncertainty source	$\Delta\mu$	
$t\bar{t}+ \geq 1b$ modelling	+0.46	-0.46
Statistical uncertainty of background model	+0.29	-0.31
b-tagging efficiency and mis-tag rates	+0.16	-0.16
Jet energy scale and resolution	+0.14	-0.14
$t\bar{t}H$ signal modelling	+0.22	-0.05
$t\bar{t}+ \geq 1c$ modelling	+0.09	-0.11
JVT and pile-up modelling	+0.03	-0.05
Other background modelling	+0.08	-0.08
$t\bar{t}+$ light modelling	+0.06	-0.03
Luminosity	+0.03	-0.02
Light lepton (e, μ) ID, isolation & trigger uncertainty	+0.03	-0.04
Total systematic uncertainty	+0.57	-0.54
$t\bar{t}+ \geq 1b$ normalisation	+0.09	-0.10
$t\bar{t}+ \geq 1c$ normalisation	+0.02	-0.03
Intrinsic statistical uncertainty	+0.21	-0.20
Total statistical uncertainty	+0.29	-0.29
Total uncertainty	+0.64	-0.61

Table 11.1: List of sources of systematic uncertainty in the $t\bar{t}H(H \to b\bar{b})$ search ranked in decreasing order by their impact on the signal strength uncertainty $\Delta\mu$ [4]. The 'statistical uncertainty of the background model' refers to the statistical uncertainties in the number of simulated events and in the estimation of the non-prompt and fake lepton contribution in the single lepton channel.

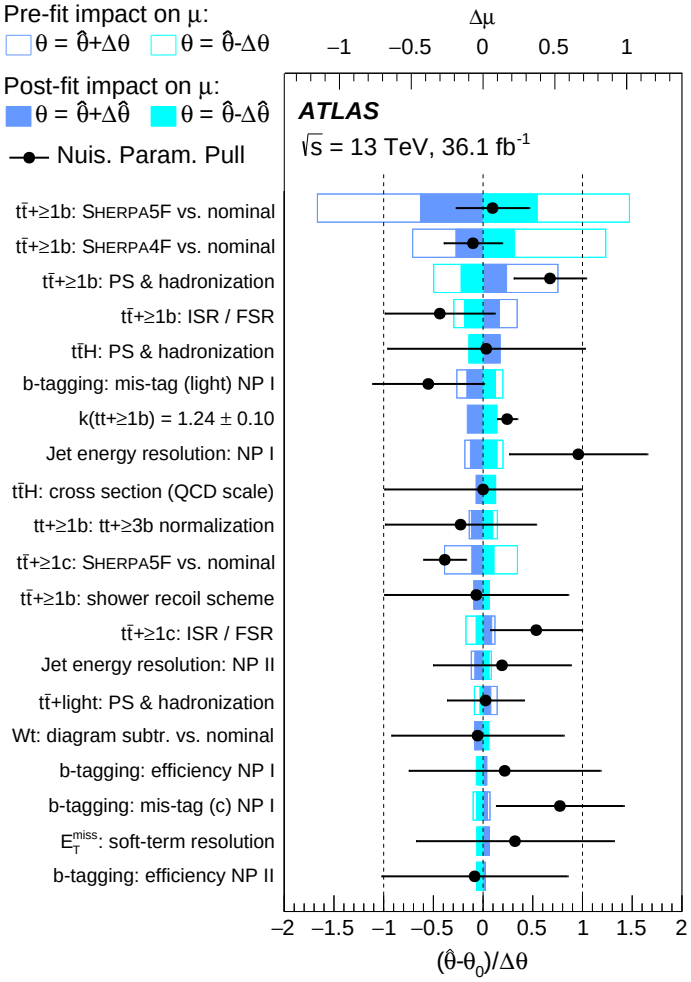

Figure 11.10: List of the most significant nuisance parameters that encode systematic uncertainties that impact the combined signal strength parameter μ during the combined fit [4]. The parameters are ranked in decreasing order by their contribution to the total uncertainty, but only the twenty parameters with the highest impact are presented. The empty (filled) blue rectangle represents the pre-fit (post-fit) contribution to $\Delta\mu$ that shift the nuisance parameter upwards, while the teal rectangles show the corresponding downward shift. The black points indicate the relative pulls of the nuisance parameters. In the bottom x-axis label, $\hat{\theta}$ (θ_0) refers to the best-fit (nominal) parameter value and $\Delta\theta$ is the pre-fit uncertainty of the parameter. $k\,(t\bar{t}+\geq 1b)$ is freely floating in the fit with a pre-fit value and uncertainty of 1.0 and thus has no pre-fit impact on μ. Some experimental uncertainties contain the label NP I and/or NP II, which refers to the first and/or second nuisance parameter encoding this uncertainty, respectively, ordered by their impact on μ.

The most precise result, quoted in Equation 11.1, shows an excess of events over the expected SM background with an observed (expected) significance of 1.4 (1.6) standard deviations (σ). The level of significance must reach at least 5.0σ to be able to claim an observation. Using the CL_S method [133–135], a signal strength larger than 2.0 is excluded at 95% confidence level as is depicted in Figure 11.11. The expected significance and exclusion limit are determined based on the post-fit background estimate [4].

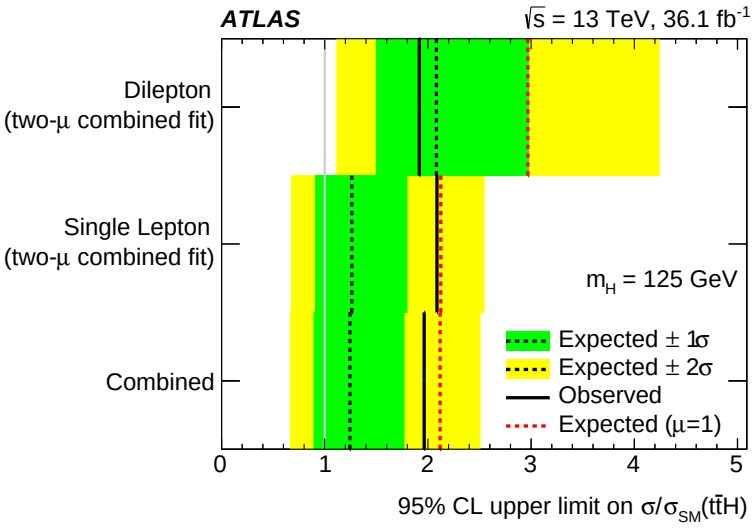

Figure 11.11: Shown are the upper limits on the signal strength at 95% CL for the individual channels and the combined result [4]. The individual limits are derived using the same approach as the quoted μ values in Equation 11.3. The black solid line represents the observed limits, while the black (red) dotted line shows the expected limits for the background-only (signal-plus-background) hypothesis. The green (yellow) area depicts the 1σ (2σ) uncertainty range on the expected limits for the background-only hypothesis.

In Figure 11.12, the event yields in data are compared to the combined post-fit prediction for all fifteen analysis regions. They are grouped together and ordered by the signal-to-background ratio in the respective bins of the final discriminants that enter the fit. The predictions are shown for the two fit hypotheses, namely the background-only scenario and the signal-plus-background scenario. The $t\bar{t}H$ signal contribution is scaled to either the observed signal strength or to the upper limit quoted above [4].

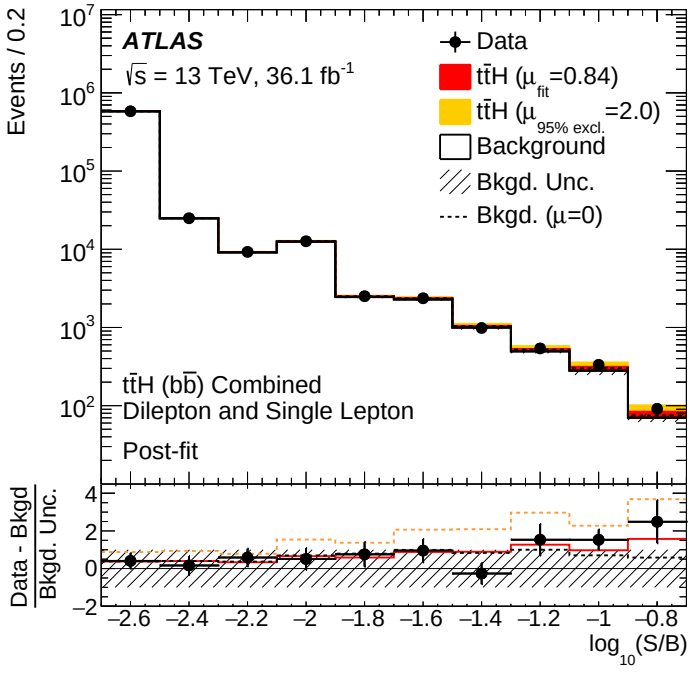

Figure 11.12: The combined post-fit event yields of both the dilepton and single lepton channels as a function of $\log_{10}(S/B)$ [4]. S (B) corresponds to the number of observed signal (background) events after the fit. In this histogram, the final-discriminant bins in all dilepton and single lepton analysis categories are combined into bins of $\log_{10}(S/B)$ in which the signal is normalised to the SM prediction used to determine $\log_{10}(S/B)$. The red (orange) area represents the number of signal events scaled to the best-fit value (95% CL exclusion limit) and is stacked on top of the expected background. In addition to this, the ratio of the difference between observed data events and B with respect to the uncertainty on B, in other words the pull on data relative to B, is compared to the pulls from the signal-plus-background hypothesis predictions. These are represented by the solid red line (dashed orange line) referring to the presence of a $t\bar{t}H$ signal with $\mu = 0.84$ ($\mu = 2.0$). The dashed black line shows the pull of B for the background-only hypothesis. The error bars on the data events are purely statistical.

CHAPTER 12

$t\bar{t} + b\bar{b}$ and $t\bar{t}H$ modelling studies

In the previous chapter, the results from the combined profile likelihood fit to data in the search for the $t\bar{t}H(H \to b\bar{b})$ process at $\sqrt{s} = 13$ TeV are presented. The leading five nuisance parameters in terms of their impact on the uncertainty on the $t\bar{t}H$ signal strength, as shown in Figure 11.10, are the following:

1. The comparison between the nominal $t\bar{t}$ +jets generator Powheg +Pythia 8 and an alternative five flavour generator Sherpa 5F. The difference between these two generators lies in the ME calculation and the PS and hadronisation model as well as the corresponding matching scheme. Along with the comparison to Powheg interfaced to Herwig 7, the uncertainty on the chosen model to calculate the ME can be assessed.

2. The comparison between Powheg +Pythia 8 and a four flavour scheme generator setup which is labelled Sherpa 4F and used to calculate the $t\bar{t}+b\bar{b}$ ME with NLO precision. This alternative prediction is used to model the $t\bar{t}+ \geq 1b$ background with the best available precision, namely by reweighting the sub-categories ($t\bar{t}+b$, $t\bar{t}+B$, $t\bar{t}+b\bar{b}$ and $t\bar{t}+ \geq 3b$) of the nominal prediction to this four flavour scheme setup. Further studies about this reweighting procedure and several tests are described in the next section.

3. The comparison between Powheg +Pythia 8 and an alternative prediction from Powheg +Herwig 7. Here, the ME is calculated with the same Powheg setup, but the PS and hadronisation models are different between the two setups, in order to evaluate the uncertainty originating from the choice of the PS and hadronisation models.

4. The comparison between Powheg +Pythia 8 and two alternative predictions from Powheg +Pythia 8 where ISR and FSR are either increased or decreased

in order to estimate the impact of the associated parameter choices.

5. The comparison between the nominal $t\bar{t}H$ generator MadGraph5_aMC@NLO interfaced to Pythia 8 and an alternative setup that is MadGraph5_aMC@NLO interfaced to Herwig++. Here, only the PS and hadronisation model is different between the two predictions which allows to assess the uncertainty related to the choice of the PS and hadronisation models.

The first four nuisance parameters affect the modelling of the $t\bar{t}+\geq 1b$ background, while the last impacts the $t\bar{t}H$ signal modelling. Since the total uncertainty on the final result is not dominated by the statistical, but rather the systematic component, it is clear that the physics modelling is the limiting factor of the $t\bar{t}H(H\rightarrow b\bar{b})$ analysis. The total uncertainty can be reduced from collecting more data as well as upgrading the Atlas detector and improving our understanding of it in the near future. However, this will not have a significant impact on the expected significance of the $t\bar{t}H(H\rightarrow b\bar{b})$ signal under the signal-plus-background hypothesis. Instead, it is indispensable to improve our understanding of the $t\bar{t}H$ signal and especially the $t\bar{t}+\geq 1b$ background processes in order to create more precise predictions not only in terms of their cross-sections, but, more importantly, event properties and kinematics.

In the next section, the reweighting procedure of the nominal $t\bar{t}$ +jets simulation to the Sherpa 4F prediction is described. In the sections thereafter, efforts within the Atlas collaboration to improve the modelling of the $t\bar{t}H$ signal and $t\bar{t}+b\bar{b}$ background processes are presented.

12.1 Reweighting of kinematic distributions

As described in Section 7.3.1, a reweighting strategy is employed in the $t\bar{t}H(H\rightarrow b\bar{b})$ analysis to improve the modelling of the $t\bar{t}+\geq 1b$ background. This strategy prescribes to take the nominal $t\bar{t}$ +jets prediction generated with the five flavour setup, Powheg interfaced to Pythia 8, and scale the event fractions of the $t\bar{t}+\geq 1b$ sub-categories, namely $t\bar{t}+b$, $t\bar{t}+B$, $t\bar{t}+b\bar{b}$ and $t\bar{t}+\geq 3b$, to the predicted fractions of the four flavour setup Sherpa 4F. The $t\bar{t}+b$(MPI/FSR) component is not reweighted in this approach, because it is not present in the Sherpa 4F setup at NLO precision. The difference in these event fractions between the two setups is depicted in Figure 7.3. In this procedure, the weights of the events generated by Powheg +Pythia 8 are scaled by the ratio in the lower panel of Figure 7.3 in the corresponding sub-categories. This reweighting improves the background model significantly as it increases the agreement between MC and data. On the other hand, the nuisance parameter associated to the uncertainty on the reweighting ranks second in the list of the variables with the highest impact on the fit, which is shown in Figure 11.10. Thus, having a model that does not require such a reweighting would reduce the modelling uncertainty.

One additional concept has been investigated based on the assumption that the SHERPA 4F not only predicts more precise event fractions of the $t\bar{t}+ \geq 1b$ sub-categories than the nominal setup, but also provides a better prediction of the kinematic observables of the final state particles, in particular the additional $b\bar{b}$ pair not coming from the $t\bar{t}$ decay. The idea is thus to additionally apply the reweighting procedure to crucial observables in events falling into the respective $t\bar{t}+ \geq 1b$ sub-categories. These variables are listed in the following, ordered in the sequence they are reweighted in:

1. The p_{T} of the $t\bar{t}$ system;
2. The p_{T} of the individual top and antitop quarks;
3. The p_{T} of the leading additional b-jet (for the $t\bar{t}+b$ and $t\bar{t}+B$ categories) or the two leading additional b-jets (for the $t\bar{t}+b\bar{b}$ and $t\bar{t}+ \geq 3b$ categories); and
4. ΔR between the two leading additional b-jets for the $t\bar{t}+b\bar{b}$ and $t\bar{t}+ \geq 3b$ categories.

These kinematic variables are expected to be modelled more precisely by the SHERPA 4F setup within the $t\bar{t}+ \geq 1b$ phase space and also play an important role in describing the event kinematics, thus affecting the event selection in the $t\bar{t}+ \geq 1b$ analysis regions. Figures 12.1–12.5 present the event fractions and observables considered in this shape reweighting procedure as they are predicted by the various $t\bar{t}$ generator setups employed in this analysis. All the distributions are shown before the shape reweighting and in the corresponding $t\bar{t}+ \geq 1b$ sub-categories, except for the primary heavy flavour classification (HFC) event fractions. Figures 12.6–12.10 depict the corresponding distributions after the shape reweighting, excluding the event fractions which are unaffected by the shape reweighting.

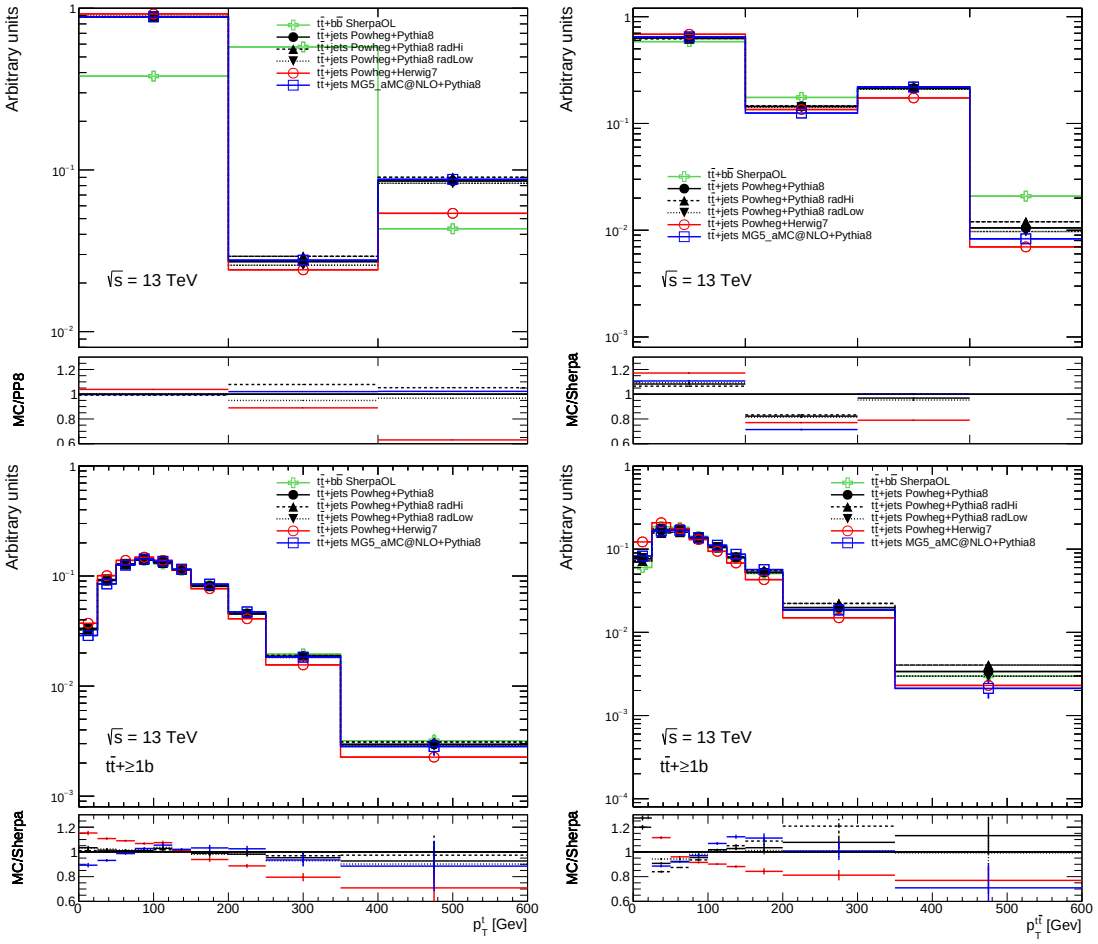

Figure 12.1: Predictions by the various $t\bar{t}$ setups considered in the $t\bar{t}H(H \rightarrow b\bar{b})$ search, namely the Sherpa 4F setup (green), nominal Powheg +Pythia 8 (solid black), Powheg +Pythia 8 with increased and decreased radiation (dashed black and dotted black), Powheg +Herwig 7 (red) and MadGraph5_aMC@NLO +Pythia 8 (blue). Only the Sherpa 5F setup is excluded. The predicted distributions are all normalised to unity and the ratio in the lower panel is with respect to the $t\bar{t}+b\bar{b}$ prediction from the Sherpa 4F generator, such that the relative event fractions i.e. the shape can be compared between the predictions. Shown are the general HFC categories (top left), the $t\bar{t}+\geq 1b$ sub-categories (top right), the p_{T} of the individual top quarks (bottom left) and the p_{T} of the $t\bar{t}$ system (bottom right), the latter three in the $t\bar{t}+\geq 1b$ category. All distributions are shown before the shape reweighting. The error bars reflect only the statistical uncertainty from simulation.

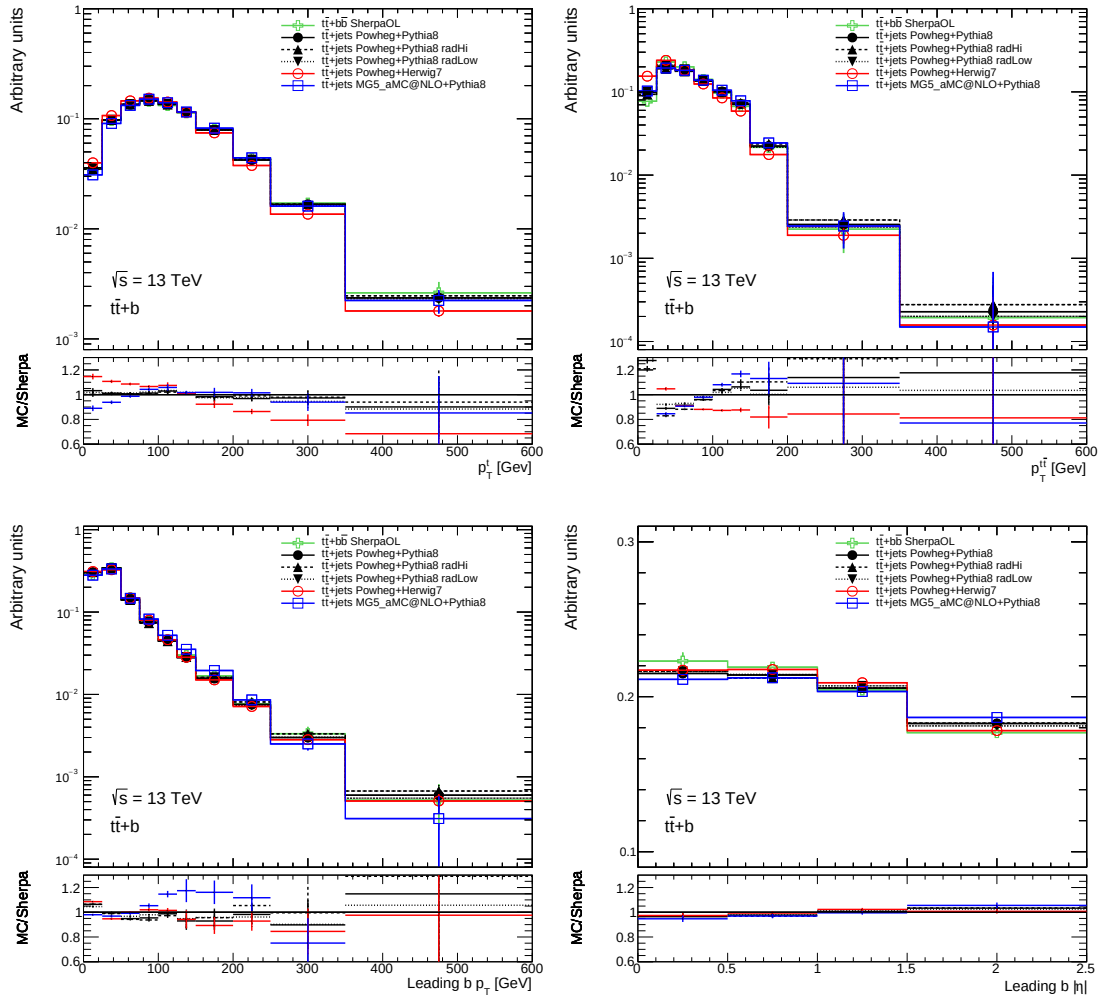

Figure 12.2: Predictions by the various $t\bar{t}$ setups considered in the $t\bar{t}H(H \rightarrow b\bar{b})$ search, namely the SHERPA 4F setup (green), nominal POWHEG +PYTHIA 8 (solid black), POWHEG +PYTHIA 8 with increased and decreased radiation (dashed black and dotted black), POWHEG +HERWIG 7 (red) and MADGRAPH5_aMC@NLO +PYTHIA 8 (blue). Only the SHERPA 5F setup is excluded. The predicted distributions are all normalised to unity and the ratio in the lower panel is with respect to the $t\bar{t} + b\bar{b}$ prediction from the SHERPA 4F generator, such that the relative event fractions i.e. the shape can be compared between the predictions. Shown are the p_{T} of the individual top quarks (top left), the p_{T} of the $t\bar{t}$ system (top right), the p_{T} of the leading additional b-jet (bottom left) and $|\eta|$ of the leading additional b-jet (bottom right), all in the $t\bar{t}+b$ category. All distributions are shown before the shape reweighting. The error bars reflect only the statistical uncertainty from simulation.

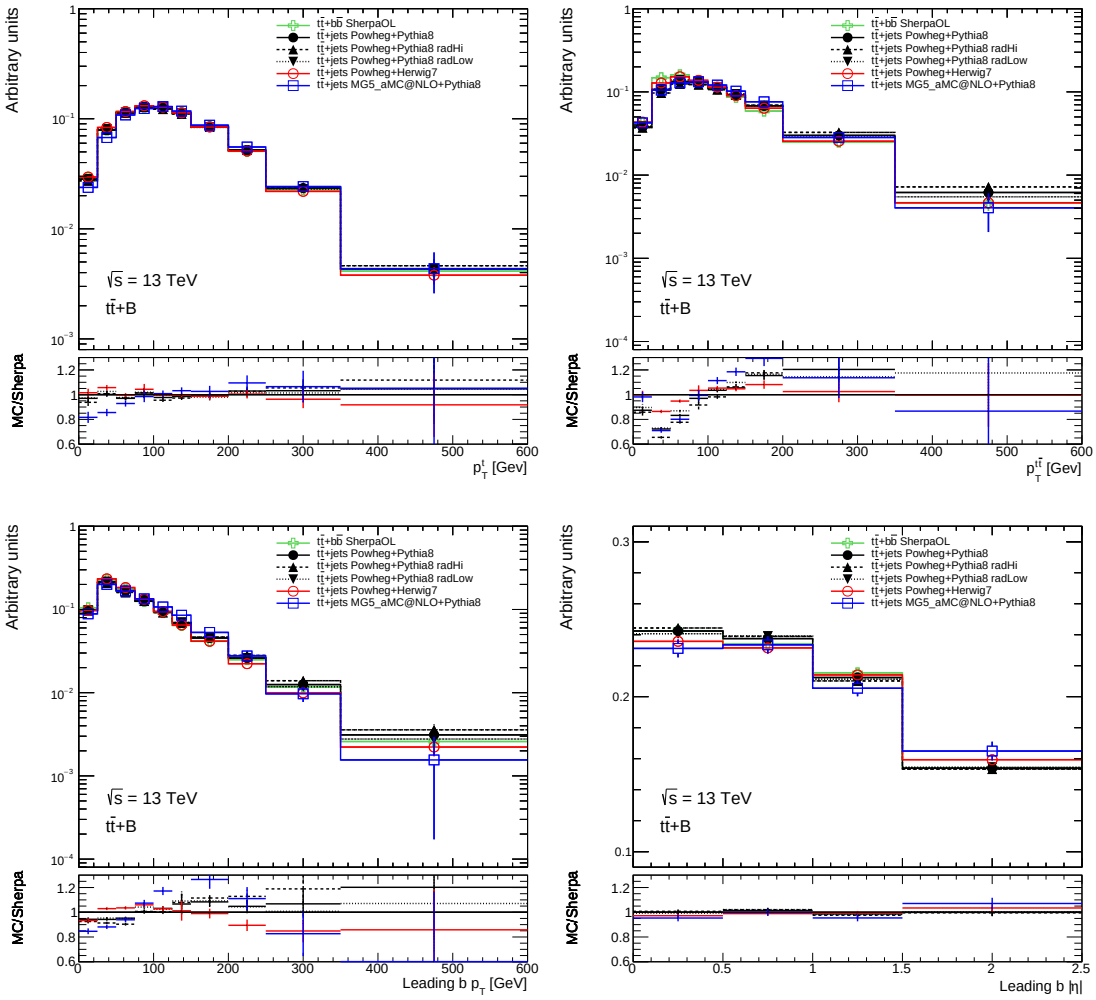

Figure 12.3: Predictions by the various $t\bar{t}$ setups considered in the $t\bar{t}H(H \rightarrow b\bar{b})$ search, namely the Sherpa 4F setup (green), nominal Powheg +Pythia 8 (solid black), Powheg +Pythia 8 with increased and decreased radiation (dashed black and dotted black), Powheg +Herwig 7 (red) and MadGraph5_aMC@NLO +Pythia 8 (blue). Only the Sherpa 5F setup is excluded. The predicted distributions are all normalised to unity and the ratio in the lower panel is with respect to the $t\bar{t}+b\bar{b}$ prediction from the Sherpa 4F generator, such that the relative event fractions i.e. the shape can be compared between the predictions. Shown are the p_{T} of the individual top quarks (top left), the p_{T} of the $t\bar{t}$ system (top right), the p_{T} of the leading additional b-jet (bottom left) and $|\eta|$ of the leading additional b-jet (bottom right), all in the $t\bar{t}+B$ category. All distributions are shown before the shape reweighting. The error bars reflect only the statistical uncertainty from simulation.

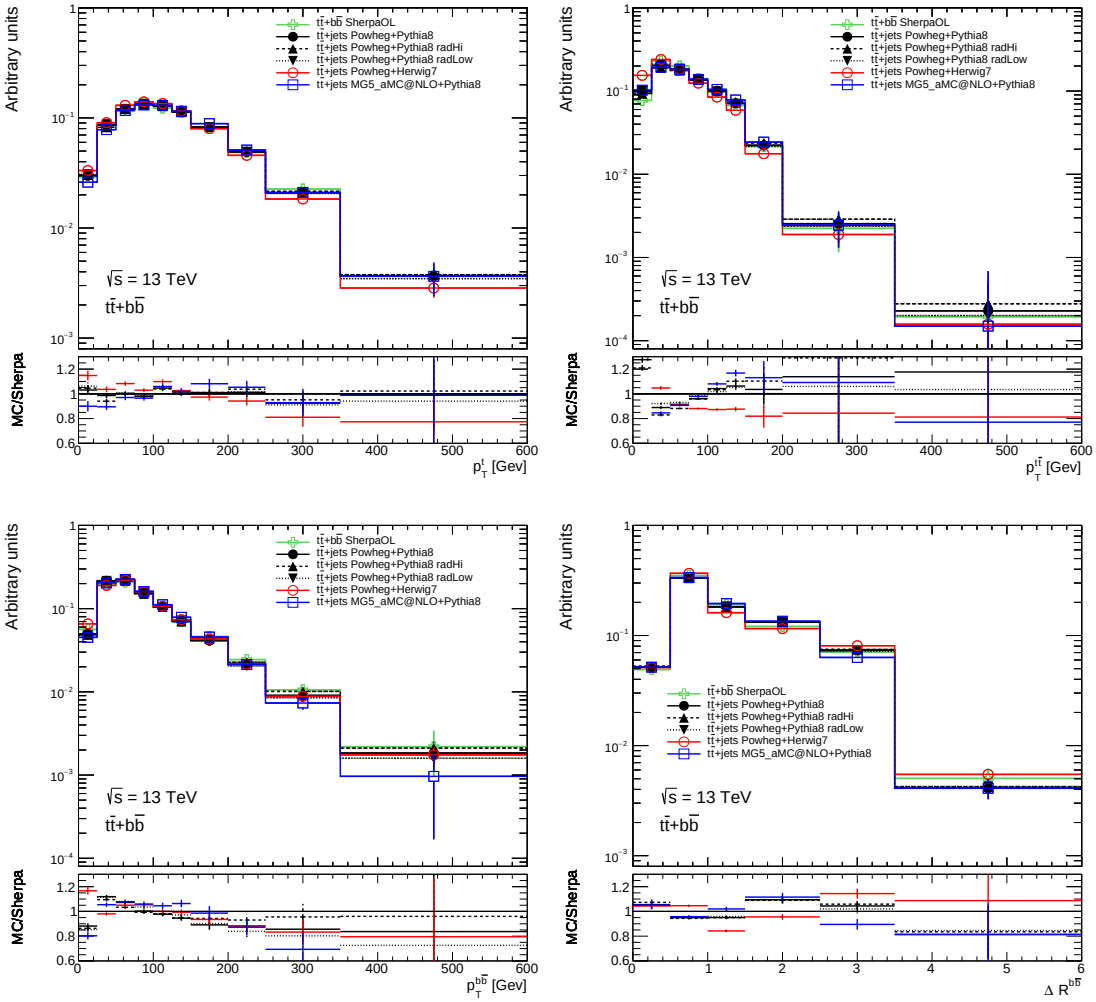

Figure 12.4: Predictions by the various $t\bar{t}$ setups considered in the $t\bar{t}H(H \rightarrow b\bar{b})$ search, namely the SHERPA 4F setup (green), nominal POWHEG +PYTHIA 8 (solid black), POWHEG +PYTHIA 8 with increased and decreased radiation (dashed black and dotted black), POWHEG +HERWIG 7 (red) and MADGRAPH5_aMC@NLO +PYTHIA 8 (blue). Only the SHERPA 5F setup is excluded. The predicted distributions are all normalised to unity and the ratio in the lower panel is with respect to the $t\bar{t} + b\bar{b}$ prediction from the SHERPA 4F generator, such that the relative event fractions i.e. the shape can be compared between the predictions. Shown are the p_{T} of the individual top quarks (top left), the p_{T} of the $t\bar{t}$ system (top right), the p_{T} of the leading additional b-jet pair (bottom left) and ΔR between the leading additional b-jet pair (bottom right), all in the $t\bar{t} + b\bar{b}$ category. All distributions are shown before the shape reweighting. The error bars reflect only the statistical uncertainty from simulation.

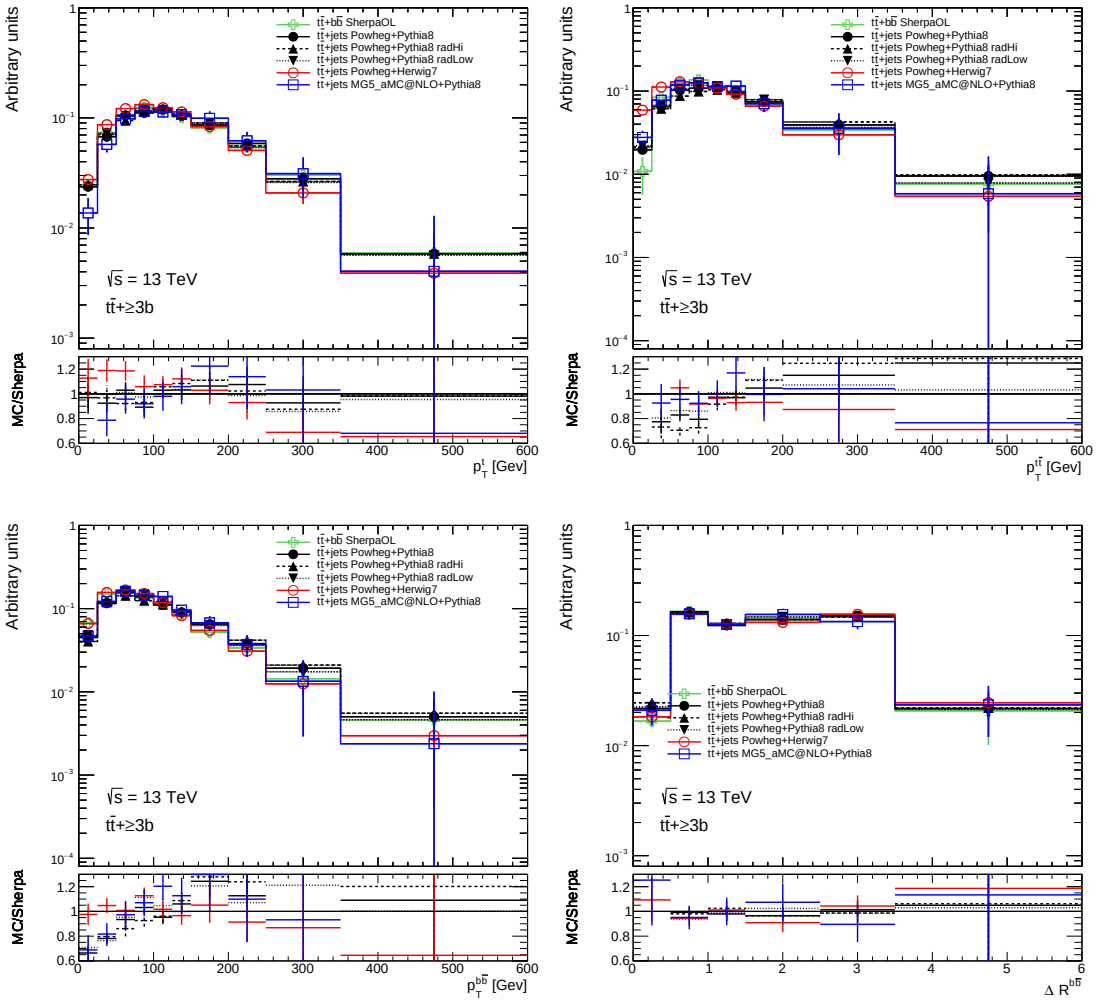

Figure 12.5: Predictions by the various $t\bar{t}$ setups considered in the $t\bar{t}H(H \rightarrow b\bar{b})$ search, namely the Sherpa 4F setup (green), nominal Powheg +Pythia 8 (solid black), Powheg +Pythia 8 with increased and decreased radiation (dashed black and dotted black), Powheg +Herwig 7 (red) and MadGraph5_aMC@NLO +Pythia 8 (blue). Only the Sherpa 5F setup is excluded. The predicted distributions are all normalised to unity and the ratio in the lower panel is with respect to the $t\bar{t} + b\bar{b}$ prediction from the Sherpa 4F generator, such that the relative event fractions i.e. the shape can be compared between the predictions. Shown are the p_{T} of the individual top quarks (top left), the p_{T} of the $t\bar{t}$ system (top right), the p_{T} of the leading additional b-jet pair (bottom left) and ΔR between the leading additional b-jet pair (bottom right), all in the $t\bar{t}+ \geq 3b$ category. All distributions are shown before the shape reweighting. The error bars reflect only the statistical uncertainty from simulation.

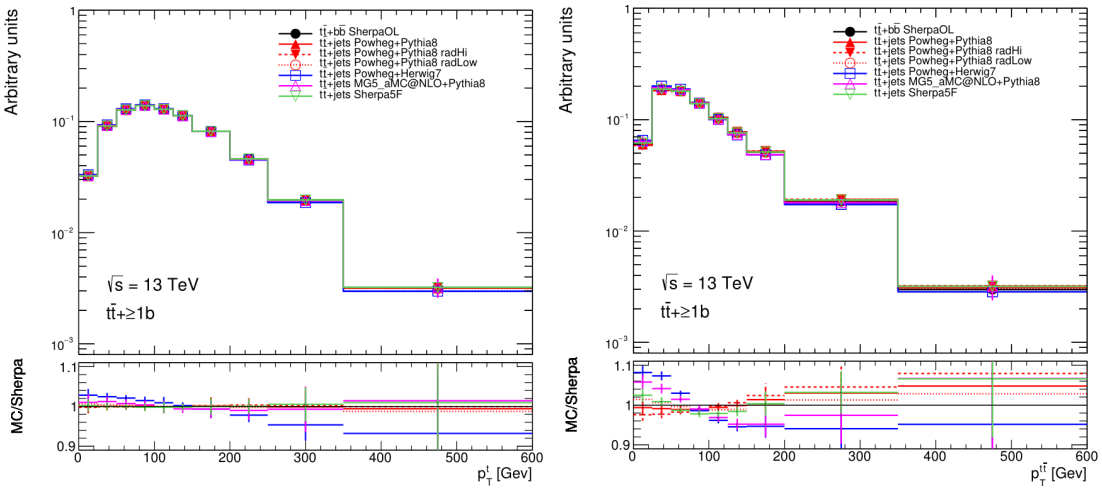

Figure 12.6: Predictions by the various $t\bar{t}$ setups considered in the $t\bar{t}H(H \rightarrow b\bar{b})$ search, namely the SHERPA 4F setup (black), nominal POWHEG +PYTHIA 8 (solid red), POWHEG +PYTHIA 8 with increased and decreased radiation (dashed red and dotted red), POWHEG +HERWIG 7 (blue), MADGRAPH5_aMC@NLO +PYTHIA 8 (pink), and SHERPA 5F (green). The predicted distributions are all normalised to unity and the ratio in the lower panel is with respect to the $t\bar{t} + b\bar{b}$ prediction from the SHERPA 4F generator, such that the relative event fractions i.e. the shape can be compared between the predictions. Shown are the p_{T} of the individual top quarks (left) and the p_{T} of the $t\bar{t}$ system (right), both in the $t\bar{t}+ \geq 1b$ category. All distributions are shown after reweighting the $t\bar{t}$ p_{T} followed by the top quark p_{T} to the SHERPA 4F prediction. The error bars reflect only the statistical uncertainty from simulation.

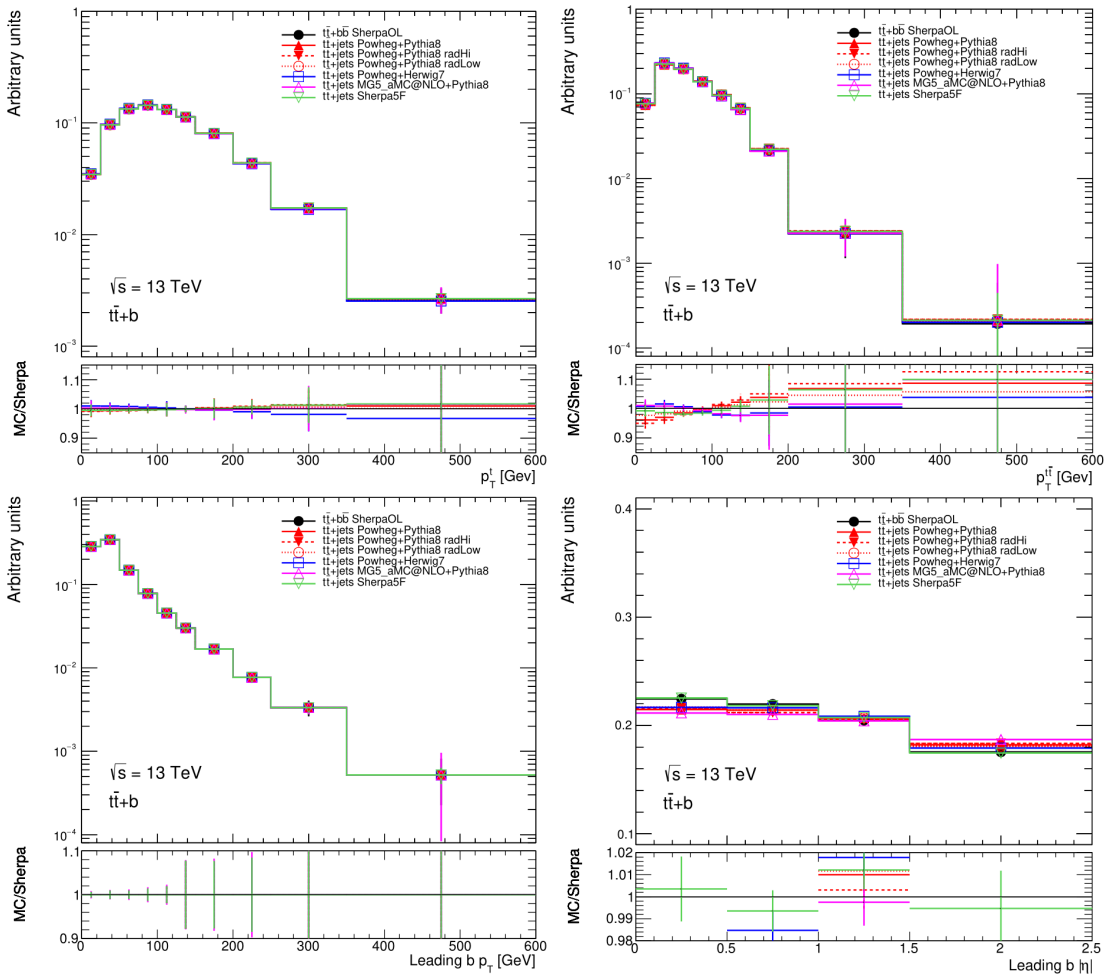

Figure 12.7: Predictions by the various $t\bar{t}$ setups considered in the $t\bar{t}H(H \rightarrow b\bar{b})$ search, namely the Sherpa 4F setup (black), nominal Powheg +Pythia 8 (solid red), Powheg +Pythia 8 with increased and decreased radiation (dashed red and dotted red), Powheg +Herwig 7 (blue), MadGraph5_aMC@NLO +Pythia 8 (pink), and Sherpa 5F (green). The predicted distributions are all normalised to unity and the ratio in the lower panel is with respect to the $t\bar{t}+b\bar{b}$ prediction from the Sherpa 4F generator, such that the relative event fractions i.e. the shape can be compared between the predictions. Shown are the p_{T} of the individual top quarks (top left), the p_{T} of the $t\bar{t}$ system (top right), the p_{T} of the leading additional b-jet (bottom left) and $|\eta|$ of the leading additional b-jet (bottom right), all in the $t\bar{t}+b$ category. All distributions are shown after the full shape reweighting to the Sherpa 4F prediction. The reweighted variables in order are: $t\bar{t}$ p_{T}; top quark p_{T}; leading additional b-jet p_{T}. The error bars reflect only the statistical uncertainty from simulation.

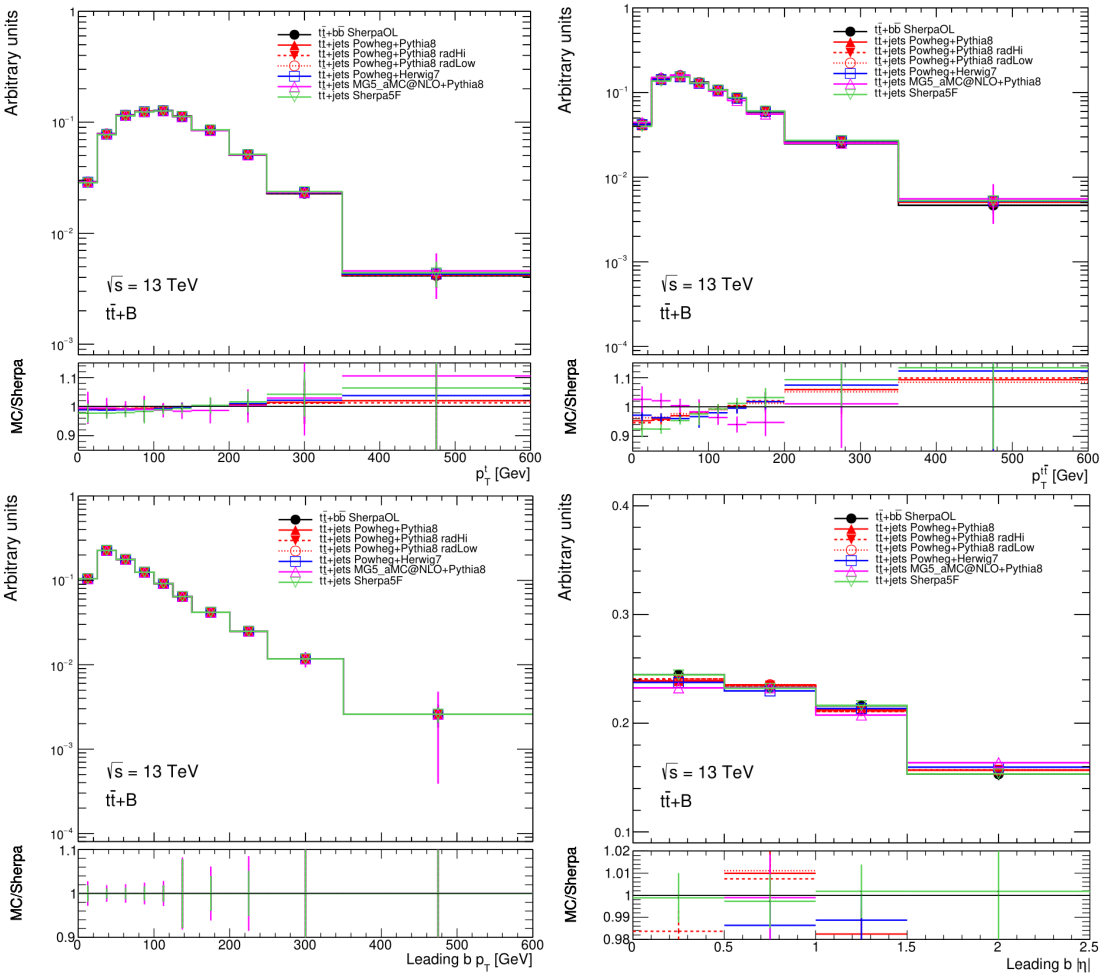

Figure 12.8: Predictions by the various $t\bar{t}$ setups considered in the $t\bar{t}H(H \to b\bar{b})$ search, namely the SHERPA 4F setup (black), nominal POWHEG +PYTHIA 8 (solid red), POWHEG +PYTHIA 8 with increased and decreased radiation (dashed red and dotted red), POWHEG +HERWIG 7 (blue), MADGRAPH5_aMC@NLO +PYTHIA 8 (pink), and SHERPA 5F (green). The predicted distributions are all normalised to unity and the ratio in the lower panel is with respect to the $t\bar{t} + b\bar{b}$ prediction from the SHERPA 4F generator, such that the relative event fractions i.e. the shape can be compared between the predictions. Shown are the p_{T} of the individual top quarks (top left), the p_{T} of the $t\bar{t}$ system (top right), the p_{T} of the leading additional b-jet (bottom left) and $|\eta|$ of the leading additional b-jet (bottom right), all in the $t\bar{t}+B$ category. All distributions are shown after the shape reweighting to the SHERPA 4F prediction. The reweighted variables in order are: $t\bar{t}$ p_{T}; top quark p_{T}; leading additional b-jet p_{T}. The error bars reflect only the statistical uncertainty from simulation.

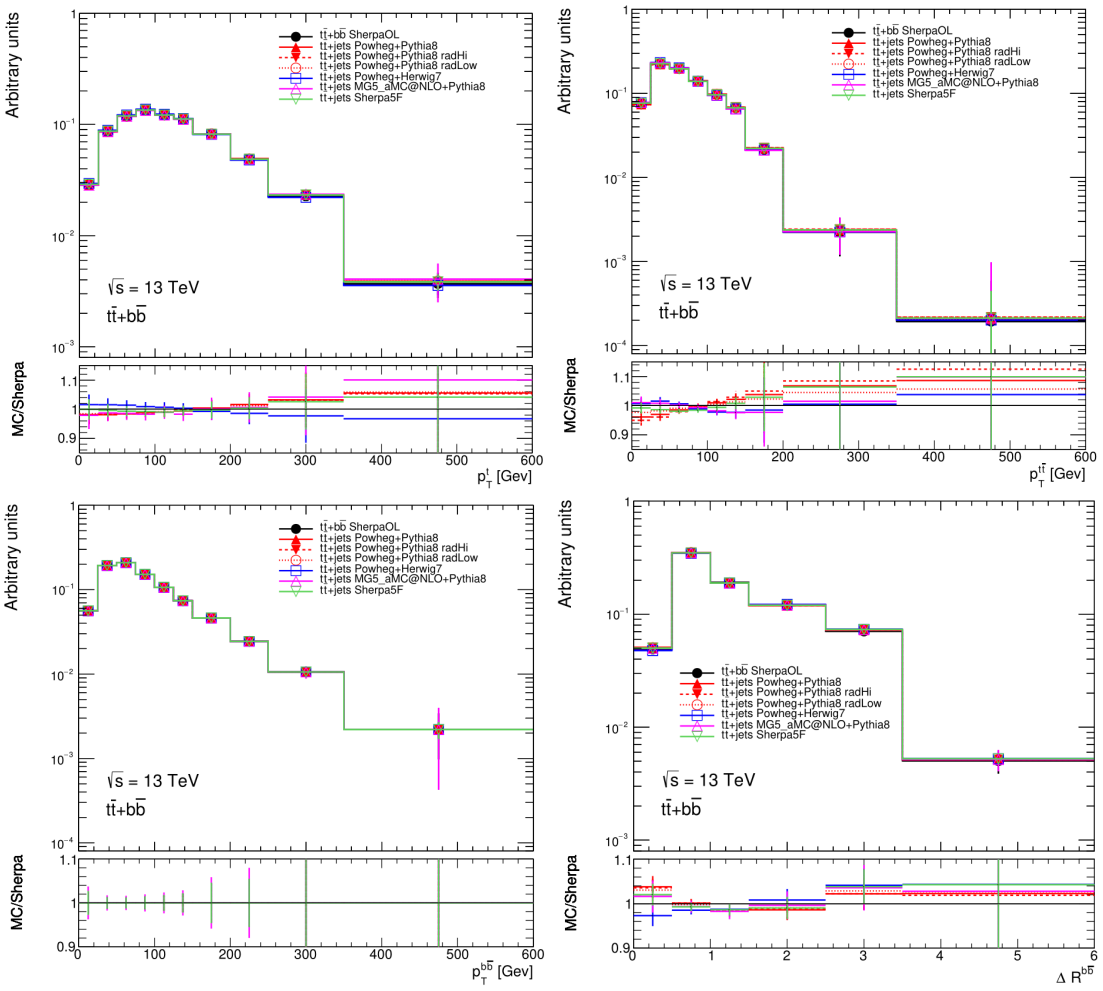

Figure 12.9: Predictions by the various $t\bar{t}$ setups considered in the $t\bar{t}H(H \rightarrow b\bar{b})$ search, namely the Sherpa 4F setup (black), nominal Powheg +Pythia 8 (solid red), Powheg +Pythia 8 with increased and decreased radiation (dashed red and dotted red), Powheg +Herwig 7 (blue), MadGraph5_aMC@NLO +Pythia 8 (pink), and Sherpa 5F (green). The predicted distributions are all normalised to unity and the ratio in the lower panel is with respect to the $t\bar{t}+b\bar{b}$ prediction from the Sherpa 4F generator, such that the relative event fractions i.e. the shape can be compared between the predictions. Shown are the p_{T} of the individual top quarks (top left), the p_{T} of the $t\bar{t}$ system (top right), the p_{T} of the leading additional b-jet pair (bottom left) and ΔR between the leading additional b-jet pair (bottom right), all in the $t\bar{t}+b\bar{b}$ category. All distributions are shown after the shape reweighting to the Sherpa 4F prediction. The reweighted variables in order are: $t\bar{t}$ p_{T}; top quark p_{T}; leading additional b-jet pair p_{T}; leading additional b-jet pair ΔR. The error bars reflect only the statistical uncertainty from simulation.

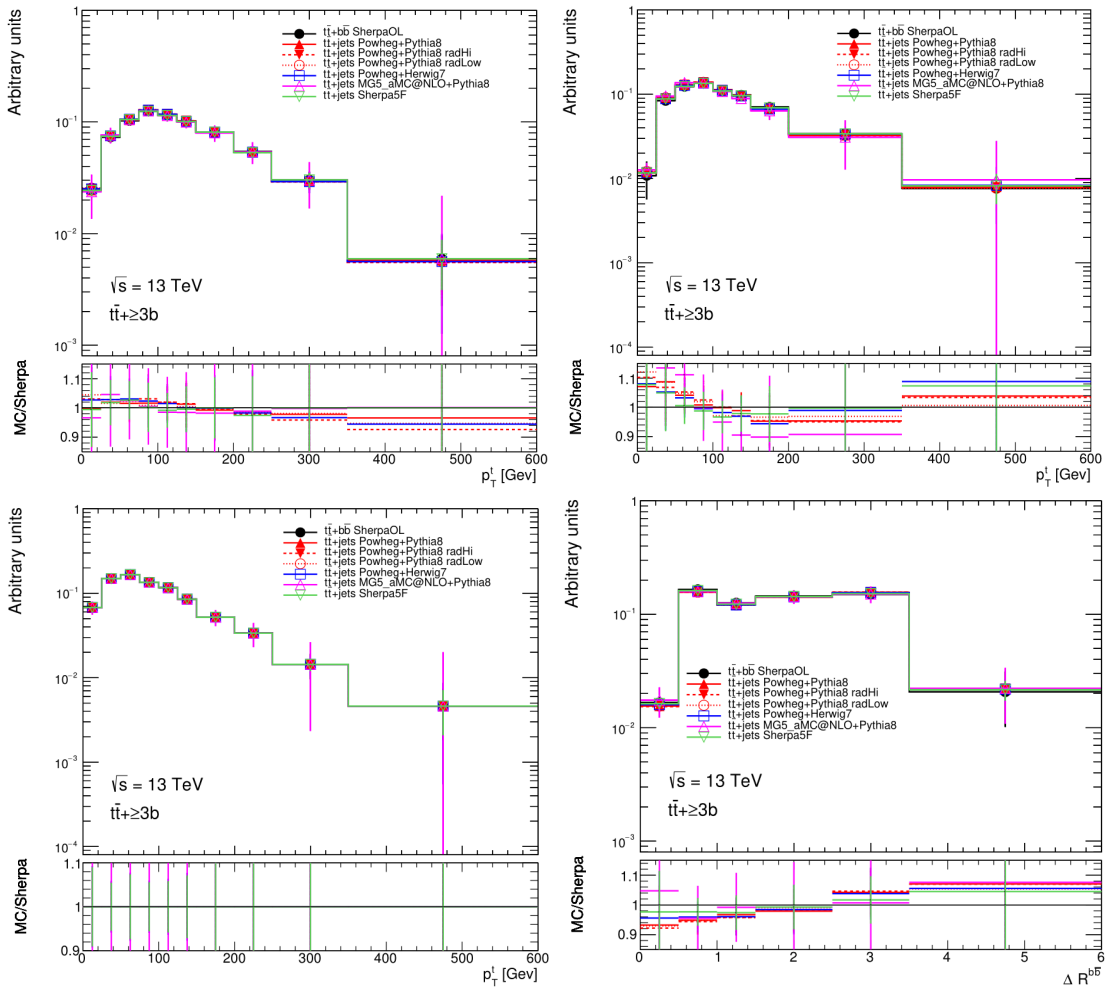

Figure 12.10: Predictions by the various $t\bar{t}$ setups considered in the $t\bar{t}H(H \rightarrow b\bar{b})$ search, namely the SHERPA 4F setup (black), nominal POWHEG +PYTHIA 8 (solid red), POWHEG +PYTHIA 8 with increased and decreased radiation (dashed red and dotted red), POWHEG +HERWIG 7 (blue), MADGRAPH5_aMC@NLO +PYTHIA 8 (pink), and SHERPA 5F (green). The predicted distributions are all normalised to unity and the ratio in the lower panel is with respect to the $t\bar{t} + b\bar{b}$ prediction from the SHERPA 4F generator, such that the relative event fractions i.e. the shape can be compared between the predictions. Shown are the p_{T} of the individual top quarks (top left), the p_{T} of the $t\bar{t}$ system (top right), the p_{T} of the leading additional b-jet pair (bottom left) and ΔR between the leading additional b-jet pair (bottom right), all in the $t\bar{t}+ \geq 3b$ category. All distributions are shown after the shape reweighting to the SHERPA 4F prediction. The reweighted variables in order are: $t\bar{t}$ p_{T}; top quark p_{T}; leading additional b-jet pair p_{T}; leading additional b-jet pair ΔR. The error bars reflect only the statistical uncertainty from simulation.

The figures after the shape reweighting show an improved agreement between the nominal $t\bar{t}$ +jets and the $t\bar{t}+b\bar{b}$ SHERPA 4F predictions, which would motivate implementing this procedure into the main analysis. Especially the observables which are reweighted last are expected to match perfectly. However, all shown distributions are obtained using truth information exclusively. No significant difference can be observed between distributions before and after the shape reweighting for events at the reconstruction level. The main impact of the reweighting procedure originates from the normalisation i.e. from reweighting the event fractions. This is illustrated in Figures 12.11 and 12.12 which compare the impact of the full shape reweighting procedure on the classification BDT and $H_{\mathrm{T}}^{\mathrm{had}}$ distributions in various single lepton SRs and CRs. Aside from this, the variables that are reweighted first often do not keep their good agreement after the other variables are reweighted. In addition to that, the uncertainties of the shape reweighting are often negligible. For all of these reasons, the shape reweighting is not employed in the $t\bar{t}H(H \to b\bar{b})$ search.

Multiple additional tests have been performed regarding the $t\bar{t}+b\bar{b}$ reweighting. These involve primarily the choice of variables to reweight as well as their order. Additional variables considered were the invariant mass of the additional b-jet pair, the absolute pseudorapidity of the leading additional b-jet and the scalar sum of the transverse momenta of all final state particles, H_{T}. These variables proved to be either insignificant or to cause significant discrepancies between the predictions of the other reweighted kinematic distributions. Furthermore, the shape reweighting was tested using a different order of the considered variables, which all led to results worse than the order presented above. Another test involved performing a two-dimensional reweighting, first of the top quark and $t\bar{t}$ system p_{T} and afterwards the observables related to the additional b-jet or $b\bar{b}$ pair. While the agreement after the reweighting improved, this effect is found to be insignificant with respect to the one-dimensional reweighting sequence and is at the cost of increased statistical uncertainties due to the difficulty of populating two-dimensional distributions in the $t\bar{t}+\geq 1b$ sub-categories.
Finally, another test has been done in which the HFC scheme was changed by modifying the p_{T} of the jet or the corresponding hadron it is matched to or adding a requirement on the ratio of both transverse momenta. The event fractions of the $t\bar{t}$ +jets categories as well as the $t\bar{t}+\geq 1b$ sub-categories behave as expected under these changes while not providing any significant advantages. The differences between the generators are stable within the considered systematic uncertainties.

12.2 Studies on the modelling of the $t\bar{t}+b\bar{b}$ process

One argument why the $t\bar{t}+b\bar{b}$ process is modelled with suboptimal precision in the $t\bar{t}H(H \to b\bar{b})$ analysis and why the reweighting strategy discussed in the previous section is employed is that the 4F $t\bar{t}+b\bar{b}$ calculation would have to be merged with the inclusive $t\bar{t}$ +jets sample while removing overlapping events. This procedure is not implemented in current state-of-the-art MC generators used in the ATLAS experiment.

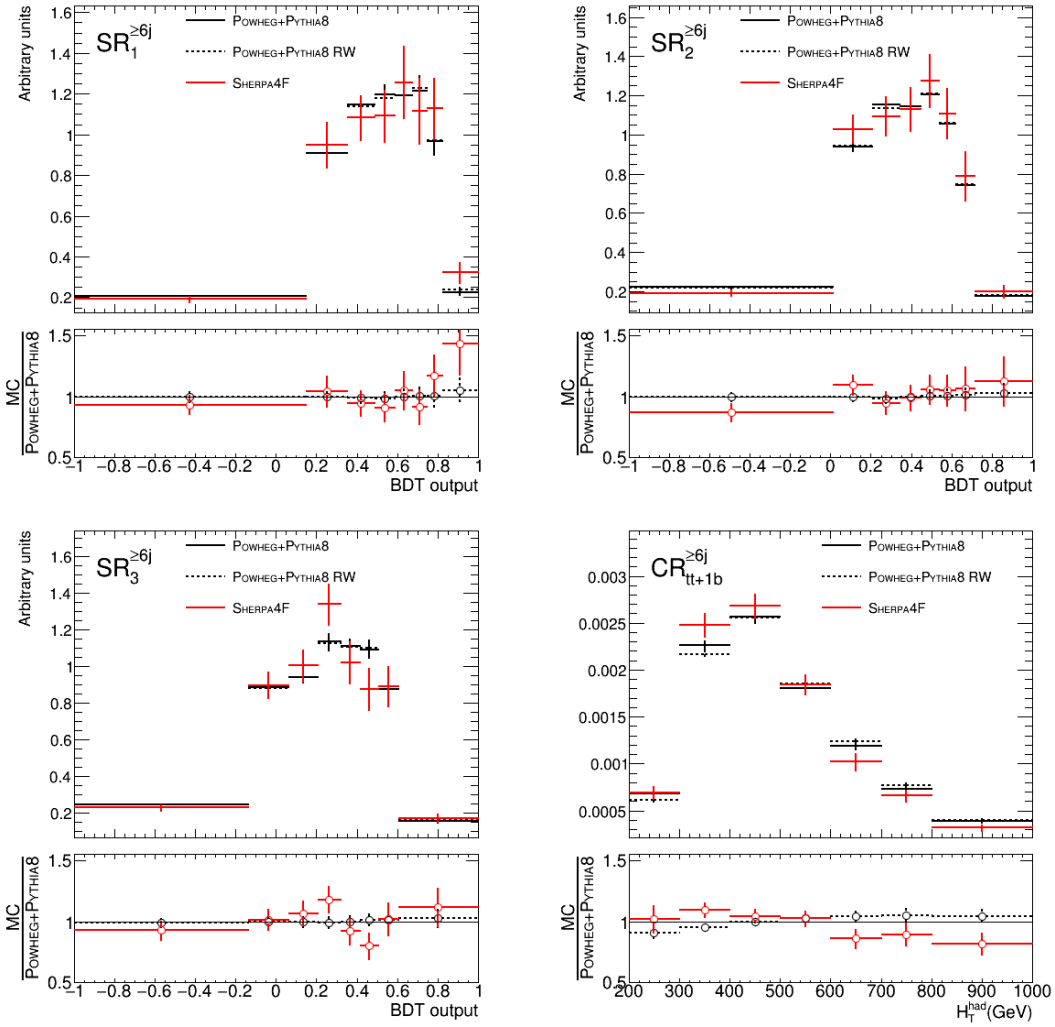

Figure 12.11: Comparison of the nominal POWHEG +PYTHIA 8 prediction (solid black) to the fully shape reweighted POWHEG +PYTHIA 8 setup (dashed black) and the SHERPA 4F prediction (solid red) for the single lepton channel with at least six jets [142]. Shown are the distributions of the classification BDT output in the SR_1^{6j} (top left), SR_2^{6j} (top right) and SR_3^{6j} (bottom left) as well as $H_\mathrm{T}^\mathrm{had}$ in $\mathrm{CR}_{t\bar{t}+b}^{6j}$ (bottom right). The ratio in the lower panel shows the predictions with respect to the nominal POWHEG +PYTHIA 8 setup. The error bars reflect only the statistical uncertainty from simulation.

Aside from the SHERPA 4F setup, another generator is able to simulate the $t\bar{t} + b\bar{b}$ process in the ME at NLO precision, namely MADGRAPH5_aMC@NLO which is also used to generate a dedicated $t\bar{t} + c\bar{c}$ sample for this analysis. However, this generator has one critical disadvantage which lies in its matching of the ME to the PS. The merged

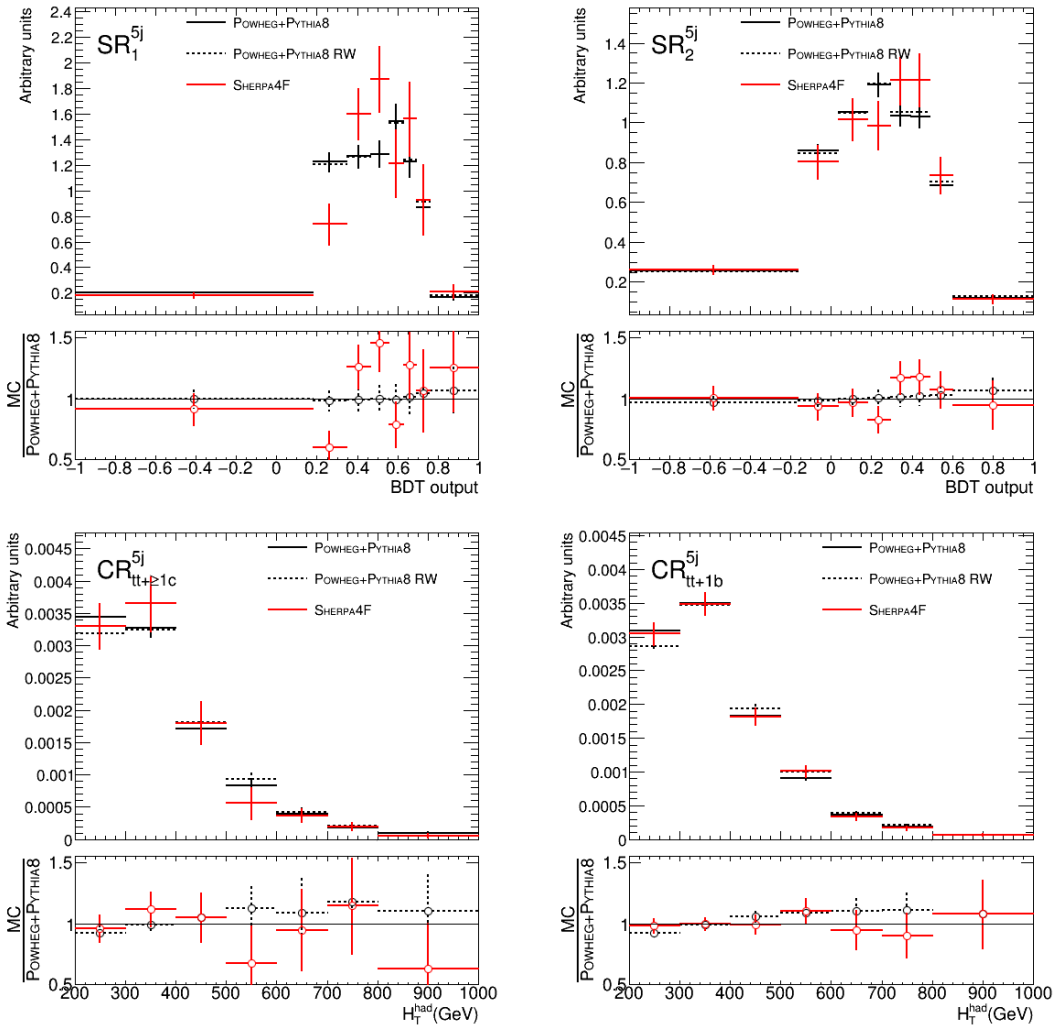

Figure 12.12: Comparison of the nominal POWHEG +PYTHIA 8 prediction (solid black) to the fully shape reweighted POWHEG +PYTHIA 8 setup (dashed black) and the SHERPA 4F prediction (solid red) for the single lepton channel with exactly five jets [142]. Shown are the distributions of the classification BDT output in the SR_1^{5j} (top left) and SR_2^{5j} (top right) as well as $H_\mathrm{T}^\mathrm{had}$ in $\mathrm{CR}^{5j}_{t\bar{t}+\geq 1c}$ (bottom left) and $\mathrm{CR}^{5j}_{t\bar{t}+b}$ (bottom right). The ratio in the lower panel shows the predictions with respect to the nominal POWHEG +PYTHIA 8 setup. The error bars reflect only the statistical uncertainty from simulation.

MADGRAPH5_aMC@NLO +PS setups in ATLAS, where the PS is either PYTHIA or HERWIG, generate a significant amount of events with negative weight factors to compensate certain other events with positive weight factors [88]. In tests, this fraction of

negative weights has been found to be roughly 20-30%. The consequence is that the effective number of events generated, given by the sum of all event weights, is significantly smaller than for a sample with the same number of events produced with POWHEG interfaced to PYTHIA 8, which practically only generates positive event weights. Thus, in order to have the same effective number of events, one would have to generate many more events which requires a large amount of CPU time. Otherwise, analyses relying on this prediction are limited by MC related statistical uncertainties. The SHERPA generator also generates events with negative weights, but significantly fewer than the MADGRAPH5_aMC@NLO +PS setups, namely around 5%.

At the time of writing this, new methods and generator setups have been developed to simulate the $t\bar{t} + b\bar{b}$ process with improved precision. The most important one for the ATLAS collaboration and its future $t\bar{t}H(H \to b\bar{b})$ analyses is a new setup in which the POWHEG generator is interfaced to OPENLOOPS to generate the $t\bar{t} + b\bar{b}$ ME at NLO precision using the four flavour scheme [143]. This setup is then interfaced to PYTHIA 8 for the PS and hadronisation and will thus be referred to as POWHEG +PYTHIA 8 in the remainder of this section. Since the nominal MC setup to generate inclusive $t\bar{t}$ events is still POWHEG +PYTHIA 8, both setups can be compared in a more controlled and comprehensible way. By using the POWHEG +PYTHIA 8 setup to generate both the inclusive $t\bar{t}$ +jets and the dedicated $t\bar{t} + b\bar{b}$ processes at NLO, there is no longer a motivation to reweight the event fractions of the $t\bar{t}+ \geq 1b$ sub-categories to another completely different generator setup. As this reweighting resulted in one of the most significant systematic uncertainties in the $t\bar{t}H(H \to b\bar{b})$ analysis, it is expected that the new $t\bar{t} + b\bar{b}$ setup will reduce this uncertainty considerably.

It is important to note that a MC generator setup based on POWHEG +PYTHIA 8 is generally preferred within the ATLAS collaboration, hence the reweighting strategy of the event fractions in the $t\bar{t}+ \geq 1b$ sub-categories was employed, instead of using the SHERPA 5F and SHERPA 4F samples as nominal predictions. This has multiple reasons, of which most have already been discussed:

- The main reason is that samples generated with POWHEG +PYTHIA 8 contain a full set of modelling associated uncertainties which allows to perform systematic variations of most simulation parameters via the implemented weights. These include, for example, the variations of μ_R, μ_F and the A14 tune parameters to control ISR and FSR as well as systematic variations of the PDF set or the α_S value used to evaluate the PDF. Most other current state-of-the-art generator setups require additional samples that incorporate these variations to be generated at the cost of significant CPU and working time.

- The increased effective number of events generated with POWHEG +PYTHIA 8 compared to MADGRAPH5_aMC@NLO +PYTHIA 8 and SHERPA due to negative event weights being used by the latter two setups.

- Despite some processes and observables being modelled worse by POWHEG in-

terfaced to Pythia 8 compared to other generator setups, the vast majority is modelled very well or even better than with other setups, for example the jet multiplicity in $t\bar{t}$ events. Those that are known to be modelled sub-optimally can be simulated by other setups and possibly used for a reweighting of events, similarly to the $t\bar{t}+\geq 1b$ reweighting.

Nonetheless, more new setups have been developed that are able to simulate the $t\bar{t}+b\bar{b}$ process in the ME at NLO using the 4F scheme. These will be compared to each other in the remainder of this section. The 4F $t\bar{t}+b\bar{b}$ ME generated with the Powheg +OpenLoops setup described above can also be interfaced to the Herwig 7 generator for PS and hadronisation. In addition to this, a new setup of MadGraph5_aMC@NLO is available which uses the version 2.6.0 and changes the value of the resummation scale μ_Q that sets an upper limit on the hardness of the radiation at the beginning of the showering process. This change is believed to better predict the behaviour of additional radiation of light-jets in the PS. This new ME is then interfaced to Pythia 8 for the PS and hadronisation modelling. Finally, the 4F $t\bar{t}+b\bar{b}$ ME has been generated with an updated version of the Sherpa generator, namely v2.2.1, in contrast to v2.1.1 used in the $t\bar{t}H(H\rightarrow b\bar{b})$ search presented here.
All of these updated setups are believed to model the $t\bar{t}+b\bar{b}$ process more precisely than what was available for the analysis presented in this thesis. Thus, it is important to study and compare their predictions to potentially improve future $t\bar{t}H(H\rightarrow b\bar{b})$ searches. Figures 12.13–12.17 show such comparisons, namely of the event fractions of HFC categories and $t\bar{t}+\geq 1b$ sub-categories in addition to various observables in the $e\mu$ decay channel of the $t\bar{t}$ pair at truth level, simulated by the four new $t\bar{t}+b\bar{b}$ predictions mentioned above.

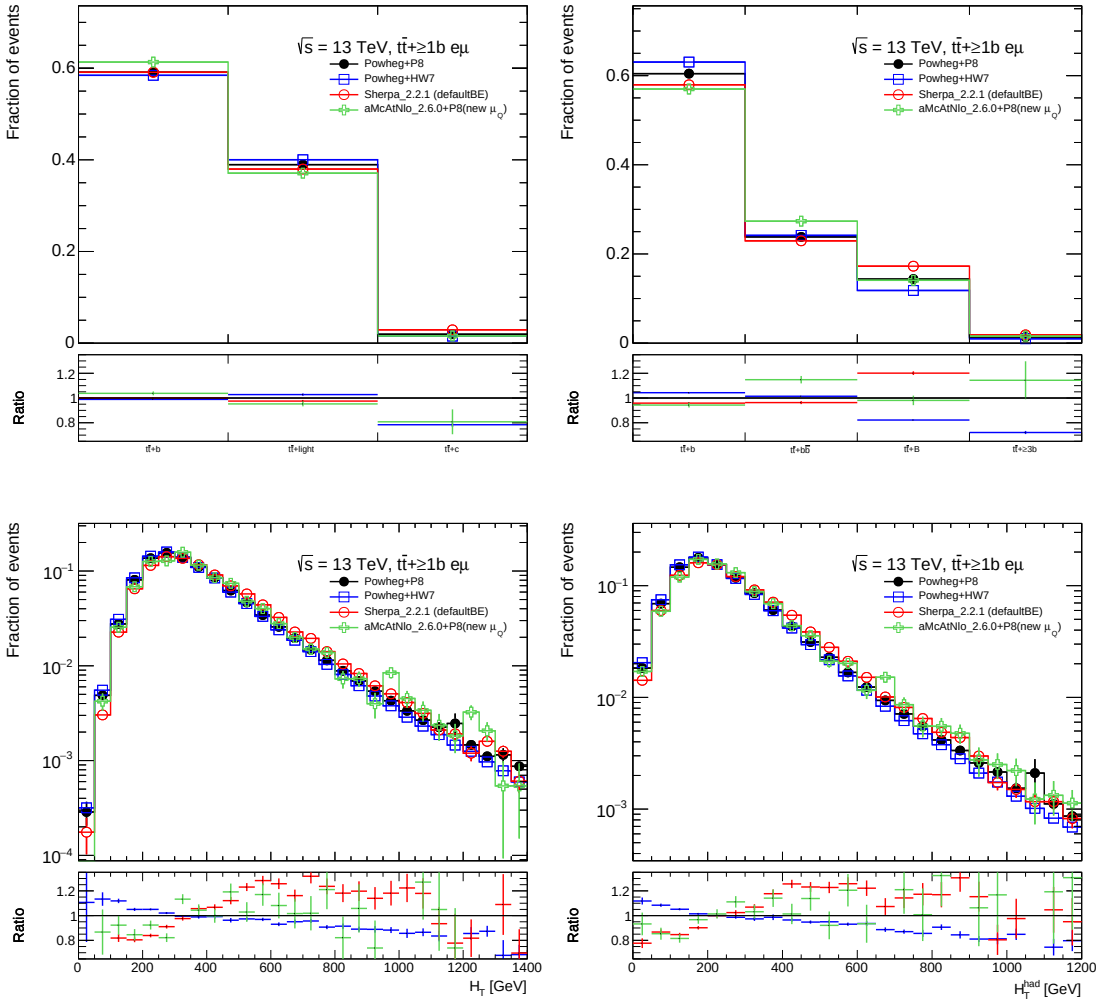

Figure 12.13: Comparison of predictions of the new 4F $t\bar{t} + b\bar{b}$ generator setups. The distributions are simulated by Powheg +Pythia 8 (black), Powheg +Herwig 7 (blue), Sherpa v2.2.1 (red) and MadGraph5_aMC@NLO +Pythia 8 (green). The event selection requires a dileptonic $t\bar{t}$ decay where the W bosons decay into an electron (positron) and an antimuon (muon), respectively. All predictions are normalised to unity in order to assess the shape differences. The lower panel represents the ratio of the respective MC prediction with respect to the Powheg +Pythia 8 prediction. Shown are the fraction of events in the primary HFC categories (top left) and in the $t\bar{t}+ \geq 1b$ sub-categories (top right) in addition to the distributions of the scalar sum of the transverse momenta either of all reconstructed jets and charged leptons (bottom left) or of all reconstructed jets (bottom right), the latter three all in the $t\bar{t}+ \geq 1b$ category. The error bars reflect only the statistical uncertainty from simulation.

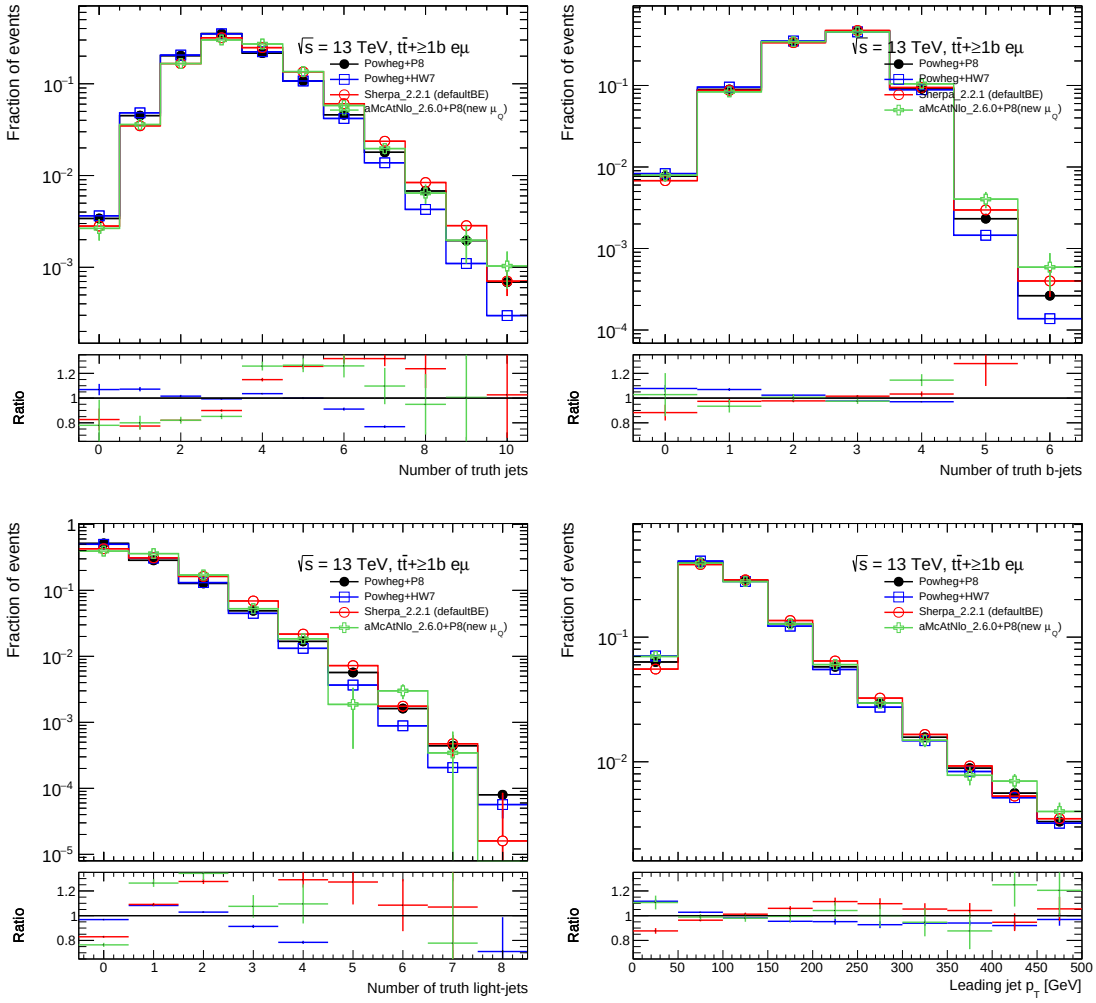

Figure 12.14: Comparison of predictions of the new 4F $t\bar{t}+b\bar{b}$ generator setups. The distributions are simulated by Powheg +Pythia 8 (black), Powheg +Herwig 7 (blue), Sherpa v2.2.1 (red) and MadGraph5_aMC@NLO +Pythia 8 (green). The event selection requires a dileptonic $t\bar{t}$ decay where the W bosons decay into an electron (positron) and an antimuon (muon), respectively, with at least one additional b-jet not coming from the top at truth level. All predictions are normalised to unity in order to assess the shape differences. The lower panel represents the ratio of the respective MC prediction with respect to the Powheg +Pythia 8 prediction. Shown are the distributions of the number of jets (top left), number of b-jets (top right), number of light-jets (bottom left) and the p_{T} of the leading jet (bottom right) in the $t\bar{t}+\geq 1b$ category. The error bars reflect only the statistical uncertainty from simulation.

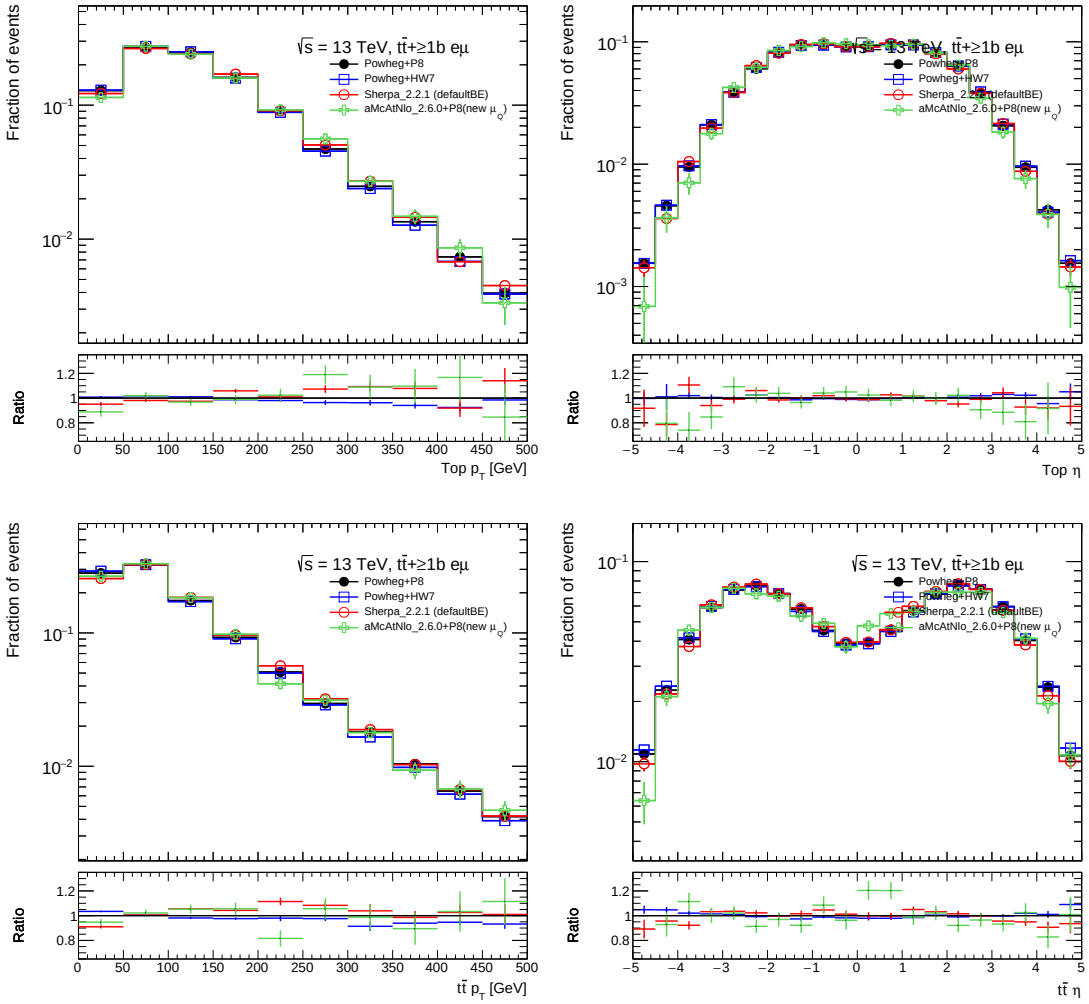

Figure 12.15: Comparison of predictions of the new 4F $t\bar{t} + b\bar{b}$ generator setups. The distributions are simulated by Powheg +Pythia 8 (black), Powheg +Herwig 7 (blue), Sherpa v2.2.1 (red) and MadGraph5_aMC@NLO +Pythia 8 (green). The event selection requires a dileptonic $t\bar{t}$ decay where the W bosons decay into an electron (positron) and an antimuon (muon), respectively, with at least one additional b-jet not coming from the top at truth level. All predictions are normalised to unity in order to assess the shape differences. The lower panel represents the ratio of the respective MC prediction with respect to the Powheg +Pythia 8 prediction. Shown are the distributions of the top quark p_{T} (top left), top quark η (top right), the $t\bar{t}$ system p_{T} (bottom left) and the $t\bar{t}$ system η (bottom right) in the $t\bar{t}+ \geq 1b$ category. The error bars reflect only the statistical uncertainty from simulation.

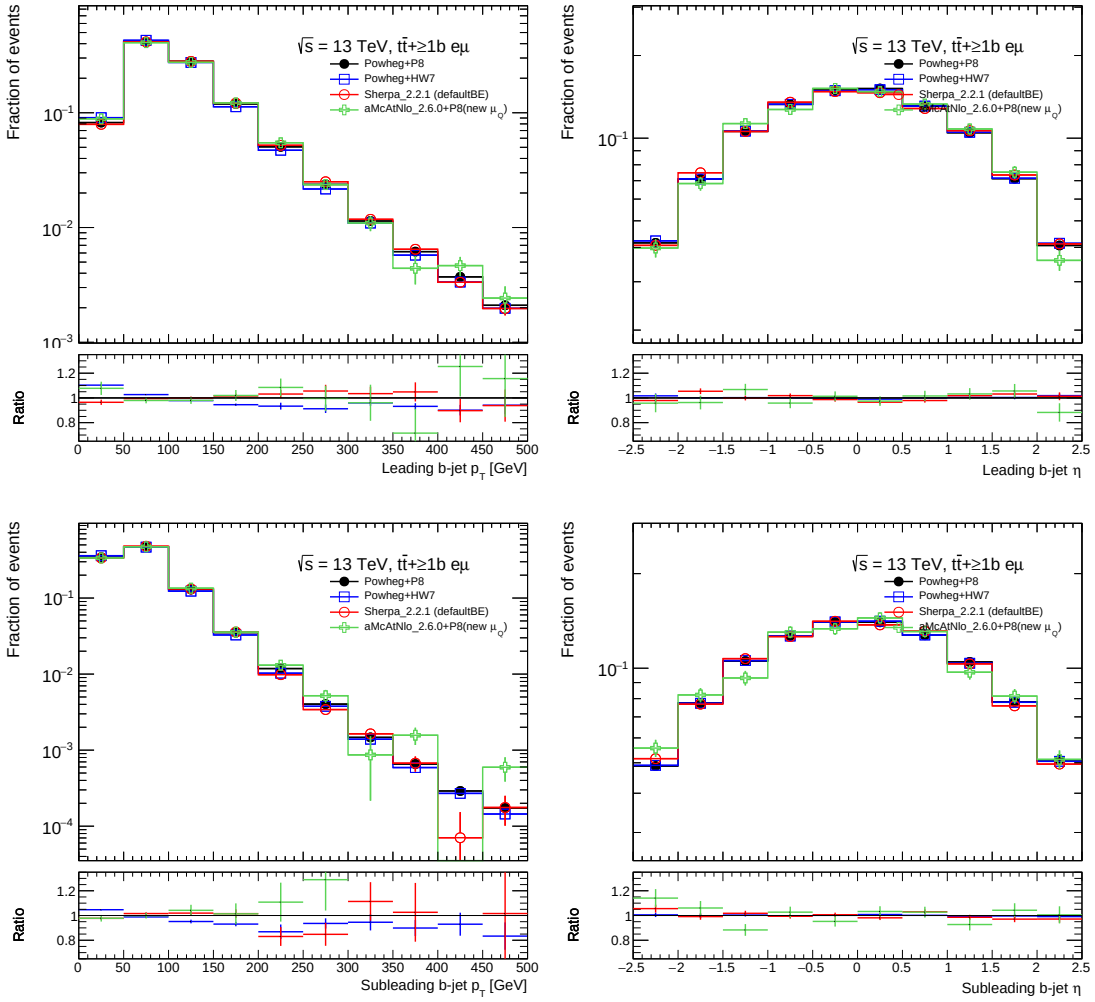

Figure 12.16: Comparison of predictions of the new 4F $t\bar{t}+b\bar{b}$ generator setups. The distributions are simulated by Powheg +Pythia 8 (black), Powheg +Herwig 7 (blue), Sherpa v2.2.1 (red) and MadGraph5_aMC@NLO +Pythia 8 (green). The event selection requires a dileptonic $t\bar{t}$ decay where the W bosons decay into an electron (positron) and an antimuon (muon), respectively, with at least one additional b-jet not coming from the top at truth level. All predictions are normalised to unity in order to assess the shape differences. The lower panel represents the ratio of the respective MC prediction with respect to the Powheg +Pythia 8 prediction. Shown are the leading b-jet p_{T} (top left) and η (top right), and the subleading b-jet p_{T} (bottom left) and η (bottom right) in the $t\bar{t}+\geq 1b$ category. The error bars reflect only the statistical uncertainty from simulation.

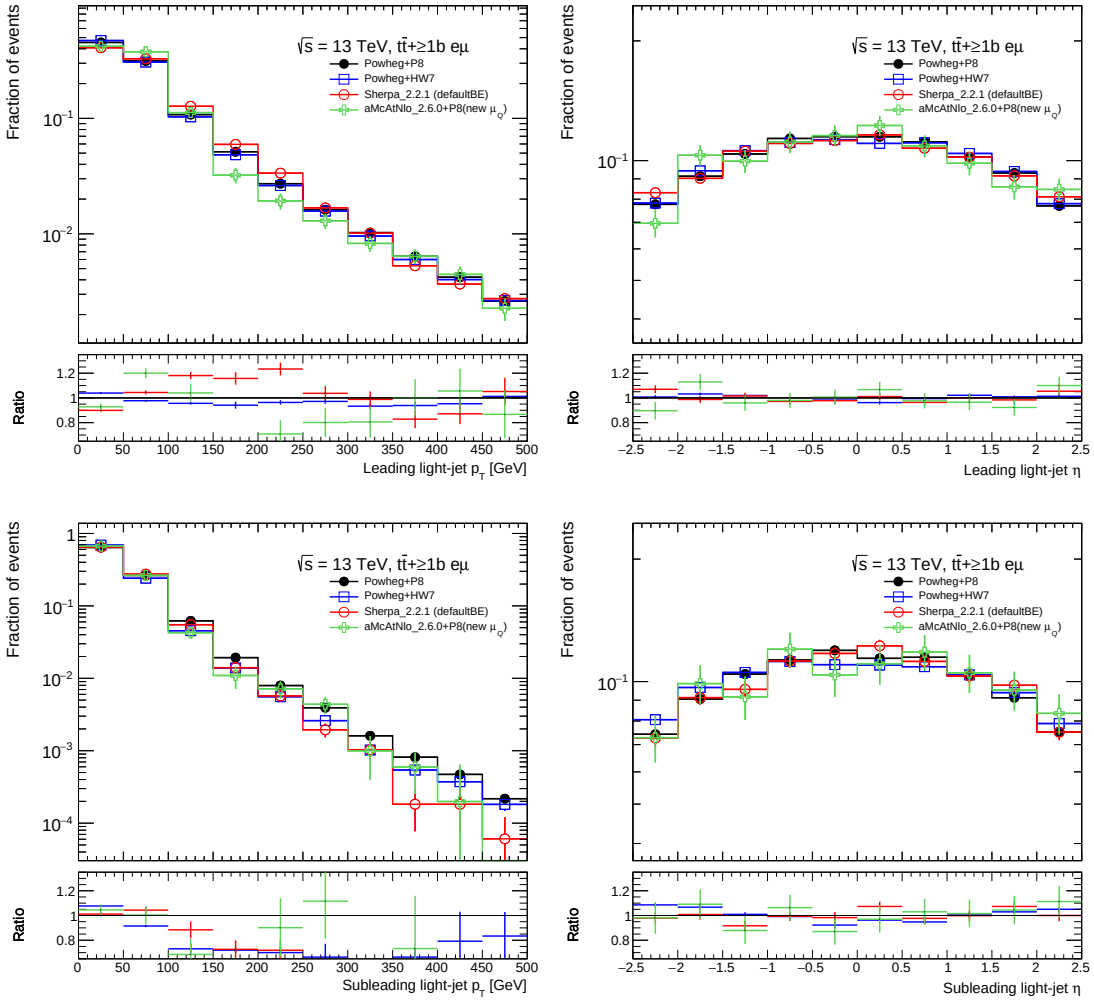

Figure 12.17: Comparison of predictions of the new 4F $t\bar{t} + b\bar{b}$ generator setups. The distributions are simulated by Powheg +Pythia 8 (black), Powheg +Herwig 7 (blue), Sherpa v2.2.1 (red) and MadGraph5_aMC@NLO +Pythia 8 (green). The event selection requires a dileptonic $t\bar{t}$ decay where the W bosons decay into an electron (positron) and an antimuon (muon), respectively, with at least one additional b-jet not coming from the top at truth level. All predictions are normalised to unity in order to assess the shape differences. The lower panel represents the ratio of the respective MC prediction with respect to the Powheg +Pythia 8 prediction. Shown are the leading light-jet p_{T} (top left) and η (top right), and the subleading light-jet p_{T} (bottom left) and η (bottom right) in the $t\bar{t}+ \geq 1b$ category. The error bars reflect only the statistical uncertainty from simulation.

As a general trend, the samples generated with Sherpa and MadGraph5_aMC@NLO interfaced to Pythia 8 show an overall larger statistical uncertainty than those generated with Powheg. This does not only result from the fraction of negative event weights, but also from slightly smaller sample sizes due to limited CPU capacities.
The event fractions of the primary $t\bar{t}+ \geq 1b$ and $t\bar{t}+$ light categories agree reasonably well between the generators, while the $t\bar{t}+ \geq 1c$ predictions show significant discrepancies. However, the expected impact from these differences is not significant. More importantly, the event fractions in the $t\bar{t}+ \geq 1b$ sub-categories predicted by Powheg +Pythia 8 and Sherpa are similar to those shown in Figure 7.3. On the other hand, changing the ME generator from Powheg to MadGraph5_aMC@NLO results in an increased fraction of $t\bar{t}+b\bar{b}$ and $t\bar{t}+ \geq 3b$ events and reduced fraction of $t\bar{t}+b$ events. Most notably, changing the PS and hadronisation model from Pythia 8 to Herwig 7 shows a clear trend in an increased fraction of $t\bar{t}+b$, but reduced fraction of $t\bar{t}+B$ and $t\bar{t}+ \geq 3b$ events.
The H_{T} and $H_{\mathrm{T}}^{\mathrm{had}}$ distributions show very similar features such as large discrepancies primarily at very low and very high values, where the amount of MC events is strongly limited leading to high statistical uncertainties. Furthermore, discrepancies at very low and high values are less significant since the majority of predicted events fall into the medium range around 200 GeV(250 GeV) for $H_{\mathrm{T}}^{\mathrm{had}}$ (H_{T}), where the deviations are less critical and, at least to some extent, covered by the statistical uncertainty. It is interesting to note that changing the PS and hadronisation model from Pythia 8 to Herwig 7 tends to increase the number of events with smaller H_{T} and $H_{\mathrm{T}}^{\mathrm{had}}$ values, while changing the ME generator from Powheg to MadGraph5_aMC@NLO tends to do the opposite. The complete setup change from Powheg +Pythia 8 to Sherpa follows the same trend as only changing the ME calculation, which intuitively is expected, since these observables are affected mainly by the final state kinematics which are determined primarily by the NLO $t\bar{t}+b\bar{b}$ ME generation rather than the PS and hadronisation modelling.
The number of all jets, b-jets and light-jets at truth level as well as the kinematic properties of the top quarks and leading jets can be summarised as follows: the Sherpa and MadGraph5_aMC@NLO +Pythia 8 setups tend, on average, to generate events with more additional jets of any flavour with respect to Powheg +Pythia 8, while the Powheg +Herwig 7 setup tends to do the opposite. In addition to this, the same comparison can be made in the sense that Sherpa and MadGraph5_aMC@NLO +Pythia 8 tend, on average, to produce harder (leading) jets, while Powheg +Herwig 7 is more likely to produce softer jets. Overall, the agreement between the generators for the predicted jet multiplicities and the kinematic properties of the final state particles is very good and well within the statistical uncertainties in the most significant bins i.e. those that contain the vast majority of simulated events.
In conclusion, the systematic uncertainties originating from the comparison of the nominal $t\bar{t}+b\bar{b}$ generator to alternative setups is expected to be reduced for future $t\bar{t}H(H \rightarrow b\bar{b})$ searches and other analyses studying the $t\bar{t}+b\bar{b}$ process. A recent publication by the Atlas collaboration highlights this aspect by comparing various generator setups, including different 4F and 5F predictions of $t\bar{t}+b\bar{b}$, to 36.1 fb^{-1} of Atlas data recorded

at $\sqrt{s} = 13$ TeV [144].

Lastly, in the following, the focus lies on the comparison between the new POWHEG +PYTHIA 8 and SHERPA v2.2.1 setups used to simulate the $t\bar{t} + b\bar{b}$ process with the 4F scheme and in particular the $t\bar{t}+ \geq 1b$ sub-categories. While they were not available for the $t\bar{t}H(H \to b\bar{b})$ analysis, the POWHEG +PYTHIA 8 and SHERPA setups were used extensively for the simulation of various physics processes in the search and thus, understanding the differences in their respective modelling of the $t\bar{t} + b\bar{b}$ process is important for future analyses. Figures 12.18 and 12.19 show plots generated with a RIVET routine [145] prepared by the LHC Higgs $t\bar{t}H/tH$ subgroup that allows to make comparisons of different $t\bar{t} + b\bar{b}$ predictions in the dileptonic channel at truth level [143]. The distributions are normalised to their respective cross-sections which for POWHEG +PYTHIA 8 and SHERPA are 2.809 fb and 2.755 fb, respectively.

Figure 12.20 shows several kinematic variables of the additional $b\bar{b}$ pair in the $t\bar{t}+b\bar{b}$ event in the semileptonic and dileptonic channels combined (labelled 'nonallhad') at truth level. These distributions are normalised to their respective cross-sections as well, namely $11.669 + 2.809 = 14.478$ fb for POWHEG +PYTHIA 8 and $11.462 + 2.775 = 14.217$ fb for SHERPA.

The conclusion of these comparisons is that the behaviour observed in Figures 12.13–12.17 can generally be confirmed. That means that the b-jet kinematics and b-jet multiplicities agree very well within statistical uncertainties, suggesting a consensus on the understanding of the additional $b\bar{b}$ pair in the $t\bar{t} + b\bar{b}$ ME. On the other hand, the simulated additional light-jets behave in a significantly different way. First of all, the number of predicted additional light-jets is much higher in the SHERPA sample compared to POWHEG +PYTHIA 8, but also the p_{T} of these additional light-jets is larger in the SHERPA prediction. This suggests that, in general, the radiation pattern in the PS of the $t\bar{t} + b\bar{b}$ process is harder when simulated with SHERPA compared to the one in POWHEG +PYTHIA 8. Additional studies may be worthwhile to investigate possible differences in the $t\bar{t} +$ light and $t\bar{t}+ \geq 1c$ backgrounds resulting from the previous assumption.

All in all, the studies presented in this section allow to expect a significantly improved modelling of the $t\bar{t}+ \geq 1b$ process in future analyses involving the $t\bar{t} + b\bar{b}$ process. Thus, future $t\bar{t}H(H \to b\bar{b})$ searches should greatly benefit from these newly developed generator setups and be less limited by the modelling uncertainty.

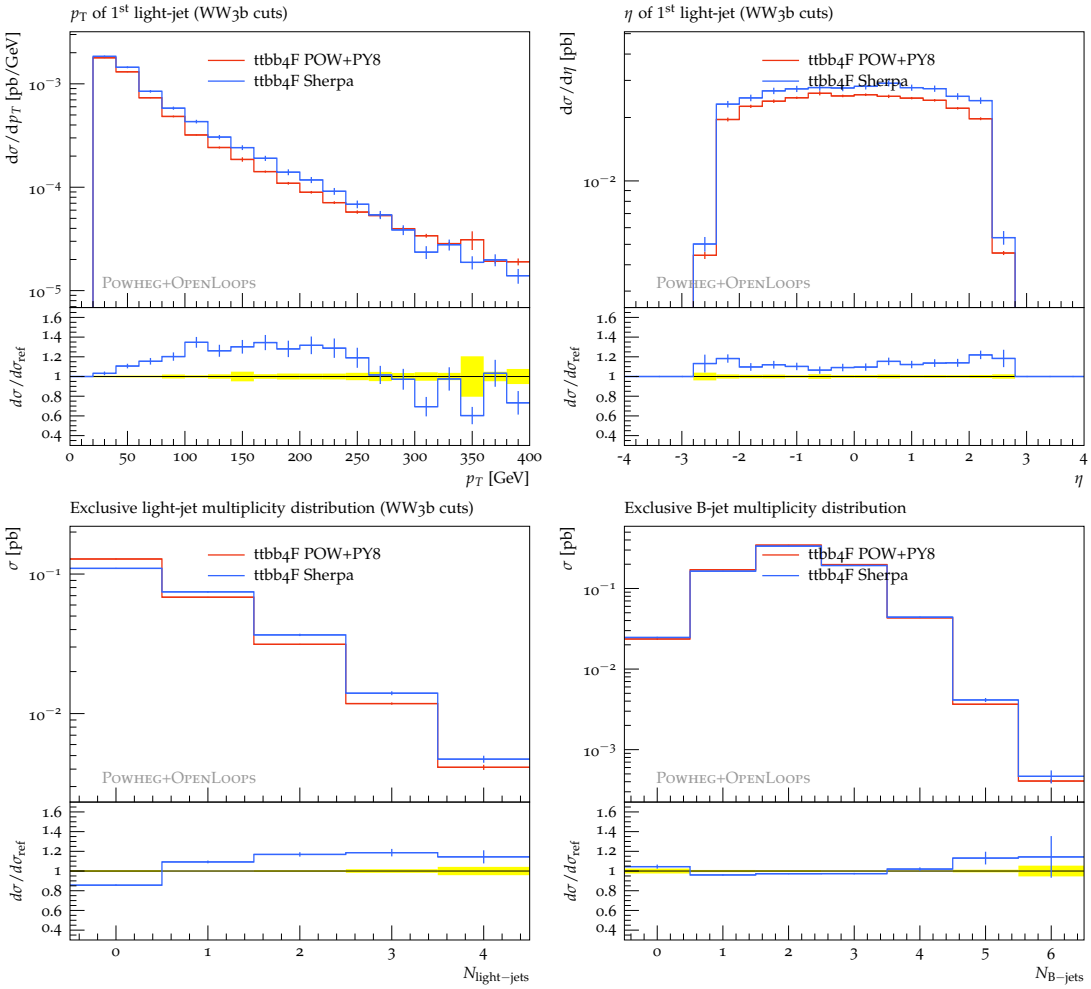

Figure 12.18: Comparison of predictions of the new 4F $t\bar{t}+b\bar{b}$ generator setups. The distributions are simulated by POWHEG +PYTHIA 8 (red) and SHERPA v2.2.1 (blue). The event selection requires a dileptonic $t\bar{t}$ decay with at least one additional b-jet not coming from the top at truth level. The distributions are normalised to their predicted cross-sections. The lower panel represents the ratio of the SHERPA prediction with respect to the POWHEG +PYTHIA 8 prediction. Shown are the leading light-jet p_T (top left) and η (top right), the exclusive light-jet multiplicity (bottom left), all in the $t\bar{t}+\geq 1b$ category, and the exclusive b-jet multiplicity (bottom right). The error bars reflect the statistical uncertainties of the respective samples, while the yellow band in the ratio panel represents the statistical uncertainty on the POWHEG +PYTHIA 8 prediction.

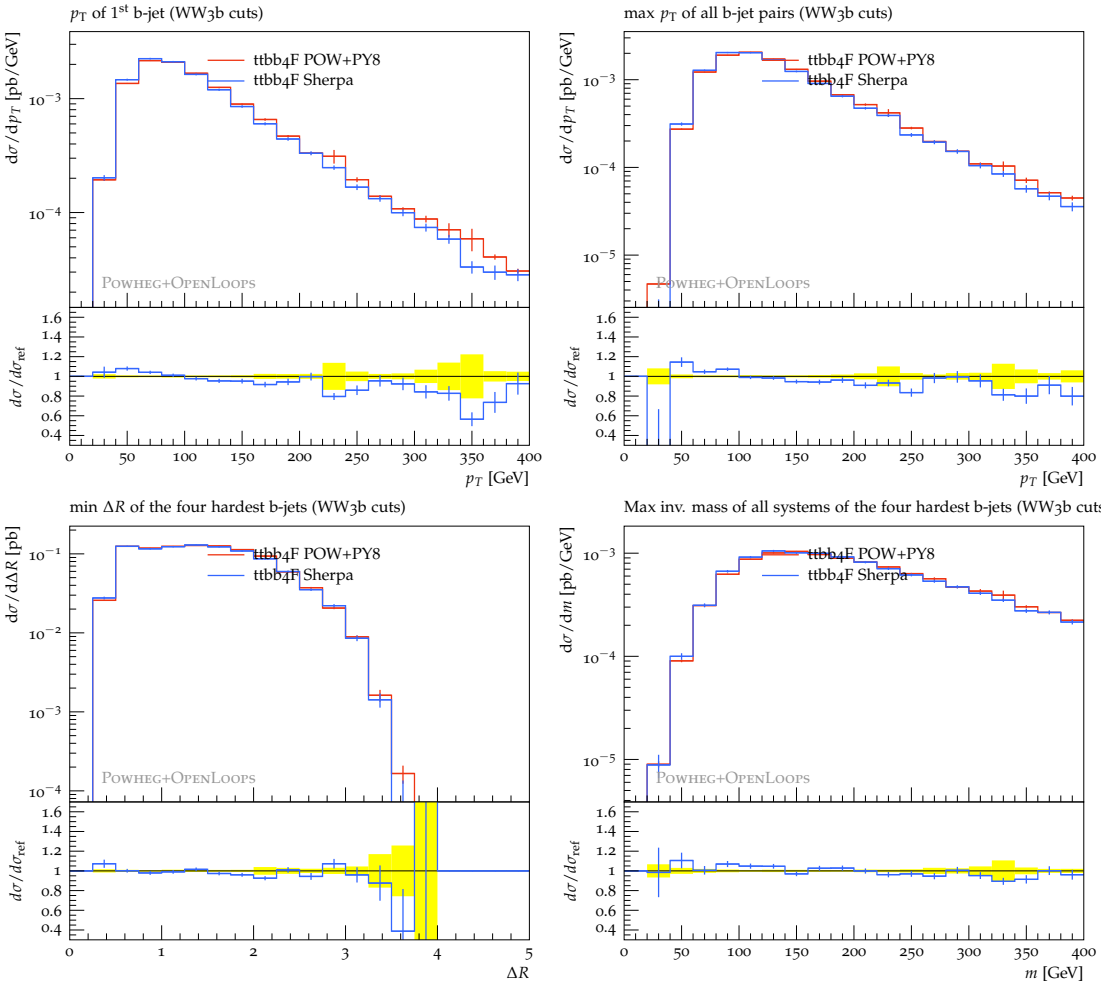

Figure 12.19: Comparison of predictions of the new 4F $t\bar{t} + b\bar{b}$ generator setups. The distributions are simulated by Powheg +Pythia 8 (red) and Sherpa v2.2.1 (blue). The event selection requires a dileptonic $t\bar{t}$ decay with at least one additional b-jet not coming from the top at truth level. The distributions are normalised to their predicted cross-sections. The lower panel represents the ratio of the Sherpa prediction with respect to the Powheg +Pythia 8 prediction. Shown are the leading b-jet p_T (top left), the maximum p_T of all b-jet pairs (top right), the minimum ΔR between all pairs built from the four hardest b-jets (bottom left) and the maximum invariant mass of all pairs built from the four hardest b-jets (bottom right), all in the $t\bar{t}+ \geq 1b$ category. The error bars reflect the statistical uncertainties of the respective samples, while the yellow band in the ratio panel represents the statistical uncertainty on the Powheg +Pythia 8 prediction.

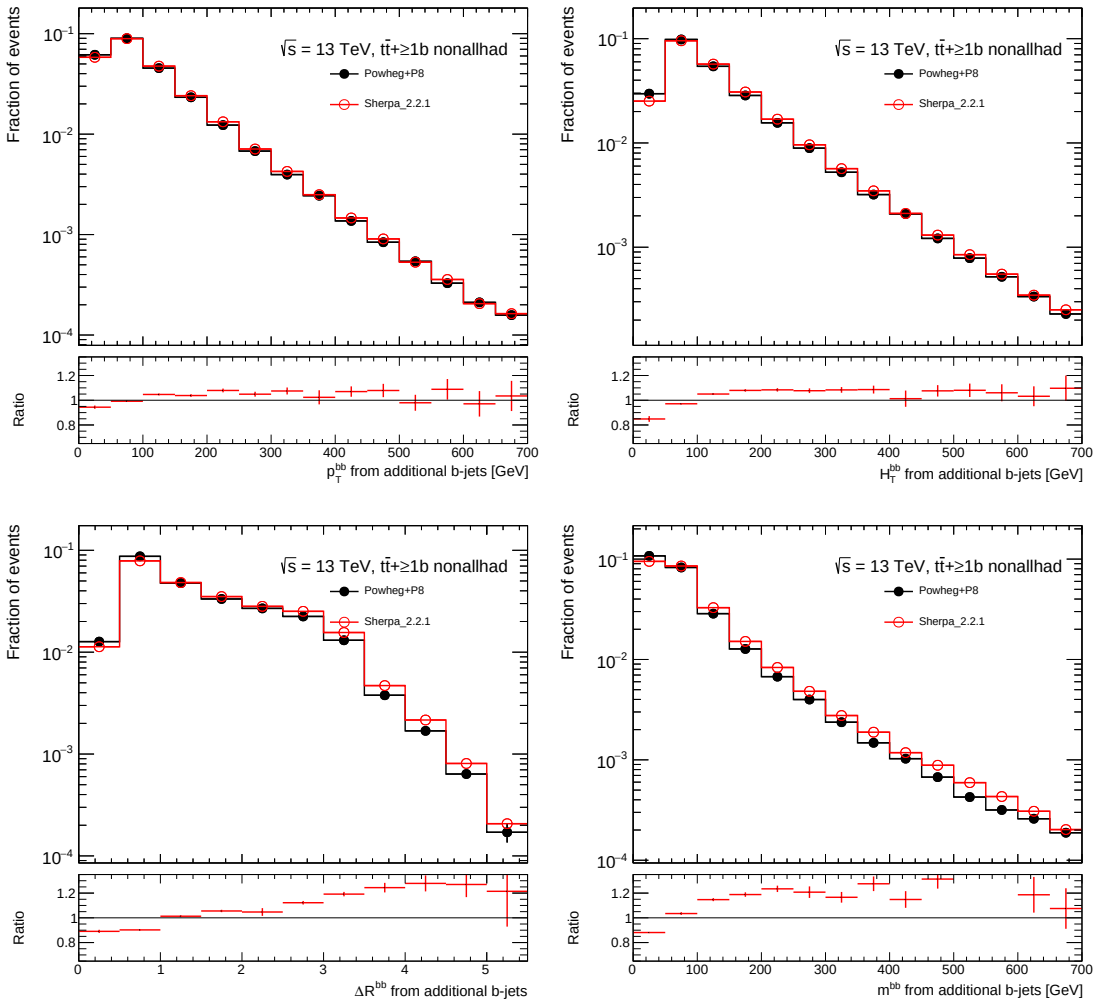

Figure 12.20: Comparison of predictions of the new 4F $t\bar{t} + b\bar{b}$ generator setups. The distributions are simulated by POWHEG +PYTHIA 8 (black) and SHERPA v2.2.1 (red). The event selection requires a dileptonic $t\bar{t}$ decay with at least one additional b-jet not coming from the top at truth level. The distributions are normalised to their predicted cross-sections. The lower panel represents the ratio of the SHERPA prediction with respect to the POWHEG +PYTHIA 8 prediction. Shown are several kinematic properties of the additional $b\bar{b}$ pair in the $t\bar{t}+ \geq 1b$ region at truth level. These are the vector sum p_{T} (top left), scalar sum p_{T} (top right), ΔR (bottom left) and invariant mass (bottom right) of the additional $b\bar{b}$ pair in the event. The error bars reflect only the statistical uncertainty from simulation.

12.3 Studies on the modelling of the $t\bar{t}H$ signal process

The $t\bar{t}H$ signal process in the analysis presented in this thesis is modelled with MadGraph5_aMC@NLO interfaced to Pythia 8, which is described in Section 7.2. To assess the modelling uncertainty of the PS and hadronisation model choice, an alternative sample is produced using MadGraph5_aMC@NLO +Herwig++. This comparison is the fifth systematic uncertainty in the combined fit, ranked by its impact on the signal strength parameter uncertainty, the first four being related to the $t\bar{t}+\geq 1b$ background modelling. However, in contrast to the $t\bar{t}$ +jets background process, no other ME calculation is considered in order to evaluate the uncertainty resulting from choosing this model. At the time this analysis was published, the other state-of-the-art generators used in Atlas were not able or properly setup to simulate the $t\bar{t}H$ process at NLO precision. But, at the time of writing this thesis, the Powheg +Pythia 8 setup can be used to do so. The difference with respect to the $t\bar{t}$ production is the value of the h_{damp} parameter, which is not $1.5 \cdot m_{\text{top}} = 258.75$ GeV, but instead set to $1.5 \cdot (2 \cdot m_{\text{top}} + m_H)/2 = 352.5$ GeV.
In addition to this, an updated configuration of MadGraph5_aMC@NLO v2.6.0 with a different resummation scale μ_Q interfaced to Pythia 8 can be used to produce $t\bar{t}H$ events, identical to the setup used for the $t\bar{t}+b\bar{b}$ production in the previous section. However, the corresponding setup of Powheg +Herwig 7 is not yet available in Atlas to simulate the $t\bar{t}H$ process.

Figures 12.21–12.23 show the comparison of the four different normalised predictions mentioned above for various important kinematic variables of the top quarks, the Higgs boson and related systems as well as H_{T} and the jet multiplicity. These variables are shown at truth level in semileptonic $t\bar{t}H(H \to b\bar{b})$ events.

The most striking differences are observed when comparing the old and the new MadGraph5_aMC@NLO +Pythia 8 setups: the new configuration has an increased resummation scale μ_Q representing an upper limit on the hardness of the radiation during the PS. Clearly, this new configuration predicts, on average, more events with a higher number of jets with $p_{\text{T}} > 25$ GeV and $|\eta| < 2.5$ and thus an increased H_{T} value, the breaking point being roughly at $N_{\text{jets}} \geq 7$ and $H_{\text{T}} \geq 450$ GeV. Aside from this, the various kinematic distributions are almost unaffected by this change within the given statistical uncertainty, which is not unexpected since they are observables constructed from the $t\bar{t}H$ system and its sub-components. However, this updated simulation cannot be simply interpreted as a more precise prediction of the $t\bar{t}H$ process until further studies are performed.
A similar trend, but in the other direction, can be seen by comparing the two old MadGraph5_aMC@NLO setups, because the Herwig++ PS and hadronisation model tends to produce, on average, events with fewer $p_{\text{T}} > 25$ GeV jets and thus a softer H_{T}. However, the Herwig++ PS predicts a less central and slightly softer $t\bar{t}$ and $t\bar{t}H$ systems, while leaving the individual top quark and Higgs boson kinematics almost unaffected

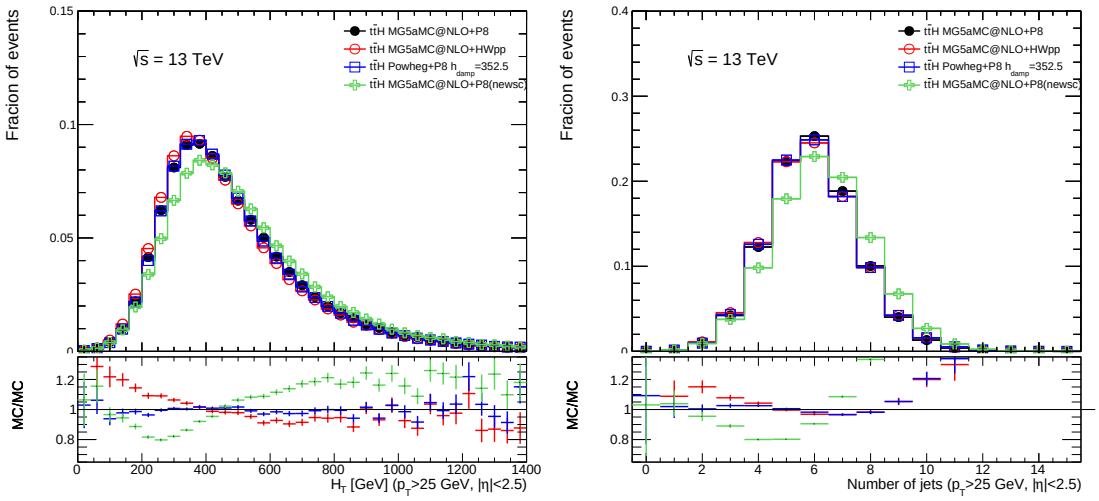

Figure 12.21: Comparison of predictions of the old and new $t\bar{t}H$ generator setups. The distributions are simulated by the old MadGraph5_aMC@NLO +Pythia 8 (black) and MadGraph5_aMC@NLO +Herwig++ (red) setups, as well as the new Powheg +Pythia 8 (blue) and MadGraph5_aMC@NLO +Pythia 8 (green) setups. The event selection requires a semileptonic $t\bar{t}$ decay and a $H \to b\bar{b}$ decay at truth level. All predictions are normalised to unity in order to assess the shape differences. The lower panel represents the ratio of the respective MC prediction with respect to the old MadGraph5_aMC@NLO +Pythia 8 prediction. Shown are the scalar sum p_{T} of all reconstructed charged leptons and jets (left) as well as the number of those jets (right) that fulfil $p_{\mathrm{T}} > 25$ GeV and $|\eta| < 2.5$. The error bars reflect only the statistical uncertainty.

within statistical uncertainties.
The most important comparison, however, is between the old nominal $t\bar{t}H$ setup and the new Powheg +Pythia 8 setup, as the latter is always preferred in Atlas simulation due to the same reasons given for the $t\bar{t} + b\bar{b}$ production in the previous section. The above plots all show that both generators agree exceptionally well in all distributions within their statistical uncertainties, with the small exception of an, on average, slightly smaller $t\bar{t}H$ system p_{T} predicted by Powheg +Pythia 8. Other bins where the agreement in these distributions is worse are not statistically significant or contain an insignificant amount out of all generated signal events. In addition to this, two additional $t\bar{t}H$ samples have been generated with Powheg +Pythia 8 but different h_{damp} values, namely one with $1.5 \cdot m_{\mathrm{top}}$ and one with ∞. The impact of this change is negligible within statistical uncertainties in all of the distributions shown above. With this, the modelling uncertainty on the $t\bar{t}H$ process originating from this comparison of shapes is expected to be very small.

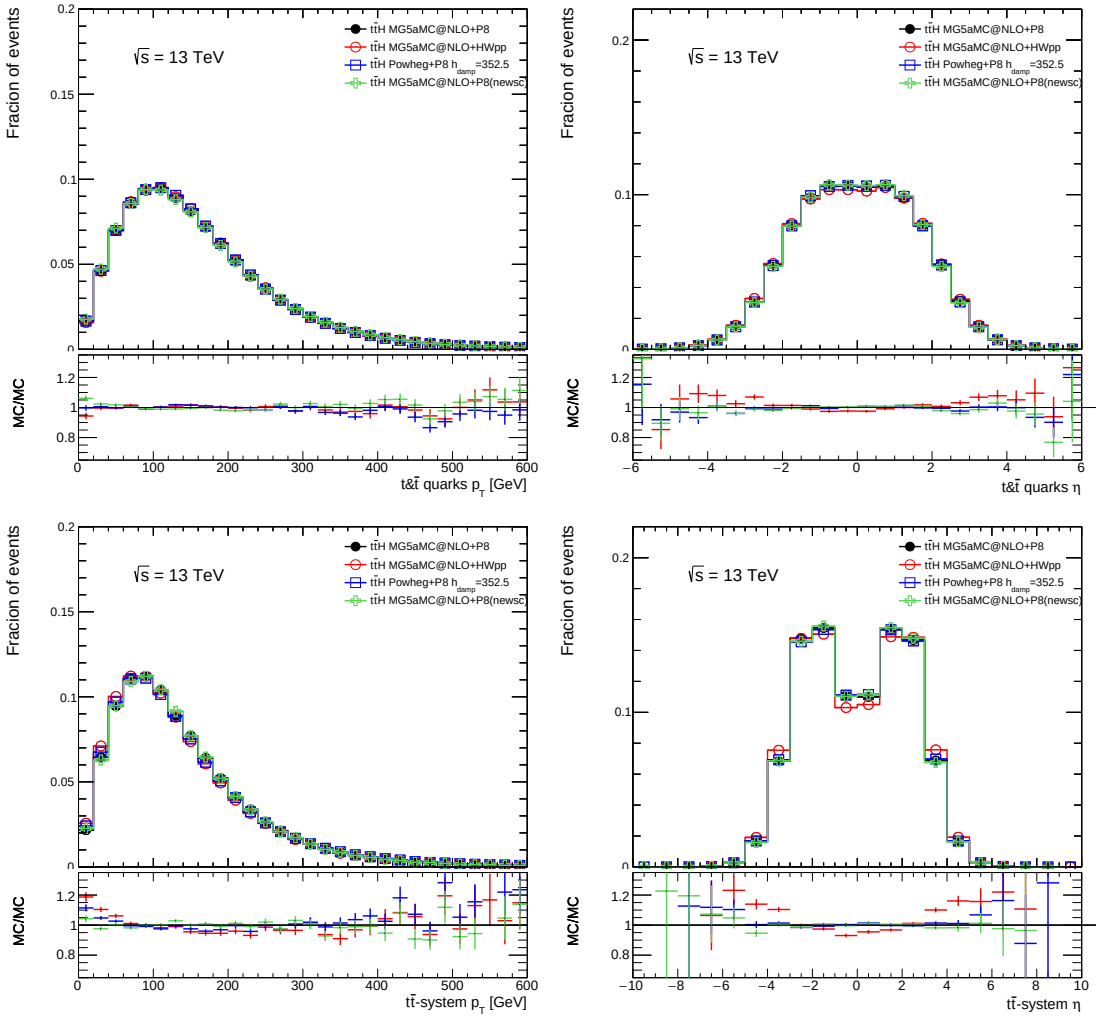

Figure 12.22: Comparison of predictions of the old and new $t\bar{t}H$ generator setups. The distributions are simulated by the old MadGraph5_aMC@NLO +Pythia 8 (black) and MadGraph5_aMC@NLO +Herwig++ (red) setups, as well as the new Powheg +Pythia 8 (blue) and MadGraph5_aMC@NLO +Pythia 8 (green) setups. The event selection requires a semileptonic $t\bar{t}$ decay and a $H \to b\bar{b}$ decay at truth level. All predictions are normalised to unity in order to assess the shape differences. The lower panel represents the ratio of the respective MC prediction with respect to the old MadGraph5_aMC@NLO +Pythia 8 prediction. Shown are the top quark p_{T} (top left) and η (top right) as well as the $t\bar{t}$ system p_{T} (bottom left) and η (bottom right). The error bars reflect only the statistical uncertainty.

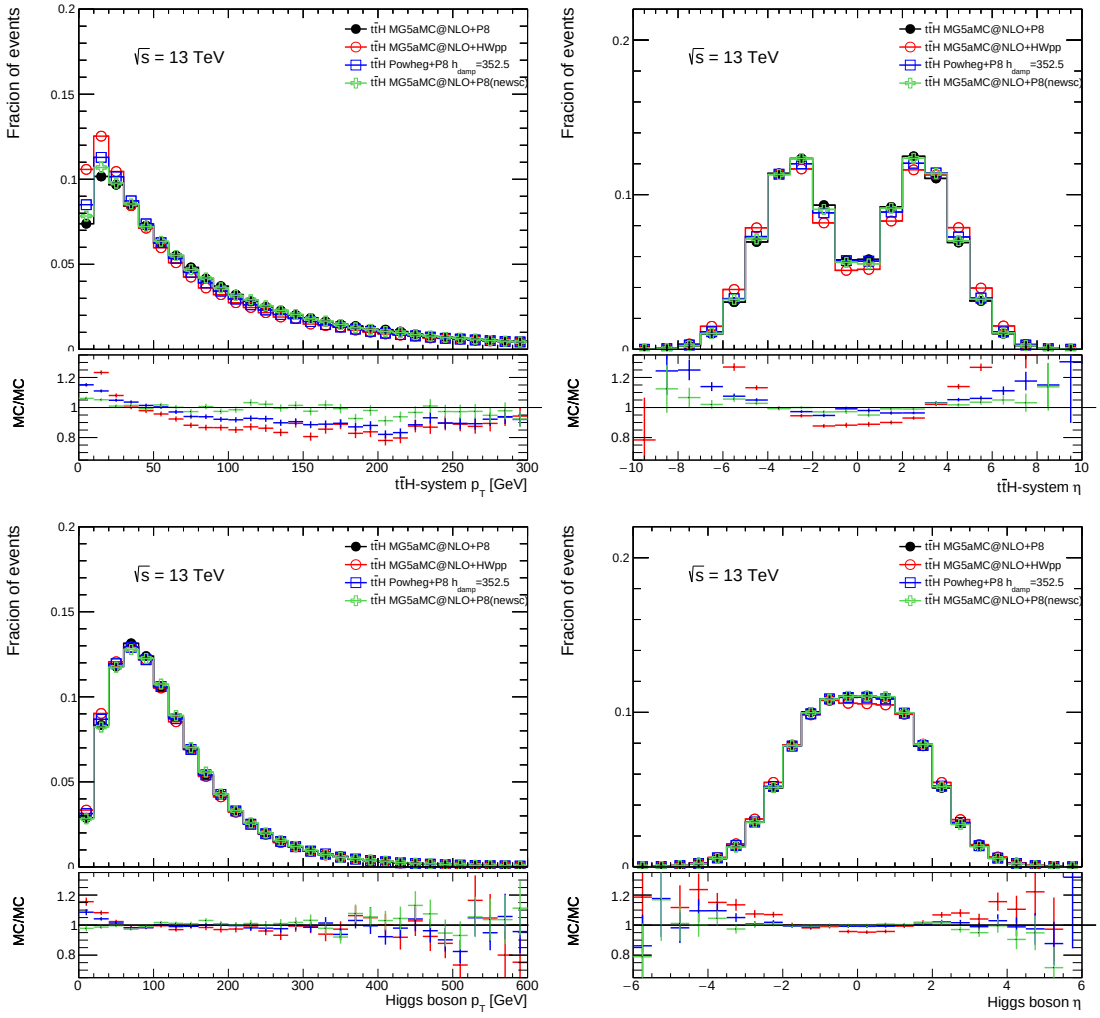

Figure 12.23: Comparison of predictions of the old and new $t\bar{t}H$ generator setups. The distributions are simulated by the old MadGraph5_aMC@NLO +Pythia 8 (black) and MadGraph5_aMC@NLO +Herwig++ (red) setups, as well as the new Powheg +Pythia 8 (blue) and MadGraph5_aMC@NLO +Pythia 8 (green) setups. The event selection requires a semileptonic $t\bar{t}$ decay and a $H \to b\bar{b}$ decay at truth level. All predictions are normalised to unity in order to assess the shape differences. The lower panel represents the ratio of the respective MC prediction with respect to the old MadGraph5_aMC@NLO +Pythia 8 prediction. Shown are the $t\bar{t}H$ system p_{T} (top left) and η (top right) as well as the Higgs boson p_{T} (bottom left) and η (bottom right). The error bars reflect only the statistical uncertainty.

Another important aspect is the event selection efficiency. When normalising the old MadGraph5_aMC@NLO +Pythia 8 sample to the new Powheg +Pythia 8 sample, the number of events predicted by Powheg +Pythia 8 in two of the most sensitive single lepton regions of the $t\bar{t}H(H \to b\bar{b})$ analysis, namely $\mathrm{SR}_1^{\geq 6j}$ and SR_1^{5j}, is $\approx 6\%$ and $\approx 5\%$ larger than the MadGraph5_aMC@NLO +Pythia 8 prediction, respectively. This significantly increased selection efficiency by the Powheg +Pythia 8 has been found to originate from the increased number of jets that are b-tagged which is illustrated in Figure 12.24. The number of signal events with at least three b-tagged jets is higher in the new setup compared to the old prediction. However, the reason for this is not entirely clear at this point and needs to be studied further.

Because of the reasons mentioned above, it is sensible to employ the Powheg +Pythia 8 generator setup to model the $t\bar{t}H$ signal process in future $t\bar{t}H$ analyses. In this case, the $t\bar{t}H(H \to b\bar{b})$ analysis would use the same generator setup for the $t\bar{t}H$ signal as well as the main backgrounds which are the inclusive $t\bar{t}$ +jets as well as the dedicated $t\bar{t} + b\bar{b}$ production processes. In addition to this, the shapes of many predicted observables from this new setup agree very well with the alternative setups used in previous analyses with the exception of the jet and b-jet multiplicities. The remaining differences with respect to other MC generator models, which are mainly acceptance effects, are understood to a certain extent, but are still subject to further investigations at this point and will continue to be in the foreseeable future. Therefore, systematic uncertainties originating from the comparison of shapes modelled by different MC generators or those originating from extrapolating information from one analysis region to another are expected to be reduced significantly when using this new setup. On the other hand, the systematic uncertainty resulting from the difference in the signal event acceptance is significant and motivates a redefinition of the assessment of modelling related systematic uncertainties in order to avoid it limiting the analyses.

Finally, with all the modelling studies presented in this chapter, the most significant uncertainties limiting the sensitivity of the $t\bar{t}H(H \to b\bar{b})$ search, which is the modelling of signal and background, can be understood, controlled and improved for future $t\bar{t}H$ analyses within the Atlas collaboration.

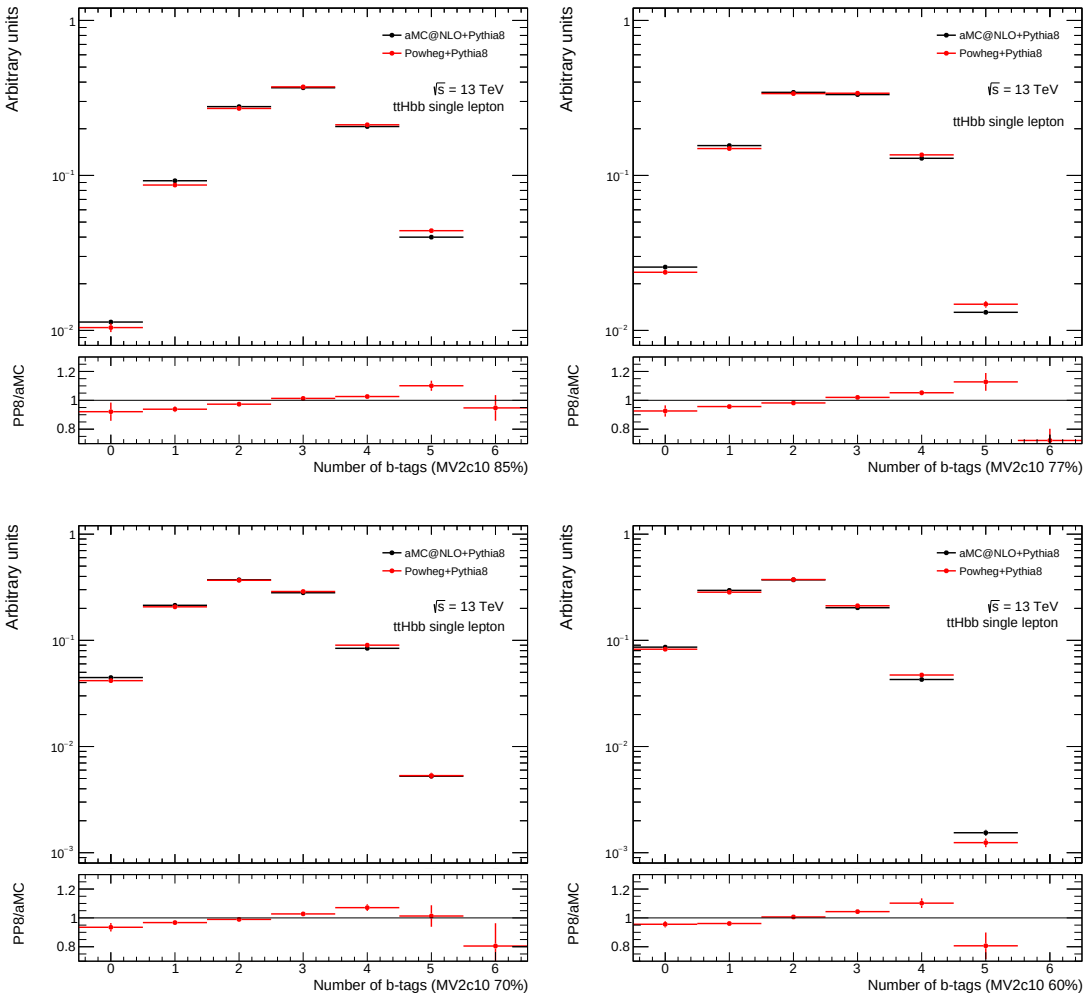

Figure 12.24: Comparison of predictions of the old nominal $t\bar{t}H$ generator, MadGraph5_aMC@NLO +Pythia 8 (black), and the new generator, Powheg +Pythia 8 (red). The event selection requires a semileptonic $t\bar{t}$ decay and a $H \to b\bar{b}$ decay at truth level. All predictions are normalised to unity in order to assess the shape differences. The lower panel represents the ratio of the new Powheg +Pythia 8 prediction with respect to the old MadGraph5_aMC@NLO +Pythia 8 prediction. Shown are the number of b-jets at truth level passing the 85% WP (top left), 77% WP (top right), 70% WP (bottom left) and 60% (bottom right). The error bars reflect only the statistical uncertainty.

CHAPTER 13

Conclusions

In this thesis, two particle physics analyses performed with the ATLAS experiment have been presented. These are the calibration of the b-tagging algorithm used in the ATLAS collaboration based on an improved likelihood method and the search for the $t\bar{t}H(H \to b\bar{b})$ process. Both analyses are based on pp collision data at $\sqrt{s} = 13$ TeV, collected with the ATLAS detector at the LHC. Their underlying ideas have been motivated and the respective analysis strategies have been explained based on physics arguments detailed in Chapter 2, while exploiting the detector features and experimental conditions given in Chapter 3 as well as the possibilities and limitations of our physics modelling capacities that were described in Chapter 4. The former analysis is crucial to the precision of the latter which contributes greatly to our fundamental understanding of particle physics. Especially the search for the $t\bar{t}H(H \to b\bar{b})$ process involves a complex and challenging final state and its observation would have a significant impact on current state-of-the-art theories regarding the Higgs boson and the top quark in the Standard Model (SM) as well as their role in possible extensions of it.

The ***b*-tagging calibration analysis** presented in Chapter 6 discusses the concept of b-tagging algorithms at hadron collider experiments and the specific method how such an algorithm is constructed, trained, and calibrated in the ATLAS experiment. This is critical, as the identification of b-jets depends on the detector properties as well as the experimental conditions at the LHC. The baseline algorithms and further optimisation studies employed during Run 2 of the LHC physics programme have been discussed [66–68]. The calibration of the ATLAS b-tagging algorithms, which are MV2c10 and DL1, is done to provide scale factors that analysers can apply to correct the b-tagging efficiencies obtained from simulation to the one observed in data. The previous calibration analysis employs a likelihood method based on probability distribution functions (PDFs) in dileptonic $t\bar{t}$ events [75]. The results from this method are strongly limited

by the systematic uncertainty originating from the physics modelling of the $t\bar{t}$ process. Nonetheless, they are used in the $t\bar{t}H(H \to b\bar{b})$ search summarised below.
The new calibration is based on an improved PDF likelihood method [76] which aims to reduce the impact from this dominating $t\bar{t}$ modelling uncertainty. The analysis is based on 80.4 ± 1.6 fb^{-1} of data [33], collected with the ATLAS detector in the years 2015, 2016, and 2017. In contrast to the previous iteration, it selects $t\bar{t}$ events where the two W bosons decay into an electron (positron) and an antimuon (muon), respectively, and exactly two jets are reconstructed. In addition to this, a new region selection as well as a new strategy to estimate non-prompt leptons have been developed. Furthermore, the new likelihood includes correction factors for the fraction of events containing certain combinations of flavours of the two jets, because either could be a b-jet or not. By fitting these fractions in the signal as well as all control regions, more information is taken from data instead of simulation and thus the calibration depends less on physics modelling and effectively decreases the systematic uncertainty related to the simulation of the $t\bar{t}$ process, on average, by a factor of two, depending on the transverse momentum (p_{T}) of the jet.
The individual datasets for 2015 and 2016 as well as 2017 are subject to different experimental and detector conditions. Despite this, the results based on these individual datasets as well as the combined results are all compatible with each other within the considered uncertainties. The quoted scale factors in all jet p_{T} bins are compatible with unity within uncertainties, while the behaviour of the fit, the results and corresponding uncertainties are well understood. The total uncertainty ranges from 8-9% at very low p_{T} to about 1% at a medium p_{T} between 100-200 GeV to 3-4% at very high p_{T}. The most dominant sources of uncertainties are related to the jet energy scale and resolution. However, the modelling of $t\bar{t}$ still contributes significantly. The latter involves primarily the choice of the parton shower (PS) and hadronisation model, the scale choice for initial (ISR) and final state radiation (FSR) and the limited amount of simulated events containing reconstructed jets with a p_{T} between 20-40 GeV. Aside from these, systematic uncertainties related to the inclusive Z boson background modelling contributes significantly to the total uncertainty as well. Lastly, the statistical uncertainty in data is becoming a significant contributor to the total uncertainty at high jet p_{T}.

In the Chapters 7–12, a search for the associated production of a SM Higgs boson and a $t\bar{t}$ pair in which the Higgs boson decays into a $b\bar{b}$ pair is presented. This search is based on 36.1 fb^{-1} data, collected with the ATLAS detector in 2015 and 2016. It focuses on the semileptonic and dileptonic decay modes of the $t\bar{t}$ pair which leads to a final state involving four b-jets, missing transverse momentum and either two charged leptons or one charged lepton and two light-jets. The most dominant background for this analysis is the inclusive $t\bar{t}$ production, in particular $t\bar{t} + b\bar{b}$ production, which is an irreducible background since its final state is identical to the signal process. The analysis strategy includes a pre-selection to place events into different categories according to their compositions, tending to enrich them either with signal events or dedicated background events. These regions are used to constrain the corresponding background contributions

and associated systematic uncertainties. Then, multivariate analysis techniques are used to discriminate between signal and background events in the signal-enriched regions and finally, a profile likelihood fit is performed simultaneously in all single lepton and dilepton regions to extract the $t\bar{t}H(H \to b\bar{b})$ cross-section.
The observed data is found to be consistent with both the background-only hypothesis as well as the signal-plus-background hypothesis assuming a SM Higgs boson with $m_H = 125$ GeV. The measured signal strength parameter from the combined fit is $\mu = \sigma^{\text{Data}}_{t\bar{t}H}/\sigma^{\text{SM}}_{t\bar{t}H} = 0.84 \pm 0.29(\text{stat.})^{+0.57}_{-0.54}(\text{syst.}) = 0.84^{+0.64}_{-0.61}$, corresponding to an observed (expected) significance of 1.4 (1.6) standard deviations. This translates into an inclusive cross-section of $\sigma_{t\bar{t}H} = 426^{+326}_{-312}$ fb when conservatively assuming the uncertainties on μ and $\sigma^{\text{SM}}_{t\bar{t}H}$ to be fully uncorrelated for simplification. A signal strength above 2.0 is excluded at 95% confidence level, compared to an expected exclusion limit of 1.2 in the absence of a signal. The total uncertainty is presently dominated by the systematic uncertainty, more specifically by the uncertainty on our theoretical knowledge of the $t\bar{t}+ \geq 1b$ process. A better understanding of this background will be crucial for future efforts to observe the $t\bar{t}H(H \to b\bar{b})$ process. Additionally, not only the background, but also the signal modelling is subject to large systematic uncertainties. And while the statistical uncertainties of data and simulation as well as the experimental uncertainties related to the detector performance will decrease in the future, the modelling uncertainties mentioned above are the limiting factor in this search and thus prevent a significant increase in sensitivity of future $t\bar{t}H(H \to b\bar{b})$ searches. This aspect motivates the direction that future studies should make significant progress in our understanding of physics of the top quark and the Higgs boson.

13.1 Outlook

The observation of the $t\bar{t}H$ process has been one of the main goals of the LHC physics programme in Run 2. It is the most sensitive process to directly measure the Yukawa-coupling of the Higgs boson to the top quark. Such a direct measurement has been successful by combining all $t\bar{t}H$ searches in the ATLAS experiment [20]. This combined result at $\sqrt{s} = 13$ TeV is $\sigma_{t\bar{t}H} = 670 \pm 90(\text{stat.})^{+110}_{-100}\text{syst.})$, corresponding to an observed (expected) significance of 5.8σ (4.9σ). The result allows to claim the observation of the $t\bar{t}H$ production process at the LHC with a significance above five standard deviations. By focusing on events where the Higgs boson decays into a $b\bar{b}$ pair, measuring the cross-section of this production mode allows to extract the Yukawa-coupling to the bottom quark as well. This effort, however, requires a lot more data to be collected and a higher efficiency to select the $t\bar{t}H(H \to b\bar{b})$ signal events as well as reject background events. This latter aspect can be achieved by significantly improving the modelling i.e. our understanding of the $t\bar{t}H$ signal as well as the $t\bar{t} + b\bar{b}$ background processes. New Monte Carlo (MC) generator setups have been developed by now to address this problem. Studies based on these new setups are presented in Chapter 12 and they give promising hints at a possible decrease of the systematic uncertainties associated to the $t\bar{t}H$ and $t\bar{t} + b\bar{b}$ modelling.

The next steps for future $t\bar{t}H(H \to b\bar{b})$ analyses in ATLAS may involve the following:

- change the current physics modelling strategies to the new MC setups mentioned above;
- include the remaining data collected during Run 2 of the LHC, namely from the years 2017 and 2018;
- include new recommendations related to the flavour tagging of jets and reconstructed objects, especially the jet energy scale and resolution, in order to reduce the corresponding systematic uncertainties;
- obtain a better understanding of how to correlate the modelling uncertainties entering b-tagging and on the $t\bar{t}$ background in the $t\bar{t}H(H \to b\bar{b})$ analysis; and
- extract the total as well as the fiducial $t\bar{t}H(H \to b\bar{b})$ cross-section and combine it with the search for the $VH(H \to b\bar{b})$ process in ATLAS to obtain a measurement of the $H \to b\bar{b}$ decay branching ratio.

From these points, the gain from implementing the new b-tagging recommendations presented in Chapter 6 should be addressed in particular. This calibration analysis is not a search for new physics or a rare SM process, but instead tries to optimise our understanding and the performance of the detector and physics modelling efforts. In the future, this analysis plans to:

- calibrate the b-tagging algorithm on data collected with the ATLAS detector during 2018;
- calibrate the new tagging algorithms which are trained using a new impact parameter based sub-algorithm that is trained with a recurrent neural network [68]; and
- include the newest recommendations regarding detector performance and physics modelling in order to reduce the associated systematic uncertainties.

In addition to this, new strategies to calibrate the b-tagging algorithms in ATLAS have been performed that use other methods to extract the b-tagging efficiency. These involve the tag & probe method, or analyses that do not select dileptonic $t\bar{t}$ events, but, for example, semileptonic $t\bar{t}$ decays or events containing two jets with a high p_{T} and a muon. The cross-check from another efficiency extraction method is a gain in confidence and potentially reduces the overall uncertainty associated to the b-tagging efficiency, but selecting non-$t\bar{t}$ events during a calibration might be of greater interest to many analyses. This reason is that $t\bar{t}$ events are studied in most SM and new physics analyses in ATLAS and thus, these efforts typically require b-tagging. However, these analyses most often select or include $t\bar{t}$ events that are used for the calibration of the b-tagging algorithms, hence there may be a correlation between the systematic uncertainties of both efforts, depending on the event selection. Using a calibration not performed on $t\bar{t}$ events would

thus be of interest to $t\bar{t}$ analyses sensitive to b-tagging related uncertainties. The same holds for mis-tagging of c-jets or light-jets, where the former mis-tag efficiency is also calibrated in $t\bar{t}$ events. Finally, an effort to develop a dedicated c-tagging algorithm is highly motivated, despite being challenging given the experimental conditions at the LHC. Many ATLAS analyses are investigating c-jets in SM precision measurements or try, for example, to reconstruct possible resonance decays of new BSM particles. These analyses, however, are dominated by events containing multiple QCD-induced light-jets and $t\bar{t}$ decays or similar processes producing a significant amount of b-jets. All of these cases impede a precise identification of c-jets at the LHC.

Given all of these points, it is clear that the observation of the $t\bar{t}H$ process is not the conclusion of the LHC physics programme. There are many points on which the $t\bar{t}H(H \rightarrow b\bar{b})$ search as well as the b-tagging calibration analysis can increase their performance and both efforts can learn from each other. In general, analyses can always be improved, uncertainties can be reduced and estimated in more sophisticated ways, and more data can be collected to facilitate the observation of rare physics processes or even new physics beyond the SM. The current SM is known to be incomplete and unable to explain everything we observe in nature. And it probably never will be. Nonetheless, trying to expand our knowledge of our Universe is essential for the future of humankind. Therefore, motivation to continue research, to understand the underlying mechanisms of the cosmos we live in, will and must always exist.

13.2 Danksagung

Ich bedanke mich herzlich bei meinem Doktorvater Prof. Dr. Arnulf Quadt für seine maßgebliche Unterstützung, nicht nur auf wissenschaftlicher Ebene, sondern noch viel mehr auf persönlicher Ebene! Deine Unterstützung hinsichtlich meiner geistigen, sowie insbesondere körperlichen Veränderungen über die Jahre in deiner Arbeitsgruppe und deren Konsequenzen haben mich immer wieder zu tiefst überrascht und beeindruckt. Dein grenzenloses Wohlwollen für deine Mitarbeiter und Studenten, mich eingeschlossen, werde ich immer in Erinnerung behalten und mir als Ziel für mein eigenes Leben setzen.
Auch bedanke ich mich herzlich bei meiner Betreuerin Dr. Elizaveta Shabalina für ihre unerschöpfliche Geduld, sowie ihr grenzenloses Wissen und ihre Fähigkeit, mich dazu zu motivieren, über mich hinaus zu wachsen. Deine direkte, aber stets lieb gemeinte Form der Kritik ist genau das, was uns junge Physiker zum Nachdenken anregt und voran treibt.
Ein großes Danke gilt auch meiner ehemaligen Betreuerin, Dr. Maria Moreno Llacer, die mich bereits seit meinem Masterstudium begleitet hat. Du hast deine Zeit stets für alle deine Studenten und Kollegen geopfert, um sie bestmöglichst zu unterstützen. So standest du auch mir dauerhaft zur Seite und hast mir viele Möglichkeiten offenbart, an spannenden Projekten teilzunehmen und Kontakte zu knüpfen. Unsere Zusammenarbeit hat mir Freude bereitet und mein Leben bereichert.
Selbstverständlich bedanke ich mich bei der ganzen ATLAS HTop-Gruppe. Es ist ein Geschenk mit einer so großen und diversen Truppe von Genies zusammenarbeiten zu dürfen, in der das Individuum nicht vergessen wird. Trotz der vielen Deadlines und schlaflosen Nächte, oder vielleicht gerade deshalb, ist sowohl die Hilfsbereitschaft als auch der Humor nie auf der Strecke geblieben.
Ein ganz besonderer Dank geht an das wundervolle, kleine und gemütliche $t\bar{t}$ PDF b-tagging Team am DESY. Ich habe einen sehr engen Kontakt zu euch aufbauen dürfen, der über das von Kollegen hinaus geht. Ich habe unglaublich viel von euch gelernt und zwar nicht nur im Bereich der Physik. Es war und ist mir eine Freude und Ehre mit euch zusammen arbeiten zu dürfen und ich hoffe, dass ihr mir meine häufigen Patzer sowie gelegentlichen inaktiveren Phasen während stressiger Zeiten verzeiht, die den engen Zeitplänen meines Studiums zu verschulden waren. Julian, Matthias und Chris, ich wünsche euch das allerbeste im Leben; hoffentlich sehen wir uns auch außerhalb der Physik in Zukunft wieder.
Darüber hinaus bedanke ich mich bei all meinen ehemaligen und jetzigen Kollegen des zweiten physikalischen Instituts in Göttingen. Ihr seid eine Truppe an wirklich vielseitigen, cleveren, witzigen und lieben Menschen. Auch wenn ich mich oft zurückgezogen habe und nicht immer klar ausgedrückt habe, was in mir vor geht, so möchte ich jetzt sagen, dass ich euch alle lieb gewonnen habe und sehr schätze. Ich habe euch viel zu verdanken, habe viel von euch gelernt, habe mit euch gelacht, mit euch Unfug angestiftet, mit euch zusammen gemeckert und bin insgesamt mit euch gewachsen. Wir werden stets in Kontakt bleiben. Ich sage nur: “five minutes outside?”.
Ein kleiner Dank geht natürlich auch an Andrea Knue. Ich kann es kaum glauben, dass

du meine ständigen, dummen Fragen nicht nur ertragen konntest, sondern auch hilfsbereit beantwortet hast. Mich mit dir und anderen Atlanten auf Dienstreisen wieder zu treffen war stets eines der großen Highlights. Die nächste Runde geht auf mich!
Außerdem möchte ich mich bei der gesamten ATLAS Kollaboration und all ihren Mitgliedern bedanken. Es war spannend, Teil einer so großen und bedeutsamen Gemeinschaft dieser Welt zu sein und dafür arbeiten zu dürfen. Trotz der Höhen und Tiefen, die ich während meiner Promotion hatte, bin ich doch froh über die vielen Möglichkeiten, die sich mir durch die Arbeit innerhalb ATLAS boten, sowie die Menschen, die ich kennenlernen durfte. Meine Arbeitszeit innerhalb ATLAS ist eine großartige und wichtige Lebenserfahrung.
Besonders im letzten Jahr meiner Promotion muss ich einen besonderen Dank an meine inzwischen zahlreichen neu gefundenen Freunde der veganen Gemeinschaft in Göttingen aussprechen. Ihr habt mir so unendlich viel auf meiner persönlichen Reise in diese wundervolle und erfüllende Lebensweise gegeben. Durch euch habe ich neue Perspektiven, Eindrücke und Sichtweisen über viele Themen dieser Welt gewinnen dürfen. Ich habe gelernt, wie wichtig und gleichzeitig erfüllend es ist, sich gemeinsam aktiv gegen die Ungerechtigkeiten und Grausamkeiten dieser Welt einzusetzen. Dass es mehr da draußen gibt, als jede Woche 40+ Stunden am Schreibtisch zu verbringen. Und ganz besonders, dass jeder einzelne von uns einen individuellen Einfluss auf diese Welt hat und jede unserer täglichen Entscheidungen zählt. Fühlt euch alle warm und herzlich umarmt.
Schließlich bedanke ich mich bei meiner Familie und meinen Freunden aus Nordhorn. Danke, dass ihr mich während meiner Promotion begleitet und ertragen habt. Danke dafür, dass, wenn ich ungefragt mit meinen dummen Problemen und Sorgen zu euch kam, ihr nie abweisend oder genervt reagiert habt, sondern mir stets ein offenes Ohr geschenkt habt und mit Mitgefühl sowie Ratschlägen zur Seite gestanden seid. Ich liebe euch!
Ich hoffe dass Du, der das hier liest, dir bewusst bist, dass du mich gerne zu jedem Thema aufsuchen und es ansprechen kannst. Jederzeit. Jeder von uns hat Sorgen, aber wir sind nicht alleine auf dieser Welt und müssen uns das Leben nicht unnötig schwerer machen. Ich möchte Dir meine alles von mir anbieten und dieser Welt zurück geben, was ich von ihr erfahren durfte.

Peace + Plants.

Bibliography

[1] The ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012) 1–29.

[2] The CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30–61.

[3] Particle Data Group, *Review of Particle Physics*, Phys. Rev. **D98** (2018) 030001.

[4] The ATLAS Collaboration, *Search for the standard model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, Phys. Rev. **D97** (2018) 072016.

[5] J. J. Thomson, *XL. Cathode Rays*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **44** (1897) 293–316.

[6] D. Griffiths, *Introduction to Elementary Particles*. WILEY-VCH Verlag, 2008.

[7] M. Thomson, *Modern Particle Physics*. Cambridge University Press, 2013.

[8] F. Halzen and A. D. Martin, *Quarks & Leptons: An Introductory Course in Modern Particle Physics*. John Wiley & Sons, 1984.

[9] R. K. Ellis, W. J. Stirling, and B. R. Webber, *QCD and Collider Physics*. Cambridge University Press, 2003.

[10] The ALEPH Collaboration, The DELPHI Collaboration, The L3 Collaboration, The OPAL Collaboration, The LEP Electroweak Working Group, *Precision electroweak measurements on the Z resonance*, Phys. Rep. **427** (2006) 257 – 454.

[11] The ALEPH Collaboration, The DELPHI Collaboration, The L3 Collaboration, The OPAL Collaboration, The LEP Electroweak Working Group, *Electroweak*

measurements in electronpositron collisions at W-boson-pair energies at LEP, Phys. Rep. **532** (2013) 119 – 244.

[12] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (1964) 321–323.

[13] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (1964) 508–509.

[14] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (1964) 585–587.

[15] The LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, CERN Yellow Reports **2** (2017).

[16] M. Kobayashi and T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, PTP **49** (1973) 652–657.

[17] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10** (1963) 531–533.

[18] The DØ Collaboration, *Observation of the top quark*, Phys. Rev. Lett. **74** (1995) 2632–2637.

[19] The CDF Collaboration, *Observation of top quark production in $\bar{p}p$ collisions*, Phys. Rev. Lett. **74** (1995) 2626–2631.

[20] The ATLAS Collaboration, *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, Phys. Lett. **B784** (2018) 173 – 191.

[21] The ATLAS Collaboration and The CMS Collaboration, *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV*, JHEP **08** (2016) 45.

[22] A. Bredenstein et al., *NLO QCD Corrections to Top Anti-Top Bottom Anti-Bottom Production at the LHC: 2. full hadronic results*, JHEP **1003** (2010) 021.

[23] S. Myers, *The LEP Collider, from design to approval and commissioning*. John Adams' Lecture. CERN, Geneva, 1991. Delivered at CERN, 26 Nov 1990.

[24] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[25] W. Herr and B. Muratori, *Concept of luminosity*, 10.5170/CERN-2006-002.361 (2006).

[26] The ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[27] The CMS Collaboration, *The CMS Experiment at the CERN LHC*, JINST **3** (2008) S08004.

[28] The ATLAS Collaboration, *ATLAS detector and physics performance: Technical Design Report, 1.* Technical Design Report ATLAS. CERN, Geneva, 1999.

[29] The ATLAS Collaboration, *ATLAS detector and physics performance: Technical Design Report, 2.* Technical Design Report ATLAS. CERN, Geneva, 1999.

[30] The ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, Tech. Rep. CERN-LHCC-2010-013. ATLAS-TDR-19, Sep, 2010.

[31] The ATLAS Collaboration, *ATLAS IBL Pixel Upgrade*, Nucl. Phys. B, Proc. Suppl. **215** (2011) 147–150.

[32] The ATLAS Collaboration, *Improved luminosity determination in pp collisions at* $\sqrt{s} = 7$ *TeV using the ATLAS detector at the LHC*, Eur. Phys. J. **C73** (2013) 2518.

[33] The ATLAS Collaboration, *Luminosity determination in pp collisions at* $\sqrt{s} = 8$ *TeV using the ATLAS detector at the LHC*, Eur. Phys. J. **C76** (2016) 653.

[34] The ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2010*, Eur. Phys. J. **C72** (2012) 1849.

[35] The ATLAS Collaboration, *The Run-2 ATLAS Trigger System*, J. Phys.: Conf. Ser. **762** (2016) 012003.

[36] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP **0902** (2009) 007.

[37] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nuclear Physics B **126** (1977) 298 – 318.

[38] A. D. Martin et al., *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189–285.

[39] M. Bahr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. **C58** (2008) 639–707.

[40] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. **C76** (2016) 196.

[41] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026.

[42] T. Sjöstrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852–867.

[43] S. Agostinelli et al., *Geant4 - a simulation toolkit*, Nucl. Instrum. Methods Phys. Res., Sect. A **506** (2003) 250 – 303.

[44] The ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. **C70** (2010) 823–874.

[45] The ATLAS Collaboration, *Luminosity Public Results*, `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2`.

[46] The ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions with the ATLAS detector*, Nucl. Instrum. Methods Phys. Res., Sect. A **824** (2016) 367 – 370, Frontier Detectors for Frontier Physics: Proceedings of the 13th Pisa Meeting on Advanced Detectors.

[47] G. Watt and R. S. Thorne, *Study of Monte Carlo approach to experimental uncertainty propagation with MSTW 2008 PDFs*, JHEP **08** (2012) 52.

[48] The ATLAS Collaboration, *ATLAS Run 1 Pythia8 tunes*, Tech. Rep. ATL-PHYS-PUB-2014-021, CERN, Geneva, Nov, 2014.

[49] The ATLAS Collaboration, *Simulation of Pile-up in the ATLAS Experiment*, Tech. Rep. ATL-SOFT-PROC-2013-030, CERN, Geneva, Oct, 2013.

[50] The ATLAS Collaboration, *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2011 LHC proton–proton collision data*, Eur. Phys. J. **C74** (2014) 2941.

[51] The ATLAS Collaboration, *Track Reconstruction Performance of the ATLAS Inner Detector at* $\sqrt{s} = 13$ *TeV*, Tech. Rep. ATL-PHYS-PUB-2015-018, CERN, Geneva, Jul, 2015.

[52] The ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, Tech. Rep. ATLAS-CONF-2016-024, CERN, Geneva, Jun, 2016.

[53] The ATLAS Collaboration, *Measurement of the muon reconstruction performance of the ATLAS detector using 2011 and 2012 LHC proton–proton collision data*, Eur. Phys. J. **C74** (2014) 3130.

[54] The ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. **C76** (2016) 292.

[55] The ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. **C77** (2017) 490.

[56] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti-k(t) jet clustering algorithm*, JHEP **0804** (2008) 063.

[57] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet user manual*, Eur. Phys. J. **C72** (2012) 1896.

[58] The 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. **D96** (2017) 072002.

[59] The ATLAS Collaboration, *Properties of jets measured from tracks in proton-proton collisions at center-of-mass energy $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Rev. **D84** (2011) 054001.

[60] The ATLAS Collaboration, *Variable Radius, Exclusive-k_T, and Center-of-Mass Subjet Reconstruction for Higgs($\to b\bar{b}$) Tagging in ATLAS*, Tech. Rep. ATL-PHYS-PUB-2017-010, CERN, Geneva, Jun, 2017.

[61] The 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. **C76** (2016) 581.

[62] The ATLAS Collaboration, *Performance of algorithms that reconstruct missing transverse momentum in $\sqrt{s} = 8$ TeV proton–proton collisions in the ATLAS detector*, Eur. Phys. J. **C77** (2017) 241.

[63] The ATLAS Collaboration, *Performance of missing transverse momentum reconstruction for the ATLAS detector in the first proton-proton collisions at at $\sqrt{s}$= 13 TeV*, Tech. Rep. ATL-PHYS-PUB-2015-027, CERN, Geneva, Jul, 2015.

[64] The ATLAS Collaboration, *Performance of b-jet identification in the ATLAS experiment*, JINST **11** (2016) P04008.

[65] The ATLAS Collaboration, *Performance of the ATLAS Secondary Vertex b-tagging Algorithm in 7 TeV Collision Data*, Tech. Rep. ATLAS-CONF-2010-042, CERN, Geneva, Jul, 2010.

[66] The ATLAS Collaboration, *Expected performance of the ATLAS b-tagging algorithms in Run-2*, Tech. Rep. ATL-PHYS-PUB-2015-022, CERN, Geneva, Jul, 2015.

[67] The ATLAS Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, Tech. Rep. ATL-PHYS-PUB-2016-012, CERN, Geneva, Jun, 2016.

[68] The ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, Tech. Rep. ATL-PHYS-PUB-2017-013, CERN, Geneva, Jul, 2017.

[69] G. Piacquadio and C. Weiser, *A new inclusive secondary vertex algorithm for b-jet tagging in ATLAS*, J. Phys.: Conf. Ser. **119** (2008) 032032.

[70] The ATLAS Collaboration, *Flavour tagging algorithms in 2018*, Tech. Rep. ATL-COM-PHYS-2018-1377, CERN, Geneva, Sep, 2018.

[71] The ATLAS Collaboration, *Commissioning of the ATLAS high-performance b-tagging algorithms in the 7 TeV collision data*, Tech. Rep. ATLAS-CONF-2011-102, CERN, Geneva, Jul, 2011.

[72] Giuseppe Cerati and others, *Kalman Filter Tracking on Parallel Architectures*, J. Phys.: Conf. Ser. **664** (2015) 072008.

[73] A. Hoecker and others, *TMVA - Toolkit for Multivariate Data Analysis*, arXiv:0703039 [physics.data-an].

[74] T. G. Rizzo, *Z' phenomenology and the LHC*, pp. , 537–575. 2006. arXiv:hep-ph/0610104 [hep-ph].

[75] The ATLAS Collaboration, *Measurements of b-jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV*, JHEP **08** (2018) 089.

[76] The ATLAS Collaboration, *Measurement of the b-jet identification efficiency with $t\bar{t}$ events using an improved likelihood method*, Tech. Rep. ATL-COM-PHYS-2018-1072, CERN, Geneva, Sep, 2018.

[77] R. J. Barlow, *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences*. John Wiley & Sons, 1989.

[78] The ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, Eur. Phys. J. **C77** (2017) 317.

[79] The ATLAS Collaboration, *Selection of jets produced in 13 TeV proton-proton collisions with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2015-029, CERN, Geneva, Jul, 2015.

[80] The ATLAS Collaboration, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, Tech. Rep. ATL-PHYS-PUB-2010-013, CERN, Geneva, Oct, 2010.

[81] P. Nason, *A New Method for Combining NLO QCD with Shower Monte Carlo Algorithms*, JHEP **11** (2004) 040–040.

[82] S. Frixione, P. Nason, and C. Oleari, *Matching NLO QCD computations with Parton Shower simulations: the POWHEG method*, JHEP **11** (2007) 070.

[83] S. Alioli et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 43.

[84] J. M. Campbell et al., *Top-pair production and decay at NLO matched with parton showers*, JHEP **04** (2015) 114.

[85] J. Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. **G43** (2016) 023001.

[86] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Methods Phys. Res., Sect. A **462** (2001) 152 – 155, BEAUTY2000, Proceedings of the 7th Int. Conf. on B-Physics at Hadron Machines.

[87] The ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling for Top2016*, Tech. Rep. ATL-PHYS-PUB-2016-020, CERN, Geneva, Sep, 2016.

[88] 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.

[89] S. Frixione et al., *Single-top hadroproduction in association with a W boson*, JHEP **2008** (2008) 029.

[90] A. Giammanco, *Single top quark production at the LHC*, Rev. Phys. **1** (2016) 1–12.

[91] The NNPDF Collaboration, *Parton distributions for the LHC run II*, JHEP **04** (2015) 40.

[92] The ATLAS Collaboration, *Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **09** (2014) 145.

[93] J. Pumplin et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **07** (2002) 012.

[94] The ATLAS Collaboration Collaboration,, *Calibration of light-flavour jet b-tagging rates on ATLAS proton-proton collision data at $\sqrt{s} = 13$ TeV*, Tech. Rep. ATLAS-CONF-2018-006, CERN, Geneva, Apr, 2018.

[95] F. James and M. Winkler, *MINUIT User's Guide*, 2004.

[96] G. Bohm and G. Zech, *Introduction to Statistics and Data Analysis for Physicists; 3rd revised*. Verlag Deutsches Elektronen-Synchrotron, Hamburg, 2017.

[97] I. Jolliffe, *Principal component analysis*. Springer Verlag, New York, 2002.

[98] The ATLAS Collaboration, *Measurement of the inclusive jet cross-section in proton–proton collisions at $\sqrt{s} = 7$ TeV using 4.5 fb^{-1} of data with the ATLAS detector*, JHEP **02** (2015) 153.

[99] The ATLAS Collaboration, *Measurement of the inclusive jet cross-sections in proton–proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **09** (2017) 020.

[100] The ATLAS Collaboration, *Measurement of inclusive jet and dijet cross-sections in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **05** (2018) 195.

[101] The ATLAS Collaboration, *Measurement of b-tagging Efficiency of c-jets in $t\bar{t}$ Events Using a Likelihood Approach with the ATLAS Detector*, Tech. Rep. ATLAS-CONF-2018-001, CERN, Geneva, Mar, 2018.

[102] R. Raitio and W. W. Wada, *Higgs-boson production at large transverse momentum in quantum chromodynamics*, Phys. Rev. **D19** (1979) 941–944.

[103] W. Beenakker et al., *NLO QCD corrections to ttH production in hadron collisions*, Nucl. Phys. B **653** (2003) 151 – 203.

[104] S. Dawson et al., *Associated Higgs boson production with top quarks at the CERN Large Hadron Collider: NLO QCD corrections*, Phys. Rev. **D68** (2003) 034022.

[105] Y. Zhang et al., *QCD NLO and EW NLO corrections to ttH production with top quark decays at hadron collider*, Phys. Lett. **B738** (2014) 1 – 5.

[106] S. Frixione et al., *Electroweak and QCD corrections to top-pair hadroproduction in association with heavy bosons*, JHEP **06** (2015) 184.

[107] M. Cacciari et al., *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, Phys. Lett. **B710** (2012) 612 – 622.

[108] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*, JHEP **12** (2012) 54.

[109] M. Czakon and A. Mitov, *NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction*, JHEP **01** (2013) 80.

[110] M. Czakon, P. Fiedler, and A. Mitov, *Total Top-Quark Pair-Production Cross Section at Hadron Colliders Through $\mathcal{O}(\alpha_S^4)$*, Phys. Rev. Lett. **110** (2013) 252004.

[111] B. Nachman et al., *Jets from jets: re-clustering as a tool for large radius jet reconstruction and grooming at the LHC*, JHEP **02** (2015) 75.

[112] P. Artoisenet et al., *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, JHEP **03** (2013) 15.

[113] A. Djouadi, J. Kalinowski, and M. Spira, *HDECAY: a program for Higgs boson decays in the Standard Model and its supersymmetric extension*, Comput. Phys. Commun. **108** (1998) 56 – 74.

[114] 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 – 2938.

[115] The ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Eur. Phys. J. **C75** (2015) 349.

[116] F. Cascioli et al., *NLO matching for $t\bar{t}b\bar{b}$ production with massive b-quarks*, Phys. Lett. **B734** (2014) 210 – 214.

[117] F. Cascioli, P. Maierhöfer, and S. Pozzorini, *Scattering Amplitudes with Open Loops*, Phys. Rev. Lett. **108** (2012) 111601.

[118] M. Guzzi et al., *CT10 parton distributions and other developments in the global QCD analysis*, arXiv:1101.0561 [hep-ph].

[119] J. Gao et al., *CT10 next-to-next-to-leading order global analysis of QCD*, Phys. Rev. **D89** (2014) 033009.

[120] The ATLAS Collaboration, *Search for flavour-changing neutral current top quark decays $t \to Hq$ in pp collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, JHEP **12** (2015) 1–65.

[121] P. Z. Skands, *Tuning Monte Carlo Generators: The Perugia Tunes*, Phys. Rev. **D82** (2010) 074018.

[122] N. Kidonakis, *Two-loop soft anomalous dimensions for single top quark associated production with a W^- or H^-*, Phys. Rev. **D82** (2010) 054018.

[123] N. Kidonakis, *Next-to-next-to-leading logarithm resummation for s-channel single top quark production*, Phys. Rev. **D81** (2010) 054028.

[124] N. Kidonakis, *Next-to-next-to-leading-order collinear and soft gluon corrections for t-channel single top quark production*, Phys. Rev. **D83** (2011) 091503.

[125] The ATLAS Collaboration, *Estimation of non-prompt and fake lepton backgrounds in final states with top quarks produced in proton-proton collisions at $\sqrt{s} = 8$ TeV with the ATLAS detector*, Tech. Rep. ATLAS-CONF-2014-058, CERN, Geneva, Oct, 2014.

[126] The ATLAS Collaboration, *Muon Combined Performance in Run 2 (25 ns runs)*, Tech. Rep. ATL-COM-MUON-2015-093, CERN, Geneva, Nov, 2015.

[127] The ATLAS Collaboration, *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*, Tech. Rep. ATL-PHYS-PUB-2015-045, CERN, Geneva, Nov, 2015.

[128] M. R. Whalley, D. Bourilkov, and R. C. Group, *The Les Houches accord PDFs (LHAPDF) and LHAGLUE*, pp. , 575–581. 2005. arXiv:hep-ph/0508110 [hep-ph].

[129] G. P. Lepage, *A new algorithm for adaptive multidimensional integration*, J. Comput. Phys. **27** (1978) 192 – 203.

[130] D. Schouten, A. DeAbreu, and B. Stelzer, *Accelerated matrix element method with parallel computing*, Comput. Phys. Commun. **192** (2015) 54 – 59.

[131] V. Barger, J. Ohnemus, and R. J. N. Phillips, *Event shape criteria for single-lepton top-quark signals*, Phys. Rev. **D48** (1993) R3953–R3956.

[132] G. C. Fox and S. Wolfram, *Observables for the Analysis of Event Shapes in e^+e^- Annihilation and Other Processes*, Phys. Rev. Lett. **41** (1978) 1581–1585.

[133] G. Cowan et al., *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. **C71** (2011) 1554.

[134] A. L. Read, *Presentation of search results: the CL_s technique*, J. Phys. G: Nucl. Part. Phys. **28** (2002) 2693.

[135] T. Junk, *Confidence level computation for combining searches with small statistics*, Nucl. Instrum. Methods Phys. Res., Sect. A **434** (1999) 435 – 443.

[136] The 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.

[137] The ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling with Sherpa and MG5_aMC@NLO*, Tech. Rep. ATL-PHYS-PUB-2017-007, CERN, Geneva, May, 2017.

[138] The ATLAS Collaboration, *Studies of tt+cc production with MadGraph5_aMC@NLO and Herwig++ for the ATLAS experiment*, Tech. Rep. ATL-PHYS-PUB-2016-011, CERN, Geneva, May, 2016.

[139] A. D. Martin et al., *Parton distributions for the LHC*, Eur. Phys. J. **C63** (2009) 189–285.

[140] The ATLAS Collaboration, *Multi-Boson Simulation for 13 TeV ATLAS Analyses*, Tech. Rep. ATL-PHYS-PUB-2016-002, CERN, Geneva, Jan, 2016.

[141] J. M. Campbell and R. K. Ellis, *$t\bar{t}W^{\pm}$ production and decay at NLO*, JHEP **07** (2012) 52.

[142] N. A. Asbah and J. Katzy, *Search for the Production of a Standard Model Higgs Boson in Association with Top-Quarks and Decaying into a Pair of Bottom-Quarks with 13 TeV ATLAS Data*, 2018. https://cds.cern.ch/record/2320703. Presented 23 May 2018.

[143] T. Jezo et al., *New NLOPS predictions for $\boldsymbol{t\bar{t}+b}$ -jet production at the LHC*, Eur. Phys. J. **C78** (2018) 502.

[144] The ATLAS Collaboration, *Measurements of fiducial and differential 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*, arXiv:1811.12113 [hep-ex].

[145] A. Buckley et al., *Rivet user manual*, Comput. Phys. Commun. **184** (2013) 2803–2819.

[146] N. D. Gagunashvili, *Pearson's chi-square test modifications for comparison of unweighted and weighted histograms and two weighted histograms*, PoS **ACAT** (2007) 060.

[147] The ATLAS Collaboration, *Measurements of Higgs boson properties in the diphoton decay channel with 36 fb^{-1} of pp collision data at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. **D98** (2018) 052005.

[148] The ATLAS Collaboration, *Measurement of the Higgs boson coupling properties in the $H \to ZZ^* \to 4l$ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **03** (2018) 95.

[149] The CMS Collaboration, *Observation of $t\bar{t}H$ Production*, Phys. Rev. Lett. **120** (2018) 231801.

[150] The ATLAS Collaboration, The CMS Collaboration and The LHC Higgs Combination Group, *Procedure for the LHC Higgs boson search combination in Summer 2011*, Tech. Rep. CMS-NOTE-2011-005. ATL-PHYS-PUB-2011-11, CERN, Geneva, Aug, 2011.

[151] The CMS Collaboration, *Search for $\mathrm{t\bar{t}}H$ production in the $H \to \mathrm{b\bar{b}}$ decay channel with leptonic $\mathrm{t\bar{t}}$ decays in proton-proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **2019** (2019) 26.

[152] The CMS Collaboration, *CMS Luminosity Measurements for the 2016 Data Taking Period*, Tech. Rep. CMS-PAS-LUM-17-001, CERN, Geneva, 2017.

[153] The CMS Collaboration, *Event generator tunes obtained from underlying event and multiparton scattering measurements*, Eur. Phys. J. **C76** (2016) 155.

[154] P. Skands, S. Carrazza, and J. Rojo, *Tuning PYTHIA 8.1: the Monash 2013 tune*, Eur. Phys. J. **C74** (2014) 3024.

[155] The CMS Collaboration, *Measurements of $t\bar{t}$ cross sections in association with b jets and inclusive jets and their ratio using dilepton final states in pp collisions at $\sqrt{s} = 13$ TeV*, Phys. Lett. **B776** (2018) 355 – 378.

[156] G. Bevilacqua, M. V. Garzelli, and A. Kardos, *$t\bar{t}b\bar{b}$ hadroproduction with massive bottom quarks with PowHel*, arXiv:1709.06915 [hep-ph].

Appendices

APPENDIX A

Estimation of non-prompt leptons in the b-tagging calibration

This chapter describes the method to estimate the background contribution from non-prompt leptons, i.e. those coming from the decay of τ-leptons or jets, and fake leptons which are, for example, photons or jets mis-identified as charged leptons. In the following, such leptons are labelled NPL.

The online reconstruction of the charged leptons was discussed in Section 6.3.1. Both, electrons and muons, must pass TIGHT identification and isolation requirements [52,54]. Consequently, the number of NPLs and thus the overall background contribution from them is assumed to be very small. This is confirmed by the low fraction of simulated events that contain one (two) NPL(s) which amounts to 0.3% ($< 0.001\%$). The downside is that this background is expected to not be well modelled in simulation. Hence, it is estimated using a data-driven technique [76].

The events with at least one NPL consist to roughly 80% of events with a prompt muon and a non-prompt electron which is labelled as 1NPel from now on. The contribution from events with two non-prompt leptons (2NPL) is compatible with zero and the contribution from events with a prompt electron and non-prompt muon (1NPmu) is considered insignificant as well. Therefore, both are extracted from simulation and the following data-driven NPL background estimation is focusing exclusively on the 1NPel component.

Most of the 1NPel events are believed to contain an electron which comes from the decay of a heavy flavour hadron or a τ-lepton. Such events are, to a large extent, modelled by physics processes with exactly one charged lepton in the ME such as a semileptonic $t\bar{t}$ decay. To estimate the 1NPel background, the selection of two oppositely charged leptons (OS) is changed such that two equally charged leptons (SS) are selected instead. The 1NPel contribution, now considered signal, is expected to be independent of the

lepton charge, while events without NPLs (0NPL), now considered background, are discarded to a very large extent by this selection. This increases the purity of the 1NPel contribution by a factor of 40, namely from 0.3% to about 12%.

The simulated SS events consist to about 45% of Diboson events, 40% $t\bar{t}$ events and 15% single top, Z+jets and W+jets, all roughly in equal amounts. Additional sources of non-prompt leptons, such as QCD-induced multijet processes, are not considered. Their contribution to events with two reconstructed TIGHT leptons is expected to be small, but it cannot be estimated properly given the limited available MC statistics. On the other hand, leaving out contributing processes only increases the final non-prompt lepton uncertainty which makes the current estimation conservative.
In the SS region, the 1NPmu component is compatible with zero within the available MC statistics which further justifies the exclusive focus on the 1NPel contribution. On the other hand, the 0NPL contribution is still significant, namely about 88% of all SS events. These events originate, to a large extent, from Diboson production where at least one lepton is lost or from dileptonic $t\bar{t}$, $Z \to \tau\tau$ and Wt events where the electron is reconstructed with the wrong charge. This latter case is referred to as charge flip. Figure A.1 shows the p_{T} and η distributions of the SS events containing a non-prompt electron. The effects contributing to the charge flip are enhanced at high values of η, which is confirmed by the bottom left plot in Figure A.1 in which the fraction of $t\bar{t}$ events increases with higher η. In addition to this, one can see the large fraction of ll events which result primarily from Diboson production, since there are no b-jets expected to be present during the hard interaction. One can also see that the data/MC ratio, which represents the agreement between data and simulation, improves for increasing non-prompt electron p_{T}. This can be explained by the 1NPel contribution decreasing with higher non-prompt electron p_{T} as such leptons are usually expected only at low p_{T}.

As the simulated 0NPL component including charge flipped electrons is expected to be well modelled, it is subtracted from both data and simulation. This assumption is based on the overall good agreement between data and simulation observed in the main analysis which contains all $e\mu$+2j events and is dominated by 0NPL events. After the subtraction, only the 1NPel component is assumed to survive and can thus be extracted. The resulting non-prompt electron p_{T} and η distributions are shown in Figure A.2.

The data/MC ratio now ranges from ≈ 3 in the first p_{T} bin ($28 - 150$ GeV) to ≈ 1 in the other p_{T} bins ($150 - 300$ GeV and ≥ 300 GeV) which are statistically limited. The data/MC ratio is slightly increased for small values of η which is attributable to fluctuations from limited MC statistics.

In the final step, the data/MC ratio of the non-prompt electron p_{T} distribution in Figure A.2 is propagated from the SS region to the OS i.e. signal region of the main analysis. This is done by applying this ratio as a SF to the MC weight of events with a prompt muon and a non-prompt electron, depending on the electron p_{T}. This propagation is

reasonable, because the 1NPel component should be independent of the electron charge, as discussed before.

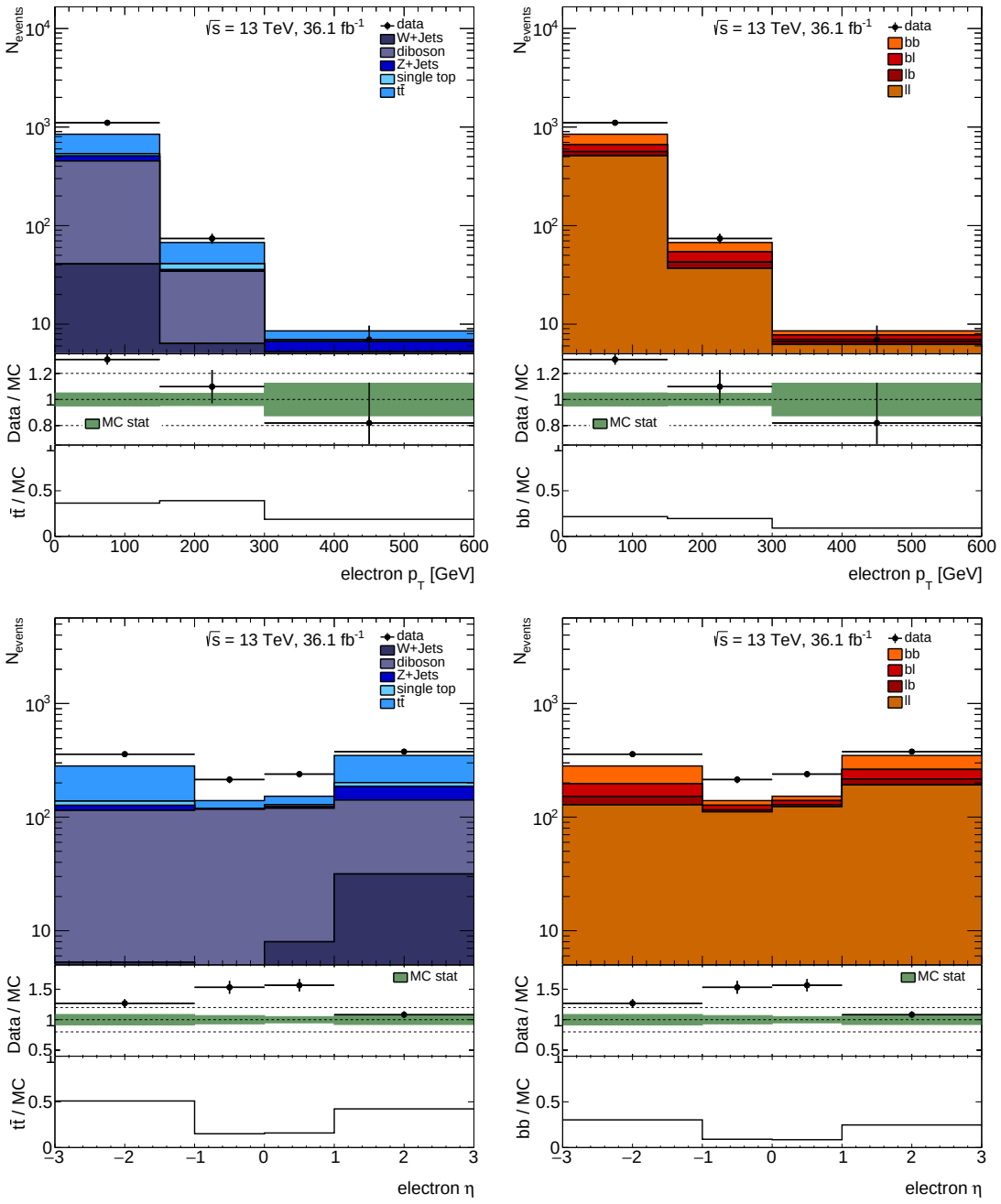

Figure A.1: The predicted distributions of the non-prompt electron p_{T} (top) and η (bottom) compared to ATLAS data from 2015 and 2016 in the SS region. The MC prediction is split either by the contributing physics processes (left) or the contributing two jet flavour compositions (right). Correspondingly, the ratio plots show, in addition to the common data/MC ratio, either the amount of $t\bar{t}$ events (left) or the amount of $b\bar{b}$ events (right) compared to all simulated events. Only statistical uncertainties from data and simulation are included in this comparison.

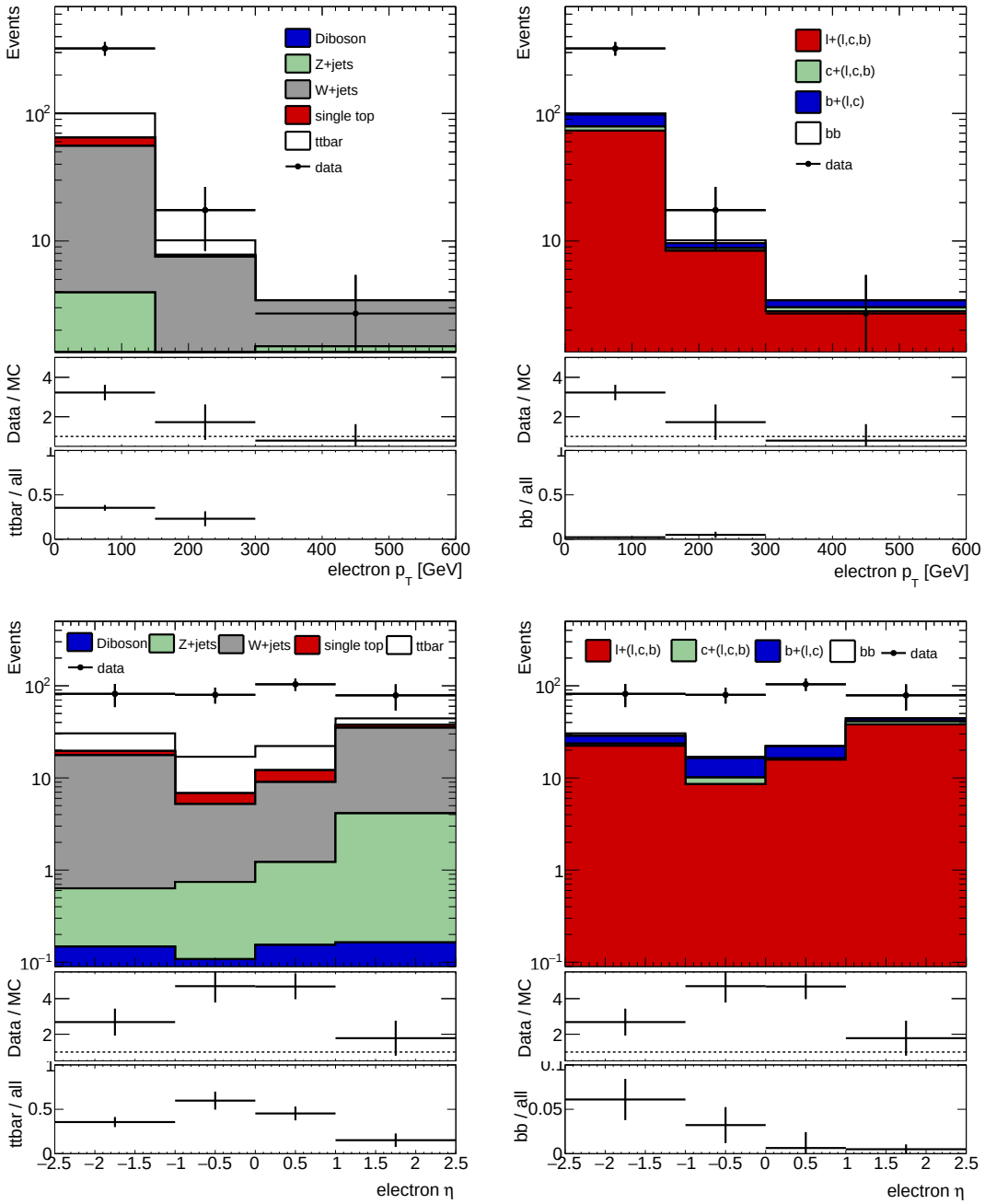

Figure A.2: The predicted distributions of the non-prompt electron p_{T} (top) and η (bottom) compared to ATLAS data from 2015 and 2016 in the SS region. Here, the simulated 0NPL component has been subtracted from both MC simulation and data. The remaining 1NPel prediction is split either by the contributing physics processes (left) or the contributing two jet flavour compositions (right). Correspondingly, the ratio plots show, in addition to the common data/MC ratio, either the amount of $t\bar{t}$ events (left) or the amount of $b\bar{b}$ events (right) compared to all simulated NPel events. Only statistical uncertainties from data and simulation are included in this comparison.

APPENDIX B

Estimation of the goodness of fit

The PDF fit method employs a procedure which simultaneously minimises 180 different LLH functions per WP, namely one for each $(p_{\mathrm{T},1}, p_{\mathrm{T},2})$ bin k for the SR and three CRs. In order to assess the agreement between data and simulation before the fit as well as the goodness-of-fit (GOF), a global χ^2 variable is constructed. It is defined as the sum of the Pearson's estimators $\chi^2_{k,\mathrm{SR}}, \chi^2_{k,\mathrm{CR}}$, which for each k bin are given by [146]:

$$\chi^2_{k,\mathrm{SR}} = \sum_{w_1,w_2} \frac{\left(n^{\mathrm{SR}}_{w_1,w_2} - \nu^{\mathrm{SR}}_{w_1,w_2}\right)^2}{\nu^{\mathrm{SR}}_{w_1,w_2}}, \chi^2_{k,\mathrm{CR}} = \sum_{CR} \frac{\left(n^{\mathrm{CR}} - \nu^{\mathrm{CR}}\right)^2}{\nu^{\mathrm{CR}}} \Rightarrow \chi^2 = \sum_k \left(\chi^2_{k,\mathrm{SR}} + \chi^2_{k,\mathrm{CR}}\right). \tag{B.1}$$

The various quantities in Equation B.1 are based on the naming conventions introduced in Section 6.2. This χ^2 estimator only considers the statistical uncertainty of the input data. It does not take into account the statistical uncertainty of simulated samples or any systematic uncertainties. Nonetheless, this estimator can be used to either compare different fit configurations or to evaluate the GOF for negligible systematic uncertainties [76].

The number of degrees of freedom (dof) of the simultaneous fit of all LLH functions is given by the sum of the dof of each LLH function in Section 6.2 minus the number of parameters extracted from the fit which are the POIs and nuisance parameters. This can be written as follows [76]:

$$N_{\mathrm{dof}} = N_k \cdot N_w^2 + N_k \cdot N_{\mathrm{CR}} - N_k \cdot N_{f_1,f_2} - N_{p_{\mathrm{T}}} \cdot (N_w - 1), \tag{B.2}$$

where $N_w = 5$ is the number of pseudo-continuous b-tagging bins, $N_{p_{\mathrm{T}}} = 9$ is the number of jet p_{T} bins, $N_k = N_{p_{\mathrm{T}}} \cdot (N_{p_{\mathrm{T}}} + 1)/2 = 45$ is the number of $(p_{\mathrm{T},1}, p_{\mathrm{T},2})$ bins, $N_{\mathrm{CR}} = 3$

is the number of CRs and $N_{f_1,f_2} = 4$ is the number of correction factors for the two jet flavour configurations, namely bb, bl, lb and ll. Inserting these values into the above equation yields $N_{\text{dof}} = 1044$ (1260) degrees of freedom after (before) the fit.

The global χ^2 defined in Equation B.1 is expected to follow a $\chi^2_{N_{\text{dof}}}$ distribution with a mean value of N_{dof} as long as only statistical uncertainties from the input data are considered. Now, to assess the agreement between data and simulation before the fit as well as the GOF, one can compute χ^2/N_{dof}, which should be around unity in case of a good agreement. In this calibration, it is 0.954 (1.12) after (before) the fit, showing not only that the agreement improves significantly after the fit, but also that the agreement overall is very good since the value is close to unity. Figure B.1 shows how the total χ^2 is distributed across all $(p_{\text{T},1}, p_{\text{T},2})$ bins. This quantity is useful to further study which p_{T} bins show a considerable disagreement between data and MC and are thus potentially problematic. In mathematical terms, this fractional χ^2 can be expressed in the following way:

$$\left(\%\chi^2\right)_k = \frac{\chi^2_{k,\text{SR}} + \chi^2_{k,\text{CR}}}{\chi^2}. \tag{B.3}$$

Since there are $N_k = 45$ different p_{T} bins, each bin should have a fraction of $1/45 \approx 0.022$ in case the χ^2 estimator is distributed evenly across all bins, meaning the disagreement between data and simulation is uniform across the bins. This, however, is not observed in all cases. Before the fit, there are five bins with a significantly large χ^2 fraction above 0.03, for example $(p_{\text{T},1}, p_{\text{T},2}) = (40 - 60\ \text{GeV}, 30 - 40\ \text{GeV})$. On the other hand, the fractions range from 0.014 to 0.037 after the fit, while only three bins have a fraction above 0.03, showing again that the fit behaves as expected and the GOF is a good estimator of the fit performance.

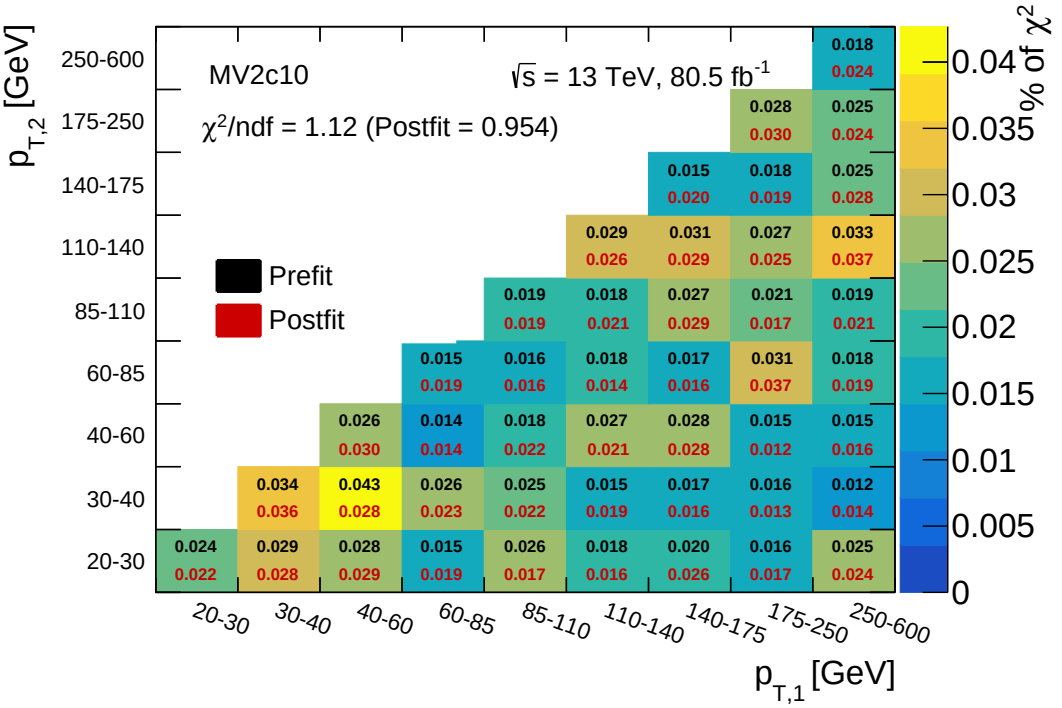

Figure B.1: The fraction of χ^2 as a function of the leading and subleading jet p_{T} is shown before (black) and after the fit (red) to data taken during the years 2015, 2016 and 2017 with the ATLAS detector. If the disagreement between data and simulation is uniform across all two dimensional p_{T} bins, the χ^2 fraction should be distributed evenly with an average value of 0.022 [76].

APPENDIX C

Results of the DL1 calibration

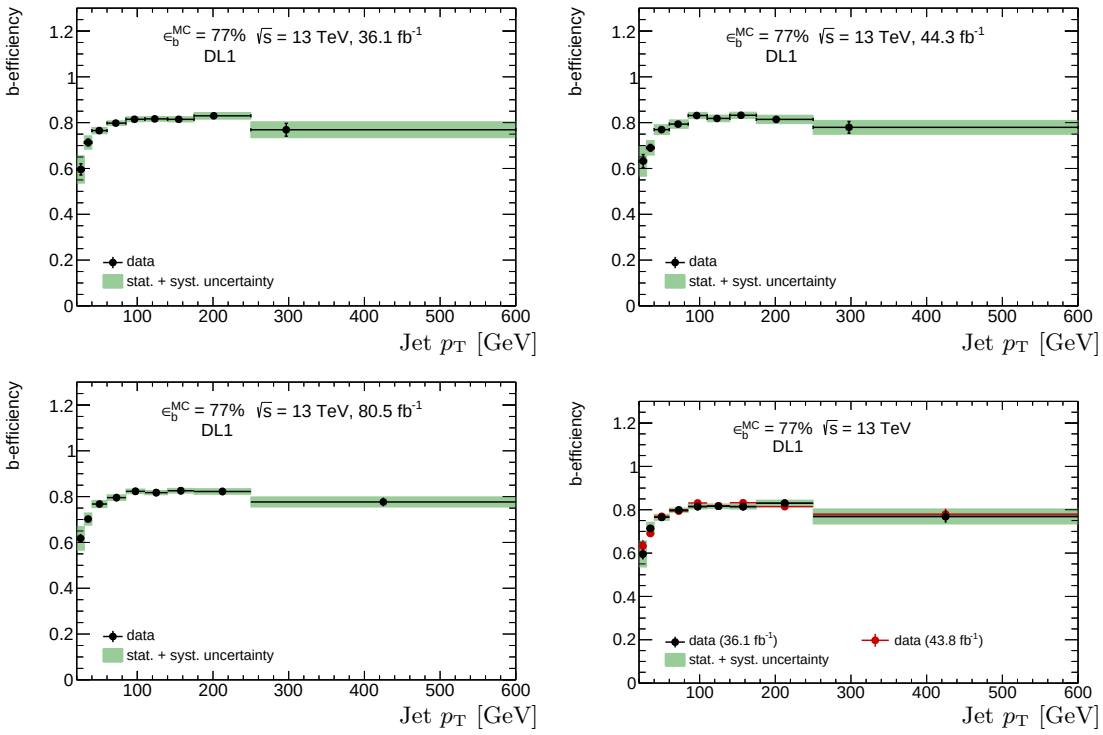

Figure C.1: The b-tagging efficiency as a function of the jet p_T extracted from the LH fit method to data for the DL1 algorithm at the 77% WP. The fit result is shown based on ATLAS data from the years 2015 and 2016 (top left), 2017 (top right) and all three years combined (bottom left). In the bottom right plot, a comparison of the 2015 and 2016 results (black) to the 2017 result (red) is depicted. The individual fit results (top left and top right) show the bin centres at the mean jet p_T value of each bin. The vertical error bars include only the statistical uncertainties from data, while the green error band represents the sum in quadrature of the statistical and systematic uncertainties.

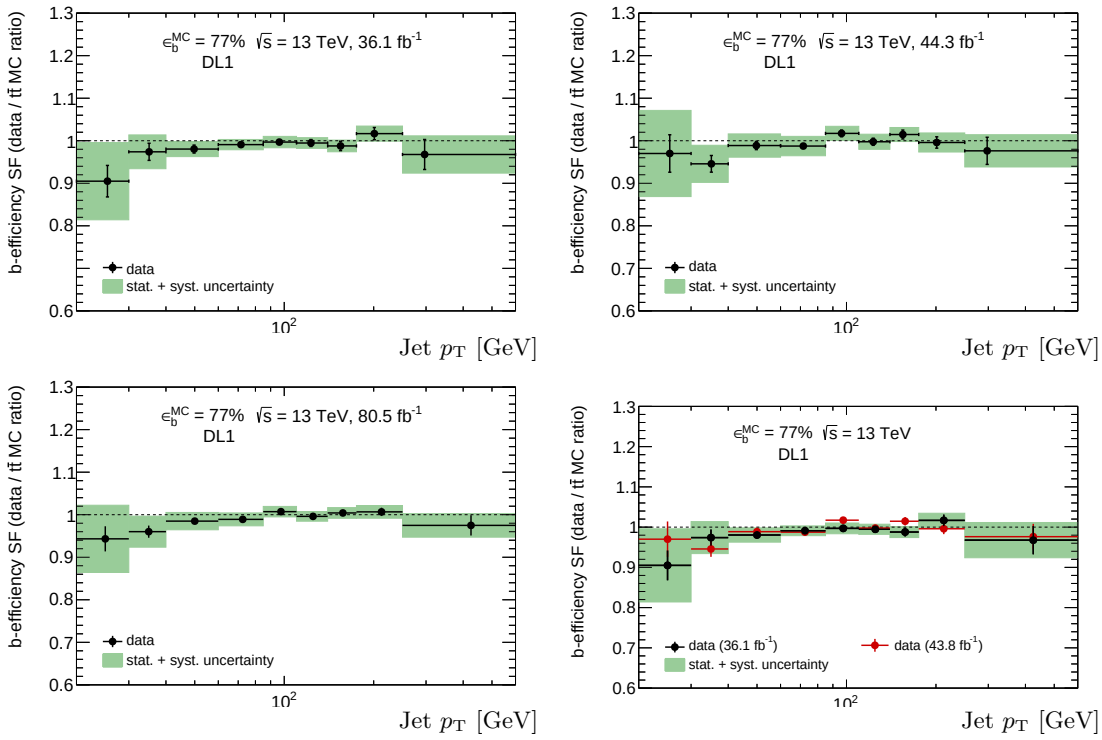

Figure C.2: The b-tagging efficiency SF as a function of the jet p_T extracted from the LH fit method to data for the DL1 algorithm at the 77% WP. The fit result is shown based on ATLAS data from the years 2015 and 2016 (top left), 2017 (top right) and all three years combined (bottom left). In the bottom right plot, a comparison of the 2015 and 2016 results (black) to the 2017 result (red) is depicted. The individual fit results (top left and top right) show the bin centres at the mean jet p_T value of each bin. The vertical error bars include only the statistical uncertainties from data, while the green error band represents the sum in quadrature of the statistical and systematic uncertainties.

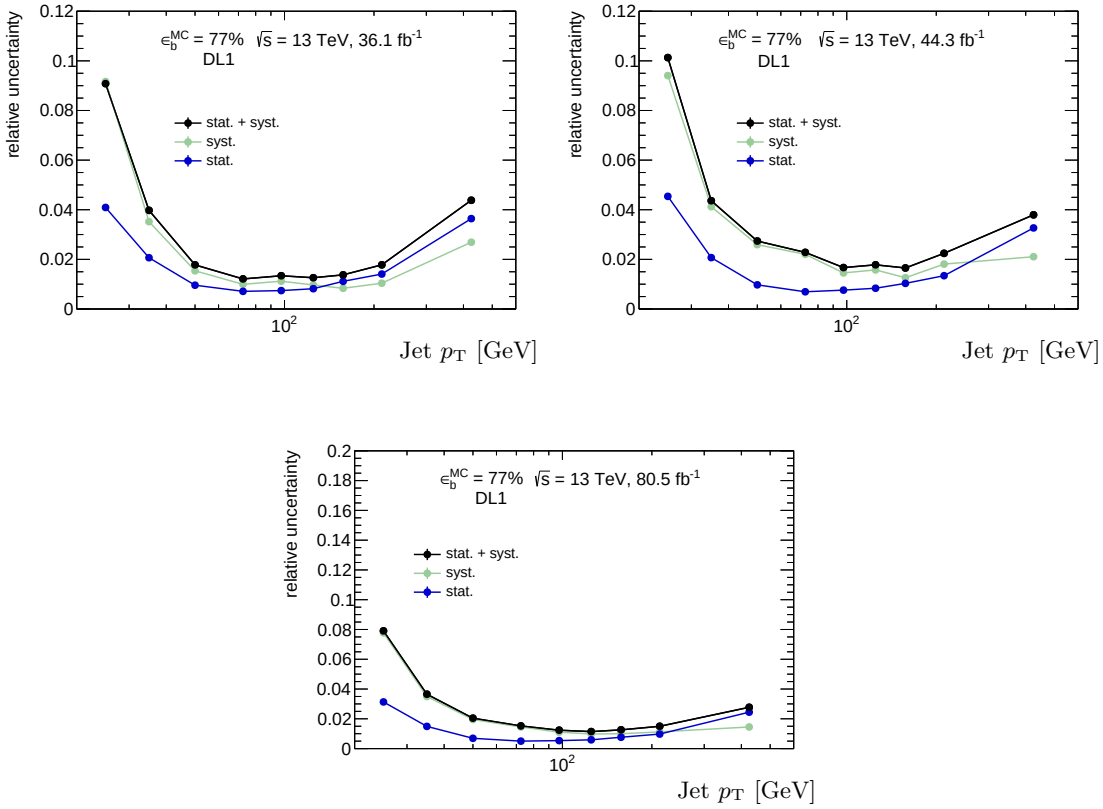

Figure C.3: The uncertainty components relative to the extracted b-tagging efficiency as a function of the jet p_T from the LH fit method to data for the DL1 algorithm at the 77% WP. Shown are the statistical uncertainty from data (blue), the total systematic uncertainty (green) and their sum in quadrature (black). The uncertainty estimations are based on the fit to ATLAS data from the years 2015 and 2016 (top left), 2017 (top right) and all three years combined (bottom).

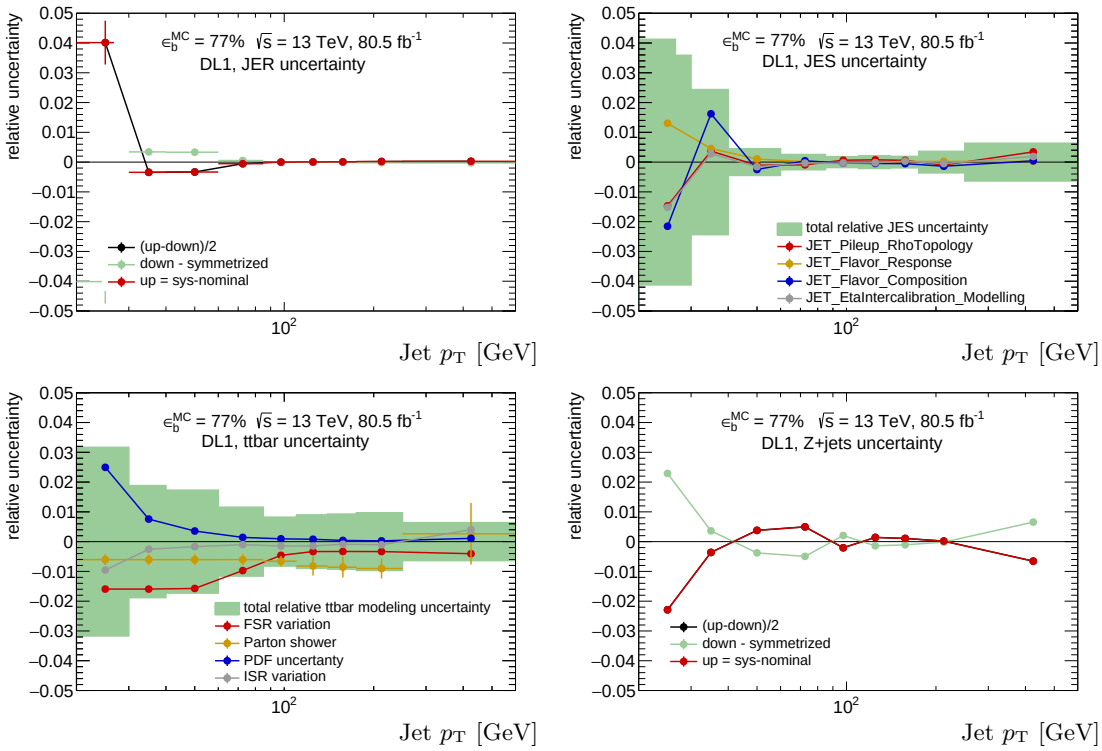

Figure C.4: The dominant systematic uncertainties as a function of the jet p_T during the b-tagging efficiency measurement for the DL1 algorithm at the 77% WP. The uncertainties shown are from the fit results based on ATLAS data from the years 2015, 2016 and 2017. They represent the jet energy resolution (top left), jet energy scale (top right), $t\bar{t}$ modelling (bottom left) and Z+jets modelling (bottom right) as a function of jet p_T.

APPENDIX D

Choice of observables

The MVA training in this analysis focuses only on how to separate the $t\bar{t}H(H \to b\bar{b})$ signal from the $t\bar{t}+b\bar{b}$ background as well as possible. Further attempts to discriminate the signal from other background sources, for example Diboson production, are not considered critical for this analysis, because the $t\bar{t}+b\bar{b}$ background is so dominant above all other background processes. In consequence, only observables that possess significant power to discriminate between them are considered and studied. But it is not necessarily a simple task to find such observables, especially if not done in a structured way. A good starting point is to motivate the choice of each observable, because this choice should always be driven by physics reasons that explain why these observables hold the desired or observed discriminating power. Doing so helps to improve the understanding of the analysed processes through which analysers can choose the optimal set of variables to train on. This allows to minimise the number of variables as well as the correlation between them, thus minimising the associated systematic uncertainties that may result, for example, from the variables' model dependence.

The properties of the Higgs boson, the top quark, the bottom quark and QCD were discussed in Chapter 2. These properties serve as the basis of this discussion. The Higgs boson coupling to other SM particles is proportional to the mass of the involved particle and thus, with the top quark being the heaviest SM particle, the Higgs boson prefers to couple to the top quark, rather than other quarks, by a large margin. This is important to consider, because the $t\bar{t}$ process at the LHC is initiated predominantly by gluons ($\approx 90\%$) and only to a lesser extent by quarks ($\approx 10\%$). Gluons do not directly couple to the Higgs boson and initial state quarks in the proton are mostly light quarks which have a significantly suppressed Yukawa coupling to the Higgs boson. Therefore, the Higgs boson in the $t\bar{t}H$ process is expected to be radiated off of one of the top quarks, rather than any initial state particle or decay product of the top quark. This

affects the $t\bar{t}$ kinematics within $t\bar{t}H$ production primarily after the hard interaction, as one of the top quarks transfers a part of its momentum to the produced Higgs boson and this propagates further into the $b\bar{b}$ pair. On the other hand, in $t\bar{t}+b\bar{b}$ production, the $b\bar{b}$ pair may be produced by gluon radiation of the initial state particles and subsequent gluon splitting. This affects the $t\bar{t}$ kinematics and event topology before the hard interaction, because the initial state particles lose a part of their momenta, which would have otherwise been transferred to the $t\bar{t}$ system, and give it to the $b\bar{b}$ pair. This is important, because the production of a $b\bar{b}$ pair in $t\bar{t}+b\bar{b}$ is mediated via QCD, while in $t\bar{t}H(H \to b\bar{b})$ it is via the Higgs boson decay. On a fundamental level, both interactions are differently structured, namely $\mathrm{SU}(3)_C$ compared to $\mathrm{SU}(2)_L \times \mathrm{U}(1)_Y$. As a result they produce different $b\bar{b}$ kinematics and an overall different event topology [7,9].

One of the first and obvious observables to consider are those related to the $b\bar{b}$ pair not originating from the $t\bar{t}$ decay. For example, in $t\bar{t}H(H \to b\bar{b})$, the additional $b\bar{b}$ pair has an invariant mass close to the Higgs mass of $m_H = 125$ GeV within detector resolution and should follow a Breit-Wigner distribution. In $t\bar{t}+b\bar{b}$ production, the invariant mass does not correspond to a high mass resonance decay and thus should follow an exponential distribution. Additionally, the invariant masses of the W boson and top quark candidates in $t\bar{t}H(H \to b\bar{b})$ events follow slightly different distributions with respect to those in $t\bar{t}+b\bar{b}$ events as well. This is because of the aforementioned $t\bar{t}$ system kinematics that are affected in a different way depending on the process which leads to different candidate reconstruction efficiencies and mass resolutions.
Another example is the angular separation, expressed by ΔR, between the two b-jets which should be much closer together if they originate from a Higgs boson decay than from gluon splitting. This is because the $b\bar{b}$ pair is expected to have a stronger boost through the decay of a very heavy particle such as the Higgs boson. The angular separation between the individual $t\bar{t}$ decay products is affected as well. If one of the top quarks radiates off a Higgs boson, its p_T is decreased significantly which propagates into the p_T of the decay products. Thus, the decay products of one top quark tend to be closer together in the centre-of-mass system, while the decay products of different top quarks are expected to be farther apart.

Lastly, for the final classification of reconstructed events, additional kinematic variables depending on the general event topology as well as variables depending on the b-tagging weights of b-tagged jets are considered. The former include the number of jets with p_T above 40 GeV, the aplanarity of the event and the scalar sum p_T of jets and leptons which is H_T. The aplanarity is the opposite to the planarity which is a measure of how much an event evolved spatially within a plane. Its definition is given in Table 9.4.
The complete lists of variables included in the respective BDT trainings are given in Section 9.1.

APPENDIX E

Event yields in the $t\bar{t}H(H \to b\bar{b})$ analysis

Listed in Tables E.1–E.6 are the event yields in the respective analysis regions of the single lepton and dilepton channels before and after the fit to data. The yields are broken down into the different signal and background categories defined in Section 7.3 and compared to the observed data yields.

E Event yields in the $t\bar{t}H(H \to b\bar{b})$ analysis

Region	SR_1^{5j}		SR_2^{5j}	
Sample	Pre-fit	Post-fit	Pre-fit	Post-fit
$t\bar{t}H$	15.9 ± 2.1	13.3 ± 9.8	40.1 ± 5.1	34 ± 25
$t\bar{t}$ + light	15 ± 33	12.5 ± 9.3	500 ± 210	393 ± 67
$t\bar{t}+ \geq 1c$	30 ± 17	28 ± 14	436 ± 92	610 ± 100
$t\bar{t}+ \geq 1b$	273 ± 53	335 ± 25	1230 ± 200	1450 ± 110
$t\bar{t} + V$	6.4 ± 1.3	6.4 ± 1.2	19.9 ± 2.9	19.7 ± 2.4
Non-$t\bar{t}$	54 ± 11	28.1 ± 8.4	269 ± 64	220 ± 52
Total	371 ± 68	423 ± 23	2440 ± 390	2724 ± 70
Data	426		2798	

Table E.1: Pre-fit and post-fit event yields including the number of selected data events for each single lepton SR with exactly five jets [4]. The uncertainties are given by the sum in quadrature of the statistical and systematic uncertainties in the yields. The post-fit yields are those from the combined fit in all channels to data and the corresponding uncertainties are calculated while accounting for correlations between the nuisance parameters and between the normalisations of the different processes. There is no pre-fit uncertainty defined for the normalisation factors $k(t\bar{t}+ \geq 1b)$ and $k(t\bar{t}+ \geq 1c)$, thus it is only included in the post-fit uncertainties. As a consequence, the pre-fit uncertainties of the $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ categories are given only by acceptance effects. The quoted $t\bar{t}H$ yields before the fit correspond to the theoretical prediction with corresponding uncertainties. After the fit, the $t\bar{t}H$ yields and uncertainties correspond to those obtained by the fit of the signal strength.

Region	$CR_{t\bar{t}+b}^{5j}$		$CR_{t\bar{t}+\geq 1c}^{5j}$		$CR_{t\bar{t}+\text{light}}^{5j}$	
Sample	Pre-fit	Post-fit	Pre-fit	Post-fit	Pre-fit	Post-fit
$t\bar{t}H$	68.0 ± 7.6	57 ± 24	18.7 ± 2.5	15 ± 12	224 ± 22	190 ± 140
$t\bar{t}$ + light	4250 ± 920	3560 ± 240	2580 ± 720	2300 ± 210	197000 ± 26000	179900 ± 4900
$t\bar{t}+ \geq 1c$	1770 ± 270	2590 ± 390	1280 ± 500	1840 ± 250	27500 ± 4300	44100 ± 5500
$t\bar{t}+ \geq 1b$	3400 ± 440	4030 ± 320	790 ± 130	944 ± 94	11300 ± 1100	13500 ± 1300
$t\bar{t} + V$	48.1 ± 5.9	46.6 ± 5.4	23.2 ± 4.1	21.3 ± 2.9	589 ± 55	584 ± 54
Non-$t\bar{t}$	960 ± 190	860 ± 160	520 ± 180	440 ± 100	21300 ± 4100	20900 ± 3200
Total	10400 ± 1300	11140 ± 290	5200 ± 1100	5560 ± 160	258000 ± 29000	259000 ± 900
Data	11095		5465		259320	

Table E.2: Pre-fit and post-fit event yields including the number of selected data events for each single lepton CR with exactly five jets [4]. The uncertainties are given by the sum in quadrature of the statistical and systematic uncertainties in the yields. The post-fit yields are those from the combined fit in all channels to data and the corresponding uncertainties are calculated while accounting for correlations between the nuisance parameters and between the normalisations of the different processes. There is no pre-fit uncertainty defined for the normalisation factors $k(t\bar{t}+ \geq 1b)$ and $k(t\bar{t}+ \geq 1c)$, thus it is only included in the post-fit uncertainties. As a consequence, the pre-fit uncertainties of the $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ categories are given only by acceptance effects. The quoted $t\bar{t}H$ yields before the fit correspond to the theoretical prediction with corresponding uncertainties. After the fit, the $t\bar{t}H$ yields and uncertainties correspond to those obtained by the fit of the signal strength.

Region	$SR_1^{\geq 6j}$		$SR_2^{\geq 6j}$		$SR_3^{\geq 6j}$	
Sample	Pre-fit	Post-fit	Pre-fit	Post-fit	Pre-fit	Post-fit
$t\bar{t}H$	62 ± 11	51 ± 38	81 ± 10	68 ± 50	85 ± 10	71 ± 52
$t\bar{t}$ + light	14 ± 10	12.1 ± 5.8	210 ± 210	96 ± 33	750 ± 370	586 ± 98
$t\bar{t}+ \geq 1c$	53 ± 33	44 ± 20	350 ± 100	473 ± 99	880 ± 350	1330 ± 190
$t\bar{t}+ \geq 1b$	1010 ± 240	1032 ± 59	1750 ± 370	1850 ± 130	2100 ± 420	2290 ± 170
$t\bar{t} + V$	25.8 ± 3.7	25.3 ± 3.2	40.8 ± 5.7	40.3 ± 4.8	51.2 ± 7.4	50.8 ± 5.9
Non-$t\bar{t}$	75 ± 20	58 ± 17	155 ± 52	134 ± 46	303 ± 82	267 ± 63
Total	1220 ± 250	1223 ± 42	2550 ± 510	2657 ± 82	4140 ± 850	4590 ± 110
Data	1222		2641		4698	

Table E.3: Pre-fit and post-fit event yields including the number of selected data events for each single lepton SR with at least six jets [4]. The uncertainties are given by the sum in quadrature of the statistical and systematic uncertainties in the yields. The post-fit yields are those from the combined fit in all channels to data and the corresponding uncertainties are calculated while accounting for correlations between the nuisance parameters and between the normalisations of the different processes. There is no pre-fit uncertainty defined for the normalisation factors $k(t\bar{t}+ \geq 1b)$ and $k(t\bar{t}+ \geq 1c)$, thus it is only included in the post-fit uncertainties. As a consequence, the pre-fit uncertainties of the $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ categories are given only by acceptance effects. The quoted $t\bar{t}H$ yields before the fit correspond to the theoretical prediction with corresponding uncertainties. After the fit, the $t\bar{t}H$ yields and uncertainties correspond to those obtained by the fit of the signal strength.

Region	$CR_{t\bar{t}+b}^{\geq 6j}$		$CR_{t\bar{t}+\geq 1c}^{\geq 6j}$		$CR_{t\bar{t}+\text{light}}^{\geq 6j}$	
Sample	Pre-fit	Post-fit	Pre-fit	Post-fit	Pre-fit	Post-fit
$t\bar{t}H$	100 ± 12	83 ± 61	102 ± 13	87 ± 64	450 ± 48	370 ± 280
$t\bar{t}$ + light	2200 ± 520	1820 ± 170	4300 ± 2000	3350 ± 430	125000 ± 34000	108200 ± 4300
$t\bar{t}+ \geq 1c$	1460 ± 330	2080 ± 300	3600 ± 1300	5300 ± 680	28400 ± 7200	45700 ± 5100
$t\bar{t}+ \geq 1b$	3670 ± 500	4080 ± 320	2660 ± 540	2950 ± 280	13100 ± 1800	14600 ± 1400
$t\bar{t} + V$	70.5 ± 8.5	67.9 ± 7.2	118 ± 21	118 ± 14	1010 ± 120	996 ± 91
Non-$t\bar{t}$	710 ± 160	600 ± 110	1060 ± 340	1000 ± 210	12600 ± 3000	11800 ± 2000
Total	8200 ± 1100	8730 ± 230	11800 ± 3200	12810 ± 260	181000 ± 39000	181690 ± 860
Data	8576		12778		181706	

Table E.4: Pre-fit and post-fit event yields including the number of selected data events for each single lepton CR with at least six jets [4]. The uncertainties are given by the sum in quadrature of the statistical and systematic uncertainties in the yields. The post-fit yields are those from the combined fit in all channels to data and the corresponding uncertainties are calculated while accounting for correlations between the nuisance parameters and between the normalisations of the different processes. There is no pre-fit uncertainty defined for the normalisation factors $k(t\bar{t}+ \geq 1b)$ and $k(t\bar{t}+ \geq 1c)$, thus it is only included in the post-fit uncertainties. As a consequence, the pre-fit uncertainties of the $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ categories are given only by acceptance effects. The quoted $t\bar{t}H$ yields before the fit correspond to the theoretical prediction with corresponding uncertainties. After the fit, the $t\bar{t}H$ yields and uncertainties correspond to those obtained by the fit of the signal strength.

Region	$SR_1^{\geq 4j}$		$SR_2^{\geq 4j}$		$SR_3^{\geq 4j}$	
Sample	Pre-fit	Post-fit	Pre-fit	Post-fit	Pre-fit	Post-fit
$t\bar{t}H$	15.6 ± 2.5	12.9 ± 9.5	29.1 ± 4.2	25 ± 18	21.9 ± 2.5	18 ± 13
$t\bar{t}$ + light	6.4 ± 9.9	11.1 ± 9.3	250 ± 110	215 ± 43	83 ± 41	95 ± 30
$t\bar{t}+ \geq 1c$	12.6 ± 9.4	25.8 ± 7.8	340 ± 210	427 ± 89	235 ± 61	313 ± 53
$t\bar{t}+ \geq 1b$	247 ± 61	263 ± 20	590 ± 96	669 ± 59	819 ± 85	917 ± 71
$t\bar{t} + V$	7 ± 56	7 ± 57	22 ± 38	22 ± 39	15 ± 35	15 ± 34
Non-$t\bar{t}$	13.6 ± 3.8	14.6 ± 3.8	115 ± 36	121 ± 29	75 ± 17	78 ± 16
Total	302 ± 85	334 ± 59	1350 ± 320	1479 ± 66	1250 ± 140	1436 ± 55
Data	319		1444		1467	

Table E.5: Pre-fit and post-fit event yields including the number of selected data events for each dilepton SR with at least four jets [4]. The uncertainties are given by the sum in quadrature of the statistical and systematic uncertainties in the yields. The post-fit yields are those from the combined fit in all channels to data and the corresponding uncertainties are calculated while accounting for correlations between the nuisance parameters and between the normalisations of the different processes. There is no pre-fit uncertainty defined for the normalisation factors $k(t\bar{t}+ \geq 1b)$ and $k(t\bar{t}+ \geq 1c)$, thus it is only included in the post-fit uncertainties. As a consequence, the pre-fit uncertainties of the $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ categories are given only by acceptance effects. The quoted $t\bar{t}H$ yields before the fit correspond to the theoretical prediction with corresponding uncertainties. After the fit, the $t\bar{t}H$ yields and uncertainties correspond to those obtained by the fit of the signal strength.

Region	$CR_{t\bar{t}+\geq 1c}^{\geq 4j}$		$CR_{t\bar{t}+\text{light}}^{\geq 4j}$		$CR_{t\bar{t}+\geq 1b}^{3j}$		$CR_{t\bar{t}+\text{light}}^{3j}$	
Sample	Pre-fit	Post-fit	Pre-fit	Post-fit	Pre-fit	Post-fit	Pre-fit	Post-fit
$t\bar{t}H$	35.3 ± 3.6	29 ± 22	114 ± 11	95 ± 70	8.7 ± 1.1	7.3 ± 5.4	32.2 ± 3.8	27 ± 20
$t\bar{t}$ + light	1730 ± 730	1410 ± 180	42500 ± 9700	37100 ± 1300	291 ± 110	255 ± 44	63100 ± 5500	59100 ± 1400
$t\bar{t}+ \geq 1c$	1410 ± 590	2160 ± 290	6300 ± 2800	10300 ± 1400	360 ± 160	536 ± 89	4800 ± 2100	7700 ± 1100
$t\bar{t}+ \geq 1b$	1080 ± 120	1240 ± 110	2850 ± 290	2510 ± 280	710 ± 140	848 ± 75	2130 ± 230	2620 ± 240
$t\bar{t} + V$	52 ± 41	50 ± 39	350 ± 180	330 ± 170	7 ± 27	7 ± 30	113 ± 31	112 ± 29
Non-$t\bar{t}$	42 ± 120	460 ± 100	4700 ± 1100	4930 ± 910	110 ± 29	112 ± 23	6300 ± 1500	6500 ± 1200
Total	4700 ± 1100	5350 ± 120	56000 ± 11000	55650 ± 420	1500 ± 260	1765 ± 60	76400 ± 6500	76010 ± 390
Data	5389		55627		1744		76025	

Table E.6: Pre-fit and post-fit event yields including the number of selected data events for each dilepton CR with at least four or exactly three jets [4]. The uncertainties are given by the sum in quadrature of the statistical and systematic uncertainties in the yields. The post-fit yields are those from the combined fit in all channels to data and the corresponding uncertainties are calculated while accounting for correlations between the nuisance parameters and between the normalisations of the different processes. There is no pre-fit uncertainty defined for the normalisation factors $k(t\bar{t}+ \geq 1b)$ and $k(t\bar{t}+ \geq 1c)$, thus it is only included in the post-fit uncertainties. As a consequence, the pre-fit uncertainties of the $t\bar{t}+ \geq 1b$ and $t\bar{t}+ \geq 1c$ categories are given only by acceptance effects. The quoted $t\bar{t}H$ yields before the fit correspond to the theoretical prediction with corresponding uncertainties. After the fit, the $t\bar{t}H$ yields and uncertainties correspond to those obtained by the fit of the signal strength.

APPENDIX F

Comparison to other analyses

F.1 Combination of Atlas results

Aside from the $t\bar{t}H(H \to b\bar{b})$ process, other decay channels of the Higgs boson in the $t\bar{t}H$ production are studied in the ATLAS experiment during Run 2. These include the search for $t\bar{t}H$ production where the Higgs boson decays into:

- a pair of photons ($H \to \gamma\gamma$) [147];
- a real Z and an off-shell Z^* boson which further decay into four charged leptons ($H \to ZZ^* \to 4l$) [148]; or
- final states involving various combinations of leptons ($H \to$ ML) such as same-sign lepton pairs or multiple hadronically decaying τ-leptons [149]. This is achieved by searching for Higgs boson decays into WW^*, $\tau\tau$ and ZZ^*. Through appropriate event selections, this analysis is kept orthogonal to the $4l$ search above as well as the $t\bar{t}H(H \to b\bar{b})$ search presented in this thesis, as was discussed in Section 8.1.

The $H \to \gamma\gamma$ and $H \to ZZ^* \to 4l$ analyses mentioned above have been performed at $\sqrt{s} = 13$ TeV based on data recorded with the ATLAS detector in the years 2015 and 2016. They have since been updated to also include data from 2017, resulting in a total integrated luminosity of 79.8 ± 1.6 fb^{-1} [33]. Furthermore, the analyses have been improved by employing dedicated analysis strategies tailored to their respective $t\bar{t}H$ final states as well as using improved reconstruction software [20]. The individual results of these two analyses have been combined with the $t\bar{t}H(H \to b\bar{b})$ and $H \to$ ML searches based on 2015 and 2016 data and this combination has then been further combined with the corresponding searches based on data recorded at $\sqrt{s} = 7$ TeV and 8 TeV during Run 1 of the LHC [20].
The combination has been performed using a profile likelihood method, following the

guidelines in Ref. [150]. Instead of combining the fit results of the signal strength parameter of each individual analysis, a new procedure is employed that simultaneously fits all signal and control regions of each analysis. As written above, the analyses are kept orthogonal by dedicated event selections. Nonetheless, there is an overlap between selected events entering the fit, but this is found to have a negligible impact [20].
The difficulty of this combination lies primarily in the correlation of the experimental as well as theoretical uncertainties across the analyses. This combination strategy results in slightly different best-fit values of the signal strength parameter associated to the individual analyses and is described in more detail in Ref. [20].

Firstly, it is interesting to look at the individual results of the analyses compared to the $t\bar{t}H(H \to b\bar{b})$ result presented in this chapter. The $t\bar{t}H(H \to \gamma\gamma)$ analysis reaches an observed (expected) significance of 4.1σ (3.7σ), while the $t\bar{t}H(H \to ZZ^* \to 4l)$ analysis has an expected significance of 1.2σ, but does not observe any events and thus sets an upper limit of $\mu < 1.77$ at 68% CL based on pseudo-experiments. The $t\bar{t}H(H \to \text{ML})$ search quotes an observed (expected) significance of 4.1σ (2.8σ). The $t\bar{t}H(H \to b\bar{b})$ search achieves an observed (expected) significance of 1.4σ (1.6σ). Out of the four analyses, the $t\bar{t}H(H \to b\bar{b})$ search is the one most strongly limited by the systematic uncertainties, more specifically those related to physics modelling. While the $t\bar{t}H(H \to ZZ^* \to 4l)$ analysis is completely limited by an insufficient amount of data events, the $t\bar{t}H(H \to \gamma\gamma)$ search is also dominated by the statistical uncertainty over the systematic one, however only by a ratio of $\approx 2:1$. The two uncertainties on μ in the $t\bar{t}H(H \to \text{ML})$ search are almost equal, while in the $t\bar{t}H(H \to b\bar{b})$ analysis their ratio is $\approx 1:2$ [20].

The dominant uncertainties in the combined result originate from the modelling of the $t\bar{t}$ + heavy flavour background in the $t\bar{t}H(H \to b\bar{b})$ analysis as well as the modelling of the $t\bar{t}H$ signal, the latter having a direct impact on the events selected by all four analyses. It is important to note here that our theoretical understanding of the $t\bar{t}$ + heavy flavour process is limited in such a way that the associated uncertainties dominate even a combination of complex analyses of which the majority is limited by the statistical uncertainty. Further important uncertainties are associated to the estimation of leptons from b- or c-jet decays, photon conversions or mis-identified jets, which are mainly attributable to the $t\bar{t}H(H \to \text{ML})$ search. In addition to this, the uncertainty on the jet energy scale and resolution in all analyses contributes significantly to the uncertainty of the combined result, since the $t\bar{t}$ decay always involves multiple jets in the final state, regardless of the Higgs boson decay. Table F.1 summarises the most important sources of uncertainties and their relative impact on the extracted $t\bar{t}H$ cross-section. More details on the considered uncertainties and their origin are given in Ref. [20].

The results for the extracted signal strength when combining these four $t\bar{t}H$ analyses at $\sqrt{s} = 13$ TeV are shown in Figure F.1. The combined best-fit result is $1.32^{+0.28}_{-0.26}$, which excludes the background-only hypothesis with an observed (expected) significance of 5.8σ (4.9σ), respectively [20]. The $t\bar{t}H$ cross-section at $\sqrt{s} = 13$ TeV is predicted to be

Uncertainty source	$\Delta\sigma_{t\bar{t}H}/\sigma_{t\bar{t}H}$ [%]
Theory uncertainties (physics modelling)	11.9
$t\bar{t}$ + heavy flavour	9.9
$t\bar{t}H$	6.0
Non-$t\bar{t}H$ Higgs boson production modes	1.5
Other background processes	2.2
Experimental uncertainties	9.3
Fake leptons	5.2
Jets, $E_{\mathrm{T}}^{\mathrm{miss}}$	4.9
Electrons, photons	3.2
Luminosity	3.0
τ-lepton	2.5
Flavour tagging	1.8
MC statistical uncertainties	4.4

Table F.1: Presented are the most significant systematic uncertainties that impact the combined signal strength parameter μ during the fit based on ATLAS data collected at $\sqrt{s} = 13$ TeV. The parameters are ranked in decreasing order by their contribution to the total uncertainty, but only those are shown that contribute at least 1% to the relative total uncertainty on μ. More details on the origin of the uncertainties are given in Ref. [20].

507^{+35}_{-50} fb (see Section 7.2) and thus, the extracted combined value for μ corresponds to a cross-section of $670 \pm 90(\text{stat.})^{+110}_{-100}(\text{syst.})$ fb. This result agrees well with the SM prediction within the associated uncertainties. The combined fit based on the analyses at $\sqrt{s} = 7$ TeV, 8 TeV and 13 TeV allows to claim the presence of a $t\bar{t}H$ signal with an observed (expected) significance of 6.3σ (5.1σ). These results are depicted in Table F.2.

Finally, the combined event yields are shown in Figure F.2 as a function of $\log_{10}(S/B)$, similar to Figure 11.12 for the $t\bar{t}H(H \to b\bar{b})$ result. S refers to the expected (observed) number of signal events corresponding to $\mu = 1.0$ ($\mu = 1.32$), while B is the fitted number of background events with freely floating signal [20]. In the right-most bins, a clear excess of $t\bar{t}H$ signal-like events can be observed.

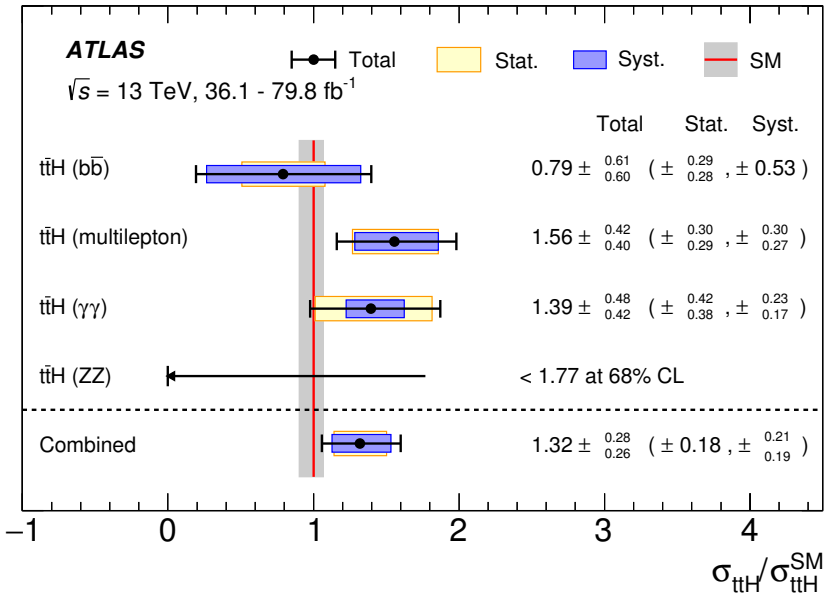

Figure F.1: The results of the best-fit values for the $t\bar{t}H$ signal strength parameters of the four analyses based on ATLAS data collected at $\sqrt{s} = 13$ TeV as well as their combined value [20]. The red line with grey uncertainty band represents the SM prediction and its associated uncertainties originating from the PDF+α_S choice and missing higher-order corrections during the calculation. The black lines show the central value and total uncertainty, while the blue (yellow) bands indicate the systematic (statistical) uncertainty component. The $t\bar{t}H(H \to ZZ^* \to 4l)$ analysis does not observe any $t\bar{t}H$ signal events and thus quotes an upper limit of $\mu < 1.77$ at 68% CL.

Analysis	Integrated luminosity [fb^{-1}]	$t\bar{t}H$ cross section [fb]	Observed sign.	Expected sign.
$H \to \gamma\gamma$	79.8	$710^{+210}_{-190}(\text{stat.})^{+120}_{-90}(\text{syst.})$	4.1σ	3.7σ
$H \to$ ML	36.1	$790 \pm 150(\text{stat.})^{+150}_{-140}(\text{syst.})$	4.1σ	2.8σ
$H \to b\bar{b}$	36.1	$400^{+150}_{-140}(\text{stat.}) \pm 270(\text{syst.})$	1.4σ	1.6σ
$H \to ZZ^* \to 4l$	79.8	< 900 (68% CL)	0σ	1.2σ
Combined (13 TeV)	$36.1 - 79.8$	$670 \pm 90(\text{stat.})^{+110}_{-100}(\text{syst.})$	5.8σ	4.9σ
Combined (7, 8, 13 TeV)	$4.5, 20.3, 36.1 - 79.8$	–	6.3σ	5.1σ

Table F.2: The results of the best-fit values for the $t\bar{t}H$ signal cross-section of the four analyses based on ATLAS data collected at $\sqrt{s} = 13$ TeV as well as their combined value [20]. The integrated luminosity, observed and expected significance of the results is given in addition to the combined value of the Run 1 and Run 2 $t\bar{t}H$ analyses. The $t\bar{t}H(H \to ZZ^* \to 4l)$ analysis does not observe any $t\bar{t}H$ signal events and thus quotes an upper limit on the cross-section at 68% CL.

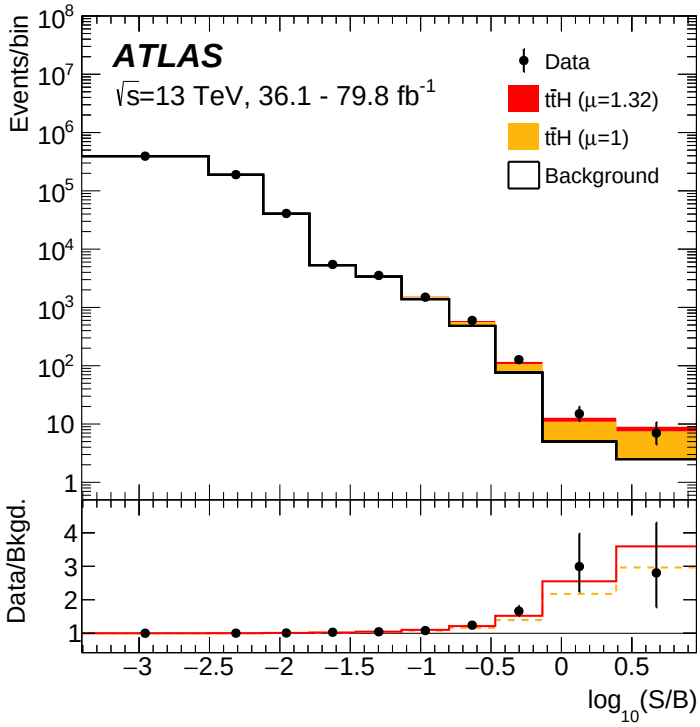

Figure F.2: The combined event yields of all four $t\bar{t}H$ searches based on ATLAS data collected at $\sqrt{s} = 13$ TeV as a function of $\log_{10}(S/B)$ [20]. S corresponds to the number of observed (expected) signal events corresponding to $\mu = 1.32$ (1.0) and B refers to the observed number of background events from the fit with freely floating signal. The bins are ordered such that $\log_{10}(S+B)$ decreases approximately linearly. The $H \to \gamma\gamma$ component only includes events in the smallest $m_{\gamma\gamma}$ window that contains 90% of the expected signal events. The ratio of the observed events in data with respect to B is compared to the expected distribution in case of a $t\bar{t}H$ signal presence given by $\mu = 1.32$ (full red line) and $\mu = 1$ (dashed orange line). The error bars on the data events are purely statistical.

F.2 Comparison to results from the Cms experiment

Comparing the $t\bar{t}H(H \to b\bar{b})$ searches

In this last section, the Atlas results are compared to the ones obtained by the Cms experiment, starting with the corresponding $t\bar{t}H(H \to b\bar{b})$ analysis [151]. This analysis is based on data recorded with the Cms detector in the years 2015 and 2016, corresponding to an integrated luminosity of 35.9($\pm$2.5%) fb^{-1} [152]. The analysers select the semileptonic and dileptonic decay channels of the $t\bar{t}$ pair and perform a simultaneous binned maximum likelihood fit to data across all analyses categories, similarly to the Atlas analysis. The result of this fit is depicted in Figure F.3 in the form of the signal strength parameters μ for the two channels as well as the combined result. The single lepton result is $\mu_{\text{single-lepton}} = 0.84^{+0.27}_{-0.26}(\text{stat.})^{+0.44}_{-0.43}(\text{syst.}) = 0.84^{+0.52}_{-0.50}$, the dileptonic result is $\mu_{\text{dilepton}} = -0.24^{+0.63}_{-0.60}(\text{stat.})^{+1.04}_{-0.95} = -0.24^{+1.21}_{-1.12}$ and the combined result is $\mu_{\text{combined}} = 0.72 \pm 0.24(\text{stat.}) \pm 0.38(\text{syst.}) = 0.72 \pm 0.45$. This corresponds to an observed (expected) significance of 1.6σ (2.2σ) of the $t\bar{t}H$ signal above the background-only hypothesis.

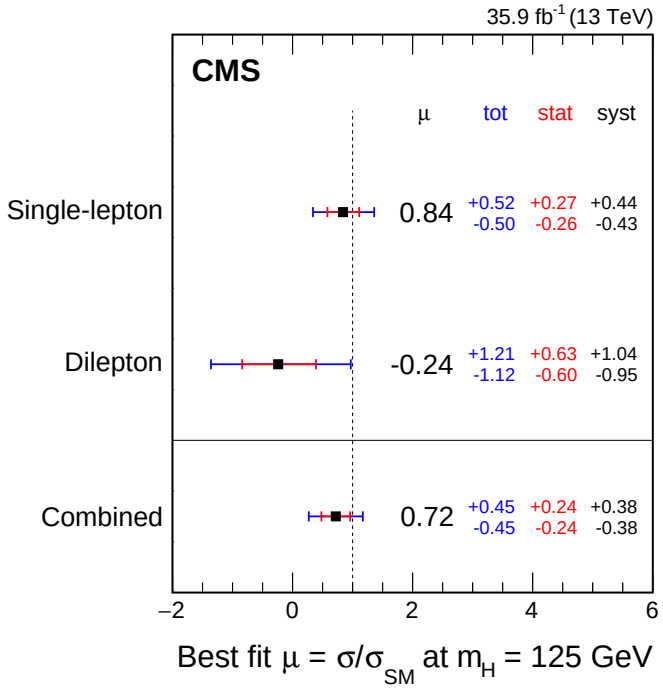

Figure F.3: The signal strength parameter values μ of the simultaneous binned maximum likelihood fit to 35.9 fb^{-1} of Cms data collected at $\sqrt{s} = 13$ TeV in all analysis regions [151]. Shown is the combined result as well as the individual results of the single lepton and dilepton channels.

The central value of the measured signal strength in the dilepton channel is strikingly similar to the corresponding measurement in ATLAS, which is $\mu_{\text{dilep,ATLAS}} = -0.24^{+0.54}_{-0.52}(\text{stat.})^{+0.87}_{-0.91}(\text{syst.}) = -0.24^{+1.02}_{-1.05}$. In addition to this, the total systematic uncertainty on μ_{combined} is significantly smaller than the one of the combined ATLAS result, which is $^{+0.57}_{-0.54}$. Since the physics modelling uncertainties are the dominating contributions to the systematic uncertainty on the ATLAS measurement, it is important to understand the origin of this observed difference. Figure F.4 lists all systematic uncertainties considered in the CMS measurement and notes whether the uncertainty affects the normalisation ('rate') or the shape of the distributions that enter the fit.

First of all, the $t\bar{t}H$ signal and the $t\bar{t}$ +jets background processes are modelled using different MC generators with respect to those employed in the ATLAS search (see Sections 7.2 and 7.3.1). In the CMS search, both processes are modelled using the same setup, namely POWHEG +PYTHIA 8 with the CUETP8M2T4 tune for the UE [153,154]. However, all samples are generated at NLO precision with the NNPDF3.0NLO PDF set in the ME calculation.
The CMS analysis also considers the $t\bar{t}+b\bar{b}$, $t\bar{t}+B$, $t\bar{t}+b$ and $t\bar{t}+c\bar{c}$ processes to be important sources of irreducible background [151]. Despite existing measurements of the $t\bar{t}$ + heavy flavour component [155] and higher-order theoretical calculations for the $t\bar{t}+b\bar{b}$ cross-section [143,156], the analysers find that the normalisation uncertainty on the backgrounds mentioned above cannot be constrained to a better accuracy than 35% [151]. Thus, a 50% uncertainty is associated to the normalisation of these background components [151].

The CMS analysis team also considers another $t\bar{t}$ + heavy flavour prediction, namely by using a sample generated with SHERPA (v2.2.2) with the 4F scheme in the ME interfaced to OPENLOOPS, which is similar to the SHERPA 4F setup used in the ATLAS analysis. By comparing the nominal $t\bar{t}$ prediction to this alternative sample, a subset of the systematic uncertainties related to the $t\bar{t}+b\bar{b}$, $t\bar{t}+B$ and $t\bar{t}+b$ background contributions considered by the ATLAS team is taken into account.
The robustness of the binned maximum likelihood fit is tested via the following two methods [151]:

1. Increasing the normalisation of the $t\bar{t}+b\bar{b}$ background by 30% according to Ref. [155]; or

2. Replacing the nominal $t\bar{t}$ + heavy flavour background predictions with the alternative 4F scheme sample from CMS.

For both scenarios, the fit is able to retrieve the injected signal within the range of a few percent which is within the uncertainties assigned to these processes [151].

In terms of further MC generator uncertainties, the uncertainty in the scheme of matching the $t\bar{t}$ ME to the PS is accounted for by comparing the nominal sample to two alternative samples in which the h_{damp} parameter is varied up or down. Additionally,

Source	Type	Remarks
Integrated luminosity	rate	Signal and all backgrounds
Lepton identification/isolation	shape	Signal and all backgrounds
Trigger efficiency	shape	Signal and all backgrounds
Pileup	shape	Signal and all backgrounds
Jet energy scale	shape	Signal and all backgrounds
Jet energy resolution	shape	Signal and all backgrounds
b tag hf fraction	shape	Signal and all backgrounds
b tag hf stats (linear)	shape	Signal and all backgrounds
b tag hf stats (quadratic)	shape	Signal and all backgrounds
b tag lf fraction	shape	Signal and all backgrounds
b tag lf stats (linear)	shape	Signal and all backgrounds
b tag lf stats (quadratic)	shape	Signal and all backgrounds
b tag charm (linear)	shape	Signal and all backgrounds
b tag charm (quadratic)	shape	Signal and all backgrounds
Renorm./fact. scales ($t\bar{t}H$)	rate	Scale uncertainty of NLO $t\bar{t}H$ prediction
Renorm./fact. scales ($t\bar{t}$)	rate	Scale uncertainty of NLO $t\bar{t}$ prediction
Renorm./fact. scales ($t\bar{t}$+hf)	rate	Additional 50% rate uncertainty of $t\bar{t}$+hf predictions
Renorm./fact. scales (t)	rate	Scale uncertainty of NLO single t prediction
Renorm./fact. scales (V)	rate	Scale uncertainty of NNLO W and Z prediction
Renorm./fact. scales (VV)	rate	Scale uncertainty of NLO diboson prediction
PDF (gg)	rate	PDF uncertainty for gg initiated processes except $t\bar{t}H$
PDF (gg $t\bar{t}H$)	rate	PDF uncertainty for $t\bar{t}H$
PDF ($q\bar{q}$)	rate	PDF uncertainty of $q\bar{q}$ initiated processes ($t\bar{t}$+W,W,Z)
PDF (qg)	rate	PDF uncertainty of qg initiated processes (single t)
μ_R scale ($t\bar{t}$)	shape	Renormalisation scale uncertainty of the $t\bar{t}$ ME generator, independent for additional jet flavours
μ_F scale ($t\bar{t}$)	shape	Factorisation scale uncertainty of the $t\bar{t}$ ME generator, independent for additional jet flavours
PS scale: ISR ($t\bar{t}$)	rate	Initial state radiation uncertainty of the PS (for $t\bar{t}$ events), jet multiplicity dependent rate uncertainty, independent for additional jet flavours
PS scale: FSR ($t\bar{t}$)	rate	Final state radiation uncertainty (for $t\bar{t}$ events), jet multiplicity dependent rate uncertainty, independent for additional jet flavours
ME-PS matching ($t\bar{t}$)	rate	NLO ME to PS matching, *hdamp* [**?**] (for $t\bar{t}$ events), jet multiplicity dependent rate uncertainty, independent for additional jet flavours
Underlying event ($t\bar{t}$)	rate	Underlying event (for $t\bar{t}$ events), jet multiplicity dependent rate uncertainty, independent for additional jet flavours
NNPDF3.0NLO ($t\bar{t}H$, $t\bar{t}$)	shape	Based on the NNPDF replicas, same for $t\bar{t}H$ and additional jet flavours
Bin-by-bin event count	shape	Statistical uncertainty of the signal and background prediction due to the limited sample size

Figure F.4: The list of the systematic uncertainties considered in the $t\bar{t}H(H \rightarrow b\bar{b})$ analysis by the CMS collaboration based on 35.9 fb^{-1} data collected at $\sqrt{s} = 13$ TeV [151]. It is also noted whether an uncertainty affects the normalisation ('rate') or the shape of the distributions that enter the fit and additional remarks are given.

uncertainties in the renormalisation and factorisation scales, μ_R and μ_F, as well as the tuned parameters are also considered by appropriate variations [151].
According to Figure F.4, the systematic uncertainties related to physics modelling are mainly considered to only affect the normalisation of the distributions of the discriminants that enter the fit. The exceptions are the variations in μ_R and μ_F in the $t\bar{t}$ +jets simulation as well as the uncertainty originating from the PDF set choice which is NNPDF3.0NLO for the $t\bar{t}H$ signal and $t\bar{t}$ +jets background predictions. However, the CMS analysis does not consider any effects on the shape of the discriminant dis-

tributions that are caused, for example, by the choice of the ISR and FSR scales, the ME to PS matching scheme or the UE modelling. In addition, systematic uncertainties related to the ME and PS models themselves are not taken into account as they are in the ATLAS analysis, for example by comparing the nominal $t\bar{t}$ +jets prediction with an alternative 5F generator setup such as MADGRAPH5_aMC@NLO +PYTHIA 8 or SHERPA 5F+OPENLOOPS. Especially the nuisance parameter associated to this latter systematic uncertainty has the highest impact on the signal strength in the ATLAS search, as is depicted in Figure 11.10. The above points may explain the significant discrepancy between the estimated systematic uncertainties of both analyses. Despite this, the results of both searches are well in agreement with each other and compatible with both the background-only as well as the signal-plus-background hypotheses assuming a SM Higgs boson with $m_H = 125$ GeV.

Comparing the combined $t\bar{t}H$ search results

Lastly, the combined results of all $t\bar{t}H$ searches by the CMS collaboration are detailed in Ref. [149]. Figure F.5 is the corresponding version of the combined ATLAS result in Figure F.2 and shows the event yields from the combined fit as a function of $\log_{10}(S/B)$, based on 35.9 fb^{-1} of data recorded at $\sqrt{s} = 13$ TeV. As in the ATLAS result, a clear excess of $t\bar{t}H$ signal-like events is visible in the right-most bins, meaning for high values of $\log_{10}(S/B)$.

The different individual best-fit signal strength parameters by the CMS analyses as well as the combined values for the analyses at $\sqrt{s} = 13$ TeV; 7 and 8 TeV; and all three results combined are presented in Figure F.6. The final result is $\mu_{t\bar{t}H} = 1.26^{+0.31}_{-0.26}$, which excludes the background-only hypothesis with an observed (expected) significance of 5.2σ (4.2σ), allowing the CMS collaboration to claim the observation of the $t\bar{t}H$ process as well. This result is well compatible with the SM expectation as well as the ATLAS result quoted above.

The uncertainty on the CMS result is decomposed into a statistical component of ± 0.16; an experimental component of $^{+0.17}_{-0.15}$; and two theoretical components for the signal ($^{+0.15}_{-0.07}$) and the background modelling ($^{+0.13}_{-0.12}$) [149]. The latter three sum up in quadrature to a total systematic uncertainty of $^{+0.26}_{-0.21}$. In contrast to this, the statistical uncertainty of the ATLAS result is ± 0.18 and thus larger than the corresponding CMS component, while the total systematic uncertainty is significantly smaller, namely $^{+0.21}_{-0.19}$. Since the ATLAS measurement uses a larger dataset, the higher statistical uncertainty may originate from including the statistical uncertainty on the freely floating fit parameters in this component. The differences in the systematic uncertainties may result from the different experimental and detector conditions as well as the assessment of simulation related uncertainties [20, 149]. This has a particularly significant impact on those uncertainties associated to the $t\bar{t}$ + heavy flavour background modelling in the $t\bar{t}H(H \to b\bar{b})$ search, which are one of the leading uncertainties of the combined result in ATLAS [20].

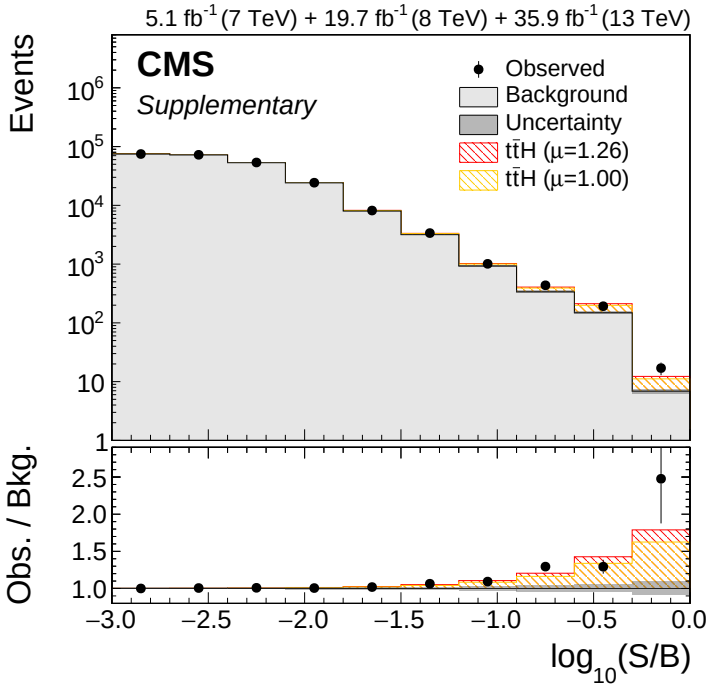

Figure F.5: The combined event yields of all CMS $t\bar{t}H$ searches based on Run 1 and Run 2 data as a function of $\log_{10}(S/B)$ [149]. S corresponds to the number of observed (expected) signal events post-fit corresponding to $\mu = 1.26$ (1.0) in the red (yellow) hatched histograms. B refers to the expected number of background events after the fit in the shaded light grey histogram. The ratio of the observed events in data with respect to B is compared to the expected distribution in case of a $t\bar{t}H$ signal presence given by $\mu = 1.26$ and the SM expectation of $\mu = 1$, respectively.

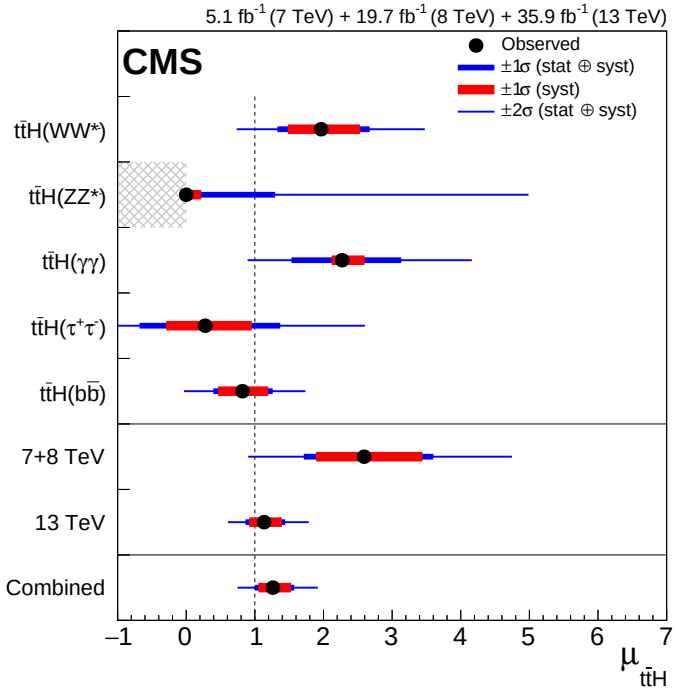

Figure F.6: Summary of the best-fit values for the $t\bar{t}H$ signal strength parameters of the CMS analyses based on $\sqrt{s} = 13$ TeV data as well as their combined value, the corresponding value for the Run 1 analyses and the total combined result from all Run 1 and Run 2 searches [149]. The thick blue (red) band represents the total (systematic) uncertainty on μ and the thin blue line indicates the 2σ uncertainty window. The dashed vertical line represents the SM expectation. The result of the $H \to ZZ^*$ analysis, $\mu_{t\bar{t}H}^{ZZ^*} = 0.00^{+1.30}_{-0.00}$, is constrained to be positive to avoid negative event yields [149].