GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Fakultät für
Physik

Master's Thesis

Modelling of the $t\bar{t} + b\bar{b}$ Process with the ATLAS Experiment

Modellierung des $t\bar{t} + b\bar{b}$ Prozesses mit dem ATLAS Experiment

prepared by

Stephen Eggebrecht

from Kassel

at the II. Physikalischen Institut

Thesis number:	II.Physik-UniGö-MSc-2021/01
Thesis period:	1st May 2020 until 30th April 2021
First referee:	Prof. Dr. Arnulf Quadt
Second referee:	Prof. Dr. Stan Lai

Abstract

The measurement of the Yukawa coupling of the Higgs boson to the top quark is an important test of the Standard Model (SM). It could be used to either find deviations from the SM or to constrain BSM theories. To study this coupling, the production of a SM Higgs boson in association with a top-quark pair is used. The decay mode where the SM Higgs boson decays to bottom quarks has the highest branching ratio. Unfortunately, it suffers from a large $t\bar{t} + b\bar{b}$ background. By reducing the modelling uncertainties of the $t\bar{t} + b\bar{b}$ background, a better sensitivity for the signal process can be achieved. The state of the art theoretical description of the $t\bar{t} + b\bar{b}$ background uses the *four flavour scheme* where the bottom quark is treated as a massive particle. To update the current set of Monte Carlo samples, newly generated samples are added. These MC samples differ in the choice of the renormalisation and factorisation scale, parton distribution function set and treatment of the bottom quark as a massive or a massless particle.

Keywords: $t\bar{t}\left(H \to b\bar{b}\right)$ Process, $t\bar{t} + b\bar{b}$ Background, Four Flavour Scheme

Zusammenfassung

Die Messung der Yukawa Kopplung des Higgs Bosons an das Top Quark ist ein wichtiger Test des Standard Modells (SM). Dieser kann verwendet werden um Abweichungen vom SM zu finden oder um BSM Theorien weiter einzuschränken. Um die Kopplung zu untersuchen, wird der Prozess verwendet, bei dem ein Higgs Boson in Assoziation mit einem Top Quark Paar erzeugt wird. Der Zerfall des Higgs Bosons mit dem größten Verzweigungsverhältnis ist der Zerfall in Bottom Quarks. Unglücklicherweise leidet dieser an einem großen Beitrag durch den $t\bar{t} + b\bar{b}$ Untergrundprozess. Durch das genaue Verstehen dieses Untergrundprozesses kann man eine bessere Sensitivität auf den Signalprozess erhoffen. Der aktuelle Stand der Beschreibung des $t\bar{t} + b\bar{b}$ Prozesses verwendet das *Four Flavour Scheme*, bei dem das Bottom Quark als massives Teilchen behandelt wird. Um den aktuellen Satz an Monte Carlo (MC) Simulationen zu erweiteren, werden neue hinzugefügt und mit den alten verglichen. Die MC Simulationen unterscheiden sich dabei durch die Wahl der Renormierung- und Faktorisierungsskala, den Partonverteilungsfunktionen und der Beschreibung des Bottom Quarks als massives oder masseloses Teilchen.

Stichwörter: $t\bar{t}\left(H \to b\bar{b}\right)$ Prozess, $t\bar{t} + b\bar{b}$ Untergrund, *Four Flavour Scheme*

Contents

1. Introduction

The Standard Model of Particle Physics (SM) is one of the most successful theories today and describes three fundamental forces which are the electromagnetic force, the weak force and the strong force. However, gravity is not included yet. In 2012, a new particle was discovered by ATLAS [1] and CMS [2]. This particle is compatible with the SM Higgs boson which is predicted by the Higgs mechanism [3–7]. The Higgs mechanism allows to generate massive gauge bosons through electroweak symmetry breaking and explains the masses of fermions as Yukawa couplings to the Higgs boson.

The Yukawa coupling is proportional to the mass of a fermion. The most massive fermions are the quarks of the third generation and the SM Higgs boson coupling to them provides an opportunity to find hints for physics beyond the SM (BSM) or to restrict BSM theories. The Yukawa coupling of a Higgs boson to a top quark, which is the most massive particle in the SM, is accessible by measurements of the SM Higgs production in association with a top quark pair $\left(t\bar{t}H\right)$. The decay mode of the SM Higgs boson to bottom quarks $H \to b\bar{b}$ is important because of its large branching ratio[1]. Both processes, $t\bar{t}H$ and $H \to b\bar{b}$, were measured successfully by ATLAS [8, 9] and CMS [10, 11]. However, the measurement of the process, where the SM Higgs boson is produced in association with a top quark pair $t\bar{t}\left(H \to b\bar{b}\right)$ and decays into a pair of bottom quarks, has a limited precision since it suffers from a large irreducible background which arises from the production of top quark pairs with additional bottom quark pairs $t\bar{t} + b\bar{b}$. Reducing the modelling uncertainties of the most dominant background will result in lower systematic uncertainties and increase the sensitivity to the signal process $t\bar{t}\left(H \to b\bar{b}\right)$.

There are two different approaches for the modelling of the $t\bar{t} + b\bar{b}$ production via the gluon splitting to bottom quarks $g \to b\bar{b}$ in theory calculations and Monte Carlo (MC) generators. In the five flavour scheme the b-quark is assumed to be massless and the splitting $g \to b\bar{b}$ is collinear and infrared divergent and therefore the splitting is gener-

[1]The decay to gauge bosons is suppressed since one particle has to be produced off-shell.

ated by the evolution of the active b-quark PDF using the DGLAP equations. The hard scattering process $gg \to t\bar{t}$ is calculated at matrix element level, while the splitting is generated at parton shower level. The current state of the art theoretical description of the $t\bar{t} + b\bar{b}$ process is the four flavour scheme. It treats the b-quark as a massive particle and decouples the b-quark PDF from the evolution. The calculation of the $gg \to t\bar{t} + b\bar{b}$ process at matrix element level is possible since the splitting is no longer collinear and infrared divergent. The full phase space of the splitting $g \to b\bar{b}$ is available and should result in a better description in the collinear and soft regions. The four flavour scheme is a fully inclusive description of the $t\bar{t} + b\bar{b}$ process at NLO accuracy [12].

In this thesis the different modelling approaches to the $t\bar{t} + b\bar{b}$ background in the five and four flavour scheme are presented and compared. For this purpose, a set of MC simulations which describe the $t\bar{t}$ inclusive process in the five flavour scheme and the $t\bar{t} + b\bar{b}$ exclusive process in the four flavour scheme is used. The MC simulations differ by the used matrix element generator and parton shower models. They also differ by their choice of the renormalisation and factorisation scale, PDF set and hadronisation model. To update the current set of ATLAS MC simulations the new version of SHERPA2.2.10 is added [13].

This thesis is organised as follows. The Chapter 2 gives a short description of the theory of the SM and describes the electroweak unification, Higgs mechanism and the properties of the SM Higgs boson. The ATLAS detector and experiment at the LHC is explained in Chapter 3 and describes the different detector types and trigger system of ATLAS. For the simulation of particle physics a large knowledge of QCD is required to cover all aspects of the process produced in a high energy particle collision. The different simulation approaches and implementations are therefore presented in Chapter 4. The main object of this thesis, different approaches to the modelling of the top pair prodcution in association with a bottom quark pair, are explained in Chapter 5. The used set of MC simulations and their parameters are presented in Chapter 6 with a description of the object identification and event selection. The last chapter summarises the results and gives an outlook for future analyses of the comparison of different MC setups in terms of flavour composition and event kinematics.

2. The Standard Model of Particle Physics

The Standard Model of particle physics is the gauge theory of the gauge groups $U(1)_Y$, $SU(2)_L$ and $SU(3)_C$ in combination with the Higgs mechanism [3–7], describing the fundamental interactions between particles: the electromagnetic, the weak, and the strong interaction. However, gravity is not included yet. The particle content of the SM is shown in Figure 2.1, and contains two types of particles: fermions and bosons [14]. Fermions are particles with half-integer spin which obey the Fermi-Dirac statistics[1]. All fermions of the SM are spin-$\frac{1}{2}$ particles and are sorted in leptons and quarks. Leptons do not interact via the strong interaction as they do not carry colour charge. Charged leptons have an electric charge of $\mp 1e$ depending on being particles or anti-particles. The charged leptons are the electron, muon, and τ-lepton. Neutrinos are the electrically neutral weak isospin[2] partner of their corresponding charged lepton. Leptons are sorted into three generations which have the same quantum numbers but differ in mass. Neutrinos do not interact via the electromagnetic force as they are electrically neutral. Quarks [15] carry colour charge and therefore interact via the strong interaction. They are organised in three generations, too. The up-type quarks carry an electric charge of $+\frac{2}{3}e$, while the down-type ones carry an electric charge of $-\frac{1}{3}e$. Up- and down-type quarks from the same generation are isospin partners. Free quarks only exist at very high energies due to the effect of asymptotic freedom [16]. At low energies quarks always hadronise due to the effect of confinement [17]. They form mesons which contain one quark and one anti-quark and baryons which contain three quarks or three anti-quarks [18, 19]. The effects of confinement and asymptotic freedom are caused by the renormalisability of the SM which leads to running coupling constants [20–22]. Electrons, neutrinos, and protons are stable. A neutron is only stable if it is bound in a nucleus. Protons and neutrons consist of first generation quarks and form with electrons ordinary matter. Free fermions

[1] If a particle obeys to the Fermi-Dirac statistics it also obeys the Paulis Principle. These statements are equivalent.

[2] The weak isospin will be introduced in Chapter 2.1.

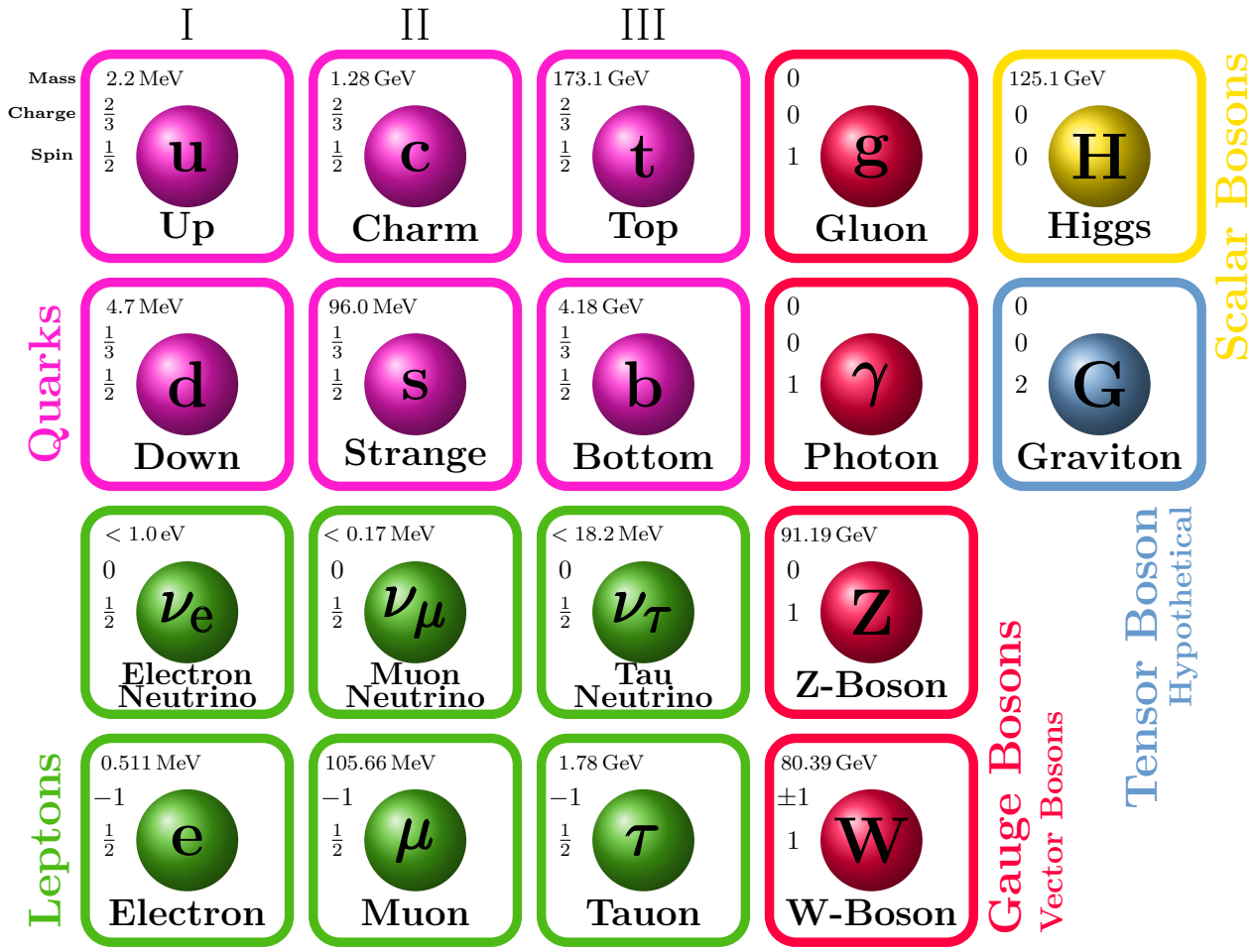

Figure 2.1.: The Standard Model is organised into three generations of quarks and leptons on the left side. The gauge bosons are shown on the right side with the Higgs boson. The values for the mass, spin, and electric charge are shown on the top left side of each particle [14]. The Graviton would complete the quantum description of the four fundamental forces. However, it is still a hypothetical particle and therefore not a part of the SM yet.

are described as spinor fields[3] ψ which satisfy the Dirac equation

$$(i\gamma^\mu \partial_\mu - m)\,\psi = 0 \tag{2.0.1}$$

with γ^μ as the Dirac matrices and m the mass of the fermion [24]. Bosons are spin integer particles and obey the Bose-Einstein statistics. The bosons are introduced by achieving gauge invariance, using Noether's theorem [25]. Every fundamental force is mediated by gauge bosons. The electromagnetic force is mediated by the massless photon. The

[3]In general, free fields are constructed using the Wigner classification. In the end, particles are unitary representation of the Poincarè group and are defined by their spin and mass [23].

weak force is mediated by the massive electric charged $W^{\pm}$ gauge bosons and the massive neutral Z^0 -boson. The strong force is mediated by colour charged massless gluons. All mentioned bosons have a spin of 1. They are called vector bosons due to their coupling nature to other particles. The Higgs boson is massive and it carries neither electric charge, colour charge nor weak isospin. Furthermore, it couples to mass. It is a scalar particle meaning its spin is 0.

2.1. Electroweak Unification

$W^{\pm}$ -bosons are maximally parity violating as shown by the Wu and Goldhaber experiments [26, 27]. $W^{\pm}$ -bosons couple exclusively to left handed particles and right handed anti-particles. The Z^0 -boson is parity violating but not maximal. It couples to both, left- and right-handed particles, but not equally. This inequality is shown by measurements of the forward-backward asymmetry [28]. In the Glashow-Weinberg-Salam theory [29–32] the electromagnetic force and weak force are unified. Left-handed particles are grouped in doublets with a weak isospin of $I = \frac{1}{2}$. Right-handed particles are grouped in singlets with a weak isospin of $I = 0$

$$\begin{bmatrix} u \\ d \end{bmatrix}_L \quad \begin{bmatrix} \nu_e \\ e \end{bmatrix}_L \quad \begin{bmatrix} c \\ s \end{bmatrix}_L \quad \begin{bmatrix} \nu_\mu \\ \mu \end{bmatrix}_L \quad \begin{bmatrix} t \\ b \end{bmatrix}_L \quad \begin{bmatrix} \nu_\tau \\ \tau \end{bmatrix}_{\mathrm{L}} \qquad u_R,\, d_R,\, e_R \quad c_R,\, s_R,\, \mu_R \quad t_R,\, b_R,\, \tau_R. \tag{2.1.1}$$

The electroweak force results from the $\mathrm{SU_L}(2) \times \mathrm{U_Y}(1)$ gauge symmetry. The first symmetry group acts only on left-handed doublets and acts trivially on right-handed singlets. The second part introduces the hypercharge Y which is defined by

$$\mathrm{Y} = 2 \cdot (q - I_3) \tag{2.1.2}$$

with the third component of the weak isospin I_3 and the electric charge q. The generators T^i for the $\mathrm{SU_L}(2)$ local gauge symmetry are the Pauli matrices $T^i = \frac{1}{2}\sigma^i$. To achieve local gauge invariance, the ordinary derivative has to be replaced by the covariant one

$$\partial_\mu \to \mathrm{D}_\mu = \partial_\mu - ig\vec{T} \cdot \vec{W}_\mu - ig' \frac{\mathrm{Y}}{2} B_\mu. \tag{2.1.3}$$

The physical field of a $W^{\pm}$ -bosons is the linear combination of the gauge field

$$W_{\mu}^{\pm} = \frac{1}{\sqrt{2}} \left(W_{\mu}^{1} \mp i W_{\mu}^{2} \right). \tag{2.1.4}$$

The physical fields for photons A_{μ} and Z^0 -bosons Z_{μ} are obtained by a rotation

$$\begin{bmatrix} A_{\mu} \\ Z_{\mu} \end{bmatrix} = \begin{bmatrix} \cos(\theta_W) & \sin(\theta_W) \\ -\sin(\theta_W) & \cos(\theta_W) \end{bmatrix} \begin{bmatrix} B_{\mu} \\ W_{\mu}^{3} \end{bmatrix} \tag{2.1.5}$$

through the Weinberg angle θ_W. The electroweak Lagrangian

$$\mathcal{L}_{\mathrm{EW}} = \sum_{f=l,q} \overline{f} i \gamma^{\mu} \mathrm{D}_{\mu} f - \frac{1}{4} W_{\mu\nu}^{k} W_{k}^{\mu\nu} - \frac{1}{4} \mathrm{B}_{\mu\nu} \mathrm{B}^{\mu\nu} \tag{2.1.6}$$

does not contain any gauge boson mass terms since these would spoil the symmetry. However, experiments show that massive gauge bosons do exist. The mass of the Z^0 -boson is $\mathrm{m}_{Z^0} = 91.2\,\mathrm{GeV}$ and the mass of the W -boson is $\mathrm{m}_{W} = 80.4\,\mathrm{GeV}$ [14]. The masses of the gauge bosons are explained by the Higgs mechanism, see Chapter 2.2.

2.2. The Higgs Mechanism

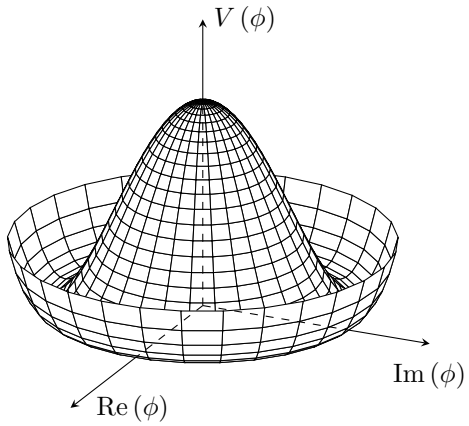

Figure 2.2.: The Higgs potential from Equation (2.2.1). To get an infinite set of minima, the parameters are required to be $\lambda > 0$ and $\mu^2 < 0$. Due to its distinctive shape it is sometimes called *mexican-hat* or *champagne-bottle* potential.

The Higgs Lagrangian is

$$\mathcal{L}_{\text{Higgs}} = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - \mu^2 \phi^\dagger \phi + \lambda \left(\phi^\dagger \phi\right)^2. \tag{2.2.1}$$

The fist term is a kinetic one. The last two terms are the Higgs potential. The first term of the Higgs potential can be interpreted as a mass term with an imaginary mass. The second term can be associated with a self-interaction term. If $\lambda > 0$ and $\mu^2 < 0$, a infinite set of minima occur at $\phi^\dagger \phi = -\frac{-\mu^2}{2\lambda} = \frac{\nu^2}{2}$. ν is called vacuum expectation value. The resulting potential is shown in Figure 2.2. The field $\phi(x)$ is an isospin doublet of complex fields[4]

$$\phi(x) = \begin{bmatrix} \phi^+(x) \\ \phi^0(x) \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \phi_1(x) + i\phi_2(x) \\ \phi_2(x) + i\phi_4(x) \end{bmatrix}. \tag{2.2.2}$$

Achieving an $U(1)_Y \times SU(2)_L$ local gauge invariance leads to gauge fields. The charged boson fields $W^\pm$ are already defined in Equation (2.1.4). The mass term for the charged boson fields determines the mass to be $m_{W^\pm} = \frac{1}{2}\nu g$. To determine the masses of the remaining two gauge bosons, the mass matrix has to be diagonalised. The resulting normalised eigenvectors are the physical gauge boson fields while eigenvalues are their masses

$$\begin{aligned} A_\mu &= \frac{g' W_\mu^3 + g B_\mu}{\sqrt{g^2 + g'^2}} \qquad & m_A &= 0 \\ Z_\mu &= \frac{g W_\mu^3 - g' B_\mu}{\sqrt{g^2 + g'^2}} \qquad & m_Z &= \frac{1}{2}\nu\sqrt{g^2 + g'^2}. \end{aligned} \tag{2.2.3}$$

The massless photon is described by the field A_μ while the massive Z^0 -boson is Z_μ. Comparing Equation (2.2.3) with Equation (2.1.5) shows the connection between the coupling constants and the Weinberg angle

$$\cos(\theta_W) = \frac{g'}{\sqrt{g^2 + g'^2}} \qquad \sin(\theta_W) = \frac{g}{\sqrt{g^2 + g'^2}}. \tag{2.2.4}$$

Furthermore the masses of the massive gauge bosons and the Weinberg angle are related

$$\cos(\theta_W) = \frac{m_{W^\pm}}{m_{Z^0}}. \tag{2.2.5}$$

[4] An isospin doublet is required, otherwise the $\text{SU}(2)_\text{L}$ will act trivially on the field.

Fermions fields ψ acquire their mass via a Yuakawa interaction

$$\mathcal{L}_{\text{Yukawa}} = \sum_{\substack{\text{fermion} \\ \text{types}}} g_{\text{Fermion}} \left(\overline{\psi}_{\text{L}} \phi \psi_{\text{R}} + \overline{\psi}_{\text{R}} \phi \psi_{\text{L}} \right) \tag{2.2.6}$$

with the Higgs field ϕ. The $\text{SU}(2)_{\text{L}}$ gauge group acts trivially on the interaction terms $\overline{\psi}_{\text{L}} \phi \psi_{\text{R}}$ and $\overline{\psi}_{\text{R}} \phi \psi_{\text{L}}$ since these are singlets. To get the eigenvalues, the coupling matrices g_{Fermion} is diagonalised. For leptons this transformation is trivial and does not change the lepton fields. For quarks this transformation mixes the weak eigenstates with the mass eigenstates. The mixing is described by the Cabibbo-Kobayashi-Maskawa (CKM) matrix and its three mixing angles [33]. The fourth parameter is a complex phase which is responsible for CP-violation. The absolute values of the CKM matrix are [14]

$$\begin{bmatrix} |V_{\text{ud}}| & |V_{\text{us}}| & |V_{\text{ub}}| \\ |V_{\text{cd}}| & |V_{\text{cs}}| & |V_{\text{cb}}| \\ |V_{\text{td}}| & |V_{\text{ts}}| & |V_{\text{tb}}| \end{bmatrix} = \begin{bmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.00014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.00016 & 0.0412^{+0.0011}_{-0.0005} \\ 0.00867^{+0.00029}_{-0.00031} & 0.0404^{+0.0011}_{-0.0005} & 0.99146^{+0.000021}_{-0.000046} \end{bmatrix}. \tag{2.2.7}$$

From the CKM matrix one can deduce that a top quark will almost always decay into a bottom quark as the CKM matrix element $|V_{\text{tb}}|$ is close to unity. In general, flavour changes will occur within the same generation since the diagonal of the CKM matrix is close to unity. The fermion masses are proportional to the coupling

$$m_{\text{Fermion}} = g_{\text{Fermion}} \frac{\nu}{\sqrt{2}} \tag{2.2.8}$$

therefore a Higgs boson will couple more strongly to a top quark than to an electron.

2.2.1. Properties of the Higgs Boson

The Standard Model Higgs boson[5] was discovered in 2012 at the LHC by the ATLAS [1] and CMS [2] experiments using the $Z^0 Z^0 \to 4\ell$, $WW \to \ell\nu_\ell \ell\nu_\ell$ and $\gamma\gamma$ decay channels. The mass of the Higgs boson is $125.10 \pm 0.42\,\text{GeV}$ [34]. The Higgs boson is a spin 0 and CP even particle [35, 36] with a lifetime of $\tau_{\text{Higgs}} \approx 10^{-22}\,\text{s}$ [14]. At the LHC with a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ three production modes are dominant. The first one with a cross-section of $\sigma_{\text{GGF,13 TeV}} = 44.1\,\text{pb}$ is the so called gluon-gluon fusion (GGF) [14]. This production mode is most dominant, because of the high number of gluons in a proton-proton collision. Since gluons are massless, Higgs bosons do not couple to them directly.

[5] In the following, the term Higgs boson always refers to the Standard Model Higgs boson.

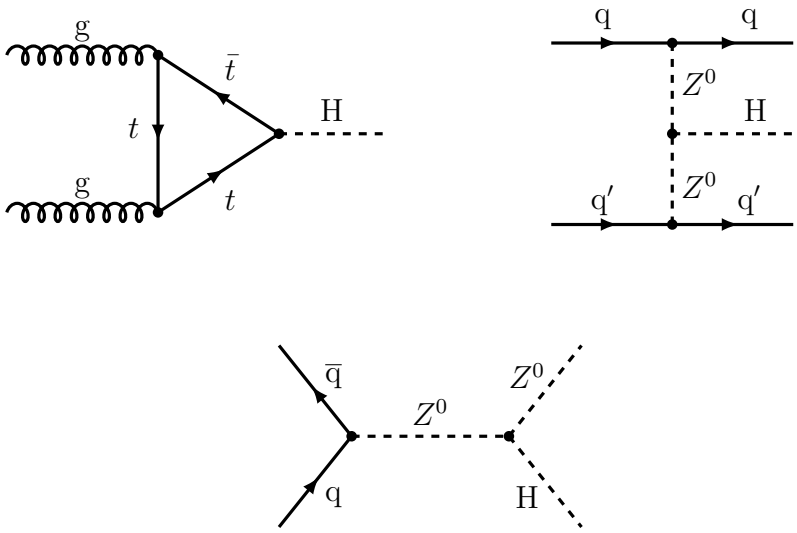

Figure 2.3.: The leading order Feynman diagrams for the production of a Higgs boson: **(a)** gluon gluon fusion (GGF), **(b)** vector boson fusion (VBF) and **(c)** Higgs strahlung. The Higgs production via gluon gluon fusion is mediated by a quark loop. A top quark loop is most likely. In the case of the vector boson fusion the production can be mediated by two W -bosons and suitable final and initial state quarks as well.

The production happens via a quark loop. Top and bottom quarks are preferred due to their high mass. The corresponding Feynman diagram is shown in Figure 2.3**(a)**. The second most dominant process at $\sqrt{s} = 13\,\text{TeV}$ is the vector boson fusion (VBF) with a cross-section of $\sigma_{\text{VBF},13\,\text{TeV}} = 3.78\,\text{pb}$ [14]. The corresponding Feynman diagram is shown in Figure 2.3**(b)**. The initial quarks radiate two vector bosons V (two Z^0 or W^+ and W^-) which merge to a Higgs boson. There are no loops at leading order so Higgs bosons are created directly. The third most dominant production mode at $\sqrt{s} = 13\,\text{TeV}$ is the Higgs-strahlung (HS) with a cross-section of $\sigma_{\text{HS},13\,\text{TeV}} = 0.88\,\text{pb}$ [14]. The corresponding Feynman diagram is shown in Figure 2.3**(c)**. A vector boson radiates a Higgs boson. For vector boson production there are always quarks in the final states. For Higgs-strahlung it depends on the decay of the vector boson. In the case of a leptonic decay $Z^0 \to \ell\ell$ or $W \to \ell\nu_\ell$, the multi-jet background can be reduced. In Figure 2.4**(a)** the predicted cross-sections of different types of Higgs productions as a function of the center-of-mass energy $\sqrt{s}$ are shown. In all cases the mass of the Higgs boson is assumed to be $m_{\text{Higgs}} = 125\,\text{GeV}$. At tree level the Higgs boson can decay into pairs of fermions and pairs of massive gauge bosons. The Higgs boson can also decay into massless particles like photons and gluons. These decays are mediated by a loop of heavy particles like top or bottom quarks or $W^\pm$ -bosons. The branching ratios for various decays as a function of the Higgs boson mass are shown in Figure 2.4**(b)** and in Table 2.1 the branching ratios for a Higgs boson mass

of $m_{\text{Higgs}} = 125\,\text{GeV}$ are written down.
The decay of $H \to \tau\tau$ was measured by ATLAS and CMS [37–41]. Searches for the decay $H \to \mu\mu$ were done by ATLAS [42] and CMS showed evidence for the decay [43]. The direct decay of a Higgs boson to an electron is not yet observed. Other decay channels like $H \to q\bar{q}$ are hard to observe due to the multi-jet background. However, in 2017 CMS and in 2018 ATLAS observed the direct decay $H \to b\bar{b}$ [9, 11]. Searches for the decay $H \to c\bar{c}$ are pursued by ATLAS and CMS in 2018 and 2020 [44, 45].

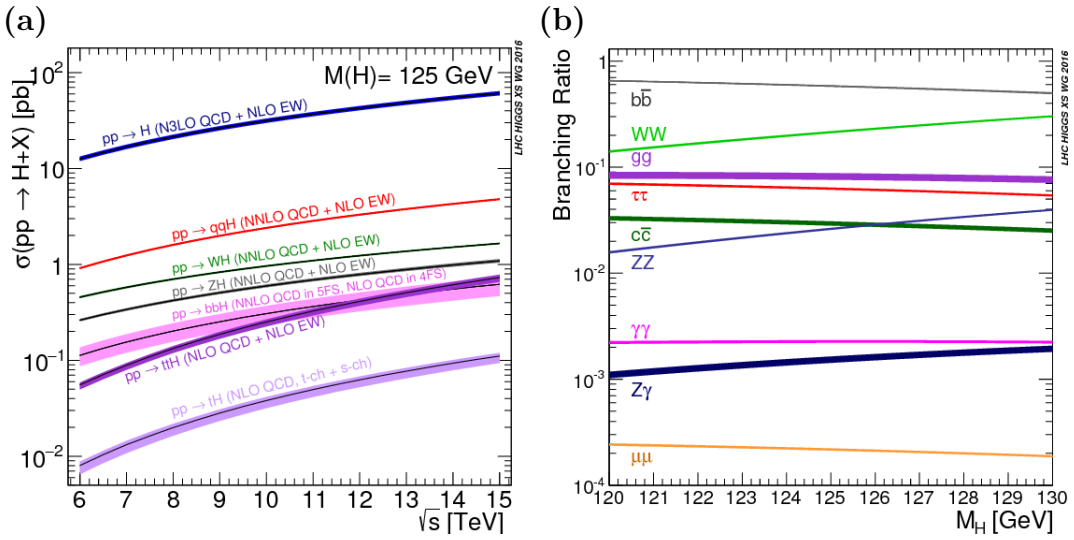

Figure 2.4.: In Figure **(a)** the production cross-sections of various production modes as a function of the center-of-mass energy $\sqrt{s}$ are shown [46]. In all cases the mass of the Higgs boson is $m_{\text{Higgs}} = 125\,\text{GeV}$. In Figure **(b)** the predicted branching ratios of various decay modes as a function of the Higgs mass are shown [46].

Table 2.1.: The branching ratios of various decay channels for a Higgs mass of $m_{\text{Higgs}} = 125\,\text{GeV}$ [46].

Decay channel	Branching ratio
$H \to \bar{b}b$	58.2 %
$H \to WW^*$	21.4 %
$H \to \text{gg}$	8.19 %
$H \to \tau\tau$	6.27 %
$H \to \bar{c}c$	2.89 %
$H \to ZZ^*$	2.62 %
$H \to \gamma\gamma$	0.227 %
$H \to \gamma Z$	0.153 %
$H \to \mu\mu$	0.022 %

3. Experimental Setup

3.1. The Large Hadron Collider

The Large Hadron Collider (LHC) [47] is a proton-proton and heavy ion synchrotron with a current center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ and a circumference of 27 km. It is located near Geneva and is operated by the European Organisation for Nuclear Research (CERN). A sketch of the accelerator complex of CERN can be seen in Figure 3.1. There are four collision points at the LHC which are occupied by four different experiments. LHCb works on CP violation measurements with hadrons that contain charm or bottom quarks [48]. The ALICE experiment uses heavy ions instead of protons as collision particles to create and study the properties of the quark-gluon plasma [49]. ATLAS and CMS are multi-purpose detectors [50, 51]. The LHC was upgraded many times in the last years. In 2010, the original center-of-mass energy was $\sqrt{s} = 7\,\text{TeV}$. After an upgrade in 2012 the center-of-mass energy was $\sqrt{s} = 8\,\text{TeV}$. The data detecting period from 2010 until 2012 is called Run 1. Run 2 began in 2015 when the center-of-mass energy was increased to $\sqrt{s} = 13\,\text{TeV}$. The LHC has the highest center-of-mass energy of all accelerators until today.

3.2. Luminosity

The number of events N for a certain process with a given cross-section σ and integrated luminosity $\int \mathcal{L}\,\mathrm{d}t$ is given by

$$N = \sigma \int \mathcal{L}\mathrm{d}t. \tag{3.2.1}$$

The instantaneous luminosity is

$$\mathcal{L} = f\frac{N_b n^2}{4\pi\sigma_\mathrm{x}\sigma_\mathrm{y}}. \tag{3.2.2}$$

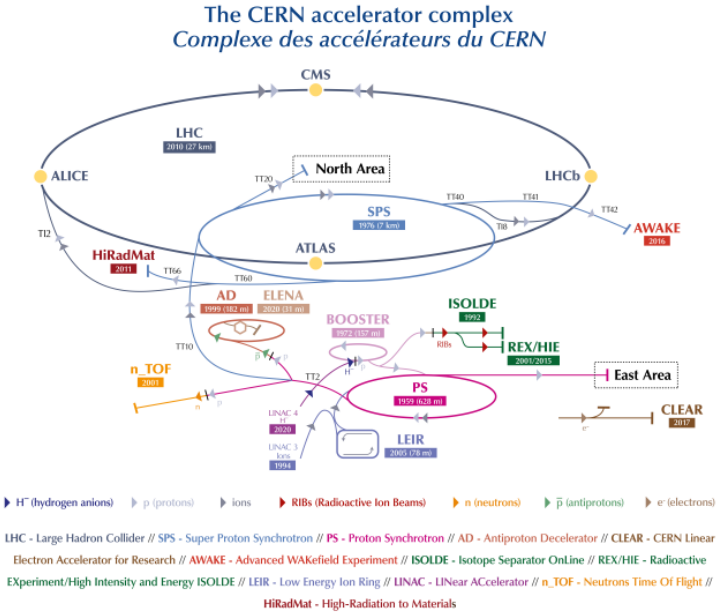

Figure 3.1.: The accelerator complex of CERN and the location of the largest experiments ALICE, LHCb, CMS and ATLAS at the LHC. ©CERN

The frequency of bunch crossings is described by f and is approximately 11 kHz, the number of protons per bunch is given by n, and N_b describes the number of bunches in the LHC. Furthermore, it is assumed that the beams have 2D-Gaussian shapes in the transverse planes, so σ_x and σ_y are the widths of the beams.

3.3. The ATLAS Coordinate System

The ATLAS coordinate system has its origin at the interaction point (IP) in the middle of the detector [50]. The x-axis points to the center of the LHC while the y-axis points upwards. The z-axis agrees with the beam direction. One variable of ATLAS is the transverse momentum which is defined by

$$p_T = \sqrt{p_x^2 + p_y^2}. \tag{3.3.1}$$

The ATLAS detector has a cylindrical shape. Therefore the azimuthal angle ϕ and polar angle θ are reasonable coordinates. The azimuthal angle is used as an ATLAS coordinate. From the polar angle, which is the angle between the z-axis and the outgoing particle the pseudorapidity η is determined. Furthermore it can be expressed by the momentum of a

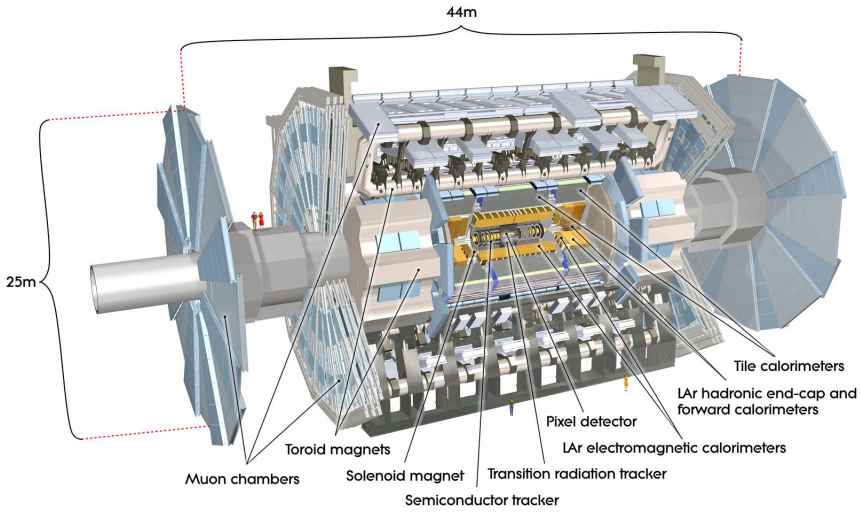

Figure 3.2.: A sketch of the ATLAS detector. The subcomponents are labelled and a human for scale, height and length are shown [50].

particle:

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right) = \frac{1}{2}\ln\left(\frac{|\vec{p}| + p_z}{|\vec{p}| - p_z}\right). \tag{3.3.2}$$

In the relativistic limit $m \ll |\vec{p}| \Rightarrow E \approx |\vec{p}|$ the pseudorapidity is equal to the rapidity

$$y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right). \tag{3.3.3}$$

The distance ΔR in the η-ϕ-plane is calculated using the Pythagorean theorem:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \tag{3.3.4}$$

The distance ΔR is invariant under boosts in the z-direction[1].

3.4. The ATLAS Detector

The ATLAS detector is 44 m long and 25 m high, making it the largest detector at CERN. However, with a weight of 7000 tonnes, it is lighter than the CMS detector. The ATLAS detector consists of many layers [50]. Each layer fulfils a specific function for the detection.

[1]The pseudorapidity η is not Lorentz invariant, but differences in pseudorapidity $\Delta\eta$ are Lorentz invariant for boosts along the z-axis. The azimuthal angle ϕ and differences in azimuthal angle $\Delta\phi$ are always Lorentz invariant since they are perpendicular to the z direction.

A sketch of the ATLAS detector is shown in Figure 3.2. Starting from the inside, the inner detector is used to detect tracks to identify charged particles. The inner detector has three smaller subcomponents. The pixel detector is the closest to the IP and allows track measurements directly after a collision takes place. The Insertable B-Layer (IBL) is the innermost layer [52] of the pixel detector. Three further pixel layers are placed in the barrel and another three layers at each end-cap. The semiconductor tracker (SCT) is the second subcomponent and is placed right after the pixel detector. It is also made up of silicon but does not consist of pixel but of strips. Information of the pixel detector and SCT are combined to track charged particles. The last subcomponent is the transition radiation tracker (TRT). The main component of the TRT is a gas mixture of Xenon, Carbon dioxide and Oxygen. If a charged particle passes the TRT, it will excite the Xenon. From the transition radiation one can distinguish electrons and hadrons. A solenoid magnet with a magnetic field of 2 T is placed around the inner detector. The Lorentz force acts perpendicular on charged particles, so that one can determine the transverse momentum from the curvature of the track. The ATLAS detector has two types of calorimeters. The function of a calorimeter is to measure the energy of particles. The electromagnetic calorimeter (ECAL) measures the energy of electrons and photons and uses Argon as active material. If a particle passes through the ECAL, it will shower and deposit energy into the active material. Two main processes occur: Bremsstrahlung and electron-positron pair production. The calorimeter is not able to distinguish between an electron or a photon. Photons do not leave a signature in the inner detector because they do not carry electric charge. By combining the information from the ECAL and the inner detector one can identify photons and electrons. The second calorimeter is the hadronic calorimeter (HCAL). Its main purpose is to measure the energy of hadrons. There are three types of HCAL in the ATLAS detector. The first one is placed in the central region of the detector and uses iron as an absorber and plastic scintillators as active material. The second HCAL is placed at the end-caps and uses copper as absorbers and liquid Argon as active material. The last part is placed in the forward region and uses Tungsten as absorber material. Muons do not shower intensively in the ECAL or HCAL. The muon spectrometer is the last component of the ATLAS detector and is used to measure the energy of muons. The transverse momentum is measured using a 4 T toroidal magnetic field. In Table 3.1 the different resolutions and coverage are shown.

Table 3.1.: The resolution and coverages for different parts of the ATLAS detector. The energy and momentum is in GeV. The table is taken from [50].

Detector Part	Resultion	Coverage
Inner Detector	$\sigma_{p_\mathrm{T}}/p_\mathrm{T} = 0.05\,\% \cdot p_\mathrm{T} \otimes 1\,\%$	$-2.5 < \eta < 2.5$
ECAL	$\sigma_E/E = 10\,\%/\sqrt{E} \otimes 3\,\%$	$-3.2 < \eta < 3.2$
HCAL		
central region, end-caps	$\sigma_E/E = 50\,\%/\sqrt{E} \otimes 3\,\%$	$-3.2 < \eta < 3.2$
forward region	$\sigma_E/E = 100\,\%/\sqrt{E} \otimes 10\,\%$	$3.1 < \lvert\eta\rvert < 4.9$
muon spectrometer	$\sigma_{p_\mathrm{T}}/p_\mathrm{T} = 0.05\,\%$ at $p_\mathrm{T} = 1\,\mathrm{TeV}$	$-2.7 < \eta < 2.7$

3.5. The Trigger System

A trigger system is needed to handle the huge amount of data recorded by the ATLAS detector caused by a high bunch crossing rate with approximately 40 MHz and the topology of proton-proton collisions. The trigger system has two levels to reject unimportant events. The first trigger system is a hardware trigger which sorts events by the number of muons, number of jets, missing transverse momentum, which is an indicator of neutrinos, or total energy [53]. The event rate can be reduced by two orders of magnitude to 100 kHz. The second level trigger is the high-level trigger (HLT) [54]. It is a software trigger which uses reconstruction and signature algorithms to apply for example a specific transverse momentum threshold. The HLT can reduce the amount of data to 1 kHz. The remaining events are then stored.

4. Particle Physics and Simulation

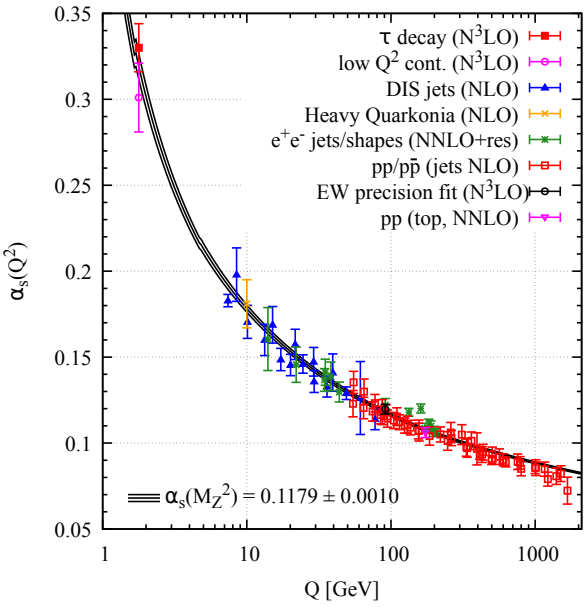

Figure 4.1.: The running coupling of the strong interaction as a function of the energy scale Q. The measurements were done by various groups [55–58]. The order of perturbative accuracy is shown in the brackets. This figure is generated by the Particle Data Group (PDG) and is taken from [14].

Simulations of collision events are necessary to estimate signal and background processes, different kinematic variables and control regions to get a better signal sensitivity. To get reasonable results, knowledge of QCD at different energy scales is important. The coupling strength α_S of the strong interaction is not a constant but a function of distance between two particles and the energy scale Q at which the interaction takes place. At high energies and small distances between two particles the coupling strength decreases. This is called *asymptotic freedom* [16]. The opposite is called *confinement*[1] and occurs at lower energies and higher distances [17]. The running coupling is a direct result of the

[1]Sometimes also described as *colour confinement*.

vacuum polarisation and is described in perturbative QCD via the *beta function* [59–63]

$$\mu^2 \frac{\partial \alpha_S}{\partial \mu^2} = \beta\left(\alpha_S\right) = -\left(\beta_0 \alpha_S^2 + \beta_1 \alpha_S^3 + ...\right) \quad \text{with } \beta_0 = \frac{11 n_c - 2 n_f}{12\pi}. \tag{4.0.1}$$

The number of quark flavours is n_f, while $n_c = 3$ is the number of quark colours. Taking just the first order, one gets the one-loop running coupling evolution equation with the solution

$$\alpha_S\left(\mu^2\right) = \frac{\alpha_S\left(\mu_0^2\right)}{1 + \frac{\alpha_S\left(\mu_0^2\right)}{4\pi}\left(11 - \frac{2 n_f}{3}\right) \ln\left(\frac{\mu^2}{\mu_0^2}\right)}. \tag{4.0.2}$$

Below the Landau pole Λ_{QCD}^2 the coupling strength is so large that perturbation theory breaks down. The value of the Landau pole depends on the used renormalisation scheme and the order of the perturbation series. Typically a value around the mass of the proton is used, so $\Lambda_{\mathrm{QCD}} \approx 1\,\mathrm{GeV}$. In Figure 4.1, the coupling strength α_S is shown as a function of the energy scale Q^2. Due to the behaviour of the coupling strength α_S at different energy scales, different approaches for simulations have to be made.

At high energies, the coupling strength α_S is low and a specific process, e.g. $gg \to t\bar{t}H$ or $gg \to t\bar{t} + b\bar{b}$, can be calculated perturbatively. After the hard interaction process the energy scale decreases. As a result the coupling strength α_S increases and processes like parton showering and hadronisation can no longer be calculated purely perturbatively, hence a different approach is needed. Parton showering happens at an energy scale of some 10 GeV and features mostly gluon radiation. Hadronisation describes partons during transition to colourless hadrons at an energy scale below the parton shower cut-off, which has a value of 1 GeV. Matrix element generation of the hard process, parton showering and hadronisation are steps of the simulation chain. The further steps are detector simulation and reconstruction. Since this thesis only describes events on parton (matrix element generation) and particle level (matrix element generation+parton shower+hadronisation) the last step is not taken into account. In Figure 4.2 a sketch of a typical process in pp collisions is shown. The hard scattering process is indicated as a dark red dot which was caused by the incoming protons. The violet part is a secondary process and is referred to as *underlying events*. The parton shower of the primary process is coloured in red and shows mostly gluon radiation and some gluon splitting to quarks. The outer region of the upper side is shown in green and describes the hadronisation to stable final states and colourless hadrons. The wiggly yellow lines are bremsstrahlung caused by photons.

The generation of matrix elements is based on perturbation theory and fixed order calcu-

lations. Of course, it is not possible to calculate an observable $\mathcal{O}$ for all orders. To control higher orders of the perturbation series, the renormalisation scale μ_R^2 and factorisation scale μ_F^2 are introduced which also fulfil the purpose to deal with ultraviolet (UV) and infrared (IR) divergences. UV divergences are caused by momentum integration for loop diagrams which occur in indirect couplings like $H \to gg$ or due to self-interaction. The renormalisation scale μ_R^2 is a cut-off parameter which cures the UV divergences. IR divergences happen in two cases: If a particle reaches a zero momentum value or if a massless particle radiates another massless particle. The first one is solved using the Konoshita-Lee-Nauenberg theorem [64, 65], while the latter one requires the introduction of the factorisation scale μ_F^2 which also separates the hard and soft QCD interactions. The PDF and the running coupling α_S are now a function of the renormalisation and factorisation scale. These parameters are non-physical and in the case of an exact expansion of the perturbation series an observable $\mathcal{O}$ is independent of these parameters

$$\frac{\mathrm{d}\mathcal{O}}{\mathrm{d}\ln(\mu^2)} = \frac{\mathrm{d}\mathcal{O}_{\mathrm{LO}}}{\mathrm{d}\ln(\mu^2)} + \frac{\mathrm{d}\mathcal{O}_{\mathrm{NLO}}}{\mathrm{d}\ln(\mu^2)} + \sum_{k=2}^{\infty} \frac{\mathrm{d}\mathcal{O}_{\mathrm{N^kLO}}}{\mathrm{d}\ln(\mu^2)} = 0. \tag{4.0.3}$$

However, since an exact calculation is not possible, observables are usually calculated at leading-order (LO), next-to-leading-order (NLO) or next-to-next-to-leading-order (NNLO). The renormalisation scale is normally chosen to be equal to the scale of the hard scattering process and is calculated individually for each event. The factorisation scale is usually set to the same value as the renormalisation scale.

4.1. Matrix Element Generation

A scattering process is described by its cross-section σ which is calculated using *Fermi's Golden Rule*

$$\sigma = \sum_{k=0}^{\infty} \int_{m+k} SF d\Omega_{m+k} \left| \sum_{l=0}^{\infty} \mathcal{M}_{m+k}^{(l)} \right|^2 \tag{4.1.1}$$

with m the number of final state particles, l the number of loop corrections and k as the number of additional emissions. S is a symmetry factor for identical final state processes and F the flux factor. The matrix element $\mathcal{M}_{m+k}^{(l)}$ is the transition amplitude between the initial state and final state. The Lorentz-invariant phase space factor is

$$d\Omega_{m+k} = \prod_{i=0}^{m+k} \frac{d^3 p_i}{(2\pi)^3 \, 2E_i}. \tag{4.1.2}$$

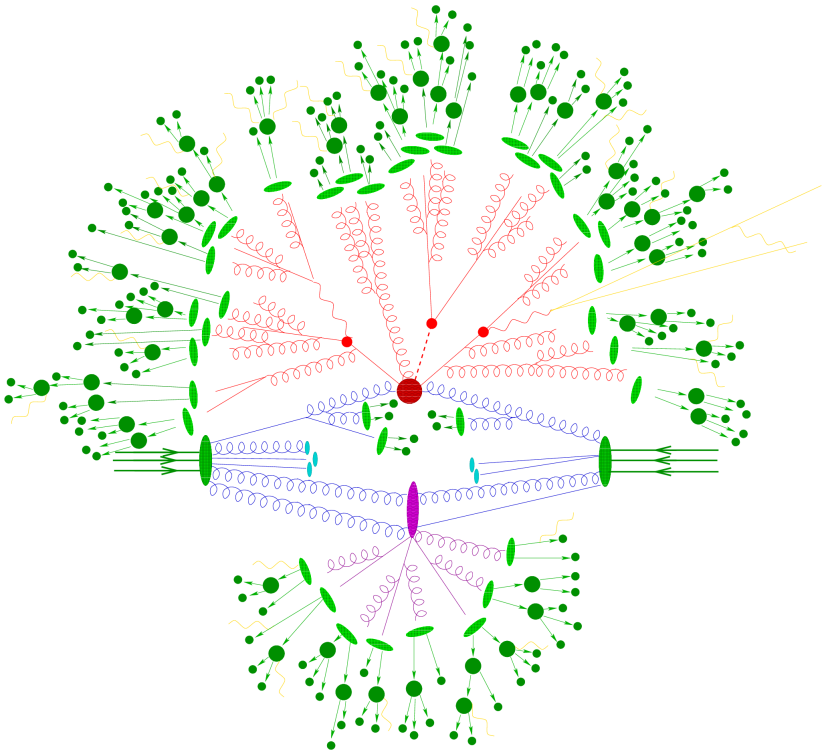

Figure 4.2.: The simulation chain for a typical process at hadron colliders, here it is a $t\bar{t}H$ production. The incoming protons cause a hard scattering process coloured in dark red. The parton shower process features mostly gluon radiation and gluon splitting and is coloured in red. The hadronisation of partons to stable final state hadrons is indicated in green in the upper side of the sketch. The yellow wiggly lines describe photon bremsstrahlung. This figure is taken from [66].

The matrix element (ME) is usually calculated at LO, NLO or NNLO.

4.1.1. Parton-Parton Interactions

Leptons are considered to be fundamental particles implying no substructure or subcomponents. Therefore the center-of-mass energy during a lepton collision is well defined. Hadrons like protons are composite particles. The substructure particles like valence quarks, sea quarks and gluons are interacting and their exact fraction of momentum and energy remains unknown. Yet, it is possible to calculate a cross section using the factorisation theorem [67]. The cross-section σ for a process $pp \to X$ is given by

$$\sigma_{pp\to X} = \sum_{\substack{\text{sum of}\\ \text{partons a,b}}} \int_0^1 \mathrm{d}x_a \int_0^1 \mathrm{d}x_b q_a\left(x_a, \mu_F^2\right) q_b\left(x_b, \mu_F^2\right) \hat{\sigma}_{ab\to X}\left(x_a p_a, x_b p_b, \mu_F^2, \mu_R^2\right) \quad (4.1.3)$$

with a and b as two interacting partons. The parton distribution function $q\left(x, \mu_F^2\right)$ describes the probability to find a specific parton in the proton, carrying a fraction x of the momentum of the proton at a momentum scale μ_F^2. The PDFs cannot be calculated due to the non-perturbative nature of QCD at low energy scales which occur inside a proton. However, the DGLAP equations [68–70] describes the PDF evolution for different energy scales Q:

$$\begin{aligned}
\frac{\mathrm{d}q_i\left(x, Q^2\right)}{\mathrm{d}\ln\left(Q^2\right)} &= \frac{\alpha_S\left(Q^2\right)}{2\pi} \int_x^1 \frac{d\xi}{\xi} \left[q_i\left(\xi, Q^2\right) P_{qq}\left(\frac{x}{\xi}\right) + g\left(\xi, Q^2\right) P_{qg}\left(\frac{x}{\xi}\right) \right] \\
\frac{\mathrm{d}g\left(x, Q^2\right)}{\mathrm{d}\ln\left(Q^2\right)} &= \frac{\alpha_S\left(Q^2\right)}{2\pi} \int_x^1 \frac{d\xi}{\xi} \left[\sum_i q_i\left(\xi, Q^2\right) P_{gq}\left(\frac{x}{\xi}\right) + g\left(\xi, Q^2\right) P_{gg}\left(\frac{x}{\xi}\right) \right]
\end{aligned} \tag{4.1.4}$$

with $q_i\left(x, Q^2\right)$ as a quark PDF, $g\left(x, Q^2\right)$ as the gluon PDF and $P\left(\frac{x}{\xi}\right)$ as the splitting functions

$$\begin{aligned}
P_{qg}(z) &= C_F \frac{1+z^2}{1-z} && C_F = \frac{4}{3} \\
P_{gq}(z) &= C_F \frac{1+(-z)^2}{x} && \\
P_{qg}(z) &= 2C_A \left[\frac{z}{1-z} + \frac{1-z}{z} + x(1-z) \right] && C_A = 3 \\
P_{qg}(z) &= T_R \left[z^2 + (1-z)^2 \right] && T_R = \frac{1}{2}.
\end{aligned} \tag{4.1.5}$$

With measurements one can determine the PDF at a given energy scale and extrapolate it with the DGLAP equations to another one. The PDF for various quarks and the gluon are shown in Figure 4.3. The PDFs are taken from the *CT14NNLO* PDF set [71]. The gluon PDF has to be rescaled by a factor of five. The PDF for gluons is, compared to a sea quark, more dense and therefore the gluon-gluon fusion production of a Higgs boson or a top quark pair is more likely.

4.2. Parton Shower

After the hard scattering, parton showering is the dominant process and is calculated with parton shower algorithms (PS). Parton showering describes the successive emission of gluons and photons[2] from the ME final state of the hard process. Using the factorisation theorem from Equation (4.1.3) in the collinear limit and without any virtual corrections

[2]The gluons can split to quark pairs, while the photons can split to gluon and charged lepton pairs.

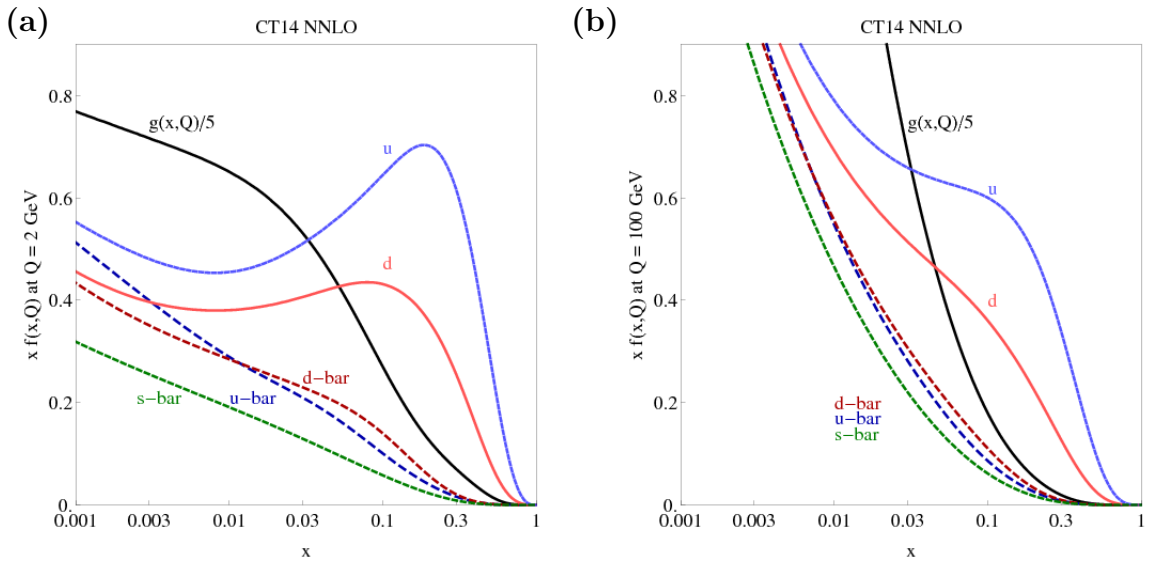

Figure 4.3.: The parton distribution function (PDF) for up, down and strange quarks and gluons. The gluon PDF is rescaled by a factor of five. In **(a)** the energy scale is $Q^2 = 4\,\mathrm{GeV}^2$ and in **(b)** it is $Q^2 = 10^4\,\mathrm{GeV}^2$ [71].

the evolution equation for the PDFs is

$$\mu_F^2 \frac{\mathrm{d}q_a\left(x,\mu_F^2\right)}{\mathrm{d}\mu_F^2} = \sum_{b\in\{q,g\}} \int_x^1 \frac{\mathrm{d}z}{z} \frac{\alpha_S}{2\pi} P_{ba}\left(z\right) q_b\left(x,\mu_F^2\right) \tag{4.2.1}$$

with the splitting functions from Equation (4.1.5). The evolution equation in Equation (4.2.1) describes the PDFs with varying factorisation scale[3]. There are no bare partons in the final state of the simulation chain. Due to confinement they will form colourless bound states which are called hadrons. This restricts the energy scale at which parton showering can happen since partons below this threshold can no longer be resolved. This restriction condition is called parton shower cut-off and is approximately 1 GeV. At this momentum scale, which is sometimes called Landau pole Λ_{QCD}, the coupling strength is too large and perturbation theory is not allowed to be used. Different approaches are needed to describe partons at lower momentum scales, see Chapter 4.3. For implementation purposes the Sudakov form factor is introduced which can be derived using Equation (4.2.1). The Sudakov form factor

$$\Delta_a\left(\mu_{F,max}^2,\mu_{F,min}^2\right) = \exp\left(-\sum_{b\in\{q,g\}} \int_{\mu_{F,min}}^{\mu_{F,max}} \int_{z_{min}}^{z_{max}} \mathrm{d}P_{ab}\left(z\right)\right) \tag{4.2.2}$$

[3]In this context the energy scale is described as the factorisation scale.

describes the probability that a parton will not split into other partons between two momentum scales $\mu^2_{F,max}$ and $\mu^2_{F,min}$. The implementation reads as follow: starting from a given momentum scale $\mu^2_{F,max}$ the parton radiates other partons which can be calculated using the Sudakov form factor from Equation (4.2.2). For the next step the lower momentum scale $\mu^2_{F,min}$ is used as the initial scale. This process is repeated until the momentum scale is below the parton shower cut-off scale of 1 GeV.

4.2.1. Matching NLO calculations with Parton Showers

The matching of matrix element calculations with the PS is not trivial and phenomena like double counting of final states can occur. There are two methods for matching: the POWHEG [72] method and the aMC@NLO [73] method. Both use a transverse momentum ordering. The hardest emission of the NLO matrix element calculation is matched with the PS first. It is worth to mention that PS calculates emissions in the collinear approximation. In general, an observable $\mathcal{O}$ with NLO matrix element calculation accuracy matched with a parton shower is given by

$$\begin{aligned}\langle\mathcal{O}\rangle = &\int \mathrm{d}\Phi_n \bar{B}\left(\Phi_n\right)\left[\mathcal{O}\left(\Phi_n\right)\Delta_{t_0} + \int \mathrm{d}\Phi_r \mathcal{O}\left(\Phi_n,\Phi_r\right)\Delta_t \frac{R\left(\Phi_n,\Phi_r\right)}{B\left(\Phi_n\right)}\right] \\ &+ \int \mathrm{d}\Phi_{n+1}\mathcal{O}\left[\Phi_{n+1}\right]\left(R\left(\Phi_{n+1}\right) - R^s\left(\Phi_{n+1}\right)\right]\end{aligned} \tag{4.2.3}$$

with B as the Born, V as the virtual emission and R as the real emission ME, while R^s is the real emission ME in the soft and collinear limit and

$$\begin{aligned}\bar{B}\left(\Phi_n\right) &= B\left(\Phi_n\right) + V\left(\Phi_n\right) + \int \mathrm{d}\Phi_r \left[R\left(\Phi_n,\Phi_r\right) - C\left(\Phi_n,\Phi_r\right)\right] \\ \Delta_t &= \exp\left(-\int \mathrm{d}\Phi'_r \frac{R\left(\Phi_n,\Phi'_r\right)}{B\left(\Phi_n\right)}\theta\left(t'-t\right)\right).\end{aligned} \tag{4.2.4}$$

The n particle final state phase space is given by Φ_n and Φ_r. In the case of a real emission the final state particle number increases to $n+1$. The main difference between the POWHEG method and aMC@NLO method lays in the different treatment of the last term in Equation (4.2.3) and the scale definition. In aMC@NLO the difference $R\left(\Phi_{n+1}\right) - R^s\left(\Phi_{n+1}\right)$ can be negative which results in negative weights. Furthermore aMC@NLO uses a resummation scale μ_Q which also defines the shower starting scale. For POWHEG the weights are always positive and a damping function which ranges from 0 to 1 controls the emission. The damping functions depends on the damping parameter h_{damp} and the p_T of the hardest emission.

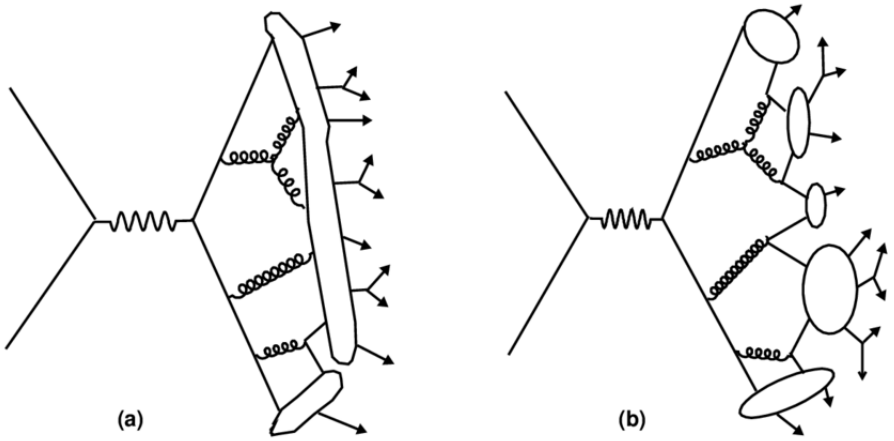

Figure 4.4.: On the left side a sketch for a typical process for the Lund string model is shown, while on the right side the cluster hadronisation model is shown. The figure is taken from [74].

4.3. Hadronisation

Hadronisation happens at an energy scale below the parton shower cut-off $\approx 1\,\mathrm{GeV}$. At small energy scales and large distances colour confinement occurs and partons will form colourless hadrons. Unfortunately, at such low energy scales the coupling constant α_S increases and is close to unity, see Figure 4.1. Perturbation theory requires a small coupling constant, therefore new methods to describe partons are necessary. Commonly, there are two hadronisation models: The Lund string model [75, 76] and the cluster model [77, 78].

Lund String Model

At low energies the potential is described as a function of the distance r between two partons

$$V_{\mathrm{QCD}} \approx -\frac{4}{3}\frac{\alpha_S}{r} + \kappa r. \tag{4.3.1}$$

At large distances the potential shows a proportional behaviour with respect to the distance r times a constant $\kappa \approx 1\frac{\mathrm{GeV}}{\mathrm{fm}}$. The Lund string model only uses the linear term of the potential. It can be interpreted as the colour flux between two partons. When the potential becomes large enough the string breaks and produces a $q\bar{q}$ pair to form hadrons. The hadron carries some momentum fraction z calculated with the Lund symmetric fragmentation function

$$f(z) = N\frac{(1-z)^a}{z}\exp\left(-\frac{bm_T^2}{z}\right) \tag{4.3.2}$$

with some normalisation constant N, the transverse mass of the produced hadron $m_T = \sqrt{m^2 + p_T^2}$ and two parameters a and b. The hadronisation process is terminated if no momentum fraction is left and if there are only on-shell hadrons in the final state. A sketch of the Lund string model is shown in Figure 4.4**(a)**.

The Cluster Model

The cluster model uses the concept of colour pre-confinement of perturbative QCD. After parton shower stops, gluons will start splitting into colour-singlet quark pairs $g \to q\bar{q}$. In general, gluons are allowed to decay into any quark flavour depending on the available phase space. The singlets are then combined into clusters depending on the distance between the quark pairs. The resultant energy and transverse momentum of the cluster is the sum of all containing particles. Heavy clusters will decay to lighter clusters, while lighter clusters decay to final state on-shell hadrons. A sketch of the cluster model is shown in Figure 4.4**(b)**.

4.4. Monte Carlo Simulations

There are several implementations for matrix element generators and parton shower algorithms which contain a hadronisation model implementation to complete the simulation chain. As mentioned before, since this thesis focuses on studies on particle level, a detector simulation is not needed. The implementations feature different methods of matching the matrix element with the parton shower, different hadronisation models, see Chapter 4.3, and underlying event tune.

Powheg-Box

POWHEG is a NLO matrix element generator [73, 79]. POWHEG stands for *positive weight hardest emission generator*, see Chapter 4.2.1. The POWHEG method is used to match the matrix element with a parton shower algorithm: The hardest parton emission is generated and matched with the NLO matrix element.

MadGraph_aMc@nlo

MadGraph_aMc@nlo is a matrix element generator which uses the helicity amplitude formalism [72, 80, 81]. Matrix elements can be calculated at LO with MadEvent and up to NLO with aMc@nlo with NLO accuracy in QCD. MadGraph_aMc@nlo features a lot of other modules like MadSpin [82, 83], to calculate the spin correlation for massive particles like the top quark, or MadDarkMatter [84] which calculates the relic density for a given Dark Matter model.

Herwig 7

Herwig 7 is a general-purpose MC event generator and is used in this thesis as a parton shower algorithm [85, 86]. The distinctive feature of Herwig 7 is the colour coherence and opening angular ordering of the parton emissions. The cluster hadronisation model is used, see Chapter 4.3.

Pythia 8

Pythia 8 is a general-purpose MC event generator and is used in this thesis as a parton shower algorithm [87, 88]. It is based on the ordering in transverse momentum of the parton emissions. Pythia 8 uses the Lund string model for hadronisation, see Chapter 4.3.

Sherpa

Sherpa is a general-purpose MC event generator. It features a matrix element generator and a parton shower algorithm [89]. The Catani-Seymour formalism is used to describe soft radiation more precisely with the Cantani-Seymour counterterm [90]. Sherpa can be combined with OpenLoops to compute loop diagrams [91–93]. For hadronisation purposes the cluster model is used.

5. Higgs Production in Association with a Top Quark Pair

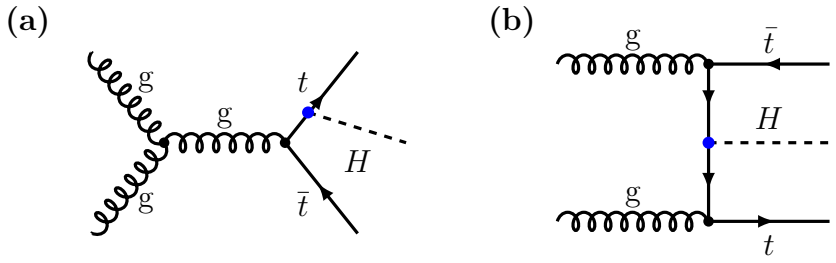

Figure 5.1.: The leading order Feynman diagrams for the production of a Higgs boson in association with a top quark pair. In **(a)** the s-channel production is shown while in **(b)** the t-channel production is shown. The Yukawa coupling is in both diagrams indicated with a blue dot.

The production of a Higgs boson in association with a top quark pair is important for the understanding of the Yukawa coupling and restrictions of Beyond SM theories. Since the Yukawa coupling is proportional to the mass, it is near to unity for the interaction between a Higgs boson and a top quark:

$$g_{\text{top}} = \sqrt{2}\frac{m_{\text{top}}}{\nu} \approx 1. \tag{5.0.1}$$

The Feynman diagram for the LO processes in the s- and t-channel are shown in Figure 5.1. The production of a top quark pair is with a cross-section of $\sigma_{t\bar{t}} = 832^{+20}_{-29}\,(\text{scale}) \pm 35\,(\text{PDF} + \alpha_S)$ pb at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ [94] one dominant process at the Lhc containing various subprocesses like $t\bar{t}Z$ or $t\bar{t}\gamma$ and the $t\bar{t}H$ process. The measured cross-section for the latter is $670 \pm 90\,(\text{stat.})^{+110}_{-100}\,(\text{syst.})$ fb at a center-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ [8] and therefore contributes only with less than 1 % to the top quark pair production. At lower center-of-mass energy of $\sqrt{s} = 7\,\text{TeV}$ and $\sqrt{s} = 8\,\text{TeV}$ the production cross-section for $t\bar{t}H$ decreases and searches by Atlas and Cms were difficult for multiple reasons [95, 96]. The decay mode of the Higgs to bottom quarks is chosen

due to its large branching ratio. It seems to be most promising to reach high statistics. However, the background is very large. There are three different decay channels of the $t\bar{t}H$ intermediate state. The corresponding LO Feynman diagram is shown in Figure 5.2. The top quark decay is mediated by the weak force into a W -boson and a b-quark since the CKM matrix element $|V_{\mathrm{tb}}|$ is close to unity. Depending on the decay of the W-boson, three decay channels are characterised:

- all-hadronic: with $\approx 46\,\%$ the all-hadronic channel has the largest branching ratio. Both W-bosons decay to quarks. The final state consists of eight jets with at least four b-jets.
- single-lepton: the second largest branching ratio, with $\approx 45\,\%$, has the single-lepton channel. It is characterised by one W -boson decaying leptonically, while the other decays hadronically. The final state consists of six jets with at least four b-jets, a charged lepton and a small amount of missing transverse momentum and energy caused by the neutrino which cannot be detected directly.
- dilepton: the smallest branching ratio of $\approx 9\,\%$ corresponds to the dilepton channel. Both W -bosons decay to leptons and their neutrinos. The final state is characterised by containing four b-jets, two charged leptons with opposite charge and a large amount of missing transverse momentum and energy.

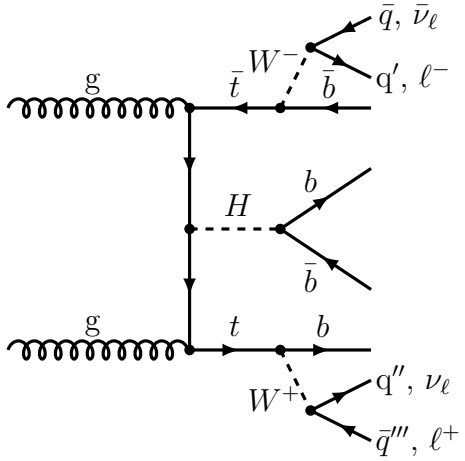

Figure 5.2.: The three different decay channels for the $t\bar{t}H$ intermediate state. Depending on the decay of the $W^{\pm}$-bosons, three different channels are identified: all-hadronic, single-lepton, and dilepton channel.

5.1. The $t\bar{t} + b\bar{b}$ Background

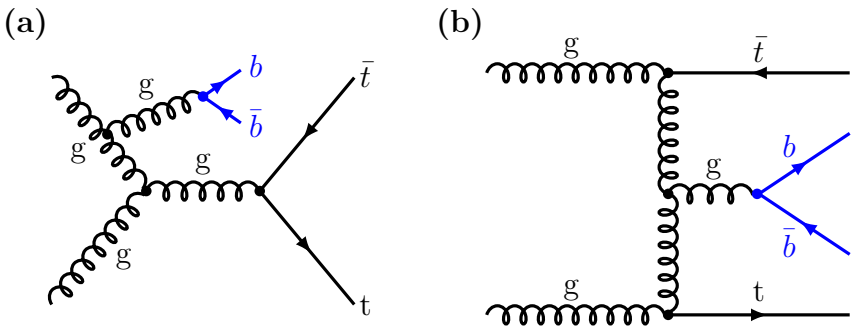

Figure 5.3.: Two leading order Feynman diagrams for the dominant background $t\bar{t}$+jets in the $t\bar{t} + b\bar{b}$ final state. In **(a)** an s-channel contribution is shown, while in **(b)** a t-channel contribution. The additional b-quarks are indicated in blue.

The search for and measurement of the Yukawa coupling in $t\bar{t}\left(H \to b\bar{b}\right)$ events suffers from large background processes. The most dominant background process is the $t\bar{t} + b\bar{b}$ background which has a two times larger cross-section and similar kinematics. Therefore it is necessary to understand the different uncertainties which arise from the theoretical modelling of the $t\bar{t} + b\bar{b}$ production. Two leading order Feynman diagrams are shown in Figure 5.3. The produced b-quarks not arising from the top quark decay and resulting jets are called *additional jets* and are indicated in blue in Figure 5.3.

At LO there are four QCD vertices which lead to an α_S^4 contribution for the calculation of the matrix element and therefore to a high sensitivity to the choice of the renormalisation scale. At LO these uncertainties are 70 % up to 80 % using factor-two scale variations. At higher orders in perturbation theory observables depend less on the renormalisation scale. Calculations at NLO are therefore necessary and reduce the uncertainties down to 20-30 % [12, 97–99].

The first calculations for the $gg \to t\bar{t}b\bar{b}$ production were done with POWHEL using the POWHEG method in the five-flavour scheme (5FS) [100, 101]. One huge disadvantage of the 5FS are collinear divergences resulting from the splitting $g \to b\bar{b}$ since the b-quark is treated as a massless particle. A different approach takes the mass of the b-quark into account. It is called four flavour scheme (4FS). In a very natural way the mass effect of the b-quark cures the infrared and collinear divergences of the splitting $g \to b\bar{b}$. The full phase space of the splitting is accessible, even regions where one b-quark would not be

resolvable. The first calculation of the $gg \to t\bar{t}b\bar{b}$ production in the 4FS was done with SHERPA+OPENLOOPS using the aMC@NLO matching method.

5.1.1. $t\bar{t} + b\bar{b}$ Production in the Five Flavour Scheme

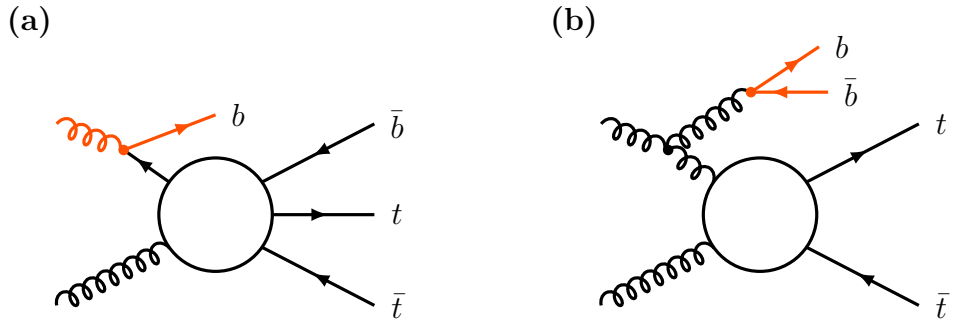

Figure 5.4.: Two typical production process for the $t\bar{t}+b\bar{b}$ in the five flavour scheme. On the left side the initial state splitting is shown, while on the right side the final state splitting is shown. The hard scattering process is calculated with ME and is indicated in black. The PS generations are shown in orange.

The typical production of $t\bar{t} + b\bar{b}$ happens at scales of $\mathcal{O}(100)$ GeV, while the produced b-quark pair tends to be collinear and soft with a transverse momentum of $\mathcal{O}(10)$ GeV. The production features mostly initial state (IS) and final state (FS) gluon splittings $g \to b\bar{b}$. The $t\bar{t} + b\bar{b}$ process is calculated with NLO matrix element generators which are merged with a parton shower algorithm.

In the five flavour scheme (5FS) the b-quark mass is neglected. The calculation of the $gg \to t\bar{t}+b\bar{b}$ process with matrix element generators is therefore complicated. Instead, the hard process $gg \to t\bar{t}g$ or $gb \to t\bar{t}\bar{b}$ will be calculated and matched with initial or final state gluon splitting. This procedure is shown in Figure 5.4. The black process is described at matrix element level, while the orange one is generated in the parton shower. On the right-hand side of Figure 5.4 the final state splitting is shown, while on the left-hand side the initial state splitting. The initial state splitting is mediated with the backwards evolution of the b-quark PDF. The resummation of large logarithmic divergences $\alpha_S \log\left(\frac{\mu_R^2}{m_b^2}\right)$ is absorbed in the PDF evolution of the b-quarks, showing a big advantage of the 5FS compared to the 4FS, see for a detailed description Chapter 5.1.2.

The interplay between the ME and PS is demonstrated in Figure 5.5. Again, the hard scattering process calculated with the matrix element generator is indicated in black, while

the parton shower is indicated in orange. For softer $b\bar{b}$ systems the PS models not only the $g \to q\bar{q}$ splitting but also the emitted soft gluon, see Figure 5.5**(a)**. Harder gluons are produced at matrix element level and the parton shower only models the splitting $g \to b\bar{b}$, as shown in Figure 5.5**(b)**. The final state b-jets will have a harder transverse momentum spectrum. The hardest spectrum corresponds to the full ME generation of the $t\bar{t}b\bar{b}$ final state, see Figure 5.5**(c)**. The PS does not contribute and therefore it does not exist in the 5FS. The different transverse momentum spectrum is a result of the different energy scales at which the ME and PS operate. ME calculations are done at much higher energy scales.

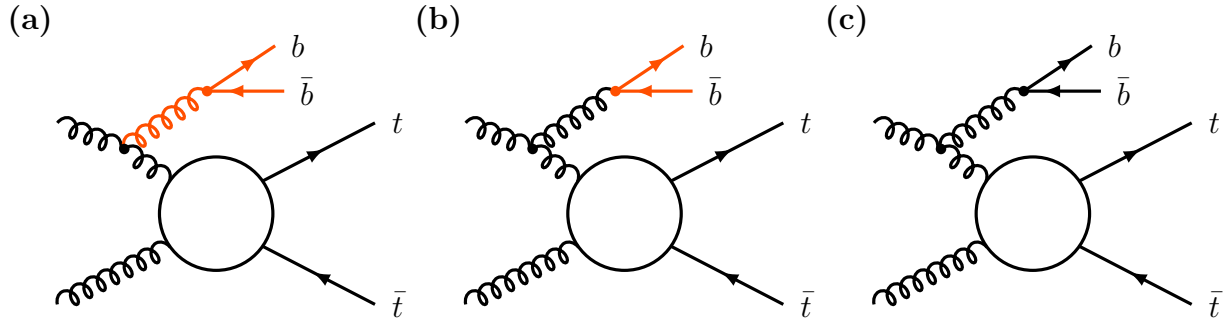

Figure 5.5.: The interplay between the ME and PS. The black process is calculated with the ME. The orange parts are done by the PS. From the left to the right the transverse momentum spectrum will get harder.

5.1.2. $t\bar{t} + b\bar{b}$ Production in the Four Flavour Scheme

To rely less on PS algorithms and their description of gluon splitting one should use the more precise calculations of the ME. To do so, the mass of the b-quark will not be neglected. As a direct result, the gluon splitting into massive particles $g \to b\bar{b}$ does not need any artificial cuts to avoid collinear and infrared divergences. Only logarithms are present in calculations: For the splitting $g \to b\bar{b}$ one has to integrate over the b-quark propagator with a mass term and at an energy scale Q^2

$$\mathcal{I}\left(Q^2\right) = \int_{t_{\max}}^{t_{\min}} \frac{\mathrm{d}t}{t - m_b^2} = \ln\left(\frac{t_{\max} - m_b^2}{t_{\min} - m_b^2}\right) = \ln\left(\frac{Q^2}{m_b^2}\right). \tag{5.1.1}$$

Instead of the hard process $gg \to t\bar{t}$, the process $gg \to t\bar{t}b\bar{b}$ with initial or final state gluon splitting will be calculated. In Figure 5.6 four typical processes for the 4FS are shown. The initial and final state gluon splittings are indicated in red. The splitting can be calculated in the entire phase space, even to the collinear regime, where the $b\bar{b}$ system

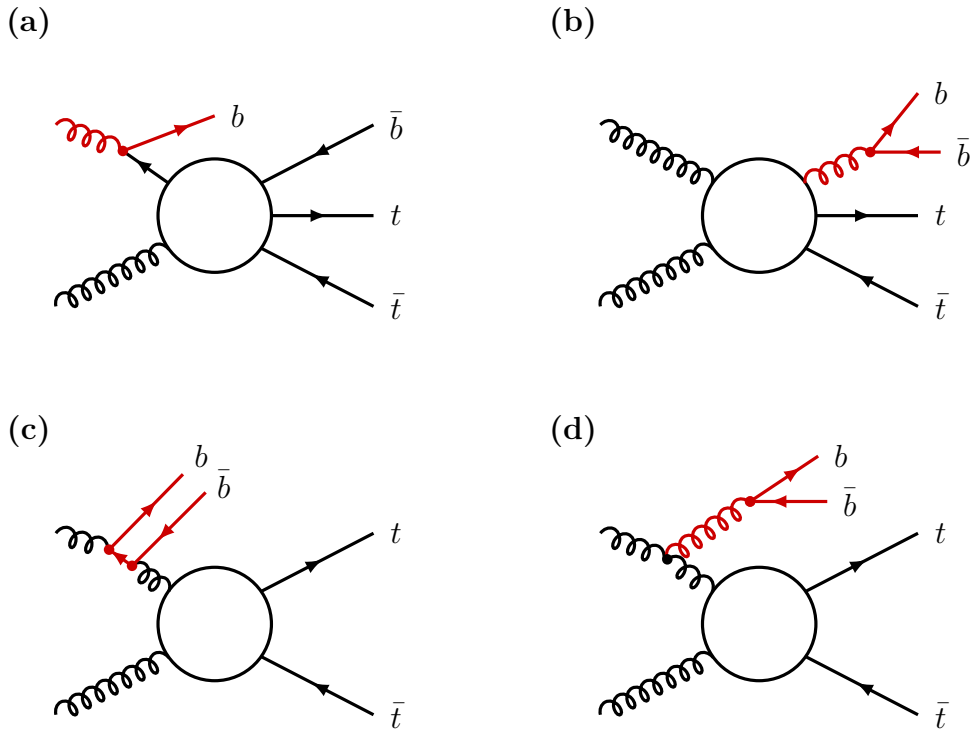

Figure 5.6.: The typical topologies for the $gg \to t\bar{t}b\bar{b}$ process in the four flavour scheme. The initial and final state splittings are indicated in red. On the left side the initial state splittings are shown and on the right side the finial state splittings. The lower processes dominate the production.

is no longer resolvable. As a result the b-quark pair will be merged into one particle jet. For the initial state splitting the spectator b-quark will be emitted to the beam direction and cannot be resolved. The IS splitting in Figure 5.6**(a)** is the same one as in the 5FS. However, it turns out that the most dominant processes for initial and final state splitting are shown in Figure 5.6**(c)** and Figure 5.6**(d)**.

The production $t\bar{t} + b\bar{b}$ in the 4FS is fully inclusive with the advantage of NLO accuracy. On the other hand, a description in the 4FS can lead to large $\alpha_S \ln\left(\frac{m_b}{Q}\right)$ for the IS splitting. In the 5FS these large logarithms will be resummed by the PDF evolution. However, these effects are quite small at the LHC [102]. In order to compare the resummation of the PDF evolution and to clarify the difference between the four and five flavour scheme, the PDF evolution of the bottom quark from Equation (4.1.4) is used

$$\frac{\mathrm{d}b\left(x,\mu^2\right)}{\mathrm{d}\ln\left(\mu^2\right)} = \frac{\alpha_S\left(\mu^2\right)}{2\pi}\int_x^1 \frac{d\xi}{\xi}\left[b\left(\xi,\mu^2\right)P_{bb}\left(\frac{x}{\xi}\right) + g\left(\xi,\mu^2\right)P_{bg}\left(\frac{x}{\xi}\right)\right]. \tag{5.1.2}$$

The splitting kernels are known from Equation (4.1.5). It is worth mentioning that the splitting kernels are calculated in the limit where the quark masses are assumed to be $m_{q_i} = 0$. The PDF in Equation (5.1.2) is coupled to the gluon PDF. The choice of the linear combination $b_{\mathrm{LC}}\left(x, \mu^2\right) = b\left(x, \mu^2\right) - \bar{b}\left(x, \mu^2\right)$ decouples the PDF from the gluon PDF. The evolution equation in the Mellin space is accessed by the Mellin transformation of the PDF, given as:

$$\tilde{b}\left(N, \mu^2\right) = \int_0^1 \mathrm{d}z z^{N-1} b\left(x, \mu^2\right). \tag{5.1.3}$$

With the Mellin transformed splitting kernel γ_{bb} the evolution equation is

$$\frac{\mathrm{d}\tilde{b}_{\mathrm{LC}}}{\mathrm{d}\ln\left(\mu^2\right)} = \frac{\alpha_S\left(\mu^2\right)}{2\pi}\gamma_{bb}\left(N, \mu^2\right)\tilde{b}\left(N, \mu^2\right). \tag{5.1.4}$$

The solution for Equation (5.1.4) is [102]

$$\begin{aligned}\tilde{b}_{\mathrm{LC}}\left(N, \mu^2\right) &= \tilde{b}_{\mathrm{LC}}\left(N, \mu_0^2\right)\exp\left[\frac{\gamma_{bb}}{2\pi\beta}\left(1 + \alpha_S\left(\mu_0^2\right)\beta\ln\left(\frac{\mu^2}{\mu_0^2}\right)\right)\right] \\ &= \tilde{b}_{\mathrm{LC}}\left(N, \mu_0^2\right)\exp\left(\frac{\gamma_{bb}}{2\pi\beta}\right)\sum_{n=0}^{\infty}\left(\frac{\gamma_{bb}}{2\pi\beta}\right)^n\left[\alpha_S\left(\mu_0^2\right)\ln\left(\frac{\mu^2}{\mu_0^2}\right)\right]^n.\end{aligned} \tag{5.1.5}$$

The PDF evolution from Equation (5.1.5) is a resummation of logarithms of the kind from equation (5.1.1). For total inclusive rates. the five flavour scheme is more favourable and shows due to the resummation a smaller sensitivity to the scale choice. The four flavour scheme is more suited to describe the kinematics of the additional jets [102, 103].

6. Monte Carlo Samples

In the following, a description of the Monte Carlo samples which are used in this thesis is presented. The samples are split into $t\bar{t}$ inclusive samples generated in the 5FS and $t\bar{t} + b\bar{b}$ samples generated in the 4FS.

- The nominal Monte Carlo sample used in ATLAS and also in this thesis is a $t\bar{t}$ inclusive sample in the 5FS. It is generated with the POWHEGBOX v2 generator [72, 73, 104, 105] which calculates the matrix element up to NLO order for the strong interaction production of top-quark pairs. The *NNPDF3.0NLO* PDF [106] set is used for the matrix element. The factorisation scale μ_F and renormalisation scale μ_R are set to the transverse mass of the top quark $m_{T,t} = \sqrt{m_t^2 + p_T^2}$. The transverse mass is calculated with the mass m_t and transverse momentum p_T of the top quark before any radiation happened. The h_{damp} parameter which controls the transverse momentum of the first additional emission beyond the Born configuration is set to $1.5m_t$ [107, 108]. The ME generator is matched with PS generator PYTHIA 8.230 [88], which uses the A14 set of tuned parameters [108] and the *NNPDF2.3lo* set [109]. The decays of the charm and bottom quarks are generated by EVTGEN 1.6.0 [110]. This sample is referred to as POWHEG+PYTHIA8 5FS.

- Another $t\bar{t}$ inclusive 5FS sample is generated with the same POWHEG settings, mentioned above. But this sample is matched with the HERWIG version HERWIG7.1.3 [85, 86] which uses the default HERWIG7.1 set of tuned parameters [86, 111] and the *MMHT2014LO* PDF set [112]. The decays of the charm and bottom quarks are generated by EVTGEN 1.6.0 [110]. This sample is referred to as POWHEG+HERWIG7.1.3.

- An alternative $t\bar{t}$ 5FS sample uses the same settings for the ME calculation like the POWHEG+HERWIG7.1.3 one, but it is interfaced with an older parton shower version HERWIG7.0.4 [85, 86]. The H7UE set of tuned parameters is used [86]. This sample is referred to as POWHEG+HERWIG7.0.4 5FS.

- An alternative $t\bar{t}$ 5FS sample uses for the ME calculation of the hard $t\bar{t}$ scattering process MADGRAPH 5_aMC@NLO v2.6.0 [80]. The *NNPDF3.0NLO* PDF set is

used [106]. MadSpin is used to model the decay of the top quarks at LO [82, 83]. The renormalisation scale μ_R and factorisation scale μ_F are set to $m_{T,t} = \sqrt{m_t^2 + p_T^2}$. The shower scale is $\mu_Q = \frac{H_T}{2}$ [113]. H_T expresses the scalar sum of transverse momenta of the outgoing partons. The events are matched to Pythia 8.230 [88] which generates PS and hadronisation. The A14 set of tuned parameters [108] and the *NNDPF2.3lo* PDF set [109] is used. This sample is referred to as aMc@nlo +Pythia8 5FS.

- Another $t\bar{t}$ inclusive 5FS sample is generated with the same settings. It is matched with Herwig7.1.3 [85, 86]. This sample is referred to as aMc@nlo +Herwig7.1.3 5FS.

- The $t\bar{t}$ inclusive 5FS sample is generated with Sherpa2.2.1. Sherpa calculates the matrix element at NLO with one additional parton, while the accuracy can be decreased to LO when calculating with up to four additional partons. The *NNPDF3.0NNLO* PDF set is used [106]. The renormalisation scale μ_R and factorisation scale μ_F are set to $\sqrt{\frac{1}{2}\left(m_{\mathrm{T},t}^2 + m_{\mathrm{T},\bar{t}}^2\right)}$. Sherpa is combined with OpenLoops [91–93]. There is no need for an additional parton shower generator, since Sherpa provides one which uses the *MEPS@NLO* prescription [114–117]. The Sherpa default settings for the tune and PS PDF are used. This sample is referred to as Sherpa2.2.1 5FS.

- An alternative $t\bar{t}$ inclusive 5FS sample is generated with the same settings as mentioned before but with the new Sherpa2.2.10 version. This sample is referred to as Sherpa2.2.10 5FS.

- A $t\bar{t}$+$b\bar{b}$ 4FS sample is generated with Powheg Box Res [12] and OpenLoops [91–93]. The $t\bar{t} + b\bar{b}$ matrix element is calculated at NLO. The *NNPDF3.0NLO* PDF set is used [106]. The renormalisation scale μ_R is $\frac{1}{2}\left(\sum_{i=t,\bar{t},b,\bar{b}} m_{T,i} + m_{T,\mathrm{Gluon}}\right)$. The factorisation scale μ_F is set to $\prod_{i=t,\bar{t},b,\bar{b}} m_{T,i}^{\frac{1}{4}}$ and the h_{damp} parameter is set to $\frac{H_\mathrm{T}}{2}$ with $H_\mathrm{T} = \frac{1}{2}\sum_{i=t,\bar{t},b,\bar{b}} E_{T,i}$ as the scalar sum of transverse energy of final state partons. The ME calculations are interfaced with Pythia8.240 [88] using the A14 set of tuned parameters [108] and the *NNPDF2.3LO* PDF set [109]. This sample is referred to as Powheg+Pythia8 4FS.

- Another $t\bar{t} + b\bar{b}$ 4FS sample is generated with Sherpa2.2.1+OpenLoops. The *CT10NNLO* PDF set is used [79, 118]. The renormalisation scale is set to $\mu_R = \prod_{i\in\{t,\bar{t},b,\bar{b}\}} m_{\mathrm{T},i}^{\frac{1}{4}}$, while the factorisation and resummation scale is set to $\mu_F = \mu_Q = \frac{H_\mathrm{T}}{2}$ with H_T as the scalar sum of transverse energy of finial state partons. This sample

is referred to as Sherpa2.2.1 4FS.

- An alternative $t\bar{t}+b\bar{b}$ 4FS sample is generated with Sherpa and the same settings as mentioned before but with the new Sherpa2.2.10 version. This sample is referred to as Sherpa2.2.10 4FS.

In all samples the top quark mass was set to $m_t = 172.5\,\text{GeV}$. In $t\bar{t} + b\bar{b}$ 4FS samples the bottom quark mass is set to $m_b = 4.75\,\text{GeV}$ for Sherpa and $m_b = 4.95\,\text{GeV}$ for Powheg+Pythia. All $t\bar{t}$ inclusive samples are normalised to the predicted $t\bar{t}$ cross-section at NNLO accuracy in perturbative QCD with resummation of next-to-next-to-leading logarithmic (NNLL) soft gluon terms [119–122]. The used cross-sections and k-factors are in Table 6.1. A summary of the sample settings which are used in this thesis is presented in Table 6.2.

6.1. Object Identification and Event Selection

The most important physical objects for this thesis are electrons, muons and b-jets. No event selection was applied, but the important physical objects have to fulfil certain selection criteria which are similar to the detector acceptance of Atlas.

- Electrons: Electrons need to have a minimal transverse momentum of 10 GeV and have to be central, meaning that the pseudorapidity is $|\eta| < 2.47$. The crack region of $1.37 < |\eta| < 1.52$ is excluded since Atlas is not able to measure precisely in that region.
- Muons: Muons need to have a minimal transverse momentum of 10 GeV and a pseudorapidity of $|\eta| < 2.5$.
- Jets: Particle jets are formed with the anti-k_t-algorithm, which is a well-defined, infrared-safe and collinear-safe algorithm [123, 124]. The radius parameter is $R = 0.4$. They are clustered from stable particles, which have a mean lifetime of more than $3 \cdot 10^{-11}\,\text{s}$. Particle jets need at least $p_\text{T} \geq 15\,\text{GeV}$ and a pseudorapidity of $|\eta| < 2.5$.

The *Heavy Flavour Simple Classification* classifies events into three non-overlapping categories. If a jet with $p_\text{T} \geq 15\,\text{GeV}$ and not originating from a top quark is identified within a radius of $\Delta\text{R} = 0.4$ around a b-hadron with $p_\text{T} \geq 5\,\text{GeV}$ this event is labelled as $t\bar{t}+ \geq 1b$. If an event contains a jet with $p_\text{T} \geq 15$, and not originating from a decay of a b-hadron or W -boson, is matched with a c-hadron with $p_\text{T} \geq 5\,\text{GeV}$ within a radius of

Table 6.1.: The k-factors and cross-sections σ for the MC sample setups. The total cross-section for the $t\bar{t}$ and $t\bar{t} + b\bar{b}$ is also given. The samples are produced with different final states. Depending on the decays of the W - boson the samples are split into dilepton and single-lepton samples. Furthermore, the POWHEG+PYTHIA8 4FS and SHERPA samples differentiate between positively (+) and negatively (-) charged leptons. The nominal sample POWHEG+PYTHIA8 5FS is a non-all hadronic sample.

Sample	Final State	k-Factor	Cross-Section σ in pb
POWHEG+PYTHIA8 5FS	all-non hadronic	1.1398	396.87
	$\sigma_{t\bar{t}} = 452.35\,\text{pb}$		
POWHEG+HERWIG7.1.3 5FS	single-lepton	1.1392	320.11
	dilepton	1.1391	77.00
	$\sigma_{t\bar{t}} = 452.38\,\text{pb}$		
POWHEG+HERWIG7.0.4 5FS	single-lepton	1.1392	320.18
	dilepton	1.1391	77.01
	$\sigma_{t\bar{t}} = 452.47\,\text{pb}$		
aMC@NLO +PYTHIA8 5FS	single-lepton	1.1691	313.27
	dilepton	1.1681	76.31
	$\sigma_{t\bar{t}} = 455.39\,\text{pb}$		
aMC@NLO +HERWIG7.1.3 5FS	single-lepton	1.1694	313.22
	dilepton	1.1681	76.29
	$\sigma_{t\bar{t}} = 455.39\,\text{pb}$		
SHERPA2.2.1 5FS	single-lepton +	1.1484	158.94
	single-lepton -	1.1484	158.99
	dilepton	1.1484	76.317
	$\sigma_{t\bar{t}} = 452.75\,\text{pb}$		
SHERPA2.2.10 5FS	single-lepton +	1.2075	151.01
	single-lepton -	1.2075	151.05
	dilepton	1.2075	72.592
	$\sigma_{t\bar{t}} = 452.39\,\text{pb}$		
POWHEG+PYTHIA8 4FS	single-lepton +	1.0	6.732
	single-lepton -	1.0	6.913
	dilepton	1.0	3.246
	$\sigma_{t\bar{t}+b\bar{b}} = 16.89\,\text{pb}$		
SHERPA2.2.1 4FS	single-lepton +	1.0	5.7283
	single-lepton -	1.0	5.7314
	dilepton	1.0	2.7549
	$\sigma_{t\bar{t}+b\bar{b}} = 14.21\,\text{pb}$		
SHERPA2.2.10 4FS	single-lepton +	1.0	7.8457
	single-lepton -	1.0	7.8458
	dilepton	1.0	3.7686
	$\sigma_{t\bar{t}+b\bar{b}} = 19.46\,\text{pb}$		

Table 6.2.: Overview of the different settings of the samples which are used in this thesis.

ME	PS	FS	μ_R	μ_F	h_{damp}	ME PDF	PS PDF	Tune
Powheg	Pythia8	5FS	$\sqrt{m_t^2 + p_T^2}$	$\sqrt{m_t^2 + p_T^2}$	$1.5m_t$	*NNPDF3.0NLO*	*NNDPF2.3lo*	A14
Powheg	Herwig7.1.3	5FS	$\sqrt{m_t^2 + p_T^2}$	$\sqrt{m_t^2 + p_T^2}$	$1.5m_t$	*NNPDF3.0NLO*	*MMHT2014LO*	Herwig deafult
Powheg	Herwig7.0.4	5FS	$\sqrt{m_t^2 + p_T^2}$	$\sqrt{m_t^2 + p_T^2}$	$1.5m_t$	*NNPDF3.0NLO*	*MMHT2014LO*	H7UE
aMc@nlo	Herwig7.1.3	5FS	$\sqrt{m_t^2 + p_T^2}$	$\sqrt{m_t^2 + p_T^2}$	-	*NNPDF3.0NLO*	*MMHT2014LO*	Herwig deafult
aMc@nlo	Pythia8	5FS	$\sqrt{m_t^2 + p_T^2}$	$\sqrt{m_t^2 + p_T^2}$	-	*NNPDF3.0NLO*	*NNDPF2.3lo*	A14
Sherpa 2.2.1	-	5FS	$\sqrt{\frac{1}{2}\left(m_{\text{T},t}^2 + m_{\text{T},\bar{t}}^2\right)}$	$\sqrt{\frac{1}{2}\left(m_{\text{T},t}^2 + m_{\text{T},\bar{t}}^2\right)}$	-	*NNPDF3.0NNLO*	Sherpa default	Sherpa default
Sherpa 2.2.10	-	5FS	$\sqrt{\frac{1}{2}\left(m_{\text{T},t}^2 + m_{\text{T},\bar{t}}^2\right)}$	$\sqrt{\frac{1}{2}\left(m_{\text{T},t}^2 + m_{\text{T},\bar{t}}^2\right)}$	-	*NNPDF3.0NNLO*	Sherpa default	Sherpa default
Powheg	Pythia8	4FS	$\frac{1}{2}\left(\sum_{i=t,\bar{t},b,\bar{b}} m_{T,i} + m_{T,\text{Gluon}}\right)$	$\prod_{i=t,\bar{t},b,\bar{b}} m_{T,i}^{\frac{1}{4}}$	$\frac{H_T}{2}$	*NNPDF3.0NLO*	*NNDPF2.3lo*	A14
Sherpa 2.2.1	-	4FS	$\prod_{i\in\{t,\bar{t},b,\bar{b}\}} m_{\text{T},i}^{\frac{1}{4}}$	$\frac{H_T}{2}$	-	*CT10nnlo*	Sherpa default	Sherpa default
Sherpa 2.2.10	-	4FS	$\prod_{i\in\{t,\bar{t},b,\bar{b}\}} m_{\text{T},i}^{\frac{1}{4}}$	$\frac{H_T}{2}$	-	*CT10nnlo*	Sherpa default	Sherpa default

$\Delta\text{R} = 0.4$, it is labelled as $t\bar{t}+ \geq 1c$. If an event contains jets, which are matched with b- and c-hadrons, it will be sorted into the category $t\bar{t}+ \geq 1b$. The remaining events are labelled as $t\bar{t} + \text{light}$. These events may contain light quarks like up, down, and strange quarks but also gluon radiation as well or no additional jets, since jets need to fulfil certain criteria. In the following, a list of kinematic variables which are used for this thesis is shown. The additional jets are ordered in transverse momentum.

- N^{ℓ}: the number of leptons.
- $p_T^{q_1}$: the transverse momentum of the first additional jet.
- m^{q_1}: the invariant mass of the first additional jet.
- η^{q_1}: the pseudorapidity of the first additional jet.
- $p_T^{q_2}$: the transverse momentum of the second additional jet.
- m^{q_2}: the invariant mass of the second additional jet.
- η^{q_2}: the pseudorapidity of the second additional jet.
- p_T^{qq}: the transverse momentum of the additional jet system.
- m^{qq}: the invariant mass of the additional jet system.
- H_{T}^{qq}: the scalar sum of transverse momentum of the additional jet system.
- ΔR^{qq}: the distance between the first and second additional jet in the η-ϕ -plane.

7. Modelling Uncertainties

To measure a process precisely a good signal background ratio has to be achieved. Therefore, understanding of the modelling of the background processes is necessary. The most dominant background is the $t\bar{t}$ + jets background in particular $t\bar{t} + b\bar{b}$.

For the modelling of high energy particle collisions, MC simulations are used. These simulations feature different assumptions and implementations and therefore will not reflect data perfectly and should be interpreted as an approximation. A detailed description of the implementation methods and various generators is presented in Chapter 4, while in Table 6.2 a summary of the parameters is presented. These simulations differ in their choice of parameters like the renormalisation and factorisation scale or the damping parameter h_{damp}. Furthermore, different PDF sets for the ME calculation and PS generation, underlying event tune, and hadronisation models are used. All these variations will result in different kinematic distributions and give an estimation for the modelling uncertainties for future comparisons with ATLAS data.

The simulations are generated in two different flavour schemes: the four and five flavour scheme. The four flavour scheme treats the bottom quark as a massive particle, while in the five flavour scheme the bottom quark is assumed to be massless. To estimate an impact of the different treatment and its effects on the kinematics, the four flavour scheme samples are compared with the corresponding five flavour scheme samples to be sure that different ME calculations, hadronisation models, PS algorithms and matching schemes do not effect the results. However, the renormalisation and factorisation scales differ and other PDF sets need to be used since the b-quark PDF is decoupled. There is a new SHERPA version, SHERPA2.2.10 4FS and 5FS, on the market. It is compared to an older version SHERPA2.2.1 4FS and 5FS to investigate improvements. The four flavour scheme samples of SHERPA2.2.10 and SHERPA2.2.1 are compared with the four flavour scheme POWHEG+PYTHIA8 setup to estimate the modelling uncertainties for different setups for the $t\bar{t} + b\bar{b}$ process.

The differences between the HERWIG versions is investigated with a direct comparison between the HERWIG7.0.4 and HERWIG7.1.3, both interfaced with POWHEG. To see the impact of the different parton shower and their modelling of the $g \to b\bar{b}$ splitting more precisely, a comparison between the same matrix element generator aMC@NLO interfaced with two different parton showers, HERWIG7.1.3 and PYTHIA8 is also studied.

7.1. The $t\bar{t}$ + jets Background

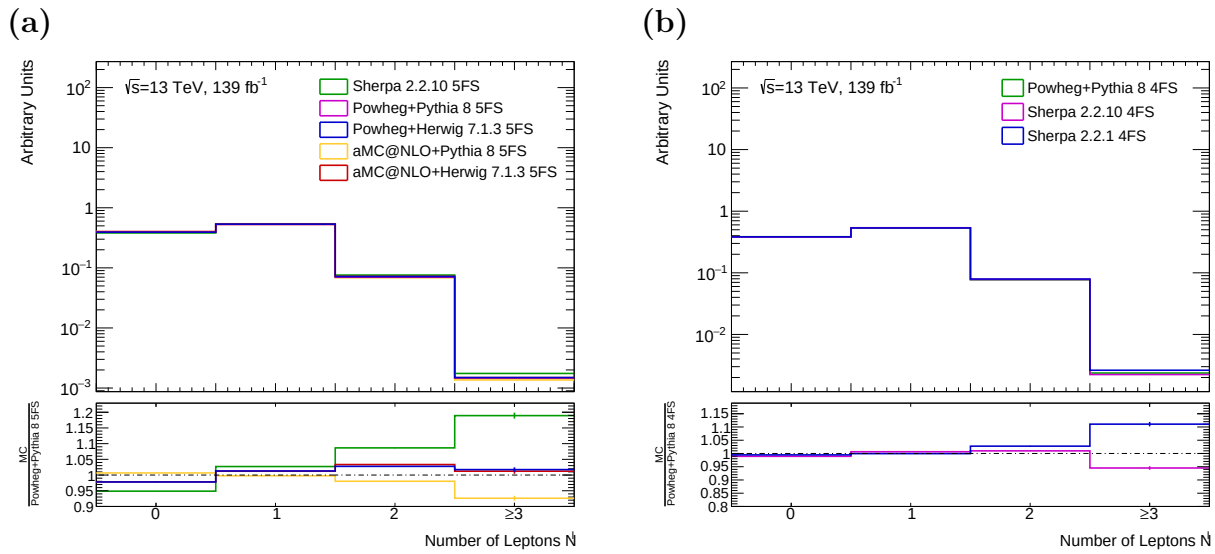

Figure 7.1.: The lepton multiplicity N^{ℓ} for the five flavour scheme samples on the left side **(a)** and for the four flavour scheme samples on the right side **(b)**. For the five flavour scheme the $t\bar{t}$ inclusive process is simulated, while for the four flavour scheme the $t\bar{t} + b\bar{b}$ process is simulated. All distributions are normalised to unity. The ratio for the five flavour ones is with respect to the POWGHEG+PYTHIA8 5FS sample. The ratio for the four flavour ones is with respect to the POWGHEG+PYTHIA8 4FS sample.

The studies of the $t\bar{t}$ + jets background are done at particle level using the object definitions from Chapter 6.1. The events are selected by the number of final state leptons $\ell \in \{e, \mu\}$ which arise from the decay of $W^{\pm}$ -bosons which originate from the top quark pair. In Figure 7.1**(a)** the lepton multiplicity for the five flavour scheme setup is shown, while in Figure 7.1**(b)** for the four flavour scheme setup. In case of the five flavour setup the ratio is calculated with respect to the nominal sample POWGHEG+PYTHIA8 5FS and for the four flavour scheme samples with respect to the POWGHEG+PYTHIA8 4FS sample.

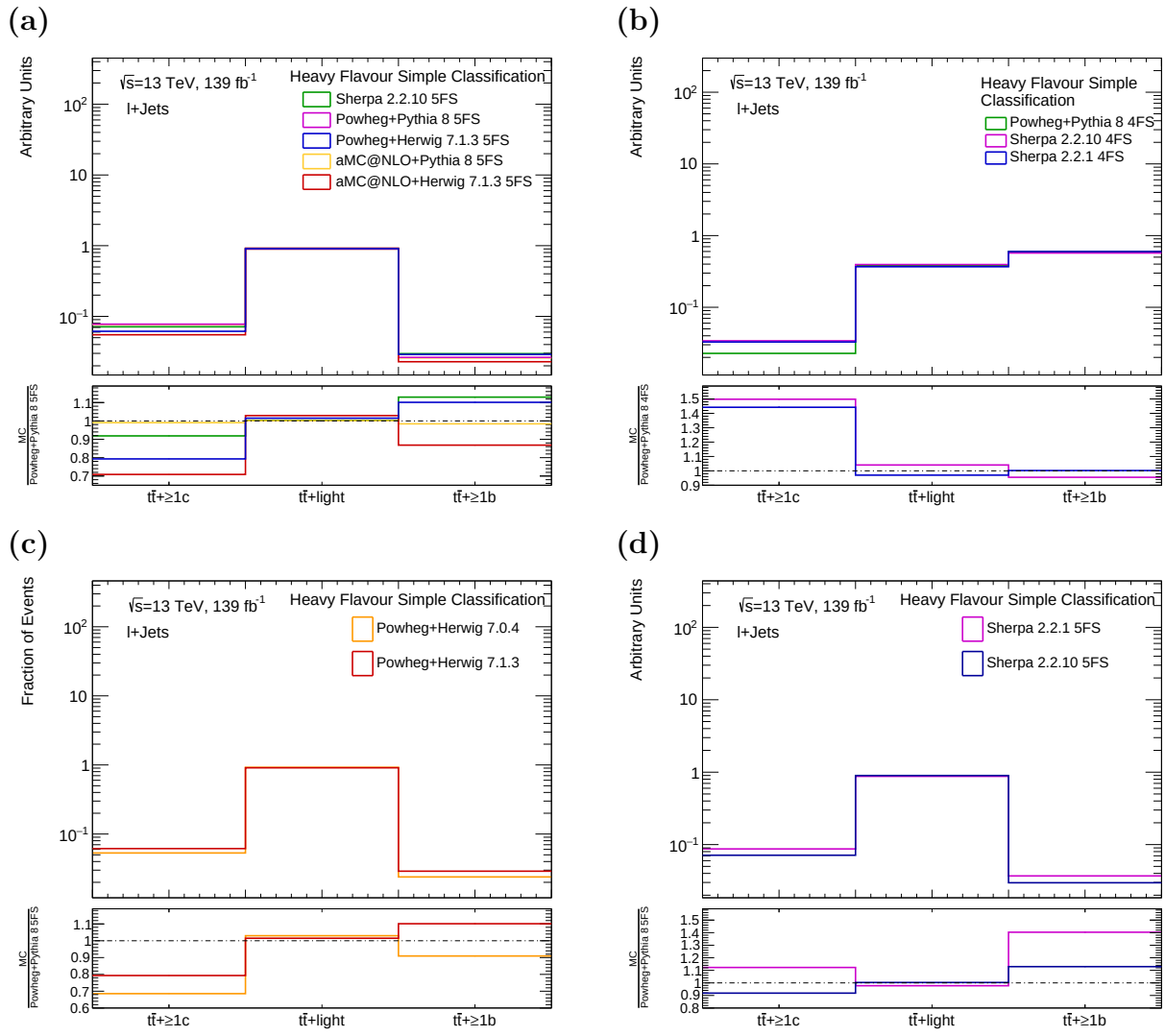

Figure 7.2.: The distributions of the *Heavy Flavour Simple Classification* for the various MC simulation setups. The definition of the categories $t\bar{t}+ \geq 1c$, $t\bar{t}+ \geq 1b$ and $t\bar{t}$ + light are described in Chapter 6.1. All distributions are normalised to unity. In **(a)** the newer versions of five flavour scheme samples setups are presented. The ratio is with respect to the nominal sample POWGHEG+PYTHIA8 5FS. In **(b)** the four flavour scheme sample setups are compared. The ratio is with respect to the POWHEG+PYTHIA8 4FS sample. In **(c)** an older version of HERWIG7.0.4 is compared with the newer one HERWIG7.1.3. Both are interfaced with POWHEG. In **(d)** the five flavour setups of SHERPA2.2.1 and SHERPA2.2.10 are compared. For the last two figures both ratios are with respect to the nominal sample POWGHEG+PYTHIA8 5FS.

Both distributions are normalised to unity.

Approximately 40 % of all events do not contain any leptons, since there are cuts for the minimal transverse momentum and pseudorapidity of a lepton, see Chapter 6.1. The largest fraction of events of approximately 50-60 % corresponds for both flavour scheme setups to events containing one lepton in the final state, which is called single-lepton channel. Since the branching ratio of the dilepton channel is five times lower, a lower fraction of events is expected. However, the fraction of the dilepton events is more than five times lower since the probability to identify zero or one out of two leptons is larger than to identify two out of two leptons from $W^{\pm}$ -boson decays. The smallest fraction corresponds to events with three or more leptons. The additional lepton in the final state may be produced by bremsstrahlung.

There are small relative uncertainties for single-lepton final state which do not exceed 5 %. The same result holds for events containing zero leptons. The dilepton channel has larger uncertainties caused by lower statistics. The largest differences corresponds to events with three or more leptons in the final state. The yields for each final state are shown in Table A.2. The five flavour scheme samples describes the $t\bar{t}$ inclusive process, while the four flavour one simulate the $t\bar{t} + b\bar{b}$. As seen from Table 6.1 the cross-section for the first one is larger compared with the latter one, resulting in a higher yield for all five flavour scheme samples. However, for both flavour setups the shape of the lepton multiplicity distribution is similar.

The kinematics of the additional jets are independent of the decay of the top quark pair and therefore a distinction between the single-lepton and dilepton channel is not necessary and would result in lower statistics and higher statistical uncertainties for the latter one. For the upcoming analysis of the $\mathrm{t\bar{t}}$ + jets and $\mathrm{t\bar{t}}+ \geq 1b$ background and the kinematic of the additional jets only events with one or two leptons in the final state are used. These events are labelled as ℓ + jets.

The $t\bar{t}$ + jets background has three contributions which describe different processes: $t\bar{t}$ + light, $t\bar{t}+ \geq 1c$ and $t\bar{t}+ \geq 1b$. These processes have different kinematics which arise from different masses. The $t\bar{t}$ inclusive samples are generated with all contributions. The additional jets of the four and five flavour scheme samples are classified by their flavour into three non-overlapping categories using the *Heavy Flavour Simple Classification* which is described in Chapter 6.1. The results for various comparisons are presented

in Figure 7.2. In Figure 7.2**(a)** the five flavour scheme samples are compared. The ratio is with respect to the nominal sample POWHEG+PYTHIA8 5FS. In Figure 7.2**(b)** the four flavour scheme samples are shown. The ratio is with respect to the POWHEG+PYTHIA8 4FS sample. The old version HERWIG 7.0.4 is compared with the version HERWIG 7.1.3 in Figure 7.2**(c)** and SHERPA2.2.1 5FS is compared with SHERPA2.2.10 5FS in Figure 7.2**(d)** while the ratio for both is with respect to POWHEG+PYTHIA8 5FS. All distributions are normalised to unity.

The fraction of the $t\bar{t}$ + light category is approximately $\approx 89\,\%$ for all $t\bar{t}$ inclusive samples and dominates all other contributions. The relative uncertainties with respect to the nominal samples are quite small. Furthermore, the nominal sample POWHEG+PYTHIA8 5FS is tuned to ATLAS data [108]. The measurements of the $t\bar{t}$ + light events are very precise. Larger differences are expected for the remaining two categories, $t\bar{t}+ \geq 1c$ and $t\bar{t}+ \geq 1b$, since measurements of these are less precise. The $t\bar{t}+ \geq 1c$ category is dominated by the aMC@NLO +PYTHIA8 5FS and the nominal samples POWHEG+PYTHIA8. Comparing them more closely the fractions of all three categories are very similar with a difference less than 2 %. This can be explained since the modelling of the additional jets in the five flavour scheme is mostly done at PS level and therefore the choice of the ME calculator should not lead to a large impact on the *Heavy Flavour Simple Classification* of the $t\bar{t}$ + jets background. Furthermore, SHERPA2.2.10 5FS has slightly more events for $t\bar{t}+ \geq 1b$. 10 % more compared with the nominal. In contrast, aMC@NLO +HERWIG7.1.3 has for both heavy flavour categories, $t\bar{t}+ \geq 1c$ and $t\bar{t}+ \geq 1b$, less events than any other MC simulation. For the the $t\bar{t}+ \geq 1c$ category 30 % and for $t\bar{t}+ \geq 1b$ 10 %. But aMC@NLO +HERWIG7.1.3 5FS has the highest fraction in the remaining $t\bar{t}$+light category.

The four flavour scheme samples describe a different process and therefore a direct comparison is not valid. The shape of the distribution is different and the $t\bar{t}+ \geq 1b$ category is dominant. A relatively high fraction of $t\bar{t}$+light events are caused by cuts on the kinematics: the pseudorapidity and minimal transverse momentum. As expected, the $t\bar{t}+ \geq 1c$ category has the lowest yield. All three four flavour scheme samples describe $t\bar{t}+ \geq 1b$ similarly. Only small differences, less than 5 % with respect to the POWHEG+PYTHIA8 4FS sample, occur. The treatment of c-quarks and therefore the $t\bar{t}+ \geq 1c$ category has the largest difference with 50 % more events for SHERPA2.2.10 4FS compared with POWHEG+PYTHIA8 4FS. The direct comparison of the different HERWIG versions in Figure 7.2**(c)** shows that the newer HERWIG version has a higher fraction of events for the $t\bar{t}+ \geq 1b$ category and less in the $t\bar{t}+ \geq 1c$ compared with the older version. This is

caused by the updated treatment of b-jets. SHERPA2.2.1 5FS produces a much higher fraction of events which are sorted into the $t\bar{t}+ \geq 1b$ category compared with all other five flavour scheme samples. However, the newer version SHERPA2.2.10 5FS has a smaller fraction which is more compatible with the other five flavour scheme samples. The relative differences are not larger than 15 % compared to the nominal one.

7.2. The $t\bar{t}+ \geq 1b$ Background

7.2.1. Heavy Flavour Categories

In this section, the $t\bar{t}+ \geq 1b$ events selected by the *Heavy Flavour Simple Classification* are studied. This is the most dominant irreducible background for the signal process $t\bar{t}\left(H \to b\bar{b}\right)$. To understand the modelling uncertainties more precisely, events will be further classified using the *Heavy Flavour Classification*. It sorts events into four non-overlapping categories which are sensitive to the number of additional jets of an event. If the event contains exactly one additional b-jet, the event is labelled as $t\bar{t} + b$. If two jets are identified it is labelled as $t\bar{t} + bb$. If a single particle jet is matched with a B-hadron pair, the event is called $t\bar{t} + B$. These events correspond mostly to additional jet productions with very collinear jets. In this case the additional jets are not resolvable and will be merged into one particle jet. The rest is labelled as $t\bar{t}+ \geq 3b$ and contain events with three or more additional jets.

In Figure 7.3 and 7.4, the results for the various sample setups for the *Heavy Flavour Classification* are shown. The first one compares the five and four flavour scheme samples directly, while the latter one compares the three four flavour samples with their corresponding five flavour ones. All distributions are normalised to unity and the ratio is calculated with respect to the nominal sample POWGHEG+PYTHIA8 5FS, except in Figure 7.3**(b)** it is calculated with respect to POWGHEG+PYTHIA8 4FS. The yields of the different categories of the *Heavy Flavour Classification* for the sample setups can be found in Table A.4.

The highest fraction of events corresponds to the $t\bar{t} + b$ category. For the four and five flavour scheme sample approximately 60 % of the total events corresponds to it, see Figure 7.3**(a)** and **(b)**. Due to cuts on the minimal transverse momentum and the restriction in pseudorapidity which are similar to the accessible ATLAS phase space one additional jet

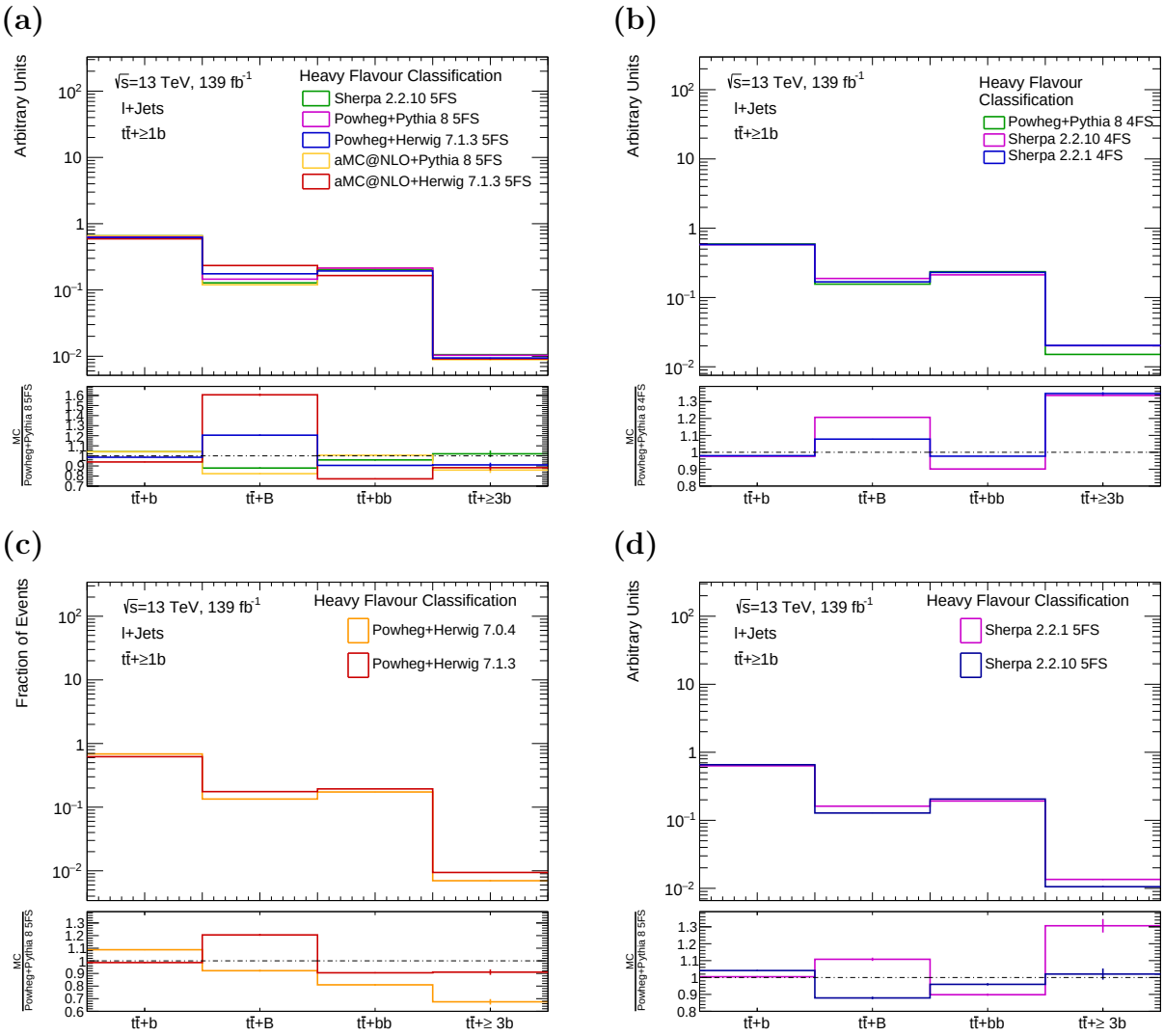

Figure 7.3.: The distributions of the *Heavy Flavour Classification* for the various MC simulation setups. All distributions are normalised to unity. In **(a)** the newer versions of five flavour scheme samples setups are presented. The ratio is with respect to the nominal sample POWGHEG+PYTHIA8 5FS. In **(b)** the four flavour scheme sample setups are compared. The ratio is with respect to the POWHEG+PYTHIA8 4FS sample. In **(c)** an older version of HERWIG7.0.4 is compared with the newer one, HERWIG7.1.3. Both are interfaced with POWHEG. In **(d)** the five flavour setups of SHERPA2.2.1 and SHERPA2.2.10 are compared. For the last two figures both ratios are with respect to the nominal sample POWGHEG+PYTHIA8 5FS.

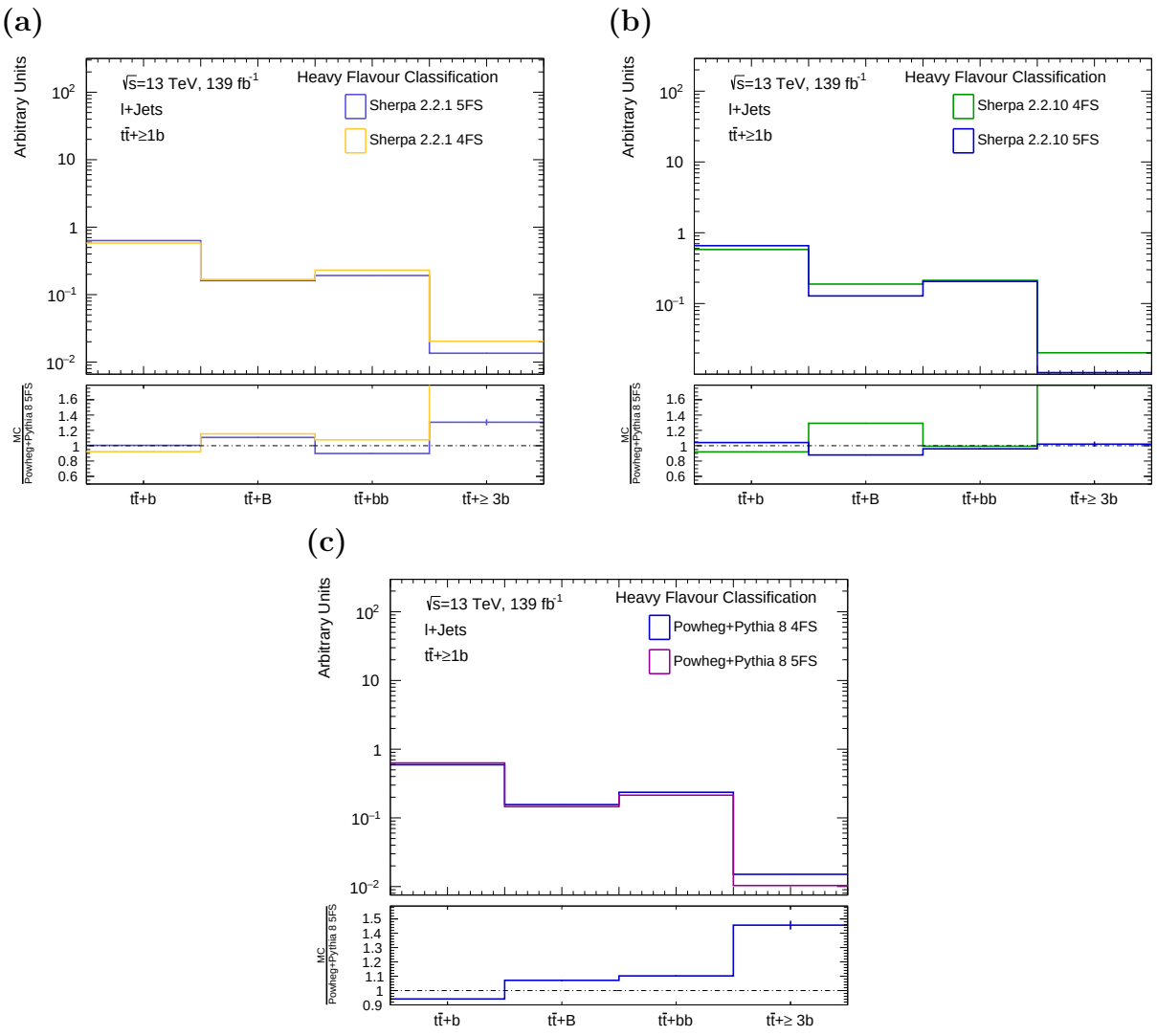

Figure 7.4.: The distributions of the *Heavy Flavour Classification* for the direct comparison of the four and five flavour scheme samples. All distributions are normalised to unity. In **(a)** the SHERPA2.2.1 versions are compared. The newer version SHERPA2.2.10 samples are compared in **(b)**. The POWGHEG+PYTHIA8 ones are shown in **(c)**. The ratios are calculated with respect to the nominal sample POWGHEG+PYTHIA8 5FS.

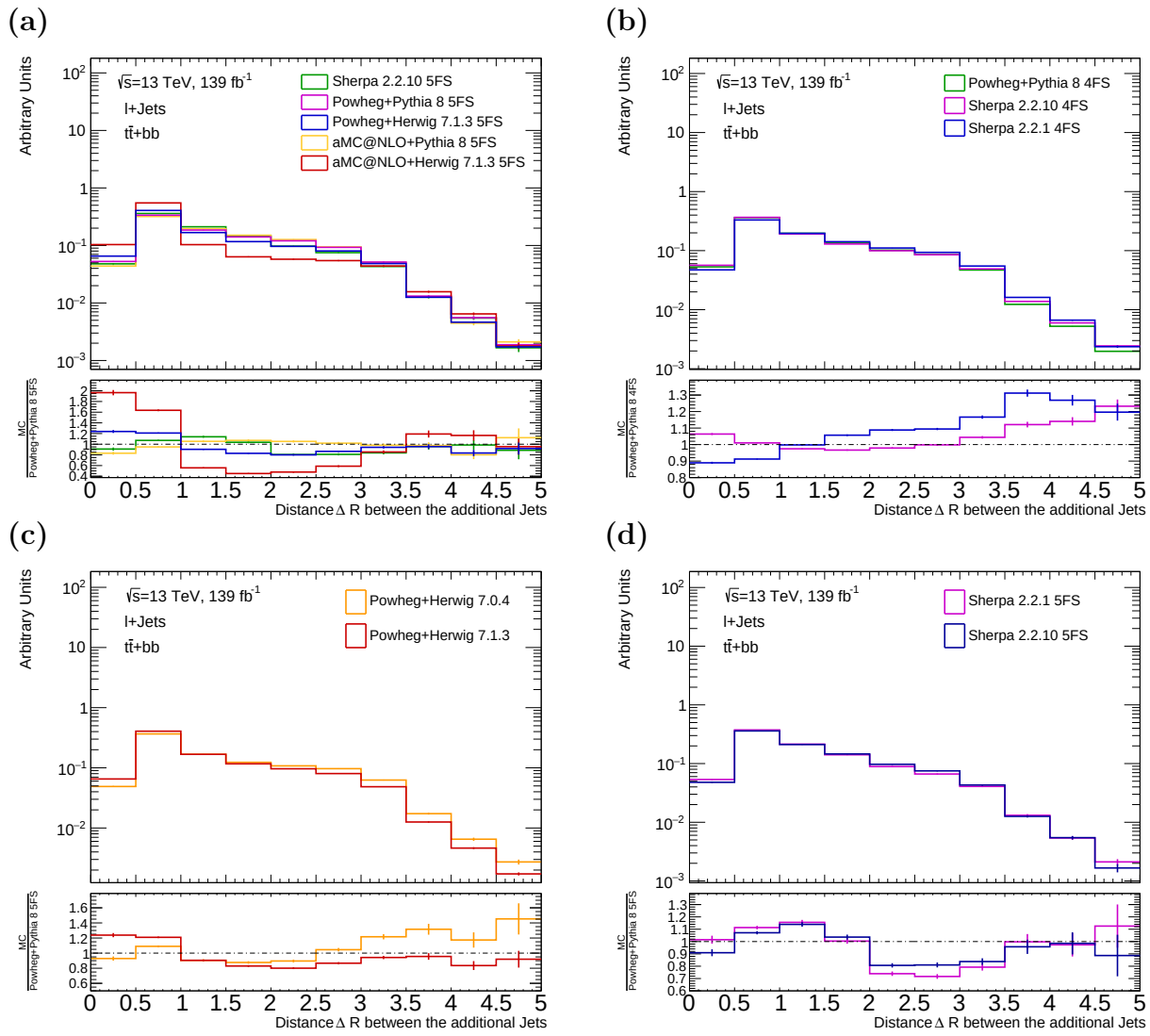

Figure 7.5.: The distributions for the distance between the first and second additional jet for the various MC simulation setups. The distance is calculated using Equation (3.3.4). All distributions are normalised to unity. In **(a)** the newer versions of five flavour scheme sample setups are presented. The ratio is with respect to the nominal sample POWGHEG+PYTHIA8 5FS. In **(b)** the four flavour scheme sample setups are compared. The ratio is with respect to the POWHEG+PYTHIA8 4FS sample. In **(c)** a version of HERWIG7.0.4 is compared with HERWIG7.1.3. Both are interfaced with POWHEG. In **(d)** the five flavour setups of SHERPA2.2.1 and SHERPA2.2.10 are compared. For the last two figures both ratios are with respect to the nominal sample POWGHEG+PYTHIA8 5FS.

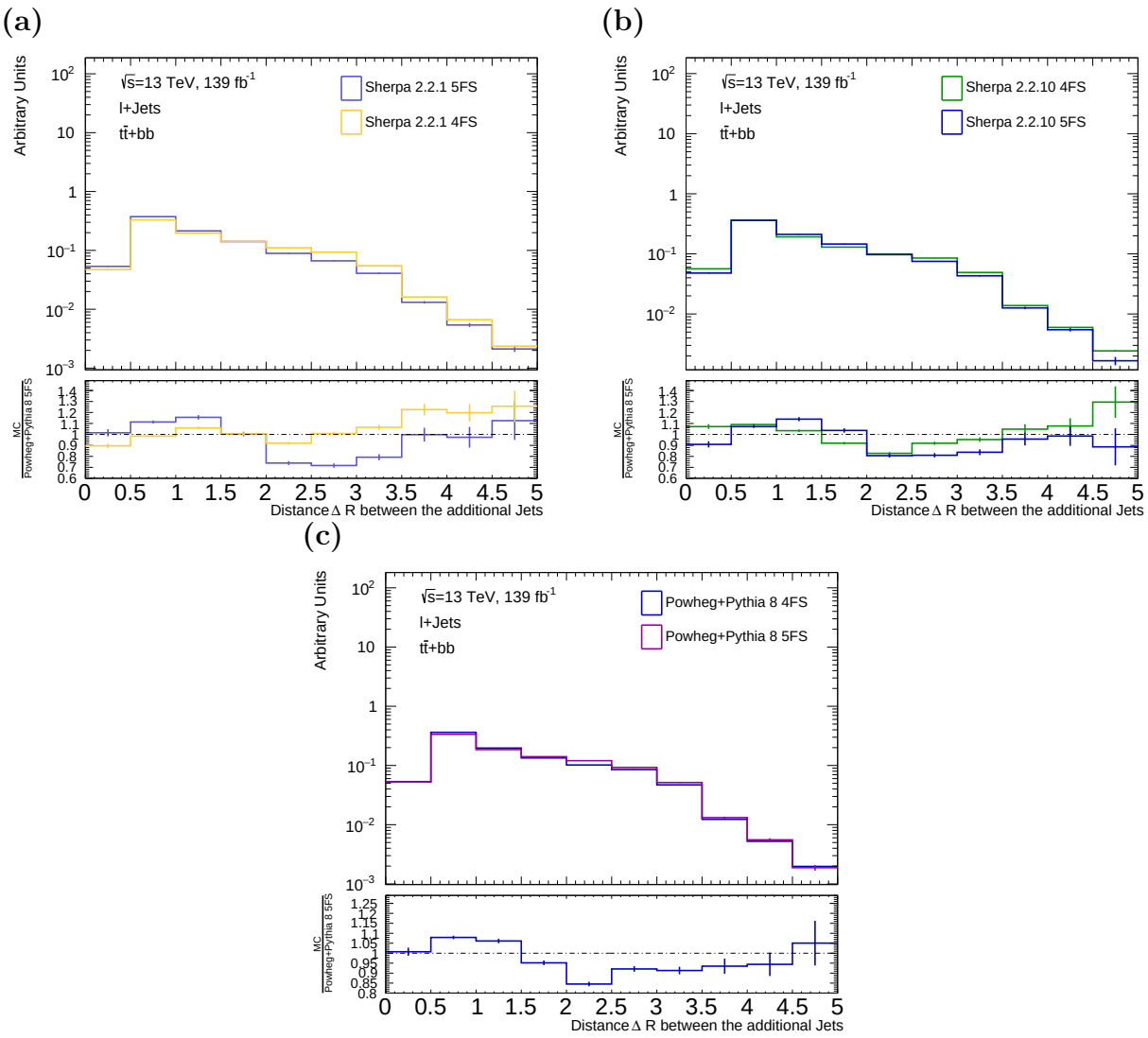

Figure 7.6.: The distributions of the distance between the first and second additional jet for the direct comparison of the four and five flavour scheme samples. All distributions are normalised to unity. In **(a)** the SHERPA2.2.1 versions are compared. The newer version SHERPA2.2.10 samples are compared in **(b)**. The POWGHEG+PYTHIA8 ones are shown in **(c)**. The ratios are calculated with respect to the nominal sample POWGHEG+PYTHIA8 5FS.

will not be identified in most events. Only small differences with respect to the nominal sample POWHEG+PYTHIA8 5FS are expected for the comparison between the five flavour scheme samples. These differences do not exceed 10 %.

The different treatments of collinear jets lead to the largest difference within the set of five flavour scheme samples for the $t\bar{t} + B$ category. The highest fraction is generated with the aMC@NLO +HERWIG7.1.3 setup. In Figure 7.5 and 7.6, the distance between the first and second additional jet of the $t\bar{t} + b\bar{b}$ category is presented. Again, the first figure compares the four and five flavour samples directly, while the latter presents a direct comparison between the flavour schemes. The distance between the additional jets ΔR^{qq} is calculated using Equation 3.3.4. All distributions are normalised to unity and the ratio is calculated in the same way like in Figure 7.3 and Figure 7.4. Figure 7.5**(a)** shows that the aMC@NLO +HERWIG7.1.3 setup produces more events with a small distance between the additional jet ΔR^{qq} in the $t\bar{t} + b\bar{b}$ category. Therefore additional jets for an event tend to be so collinear that they are not resolvable any more and will be merged into a single particle jet. This leads to a higher fraction of events for the $t\bar{t} + B$ category, as can be seen in Figure 7.3.

Comparing the remaining five flavour scheme samples a clear trend can be observed. Samples which produce more events for the $t\bar{t} + B$ category tend to have a smaller distance between the first and second additional jet ΔR^{qq}, the jets tend to be produced more collinear. This correlation is observed for the comparison of the HERWIG versions in Figure 7.3**(c)** and Figure 7.5**(c)** and the comparison of the five flavour scheme SHERPA2.2.1 and 2.2.10 in Figure 7.3**(d)** and Figure 7.5**(d)**, too. However, the comparison of the four flavour scheme samples in Figure 7.3**(b)** and Figure 7.5**(b)** shows that this correlation holds true only for SHERPA2.2.10. The older version SHERPA2.2.1 has a higher fraction of events for the $t\bar{t} + B$ category but produces more events with a larger distance between the additional jets than the POWHEG+PYTHIA8 4FS sample. All four flavour samples in Figure 7.4 have an expected higher fraction of events for the $t\bar{t} + B$ category compared with their five flavour scheme pendants.

The four flavour scheme is expected to describe the kinematics of the b-quark more precisely by considering the b-quark as a massive particle. The full phase space of the splitting $g \to b\bar{b}$ is accessible. The modelling of additional jets in the $t\bar{t} + B$ category relies strongly on this treatment since a better resolution of collinear regions results in a higher fraction of events. Very collinear additional jets cannot be produced in the five flavour scheme

because of artificial cuts on the opening angle between the additional jets to control divergences, see Figure 5.5**(a)** and **(b)**. The diagram from Figure 5.5**(c)** does not exist in the five flavour scheme. A correlation between a small distance between the additional jets in the $t\bar{t}+bb$ category, see Figure 7.6 and a higher fraction of events for the $t\bar{t}+B$ one, see Figure 7.4, is observed for the POWHEG+PYTHIA8 and SHERPA2.2.10 sample setup. However, for SHERPA2.2.1 the five flavour scheme generated more events with collinear jets than the four flavour one.

The most important category for further $t\bar{t}\left(H\rightarrow b\bar{b}\right)$ analyses is $t\bar{t}+bb$. All four flavour scheme samples have a higher fraction of $t\bar{t}+bb$ events compared with their corresponding five flavour ones, see Figure 7.4. The largest relative uncertainty occurs for the comparison between the SHERPA2.2.1 setup with 20 %. The POWHEG+PYTHIA8 4FS sample has the smallest differences with less than 2 %.

Overall, the relative uncertainties for the $t\bar{t}+bb$ category are similar to the $t\bar{t}+b$ one. The largest relative uncertainties for the $t\bar{t}+bb$ category corresponds to POWHEG+HERWIG7.1.3 with approximately 20 % with respect to the nominal sample, see Figure 7.3. However, the relative uncertainties are still smaller compared to the $t\bar{t}+B$ ones. For all three four flavour samples a higher fraction of events corresponds to $t\bar{t}+bb$ compared with their five flavour scheme pendants.

Events with more than three additional b-jets are very rare. The $t\bar{t}+\geq 3b$ category has the smallest fraction of events. Those events are not considered for the further analysis.

7.2.2. Additional Jet Kinematics

Comparison of the pseudorapidity of the additional jets shows no noticeable differences, see Figure B.3 and Figure B.5. The invariant mass spectrum of the additional jet system is expected to have no mass resonance, see Figure B.6.

The five flavour scheme samples are presented in Figure 7.7**(a)**, Figure 7.8**(a)** and Figure 7.9**(a)**. The hardest jets are generated by the aMC@NLO +HERWIG7.1.3. Comparing the POWHEG+HERWIG7.1.3 sample with the aMC@NLO +PYTHIA8 sample the kinematics are similar for the first and second additional jet. The ME generation does not have a large impact on the kinematics of the jets, because they are mostly produced at PS level. However, the transverse momentum spectrum differs largely at harder regions. This can

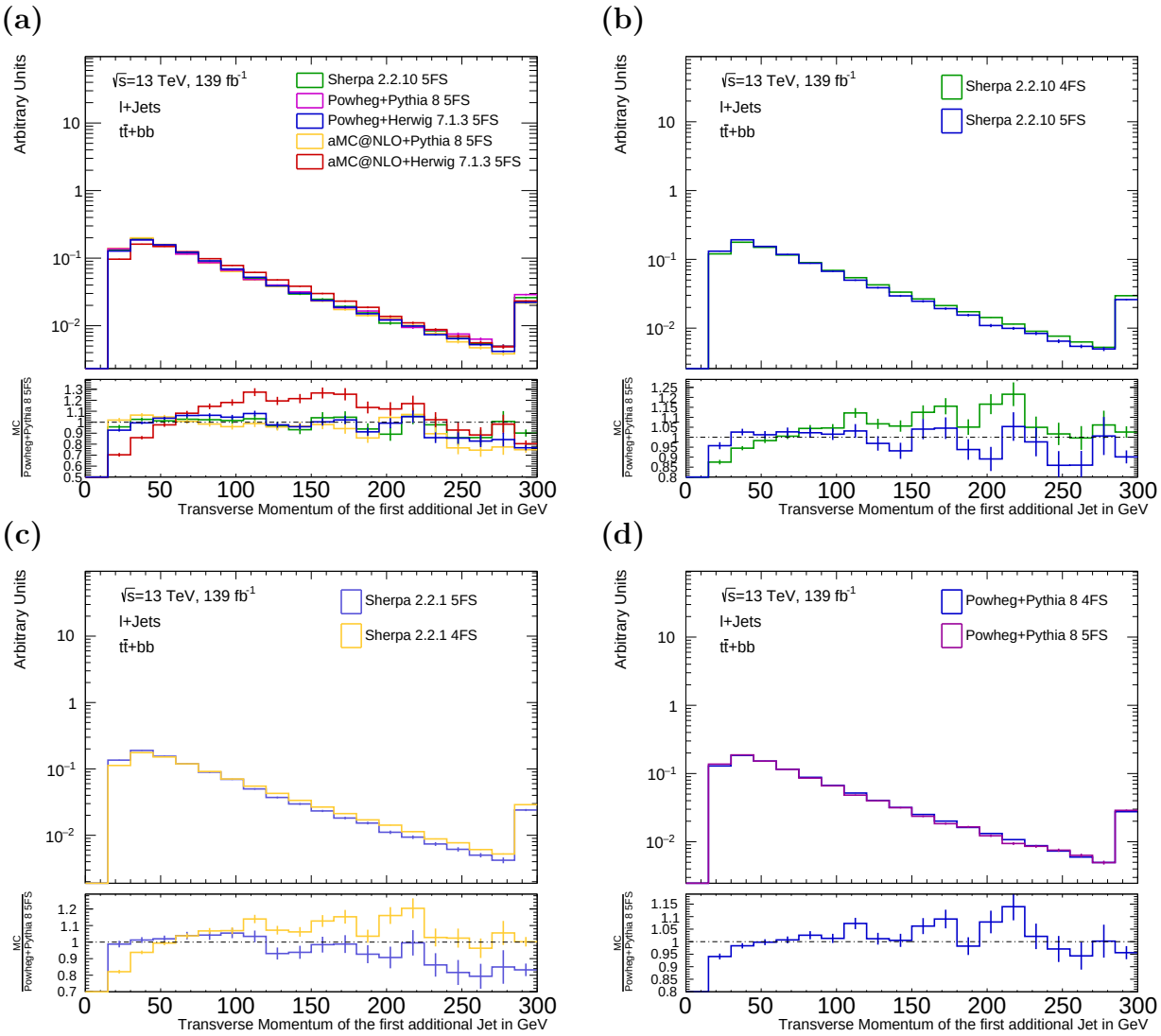

Figure 7.7.: The distributions of the transverse momentum of the first additional b-jet for the $t\bar{t} + bb$ category. In **(a)** the five flavour scheme sample setups are presented. The four flavour scheme samples are compared directly with their corresponding five flavour one: In **(b)** the SHERPA2.2.10, in **(c)** the SHERPA2.2.1 and in **(d)** the POWHEG+PYTHIA8. The ratio is with respect to the nominal sample POWGHEG+PYTHIA8 5FS. All distributions are normalised to unity.

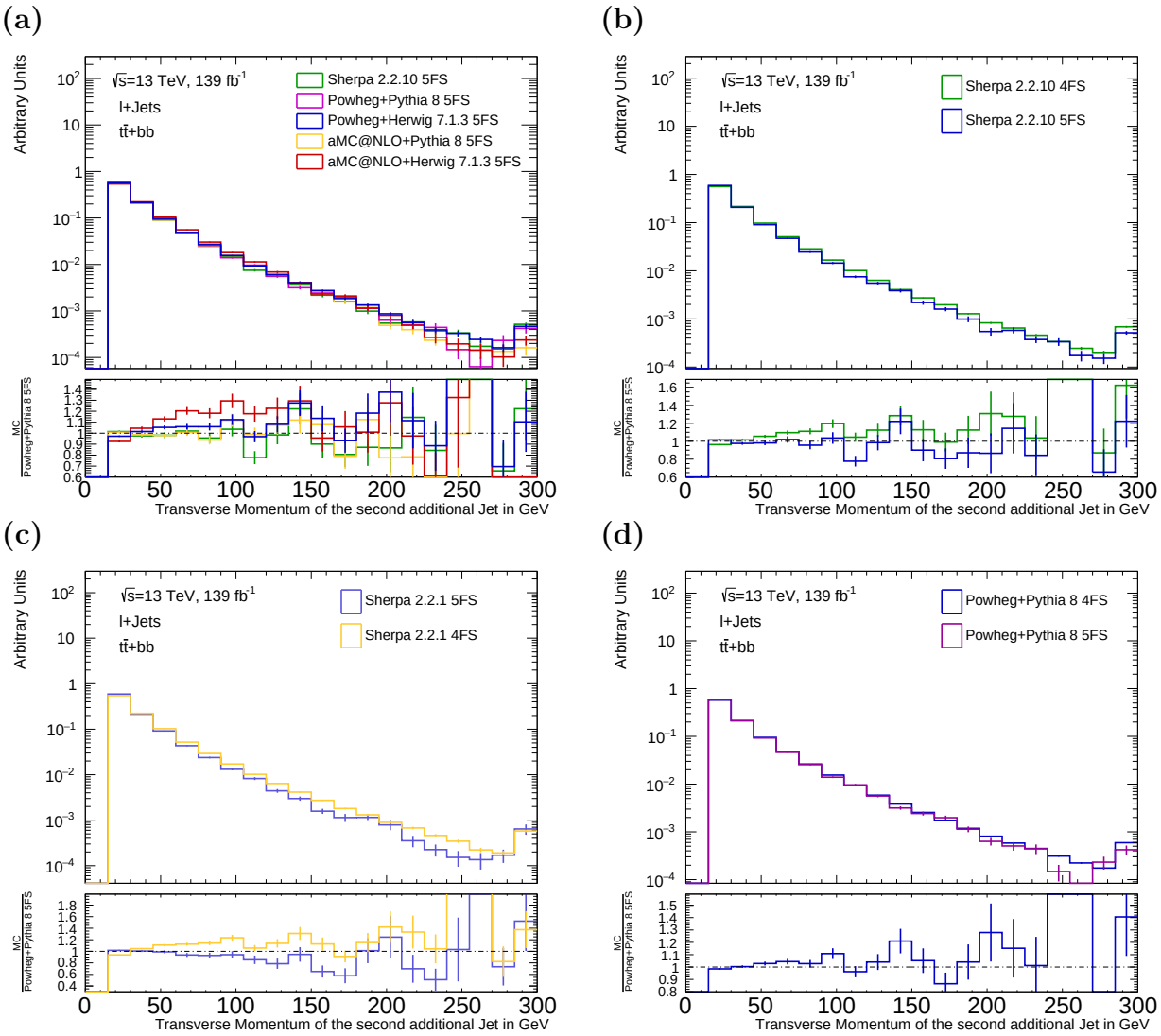

Figure 7.8.: The distributions of the transverse momentum of the second additional b-jet for the $t\bar{t}+bb$ category. In **(a)** the five flavour scheme sample setups are presented. The four flavour scheme samples are compared directly with their corresponding five flavour one: In **(b)** the SHERPA2.2.10, in **(c)** the SHERPA2.2.1 and in **(d)** the POWHEG+PYTHIA8. The ratio is with respect to the nominal sample POWHEG+PYTHIA8 5FS. All distributions are normalised to unity.

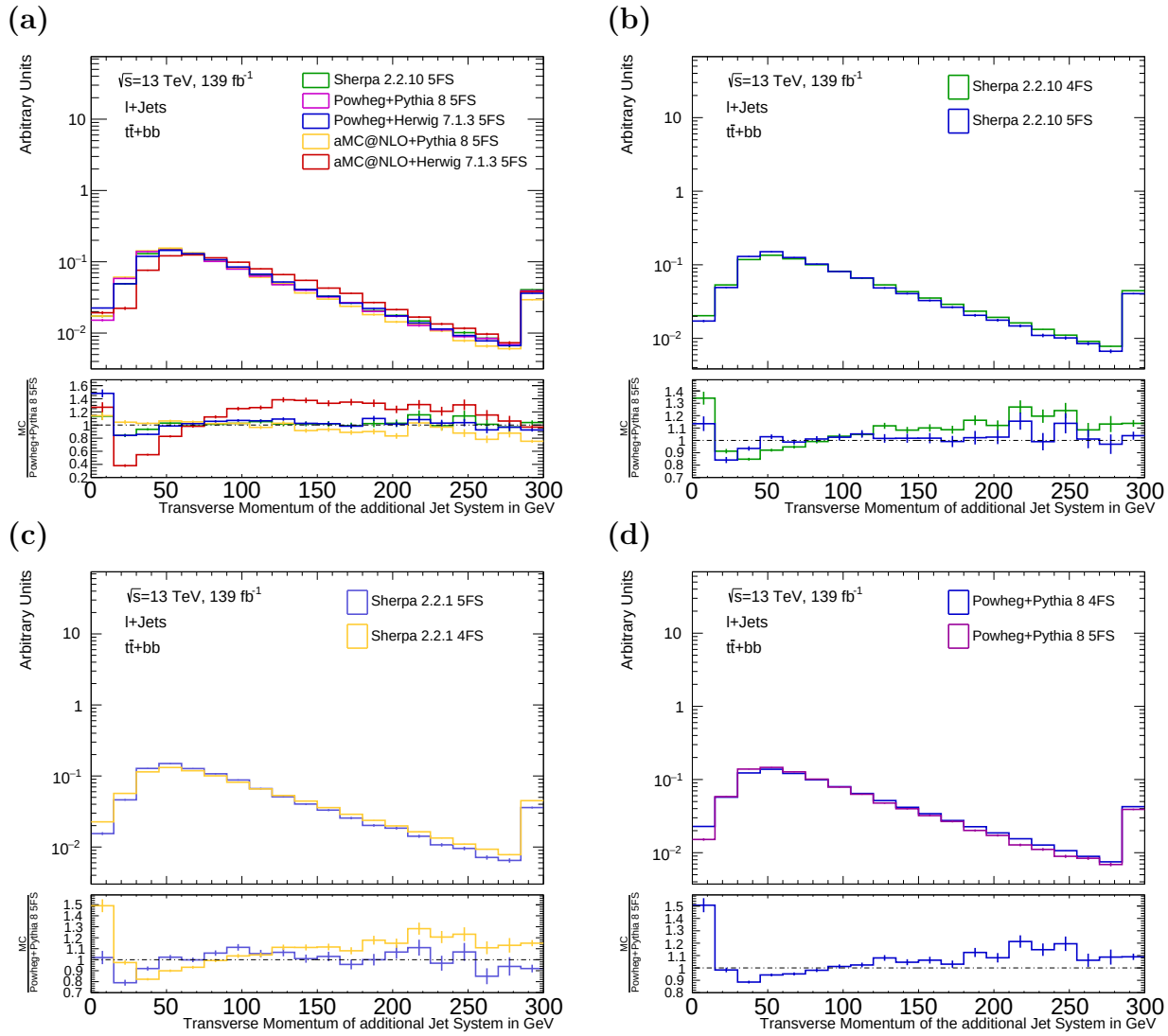

Figure 7.9.: The distributions of the transverse momentum of the additional b-jet system for the $t\bar{t} + bb$ category. In **(a)** the five flavour scheme sample setups are presented. The four flavour scheme samples are compared directly with their corresponding five flavour one: In **(b)** the SHERPA2.2.10, in **(c)** the SHERPA2.2.1 and in **(d)** the POWHEG+PYTHIA8. The ratio is with respect to the nominal sample POWHEG+PYTHIA8 5FS. All distributions are normalised to unity.

be caused by the different matching methods. The largest relative uncertainties between the five flavour samples occur in the description of the transverse momentum of the additional jet system in Figure 7.9**(a)**. aMc@nlo +Herwig7.1.3 has the hardest spectrum. Furthermore, the largest uncertainties are in the description of the soft regions with a maximum of approximately 60 % for the Powheg+Herwig7.1.3 sample. The comparison between the different Herwig versions shows only small differences for kinematic variables. However, the Herwig versions differ significantly in the description of the distance between the first and second additional jet.

The behaviour of the transverse momentum differs between the four and five flavour samples. All four flavour scheme setups have a slightly harder transverse momentum spectrum for the first and second additional jet, see Figure 7.7**(b)-(d)** and Figure 7.8**(b)-(d)**. The distributions are normalised to unity and the ratio is calculated with respect to the nominal sample Powheg+Pythia8 5FS. Jets require a minimal transverse momentum, see Chapter 6.1, so there are no jets expected with a transverse momentum lower than 15 GeV. The harder spectrum for all four flavour scheme samples can be explained by the different generation of b-quarks.

In the four flavour scheme b-quarks are produced at ME level which operates at higher energy scales compared with PS which acts on lower energy scales, see Chapter 5.1.1 and Chapter 5.1.2. Furthermore, the calculation at ME level does not require any collinear approximations, see Chapter 4.2, or constraints on the kinematics. Only energy and momentum conservation need to be fulfilled. A harder transverse momentum spectrum is also observed for the the additional jet system in Figure 7.9. There are no cuts on the transverse momentum. The four flavour scheme samples are expected to provide a better description of the lower regions of the transverse momentum spectrum. The relative differences of up to 50 % for the additional jet system with a transverse momentum lower than 15 GeV are observed. This may be due to the fact that the four flavour scheme accesses the full phase space, even in softer regions. The five flavour scheme relies on the generation of b-quarks at PS where the full phase space is not accessible.

7.3. Scale Uncertainties

The impact and uncertainties due to the renormalisation and factorisation scale choice are estimated by varying the scales by a factor of $\frac{1}{2}$, 1 and 2. This leads to nine scale

variation combinations (μ_R, μ_F). However, the simultaneous variation of both scales up and down leads to the largest effect. The same events are used for the scale variation, therefore statistical errors are not shown. The scale choices can be seen in Table 6.2.

In Figure 7.10 the two simultaneous scale variations for the POWHEG+PYTHIA8 setup in the four and five flavour scheme for the *Heavy Flavour Simple Classification* and *Heavy Flavour Classification* are presented. Figure 7.11 shows the transverse momentum of the additional jet system and the distance between the first and second additional jet. The same variation and observables for the SHERPA2.2.10 version in the four and five flavour scheme are presented in Figure 7.12 and Figure 7.13. All distributions are normalised to unity. The ratio is calculated with respect to the nominal sample with $(1\mu_R^2, 1\mu_F^2)$. The scale choice does have an impact on the shape of observables but also influences the event rate. The calculation of the $t\bar{t}$ process differs largely from the $t\bar{t} + b\bar{b}$ process. On Born level the first process is a simple $2 \to 2$ while the latter one is a $2 \to 4$ without any b-quarks in the initial state, since the b-quark PDF is decoupled. Hence, a larger effect of scale variation is expected for the four flavour scheme samples.

Comparing the scale variations of the POWHEG+PYTHIA8 5FS sample only small differences can be observed for the *Heavy Flavour Simple Classification* and *Heavy Flavour Classification*. The largest relative uncertainties occur in the $t\bar{t}+ \geq 1b$ and $t\bar{t}+B$ category with only 2 % for the five flavour scheme sample. Furthermore, the shapes of the variations are distinctive and do not show a fluctuating behaviour. The largest uncertainties of approximately 5 % corresponds to the $t\bar{t} + B$ category, see Figure 7.10**(a)** and **(b)**. For POWHEG+PYTHIA8 4FS only a small impact caused by the scale variations on the fraction of events for the *Heavy Flavour Simple Classification* can be observed, see Figure 7.10**(c)**. A down (up) variation of the scales leads to a smaller (larger) fraction of events for the $t\bar{t} + B$ category, see Figure 7.10**(d)**. For all other samples, the variation leads to an opposite effect on fractions. This can be explained by the different scale choice of the samples.

The differences between the kinematics of the additional jets caused by the scale variation are small: for the POWHEG+PYTHIA8 5FS sample the variation has only an effect in the tail of the distribution of the transverse momentum of the additional jet system, see Figure 7.11**(a)**. Moreover, it has a large effect on the distance between the additional jets, see Figure 7.11**(b)**. The POWHEG+PYTHIA8 4FS shows the smallest relative uncertainties. The distance distribution only differs within the scale variation in the very soft

regions and at the tail, see Figure 7.11**(d)**. The transverse momentum of the additional jet system differs within a range of 1.5 %, see Figure 7.11**(c)**.

The scale variation of the SHERPA2.2.10 5FS sample features the same behaviour: small differences for the fraction of events for the *Heavy Flavour Simple Classification* and *Heavy Flavour Classification*, see Figure 7.12**(a)** and **(b)**. However, the relative uncertainties are larger for the latter one. The largest uncertainties correspond to the $t\bar{t} + b\bar{b}$ category. The transverse momentum of the additional jet system has the largest difference at the tail of the distribution. The same holds true for the distance between the additional jets, see Figure 7.13**(a)** and **(b)**. In general, the relative uncertainties are larger compared with the POWHEG+PYTHIA8 5FS sample.

The largest relative uncertainties occur for the comparison of the scale variation for the SHERPA2.2.10 4FS setup, see Figure 7.12 and Figure 7.13. The treatment of very collinear additional jets in the $t\bar{t} + B$ category is very sensitive to the scale variation. Furthermore, more collinear jets are produced for the down variation of the renormalisation and factorisation scale resulting in a larger fraction of events which corresponds to $t\bar{t} + B$. The effect on the transverse momentum of the additional jet system is small compared with the five flavour pendant. However, there are large differences in the softer regions of the distribution.

Finally, the expected higher sensitivity for the variation of the scale is not observed for the POWHEG+PYTHIA8 4FS. For the POWHEG+PYTHIA8 5FS sample the scale variations play a minor role. This statement holds also true for the SHERPA2.2.10 5FS sample. However, the effects on the fraction of events for the *Heavy Flavour Simple Classification* and *Heavy Flavour Classification* are larger. The four flavour sample, POWHEG+PYTHIA8 4FS, shows a small impact due to the scale variations: They affect mostly the distance between the additional jets, see Figure 7.11**(d)**, and therefore the relative uncertainties in the $t\bar{t} + B$ category, see Figure 7.10**(d)**. The largest impact is observed for the SHERPA2.2.10 4FS sample. The relative uncertainties for the *Heavy Flavour Simple Classification* and *Heavy Flavour Classification* in Figure 7.12**(c)** and **(d)** and the distance between the additional jets in Figure 7.13**(d)** are large. However, the transverse momentum of the additional jet system is less affected by the scale variations, see Figure 7.13**(c)**.

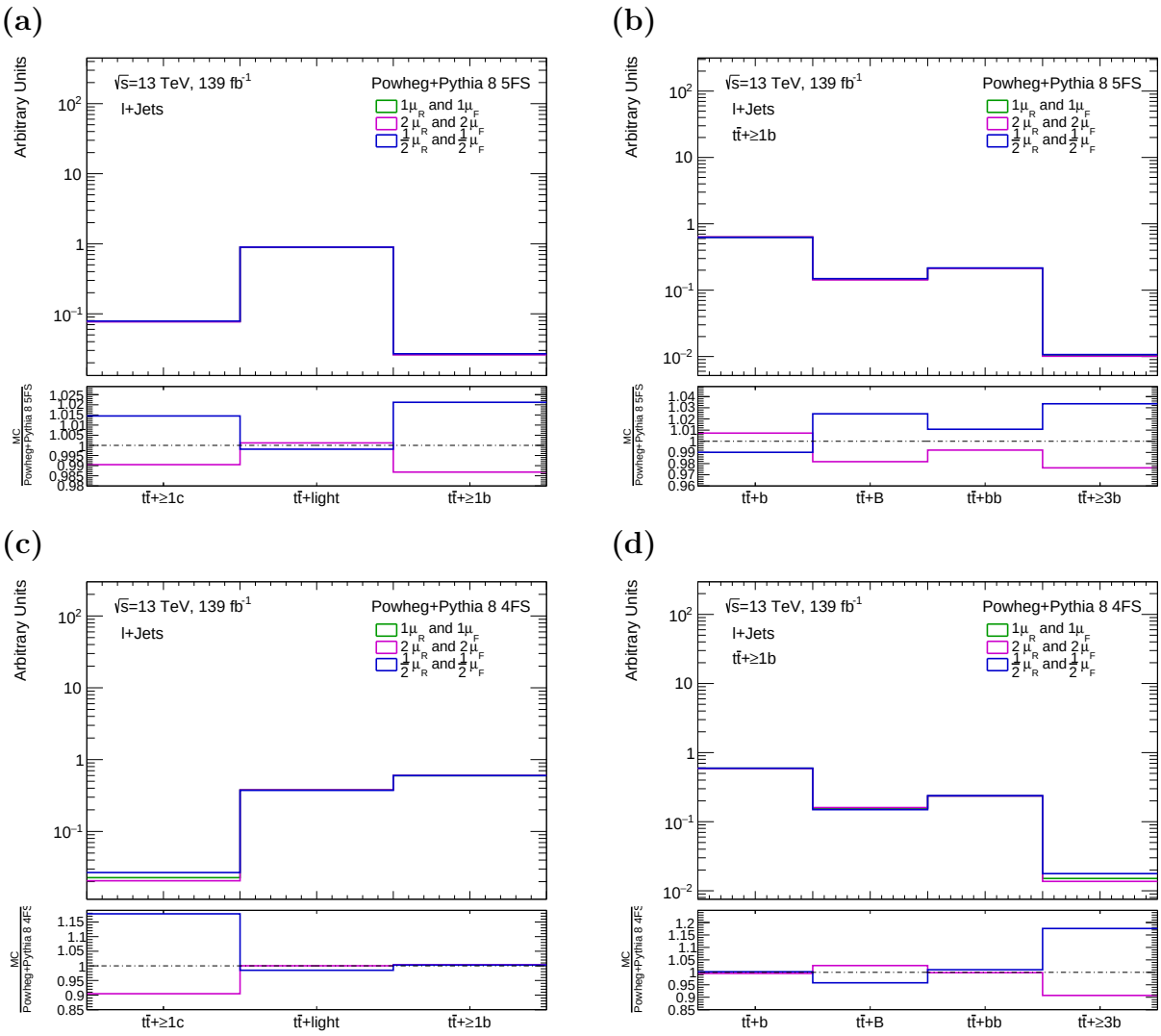

Figure 7.10.: The scale variation for the POWHEG+PYTHIA8 in the five flavour scheme (top line) and the four flavour scheme (bottom line). The simultaneous variation of the renormalisation and factorisation scales leads to the largest variation and therefore independent variations are dropped. On the left side the *Heavy Flavour Simple Classification*, while on the right the *Heavy Flavour Classification* is shown. Only events in the single-lepton or dilepton channel are considered. The distributions are normalised to unity. The ratio is calculated with respect to the nominal sample $(1\mu_R, 1\mu_F)$.

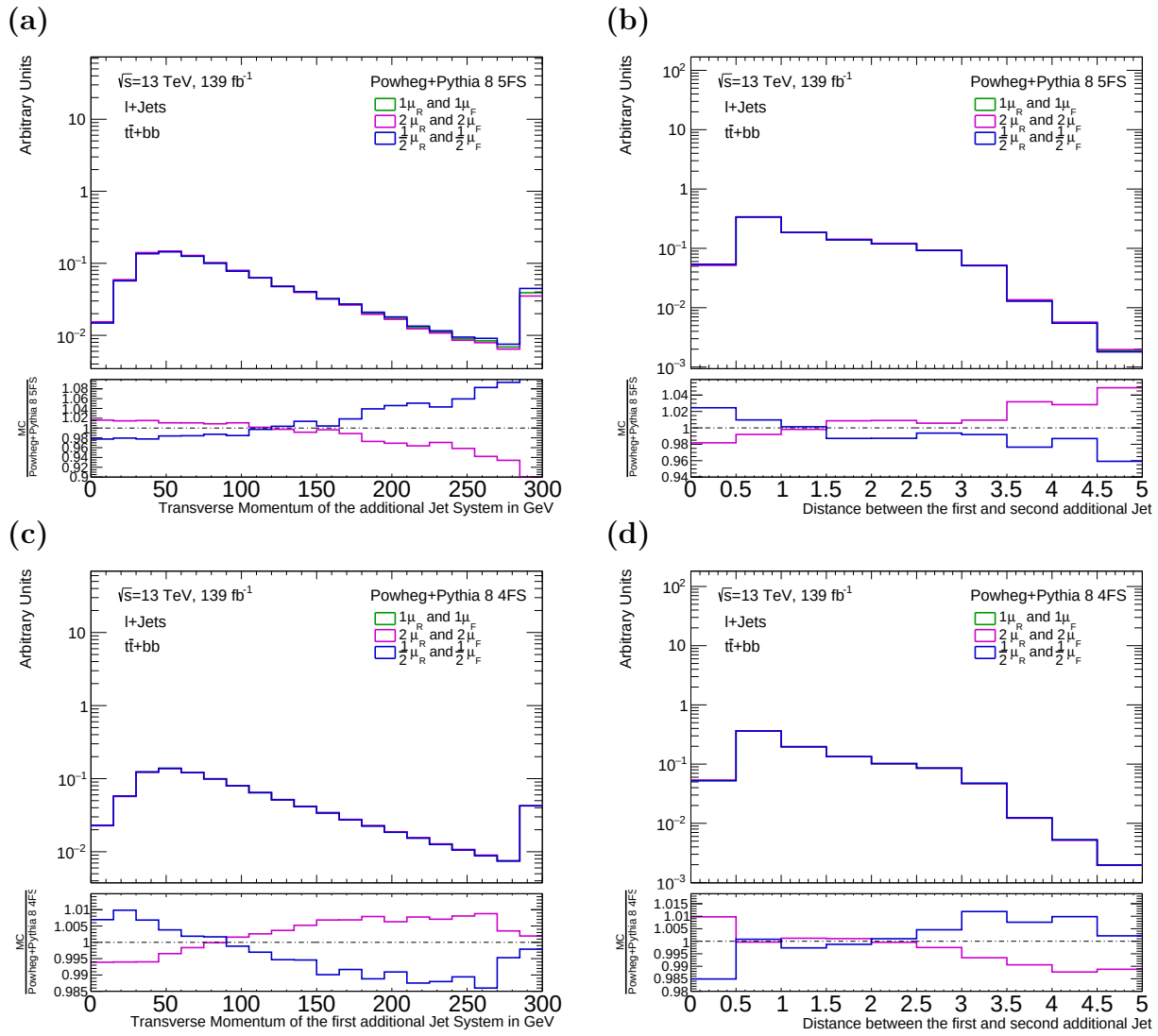

Figure 7.11.: The scale variation for the POWHEG+PYTHIA8 in the five flavour scheme (top line) and the four flavour scheme (bottom line). The simultaneous variation of the renormalisation and factorisation scales leads to the largest variation and therefore independent variations are dropped. On the left side the transverse momentum of the additional jet system is presented, while on the right side the distance between the first and second additional jet is shown. Only events in the single-lepton or dilepton channel are considered. The distributions are normalised to unity. The ratio is calculated with respect to the nominal sample $(1\mu_R, 1\mu_F)$.

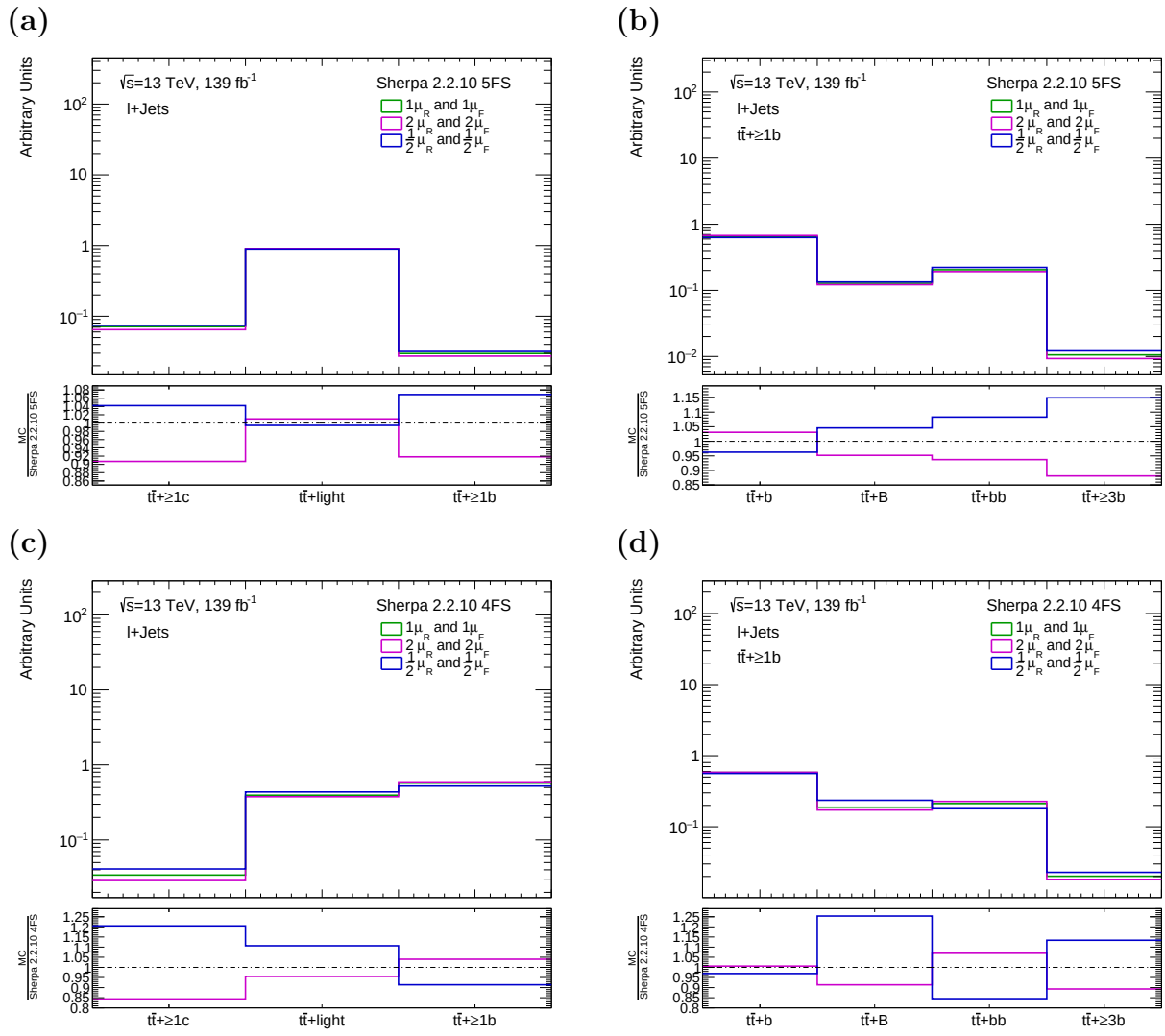

Figure 7.12.: The scale variation for the SHERPA2.2.10 samples in the five flavour scheme (top line) and the four flavour scheme (bottom line). The simultaneous variation of the renormalisation and factorisation scales leads to the largest variation and therefore independent variations are dropped. On the left side the *Heavy Flavour Simple Classification*, while on the right the *Heavy Flavour Classification* is shown. Only events in the single-lepton or dilepton channel are considered. The distributions are normalised to unity. The ratio is calculated with respect to the nominal sample ($1\mu_R, 1\mu_F$).

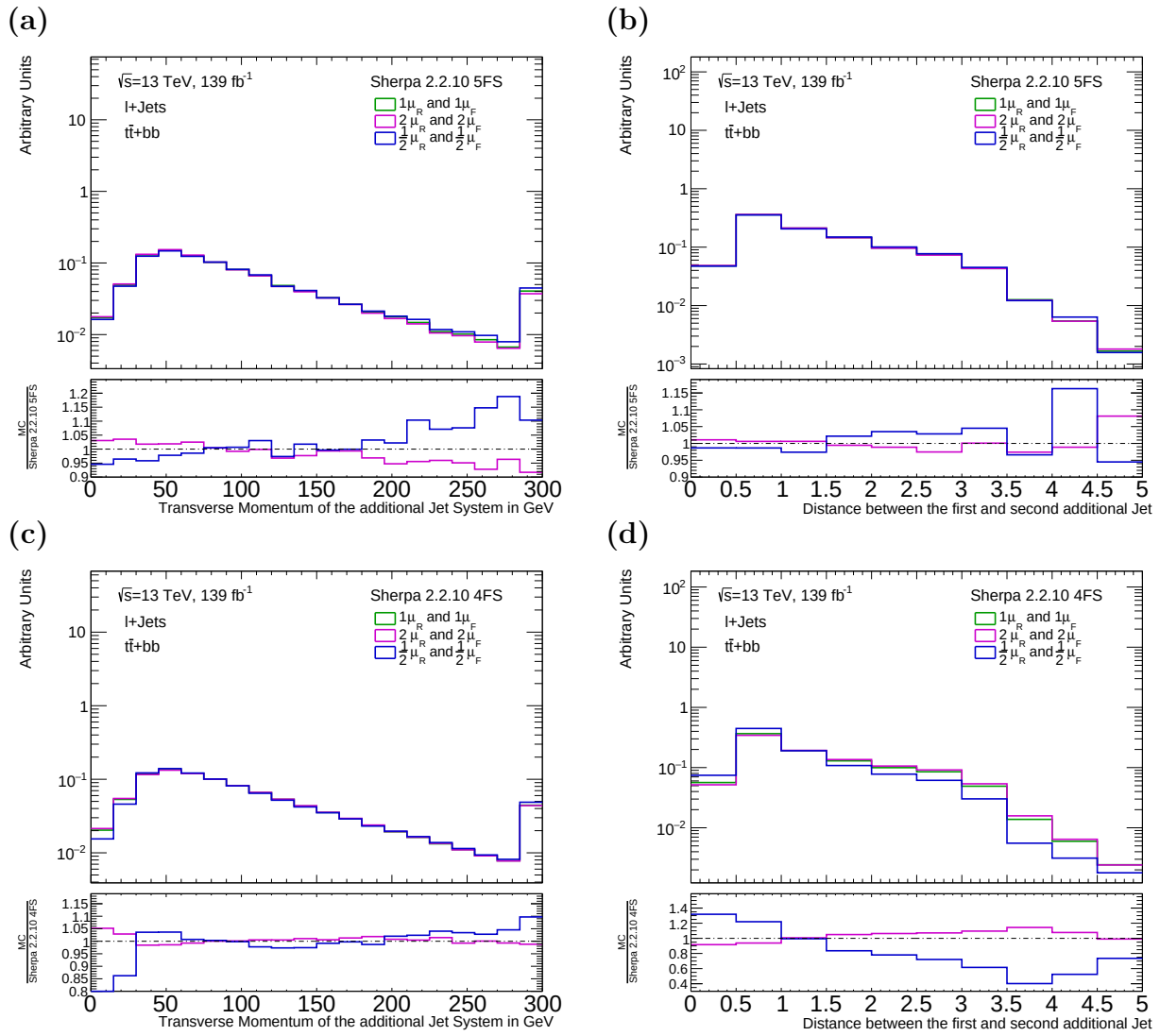

Figure 7.13.: The scale variation for the SHERPA2.2.10 in the five flavour scheme (top line) and the four flavour scheme (bottom line). The simultaneous variation of the renormalisation and factorisation scales leads to the largest variation and therefore independent variations are dropped. On the left side the transverse momentum of the additional jet system is presented, while on the right side the distance between the first and second additional jet is shown. Only events in the single-lepton or dilepton channel are considered. The distributions are normalised to unity. The ratio is calculated with respect to the nominal sample $(1\mu_R, 1\mu_F)$.

7.4. PDF Uncertainties

The PDFs are an important factor for the calculation of cross-sections and lifetimes in pp collisions. They are measured at a reference energy scale and extrapolated to other energy scales by the DGLAP equations, see Equation 4.1.4. For the PDF comparisons, a fixed set of events is used and reweighted and therefore statistical uncertainties are not shown. A short summary of the used PDF sets can be found in Table 6.2.

For this thesis, three PDF sets are compared: *NNPDF3.0NLO* [106], *MMHT2014nlo* [112] and *CT14nlo* [71]. The goal is to understand the impact on the fraction of events which corresponds to the *Heavy Flavour Simple Classification* and *Heavy Flavour Classification* and also the kinematics of the additional jets. The four flavour scheme samples use a modified PDF where the b-quark PDF is decoupled and the SHERPA setup uses the NNLO PDF set versions. For the SHERPA2.2.10 five and four flavour scheme samples the PDF set variations for the *Heavy Flavour Simple Classification* and *Heavy Flavour Classification* are shown in Figure 7.14. Moreover, the most interesting kinematic variables, the transverse momentum of the additional jet system and the distance between the jets, are presented in Figure 7.15. All distributions are normalised to unity and the ratio is with respect to the sample which uses the *NNPDF3.0NLO* PDF set.

In general, the variation of the PDF set is expected to have minor effect on the different categories of both flavour classifications. The smallest relative uncertainties corresponds to the four flavour scheme sample with less than 0.5 %, see Figure 7.14**(c)** and **(d)**. These can be explained by the different generation of the additional b-quarks. The kinematics is calculated at ME level and is therefore independent of the PDF since the b-quark PDF is decoupled, hence not active. The large relative uncertainties in the $t\bar{t}+ \geq 1c$ category are caused by the PDF set variation. The c-quark PDF is not decoupled and therefore a different PDF set should yield a different behaviour. The kinematics of the additional jets do not show any noticeable differences, see Figure 7.15**(c)** and **(d)**. However, the relative uncertainties of the transverse momentum of the additional jet system increases at the tail of the distribution up to 2 %.

The five flavour scheme has an active b-quark PDF and therefore a larger impact caused by the PDF variations is expected. The relative uncertainties of 1.5 % for the $t\bar{t}+ \geq 1b$ category are small but still larger compared with the four flavour scheme sample. This can be observed for the $t\bar{t}+B$ and $t\bar{t}+bb$ category of the *Heavy Flavour Classification*, as

well, see Figure 7.14**(a)** and **(b)**. The kinematics of the additional jets are not affected by the PDF choice, see Figure 7.15**(a)** and **(b)**. All variation have a similar shape and differ only at the tail of the transverse momentum distribution of the additional jet system.

For the internal variation of a PDF, an ensemble of PDF sets with different parameter settings is created. The internal variation is necessary to access the internal uncertainties of a PDF set since PDFs are fitted to data. The mean of the ensemble is used as the central value, while the error is the standard deviation of the PDF ensemble [125]. The *PDF4LHC15* [126] set is used for the internal variation and the results are presented in Figure 7.16 and Figure 7.17 for both flavour scheme setups. The first one shows the *Heavy Flavour Simple Classification* and *Heavy Flavour Classification* while the second one the transverse momentum of the additional jet system and the distance between the first and second additional jet.

The internal variation has a minor effect on the fraction of events for all categories. They do not exceed approximately 1 %, except for the $t\bar{t}+B$ category for the four flavour scheme sample, see Figure 7.16**(d)**. The relative uncertainties are larger and show a correlation with the distance between the additional jets. The down variation of the PDF set results in more events with a small distance between the additional jets and in a higher fraction of events for the $t\bar{t}+B$ category and vice versa. The remaining observables have no noticeable impact caused by the internal variation. For both flavour scheme setups the internal variation shows a minor effect on the transverse momentum of the additional jet system, see Figure 7.17**(a)** and **(c)**. Only at the tail the differences grow larger. However, for the four flavour scheme samples the internal variation has in the very soft regions larger differences.

Finally, the variation of PDF sets and the internal variation of one PDF set do not have a large impact on the fraction of events for the background categories and kinematics of the additional jets.

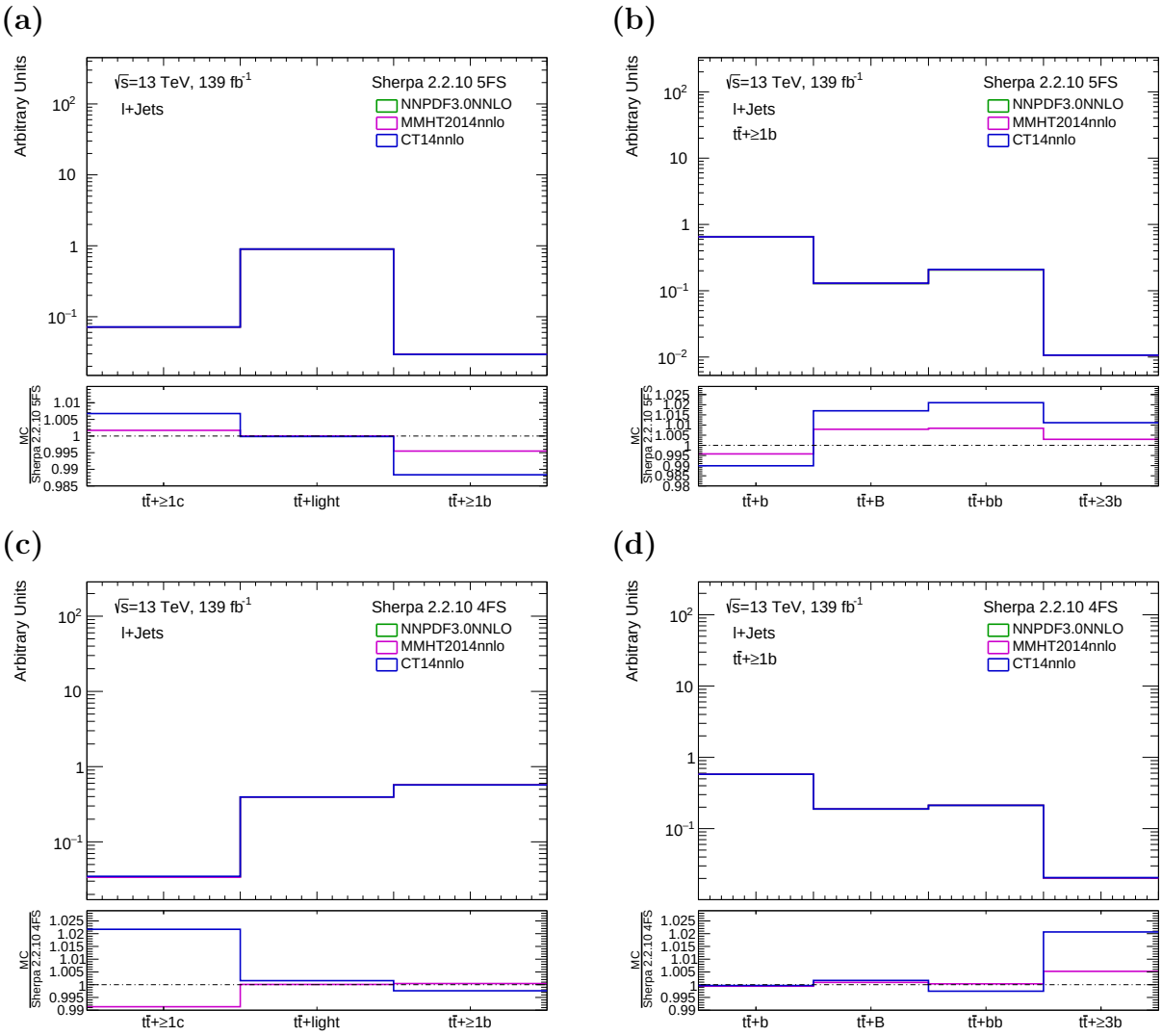

Figure 7.14.: The comparison of three different PDF sets, *NNPDF3.0NNLO*, *MMHT2014nnlo* and *CT14nnlo*, and their impact on the *Heavy Flavour Simple Classification* and *Heavy Flavour Classification*. In **(a)** and **(b)** the SHERPA2.2.10 5FS sample is used, while in **(c)** and **(d)** the SHERPA2.2.10 4FS one. All distributions are normalised to unity. The ratio is calculated with respect to the sample, which uses the *NNPDF3.0NNLO* PDF set.

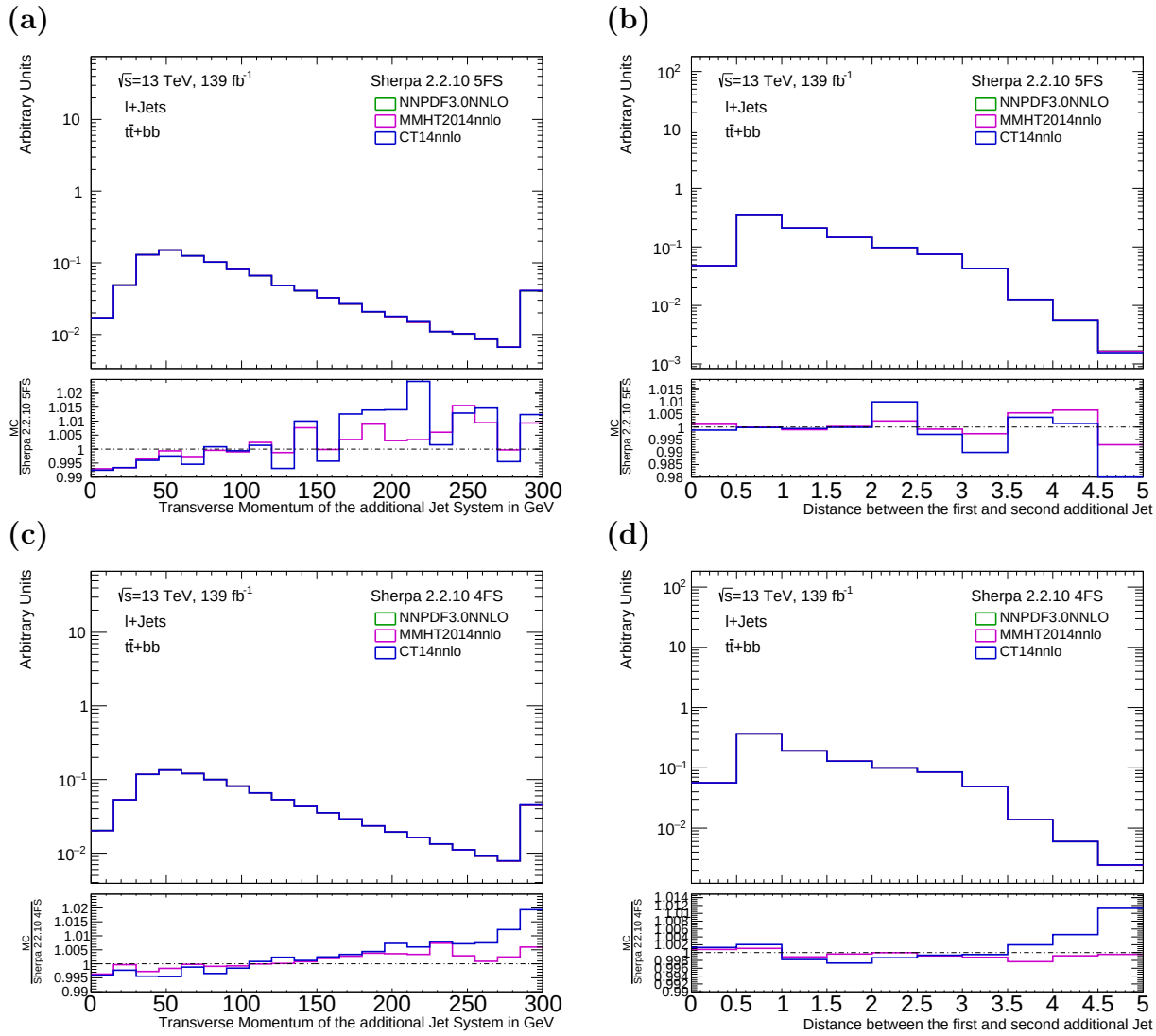

Figure 7.15.: The comparison of three different PDF sets, *NNPDF3.0NNLO*, *MMHT2014nnlo* and *CT14nnlo*, and their impact on the transverse momentum of the additional jet system and distance between the first and second additional jet. In **(a)** and **(b)** the SHERPA2.2.10 5FS sample is used, while in **(c)** and **(d)** the SHERPA2.2.10 4FS one. All distributions are normalised to unity. The ratio is calculated with respect to the sample, which uses the *NNPDF3.0NNLO* PDF set.

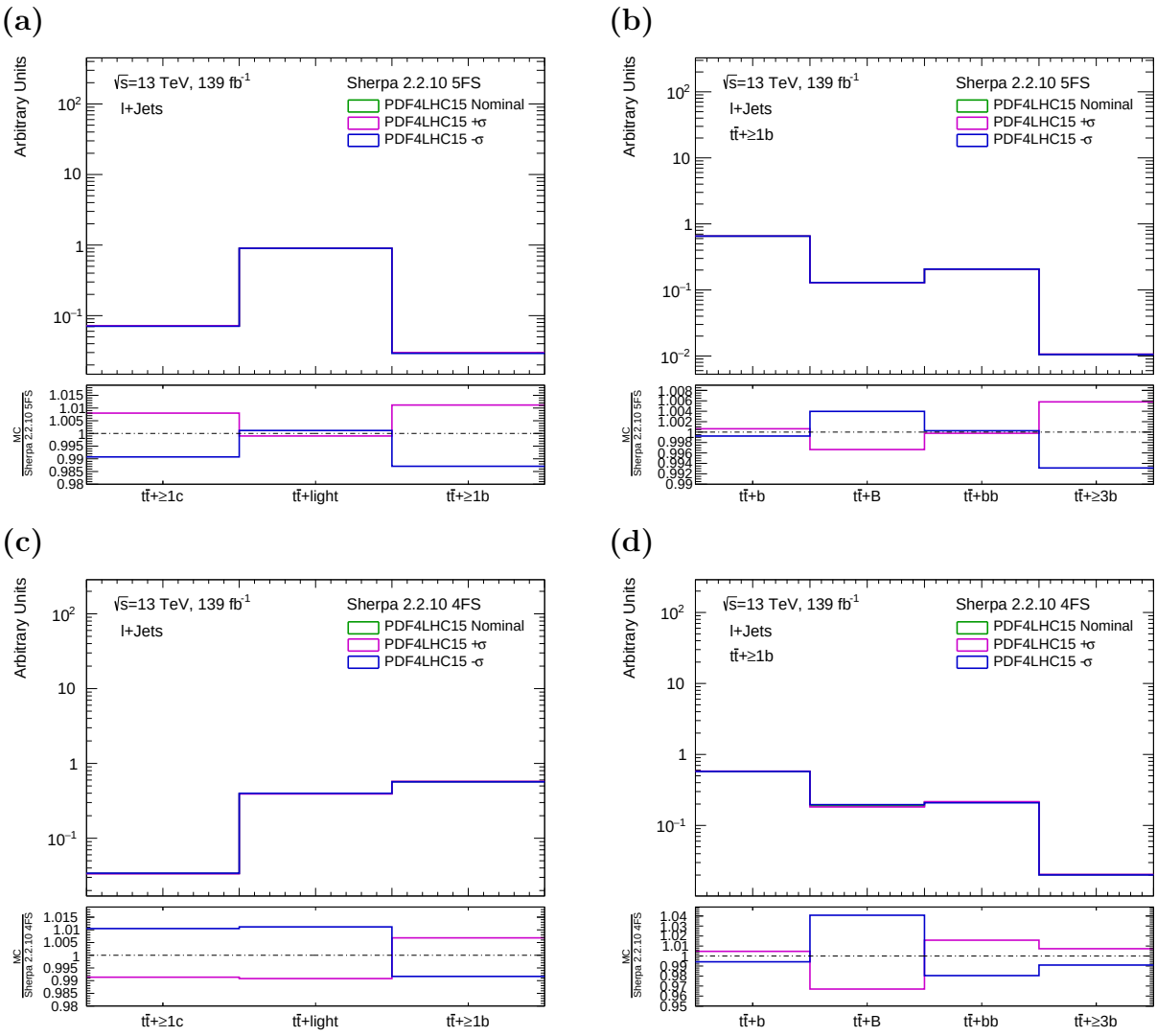

Figure 7.16.: The internal PDF variation of the *PDF4LHC15* PDF set for the four and five flavour SHERPA2.2.10 setups. On the left side the *Heavy Flavour Simple Classification* is presented, while on the right side the *Heavy Flavour Classification*. All distributions are normalised to unity and the ratio is calculated with respect to the nominal.

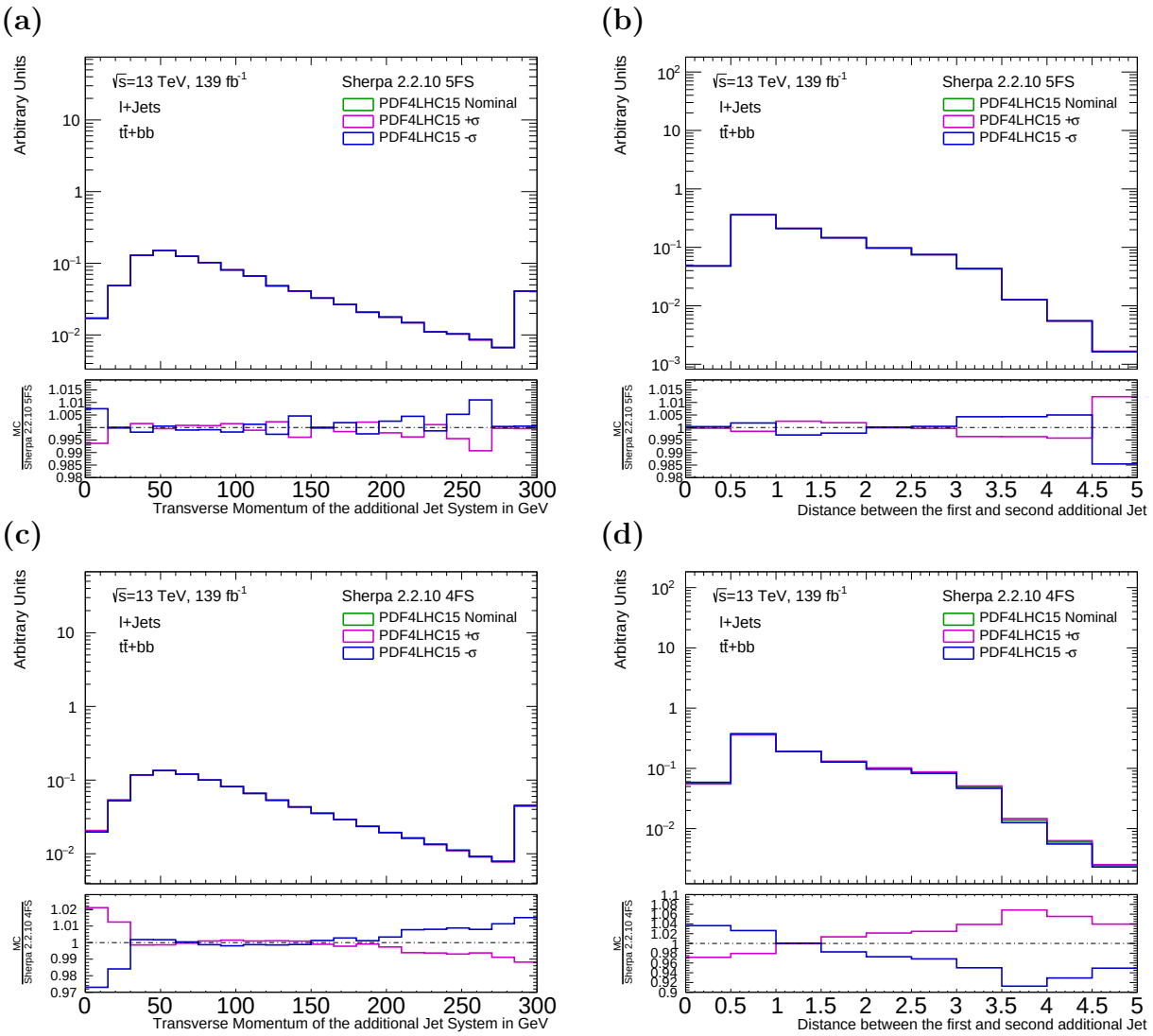

Figure 7.17.: The internal PDF variation of the *PDF4LHC15* PDF set for the four and five flavour SHERPA2.2.10 setups. On the left side the transverse momentum of the additional jet system is presented, while on the right side the distance between the additional jets. All distributions are normalised to unity and the ratio is calculated with respect to the nominal.

8. Outlook

The $t\bar{t}\left(H \to b\bar{b}\right)$ process is an opportunity to measure the Yukawa coupling of the Higgs boson to the most massive quarks and find hints for new physics or to restrict BSM predictions. Unfortunately, the $t\bar{t}\left(H \to b\bar{b}\right)$ process suffers from a large background caused by the irreducible $t\bar{t}+ \geq 1b$ background. Therefore the modelling uncertainties of this background are important. An updated set of Monte Carlo simulations with different implementation methods and settings was used to study the modelling uncertainties. Furthermore, this thesis looked at two different modelling approaches for the generation of b-quarks. In the five flavour scheme, the b-quark is assumed to be massless and is generated at parton shower level, while in the four flavour scheme the b-quark is massive and is generated at matrix element level. The different generation process results in different kinematic variable distributions and fractions of events for the various subcategories of the $t\bar{t}$ + jets background. A harder transverse momentum spectrum for the additional jets and the additional jet system was observed for the four flavour scheme samples and a more appropriate description for the softer transverse momentum regions of the additional jet system. Furthermore, a correlation between distance of the additional jets and the fraction of events for the $t\bar{t}+B$ category was observed. After the kinematic studies of the additional jets, the impact of the scale variations and PDF set choice was investigated. The scale variations only have a small impact on the fraction of events for the various subcategories of the $t\bar{t}$ + jets background. It only has a small effect on the kinematics of the additional jets. For the PDF set variation and internal PDF variation, a similar impact was observed.

For further $t\bar{t}\left(H \to b\bar{b}\right)$ measurements a full set of systematic uncertainties has to be defined. The parton shower and hadronisation model, the scale choice and the PDF choice may play an important role to estimate the systematic uncertainties. Furthermore, there are approaches to combine four and five flavour scheme samples [127]. The goal is to describe all kinematic regions properly using the different strength of the flavour schemes. The combination is not trivial. One has to avoid double counting of events and

the full phase space needs to be covered. In the end the merging should result in smooth kinematic distributions which should agree with measurements.

A. Additional Tables

A.1. Derivation Names

In this section the full names of the derivations which are used in this thesis are presented in Table A.1. The derivations are distinguished in samples which describes the $t\bar{t}$ inclusive process and the $t\bar{t} + b\bar{b}$ process. Furthermore, the samples are generated depending on the final state leptons into single-lepton ($\pm$), dilepton and non-all hadronic derivations.

Table A.1.: The derivations which are used in this thesis.

Process	Derivation Name
$t\bar{t} + b\bar{b}$	*mc16_13TeV.410323.Sherpa_NNPDF30NNLO_ttbb_lplush .deriv.DAOD_TOPQ1.e5695_a875_r9364_p4346*
$t\bar{t} + b\bar{b}$	*mc16_13TeV.410324.Sherpa_NNPDF30NNLO_ttbb_lminush .deriv.DAOD_TOPQ1.e5695_a875_r9364_p4346*
$t\bar{t} + b\bar{b}$	*mc16_13TeV.410325.Sherpa_NNPDF30NNLO_ttbb_lpluslminus .deriv.DAOD_TOPQ1.e5695_a875_r9364_p4346*
$t\bar{t} + b\bar{b}$	*mc16_13TeV.411178.PhPy8EG_A14_NNPDF30ME_ttbb_4FS_MS_dilep .deriv.DAOD_TOPQ1.e7818_a875_r9364_p4346*
$t\bar{t} + b\bar{b}$	*mc16_13TeV.411179.PhPy8EG_A14_NNPDF30ME_ttbb_4FS_MS_lplus .deriv.DAOD_TOPQ1.e7818_a875_r9364_p4346*
$t\bar{t} + b\bar{b}$	*mc16_13TeV.411180.PhPy8EG_A14_NNPDF30ME_ttbb_4FS_MS_lminus .deriv.DAOD_TOPQ1.e7818_a875_r9364_p4346*
$t\bar{t} + b\bar{b}$	*mc16_13TeV.700165.Sh_2210_ttbb_SingleLeptonP .deriv.DAOD_TOPQ1.e8263_a875_r9364_p4434*
$t\bar{t} + b\bar{b}$	*mc16_13TeV.700166.Sh_2210_ttbbSingleLeptonM .deriv.DAOD_TOPQ1.e8263_a875_r9364_p4434*
$t\bar{t} + b\bar{b}$	*mc16_13TeV.700167.Sh_2210_ttbb_dilepton .deriv.DAOD_TOPQ1.e8263_a875_r9364_p4434*

$t\bar{t}$	*mc16_13TeV.410470.PhPy8EG_A14_ttbar_hdamp258p75 _nonallhad.deriv.DAOD_TOPQ1.e6337_s3126_r9364_p4346*
$t\bar{t}$	*mc16_13TeV.410250.Sherpa_221_NNPDF30NNLO_ttbar_SingleLeptonP _MEPS_NLO.deriv.DAOD_TOPQ1.e5450_s3126_r9364_p4346*
$t\bar{t}$	*mc16_13TeV.410251.Sherpa_221_NNPDF30NNLO_ttbar_SingleLeptonM _MEPS_NLO.deriv.DAOD_TOPQ1.e5450_s3126_r9364_p4346*
$t\bar{t}$	*mc16_13TeV.410252.Sherpa_221_NNPDF30NNLO_ttbar_dilepton _MEPS_NLO.deriv.DAOD_TOPQ1.e5450_s3126_r9364_p4346*
$t\bar{t}$	*mc16_13TeV.410557.PowhegHerwig7EvtGen_H7UE_tt_hdamp258p75 _704_SingleLep.deriv.DAOD_TOPQ1.e6366_a875_r9364_p4346*
$t\bar{t}$	*mc16_13TeV.410558.PowhegHerwig7EvtGen_H7UE_tt_hdamp258p75 _704_dil.deriv.DAOD_TOPQ1.e6366_a875_r9364_p4346*
$t\bar{t}$	*mc16_13TeV.410464.aMcAtNloPy8EvtGen_MEN30NLO_A14N23LO_ttbar _noShWe_SingleLep.deriv.DAOD_TOPQ1.e6762_a875_r9364_p4346*
$t\bar{t}$	*mc16_13TeV.410465.aMcAtNloPy8EvtGen_MEN30NLO_A14N23LO_ttbar _noShWe_dil.deriv.DAOD_TOPQ1.e6762_a875_r9364_p4346*
$t\bar{t}$	*mc16_13TeV.411233.PowhegHerwig7EvtGen_tt_hdamp258p75_713_SingleLep .deriv.DAOD_TOPQ1.e7580_s3126_r9364_p4066*
$t\bar{t}$	*mc16_13TeV.411234.PowhegHerwig7EvtGen_tt_hdamp258p75_713_dil .deriv.DAOD_TOPQ1.e7580_s3126_r9364_p4066*
$t\bar{t}$	*mc16_13TeV.412116.aMcAtNloHerwig7EvtGen_MEN30NLO_ttbar_SingleLep .deriv.DAOD_TOPQ1.e7620_a875_r9364_p4346*
$t\bar{t}$	*mc16_13TeV.412117.aMcAtNloHerwig7EvtGen_MEN30NLO_ttbar_dil .deriv.DAOD_TOPQ1.e7620_a875_r9364_p4346*
$t\bar{t}$	*mc16_13TeV.700122.Sh_2210_ttbar_SingleLeptonP_maxHTavrgTopPT_SSC .deriv.DAOD_TOPQ1.e8253_s3126_r9364_p4434*
$t\bar{t}$	*mc16_13TeV.700123.Sh_2210_ttbar_SingleLeptonM_maxHTavrgTopPT_SSC .deriv.DAOD_TOPQ1.e8253_s3126_r9364_p4434*
$t\bar{t}$	*mc16_13TeV.700124.Sh_2210_ttbar_dilepton_maxHTavrgTopPT_SSC .deriv.DAOD_TOPQ1.e8253_s3126_r9364_p4434*

A.2. Yield Tables

The yields for various observables are presented in this section. The errors are all statistical ones. In Table A.2 the yields for the lepton multiplicity N^{ℓ} are presented. In Table A.3 the yields for the *Heavy Flavour Simple Classification* are shown. The finer *Heavy Flavour Classification* can be seen in Table A.4.

Table A.2.: The yields for the various MC setups for the lepton multiplicity with statistical uncertainties.

Number of Leptons N^{ℓ}	0	1	2	≥ 3
Powheg+Pythia8 5FS	25243800 ± 5024	33149100 ± 5757	4391640 ± 2096	92445 ± 304
Powheg+Herwig7.1.3 5FS	24696100 ± 4969	33590800 ± 5795	4514380 ± 2124	94078 ± 306
Powheg+Herwig7.0.4 5FS	25481900 ± 5047	33084000 ± 5751	424070 ± 2059	73167 ± 270
aMc@nlo +Pythia8 5FS	25583900 ± 5058	33296100 ± 5770	433670 ± 2081	86196 ± 293
aMc@nlo +Herwig7.1.3 5FS	24859000 ± 4985	33778900 ± 5811	4569120 ± 2137	94211 ± 307
Powheg+Pythia8 4FS	904389 ± 950	1256680 ± 1121	181227 ± 425	5551 ± 74
Sherpa2.2.1 5FS	24988800 ± 4998	33342200 ± 5774	4499500 ± 2121	102203 ± 319
Sherpa2.2.1 4FS	757027 ± 870	1056890 ± 1028	156720 ± 395	5188 ± 72
Sherpa2.2.10 5FS	23950600 ± 4893	34050000 ± 5835	4771970 ± 2184	109950 ± 331
Sherpa2.2.10 4FS	1030760 ± 1015	1453000 ± 1207	210844 ± 459	6044 ± 77

Table A.3.: The yields for the various MC setups for the Heavy Flavour Simple Classification with statistical uncertainties.

Category	$t\bar{t}$ + light	$t\bar{t}+ \geq 1c$	$t\bar{t}+ \geq 1b$
Powheg+Pythia8 5FS	33639600 ± 5800	2913180 ± 1707	988045 ± 994
Powheg+Herwig7.1.3 5FS	34655200 ± 5886	2345540 ± 1531	1104450 ± 1050
Powheg+Herwig7.0.4 5FS	34447500 ± 5869	1985240 ± 1408	893342 ± 945
aMc@nlo +Pythia8 5FS	33762400 ± 5810	2892650 ± 1700	974710 ± 987
aMc@nlo +Herwig7.1.3 5FS	35362600 ± 5946	2109180 ± 1452	876260 ± 936
Powheg+Pythia8 4FS	545051 ± 738	32704 ± 180	860154 ± 927
Sherpa2.2.1 5FS	33146600 ± 5757	3297200 ± 1815	1397910 ± 1182
Sherpa2.2.1 4FS	446272 ± 668	39808 ± 199	727534 ± 852
Sherpa2.2.10 5FS	34900900 ± 5907	2767060 ± 1663	1154090 ± 1074
Sherpa2.2.10 4FS	658094 ± 811	56840 ± 238	953211 ± 976

Table A.4.: The yields for the various MC setups for the Heavy Flavour Classification with statistical uncertainties.

Category	$t\bar{t}+b$	$t\bar{t}+B$	$t\bar{t}+bb$	$t\bar{t}+\geq 3b$
Powheg+Pythia8 5FS	556140 ± 745	128485 ± 358	188821 ± 434	9136 ± 95
Powheg+Herwig7.1.3 5FS	640727 ± 800	180927 ± 425	199689 ± 446	9716 ± 98
Powheg+Herwig7.0.4 5FS	577507 ± 760	113042 ± 336	145923 ± 381	5887 ± 76
aMc@nlo +Pythia8 5FS	564888 ± 751	103190 ± 321	185366 ± 430	7656 ± 87
aMc@nlo +Herwig7.1.3 5FS	473583 ± 688	187140 ± 432	132120 ± 363	7309 ± 85
Powheg+Pythia8 4FS	510196 ± 714	134152 ± 366	202837 ± 450	12969 ± 113
Sherpa2.2.1 5FS	835592 ± 914	212919 ± 461	253680 ± 503	17850 ± 133
Sherpa2.2.1 4FS	422834 ± 650	122306 ± 349	167607 ± 409	14787 ± 121
Sherpa2.2.10 5FS	705261 ± 839	137549 ± 370	220658 ± 469	11354 ± 106
Sherpa2.2.10 4FS	552056 ± 743	179412 ± 423	202532 ± 450	19211 ± 138

B. Additional Plots

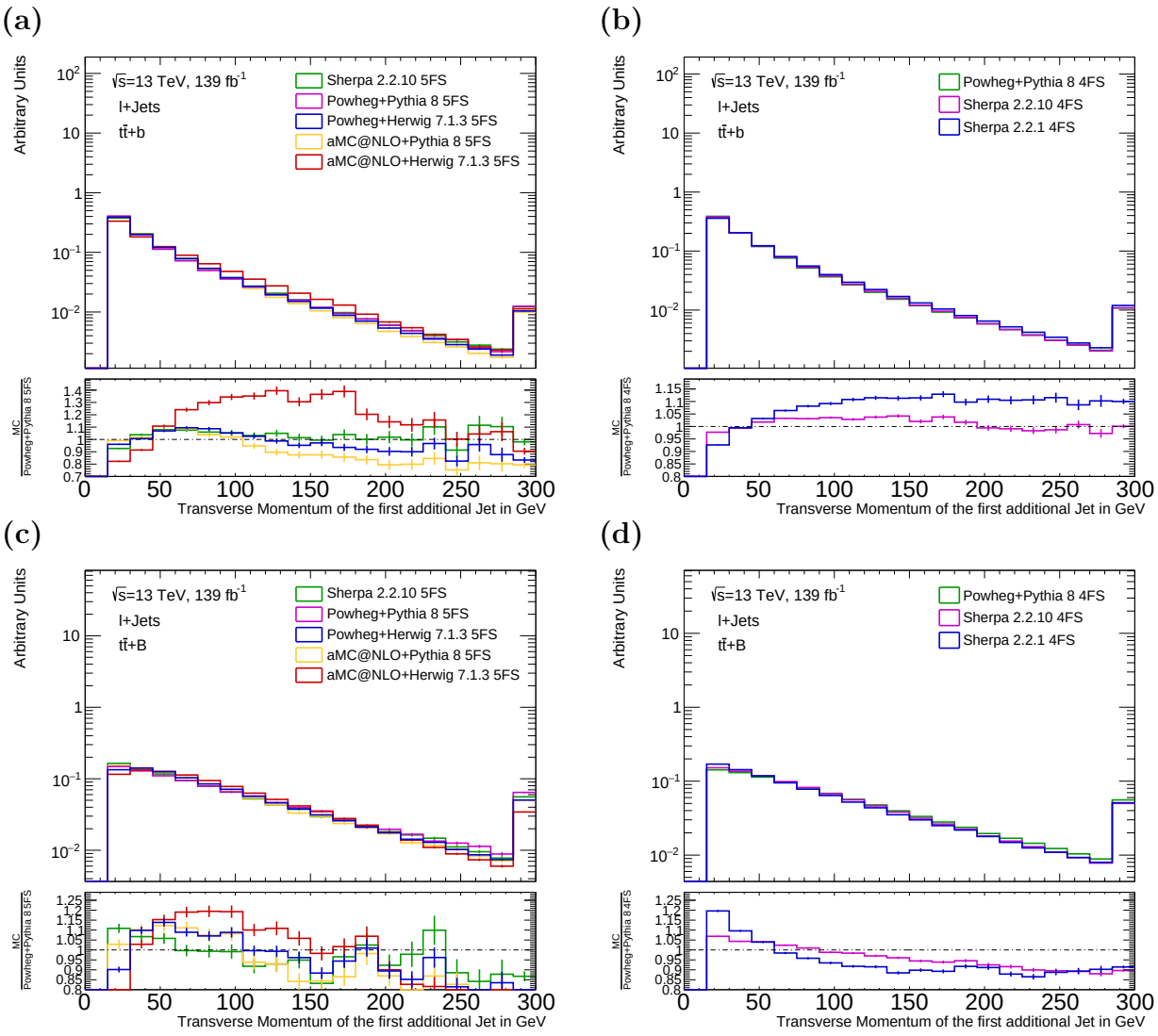

Figure B.1.: The distributions for the transverse momentum of the first additional b-jet for the $t\bar{t}+b$ category on the top side and the $t\bar{t}+B$ on the bottom side. In **(a)** and **(c)** the five flavour scheme sample setups are presented. The four flavour scheme samples are compared in **(b)** and **(d)**. All distributions are normalised to unity. The ratio is calculated for the four flavour scheme samples with respect to Powheg+Pythia8 4FS and for the five flavour ones with respect to the nominal sample Powheg+Pythia8 5FS.

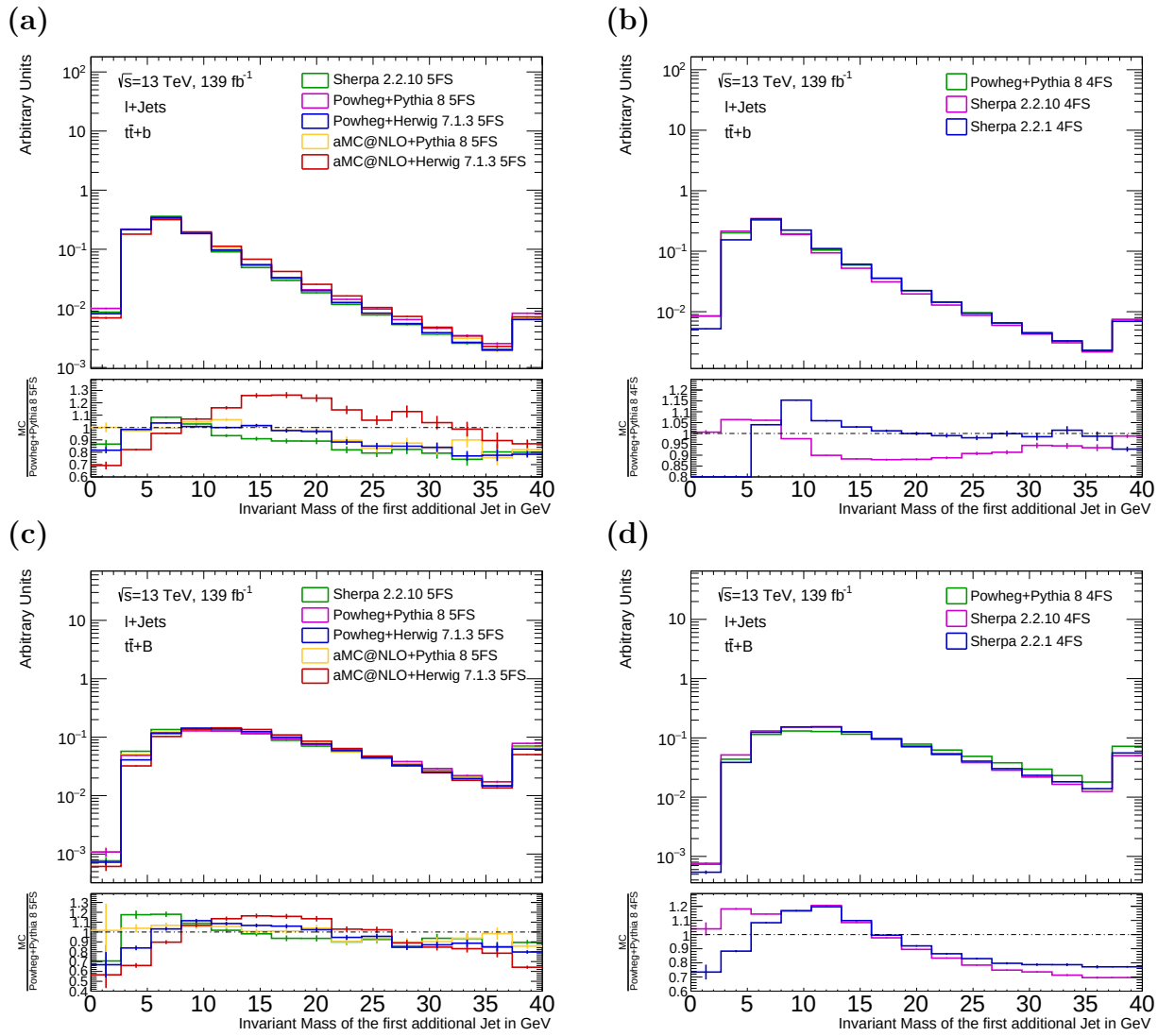

Figure B.2.: The distributions for the invariant mass of the first additional b-jet for the $t\bar{t}+b$ category on the top side and the $t\bar{t}+B$ on the bottom side. In **(a)** and **(c)** the five flavour scheme sample setups are presented. The four flavour scheme samples are compared in **(b)** and **(d)**. All distributions are normalised to unity. The ratio is calculated for the four flavour scheme samples with respect to POWHEG+PYTHIA8 4FS and for the five flavour ones with respect to the nominal sample POWHEG+PYTHIA8 5FS.

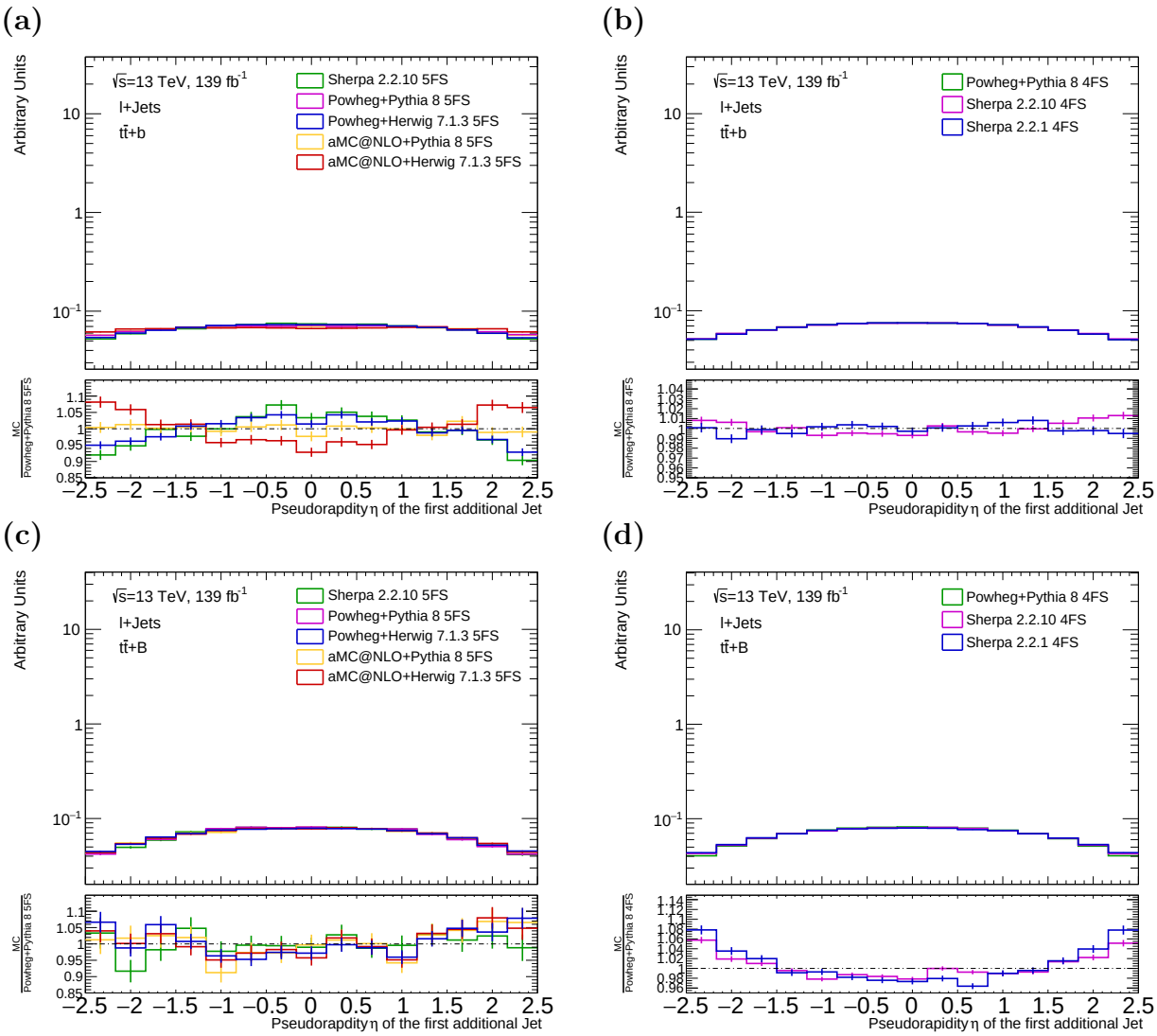

Figure B.3.: The distributions for the pseudorapidity η of the first additional b-jet for the $t\bar{t} + b$ category on the top side and the $t\bar{t} + B$ on the bottom side. In **(a)** and **(c)** the five flavour scheme sample setups are presented. The four flavour scheme samples are compared in **(b)** and **(d)**. All distributions are normalised to unity. The ratio is calculated for the four flavour scheme samples with respect to Powheg+Pythia8 4FS and for the five flavour ones with respect to the nominal sample Powheg+Pythia8 5FS. A centralised distribution is expected for all samples.

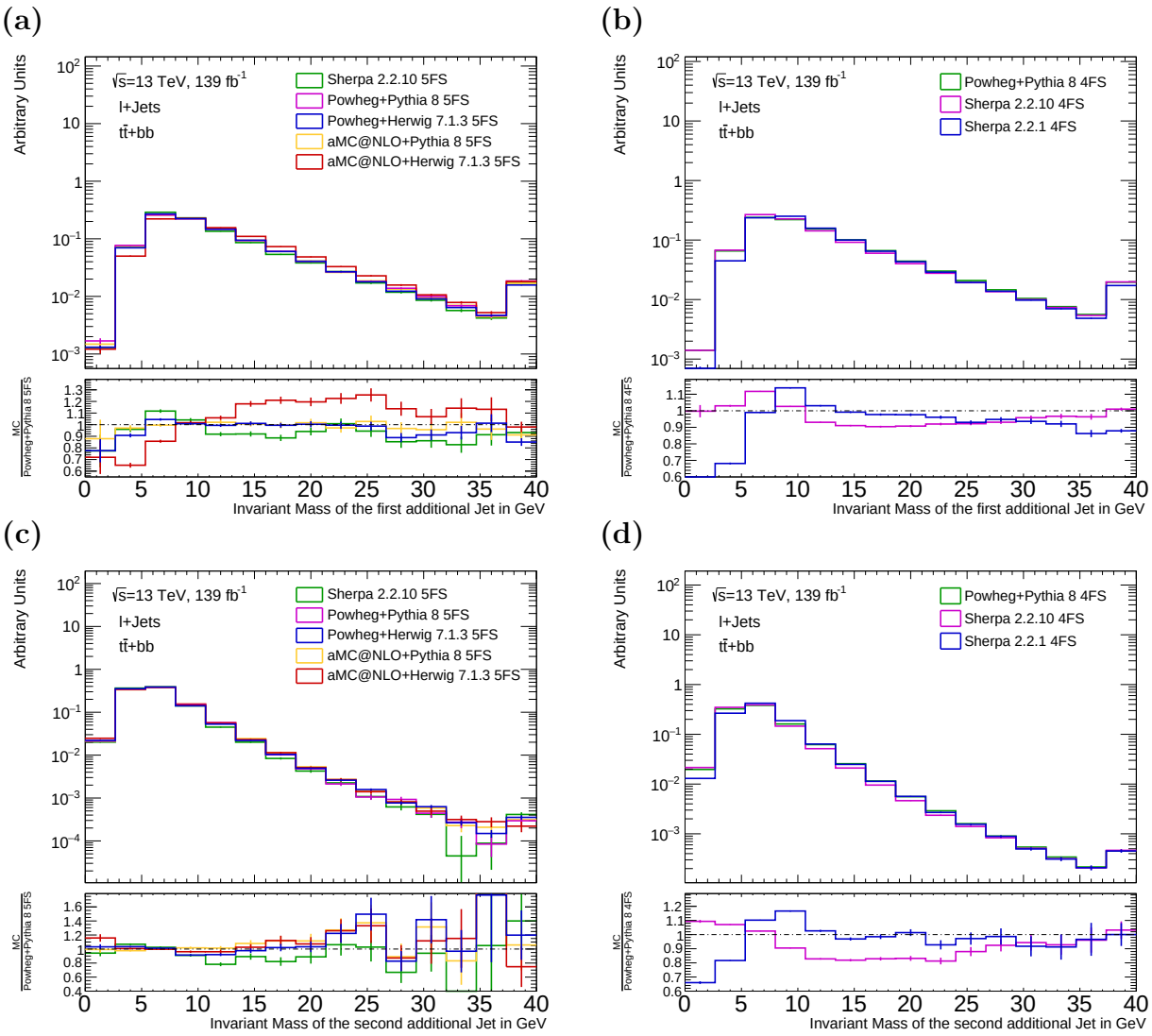

Figure B.4.: The distributions for the invariant mass of the first additional b-jet for the $t\bar{t} + bb$ category on the top side and the second additional b-jet on the bottom side. In **(a)** and **(c)** the five flavour scheme sample setups are presented. The four flavour scheme samples are compared in **(b)** and **(d)**. All distributions are normalised to unity. The ratio is calculated for the four flavour scheme samples with respect to POWHEG+PYTHIA8 4FS and for the five flavour ones with respect to the nominal sample POWHEG+PYTHIA8 5FS. A centralised distribution is expected for all samples.

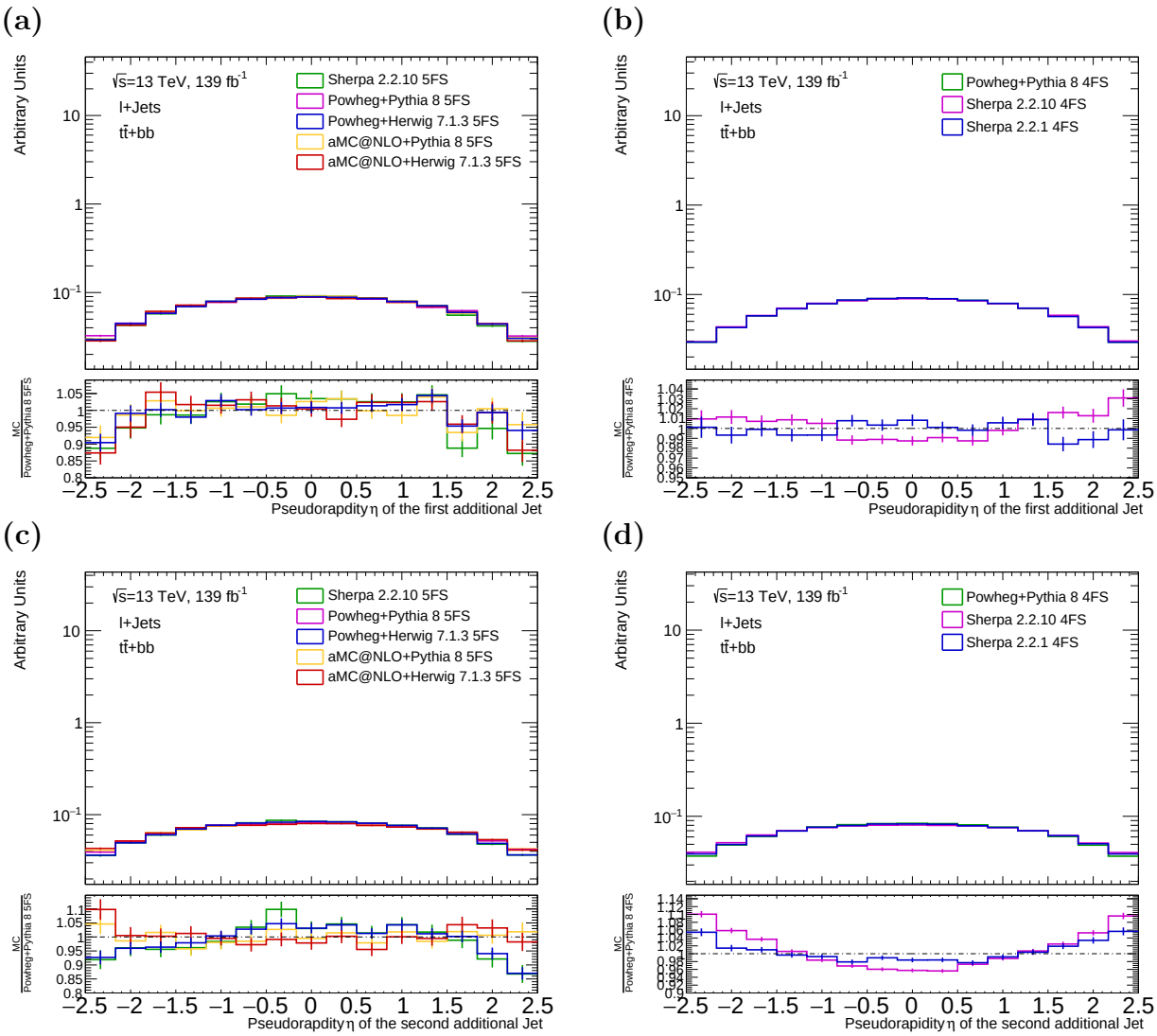

Figure B.5.: The distributions for the pseudorapidity η of the first additional b-jet for the $t\bar{t} + bb$ category on the top side and the second additional b-jet on the bottom side. In **(a)** and **(c)** the five flavour scheme sample setups are presented. The four flavour scheme samples are compared in **(b)** and **(d)**. All distributions are normalised to unity. The ratio is calculated for the four flavour scheme samples with respect to POWHEG+PYTHIA8 4FS and for the five flavour ones with respect to the nominal sample POWHEG+PYTHIA8 5FS. A centralised distribution is expected for all samples.

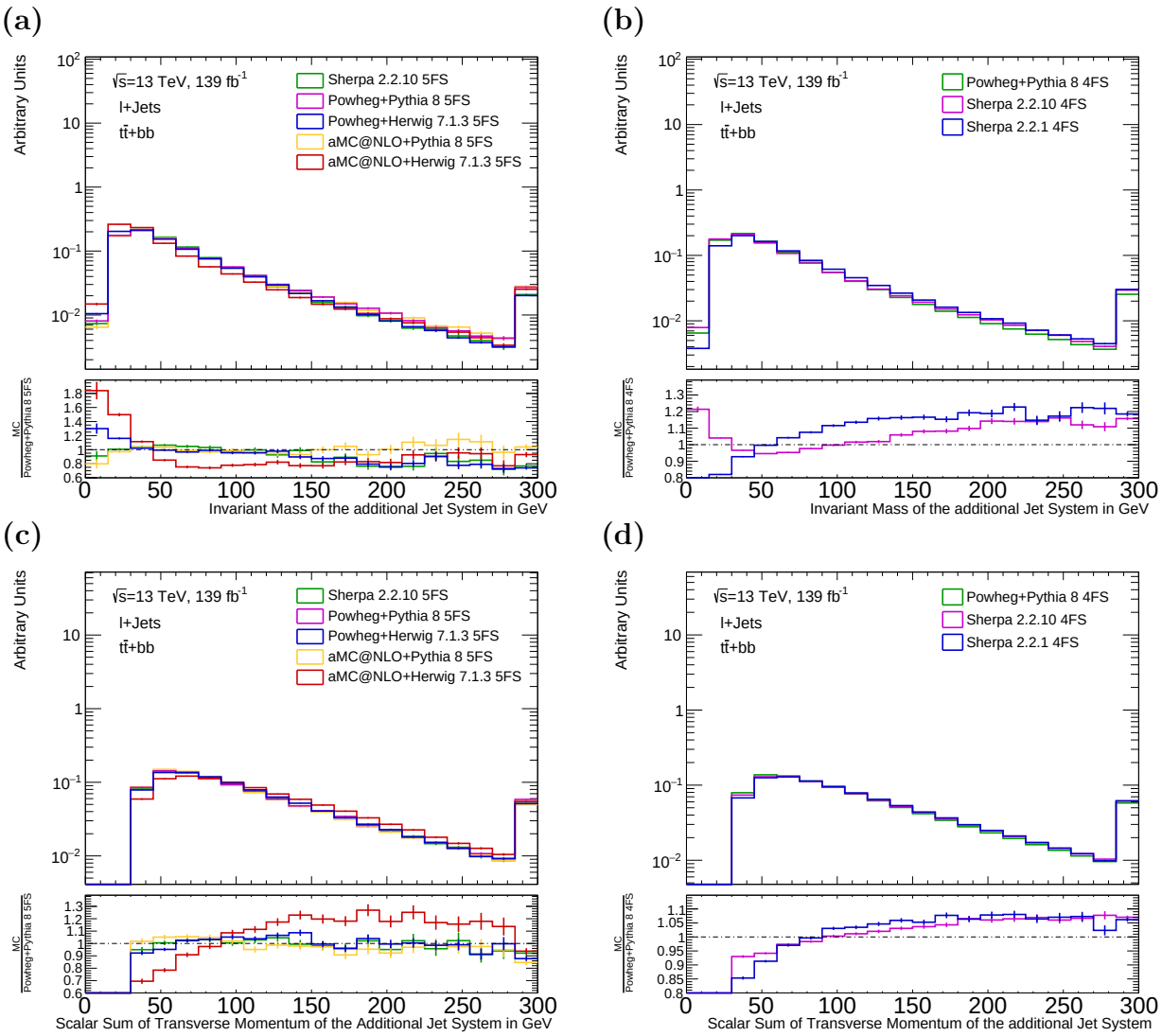

Figure B.6.: The distributions for the invariant mass of the additional jet system for the $t\bar{t} + bb$ category is presented. In **(a)** the five flavour scheme sample setups are shown. The four flavour scheme samples are compared in **(b)**. In the lower row the scalar sum of transverse momentum of the additional jet system is shown. All distributions are normalised to unity. The ratio is calculated for the four flavour scheme samples with respect to POWHEG+PYTHIA8 4FS and for the five flavour ones with respect to the nominal sample POWHEG+PYTHIA8 5FS. The invariant mass distributions do not show any mass resonances.

Bibliography

[1] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1.

[2] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30.

[3] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.

[4] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.

[5] P. W. Higgs, *Spontaneous symmetry breakdown without massless bosons*, Phys. Rev. **145** (1966) 1156.

[6] G. S. Guralnik et al., *Global conservation laws and massless particles*, Phys. Rev. Lett. **13** (1964) 585.

[7] T. W. Kibble, *Symmetry breaking in non-Abelian gauge theories*, Phys. Rev. **155** (1967) 1554.

[8] ATLAS Collaboration, *Observation of Higgs boson production in association with a top quark pair at the LHC with the ATLAS detector*, Phys. Lett. B **784** (2018) 173.

[9] ATLAS Collaboration, *Observation of $H \to b\bar{b}$ decays and VH production with the ATLAS detector*, Phys. Lett. B **786** (2018) 59.

[10] CMS Collaboration, *Observation of $t\bar{t}H$ production*, Phys. Rev. Lett. **120** (2018) no. 23, 231801.

[11] CMS Collaboration, *Observation of Higgs boson decay to bottom quarks*, Phys. Lett. B **121** (2018) 121801.

[12] S. Pozzorini et al., *New NLOPS predictions for $t\bar{t}$ + b-jet production at the LHC*,

Eur. Phys. J. C **78** (2018) 1.

[13] E. Bothmann et al., *Event generation with Sherpa 2.2.*, SciPost Phys. **7** (2019) 034.

[14] Particle Data Group, *Review of particle physics*, Prog. Theor. Exp. Phys. **2020** (2020).

[15] M. Gell-Mann, *A schematic model of baryons and mesons*, Phys. Lett. **8** (1964) 214.

[16] D. J. Gross and F. Wilczek, *Ultraviolet behavior of non-abelian gauge theories*, Phys. Rev. Lett. **30** (1973) 1343.

[17] K. G. Wilson, *Confinement of quarks*, Phys. Rev. D **10** (1974) 2445.

[18] M. Gell-Mann, *Symmetries of baryons and mesons*, Phys. Rev. **125** (1962) 1067.

[19] Y. Ne'eman, *Derivation of strong interactions from a gauge invariance*, Nucl. Phys. **26** (1961) 222.

[20] G. t'Hooft, *Renormalizable lagrangians for massive Yang-Mills fields*, Nucl. Phys. B **35** (1971) 167.

[21] G. 't Hooft and M. J. G. Veltman, *Regularization and renormalization of gauge fields*, Nucl. Phys. B **44** (1972) 189.

[22] G. 't Hooft and M. J. G. Veltman, *Combinatorics of gauge fields*, Nucl. Phys. B **50** (1972) 318.

[23] E. Wigner, *On unitary representations of the inhomogeneous Lorentz group*, Ann. Math. (1939) 149.

[24] P. A. M. Dirac, *The quantum theory of the electron*, Proc. R. Soc. Lond. A **117** (1928) 610.

[25] E. Noether, *Invariante Variationsprobleme*, Nachr. D. König. Gesellsch. D. Wiss. Zu Göttingen, Math-phys. Klasse, S. (1918) 235.

[26] C. S. Wu et al., *Experimental test of parity conservation in beta decay*, Phys. Rev. **105** (1957) 1413.

[27] M. Goldhaber et al., *Helicity of neutrinos*, Phys. Rev. **109** (1958) 1015.

[28] LHCb Collaboration, *Measurement of the forward-backward asymmetry in $Z/\gamma^* \to \mu^+\mu^-$ decays and determination of the effective weak mixing angle*, JHEP 01 (2015) 190.

[29] S. Weinberg, *A model of leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[30] S. L. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. **22** (1961) 579.

[31] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. **C68519** (1968) 367.

[32] S. L. Glashow, J. Iliopoulos, and L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. **D2** (1970) 1285.

[33] M. Kobayashi and T. Maskawa, *CP-violation in the renormalizable theory of weak interaction*, Prog. Theor. Phys. **49** (1973) 652.

[34] ATLAS Collaboration, *Measurement of the Higgs boson mass in the $H \to ZZ \to 4l$ and $H \to \gamma\gamma$ channels with $\sqrt{\mathbf{s}} = 13$ TeV pp collisions using the ATLAS detector*, Phys. Lett. B **784** (2018) 345.

[35] ATLAS Collaboration, *Evidence for the spin-0 nature of the Higgs boson using ATLAS data*, Phys. Lett. B **726** (2013) 120.

[36] CMS Collaboration, *Study of the mass and spin-parity of the Higgs boson candidate via its decays to Z boson pairs*, Phys. Rev. Lett. **110** (2013) 081803.

[37] ATLAS Collaboration, *Evidence for the Higgs-boson Yukawa coupling to tau leptons with the ATLAS detector*, JHEP **2015** (2015) no. 4, 117.

[38] CMS Collaboration, *Evidence for the 125 GeV Higgs boson decaying to a pair of τ leptons*, JHEP 05 (2014) 104.

[39] CMS Collaboration, *Observation of the Higgs boson decay to a pair of tau leptons with the CMS detector*, Phys. Lett. B **779** (2018) 283.

[40] E. Drechsler, *Measurement of the $H \to \tau^+\tau^-$ Production Cross-Section in Hadronic Final States in 13 TeV Proton-Proton Collisions with the ATLAS Detector*, PhD thesis, Georg August Universität Göttingen, II.Physik-UniGö-Diss-2018/02, 2018.

[41] A. De Maria, *Measurement of the Higgs boson production in the $H \to \tau\tau \to \tau_{lep}\tau_{had}$ decay channel at $\sqrt{s} = 13$ TeV with the ATLAS detector at the LHC*, PhD thesis, Georg August Universität Göttingen, II.Physik-UniGö-Diss-2019/04, 2019.

[42] ATLAS Collaboration, *A search for the dimuon decay of the Standard Model Higgs boson with the ATLAS detector*, Phys. Lett. B **812** (2021) 135980.

[43] CMS Collaboration, *Evidence for Higgs boson decay to a pair of muons*, JHEP

2021 (2021) no. 1, 1.

[44] ATLAS Collaboration, *Search for the Decay of the Higgs Boson to Charm Quarks with the ATLAS Experiment*, Phys. Rev. Lett. **120** (2018) no. 21, 211802.

[45] CMS Collaboration, *A search for the standard model Higgs boson decaying to charm quarks*, JHEP **2020** (2020) no. 3, 1.

[46] D. de Florian et al., *Handbook of LHC Higgs cross sections: 4. Deciphering the nature of the Higgs sector*, CERN Yellow Reports: Monographs **2/2017** (2016).

[47] O. Brüning et al., *LHC design report(Volume I, The LHC main ring)*, CERN Yellow Reports: Monographs (2004).

[48] LHCb Collaboration, *The LHCb detector at the LHC*, JINST **3** (2008) 1748.

[49] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.

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

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

[52] ATLAS Collaboration, *ATLAS Insertable B-Layer Technical Design Report*, ATLAS-TDR-19 (2010). Addendum: ATLAS-TDR-19-ADD-1, 2012.

[53] ATLAS Collaboration, *The ATLAS Level-1 Topological Trigger performance in Run 2*, J. Phys. Conf. Ser **898** (2017) 032037.

[54] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, Eur. Phys. J. C **77** (2017) 317.

[55] C. Glasman, *Precision measurements of α_S at HERA*, J. Phys. Conf. Ser. **110** (2008) no. 2, 022013.

[56] T. Biekötter, M. Klasen, and G. Kramer, *Next-to-next-to-leading order contributions to inclusive jet production in deep-inelastic scattering and determination of α_S*, Phys. Rev. D **92** (2015) no. 7, 074037.

[57] J. Blümlein, H. Böttcher, and A. Guffanti, *Non-singlet QCD analysis of deep inelastic world data at $\mathcal{O}(\alpha_S^3)$*, Nucl. Phys. B **774** (2007) no. 1-3, 182.

[58] S. Alekhin et al., *Parton distribution functions, α_S, and heavy-quark masses for*

LHC Run II, Phys. Rev. D **96** (2017) no. 1, 014011.

[59] K. G. Wilson, *Non-Lagrangian models of current algebra*, Phys. Rev. **179** (1969) no. 5, 1499.

[60] C. G. Callan Jr., *Broken scale invariance in scalar field theory*, Phys. Rev. D **2** (1970) no. 8, 1541.

[61] K. Symanzik, *Small distance behaviour in field theory and power counting*, Commun. Math. Phys. **18** (1970) no. 3, 227.

[62] N. Christ, B. Hasslacher, and A. H. Mueller, *Light-cone behavior of perturbation theory*, Phys. Rev. D **6** (1972) no. 12, 3543.

[63] Y. Frishman, *Light cone and short distances*, Phys. Rept. **13** (1974) no. 1, 1.

[64] T. Kinoshita, *Mass singularities of Feynman amplitudes*, J. Math. Phys. **3** (1962) no. 4, 650.

[65] T. Lee and M. Nauenberg, *Degenerate systems and mass singularities*, Phys. Rev. **133** (1964) no. 6B, B1549.

[66] S. Höche, *Introduction to parton-shower event generators*, in *Journeys Through the Precision Frontier: Amplitudes for Colliders: TASI 2014 Proceedings of the 2014 Theoretical Advanced Study Institute in Elementary Particle Physics*, p. 235. World Scientific, 2016.

[67] J. C. Collins, D. E. Soper, and G. Sterman, *Factorization of hard processes in QCD*, Adv. Ser. Direct. High Energy Phys. **5** (1989) 1.

[68] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nucl. Phys. B **126** (1977) 298.

[69] V. Gribov, V. Naumovich, and L. Lipatov, *Deep Inelastic ep-Scattering in Perturbation Theory*, Sov. J. Nucl. Phys. **15** (1972) 438.

[70] Y. L. Dokshitzer, *Calculation of the structure functions for deep inelastic scattering and e^+e^- annihilation by perturbation theory in quantum chromodynamics*, Sov. Phys. JETP **46** (1977) 641.

[71] S. Dulat et al., *New parton distribution functions from a global analysis of quantum chromodynamics*, Phys. Rev. D **93** (2016) 033006.

[72] S. Frixione and B.R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP 06 (2002) 029.

[73] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP 11 (2004) 040.

[74] S. Chauhan, *Search for Quark Compositeness at $\sqrt{s} = 14$ TeV at the Large Hadron Collider*. PhD thesis, Delhi University, 2010.

[75] B. Andersson et al., *Parton fragmentation and string dynamics*, Phys. Rep. **97** (1983) 31.

[76] T. Sjöstrand, *Jet fragmentation of nearby partons*, Nucl. Phys. B **248** (1984) 469.

[77] B. R. Webber, *A QCD model for jet fragmentation including soft gluon interference*, Nucl. Phys. B **238** (1984) 492.

[78] G. Marchesini and B. R. Webber, *Monte Carlo simulation of general hard processes with coherent QCD radiation*, Nucl. Phys. B **310** (1988) 461.

[79] E. Re, *Single-top Wt-channel production matched with parton showers using the POWHEG method*, Eur. Phys. J. C **71** (2011) 1547.

[80] 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) 79.

[81] O. Mattelaer, *MadGraph5_AMC@ NLO and top pair production at a Linear Collider*, Il nuovo cimento C **37** (2014) 61.

[82] P. Artoisenet et al., *Automatic spin-entangled decays of heavy resonances in Monte Carlo simulations*, JHEP **2013** (2013) 15.

[83] S. Frixione et al., *Angular correlations of lepton pairs from vector boson and top quark decays in Monte Carlo simulations*, JHEP **2007** (2007) 081.

[84] M. Backovic et al., *MadDM v. 1.0: Computation of dark matter relic abundance using MadGraph 5*, Phys. Dark Universe (2014) 18.

[85] M. Bähr et al., *Herwig++ physics and manual*, Eur. Phys. J. C **58** (2008) 639.

[86] J. Bellm et al., *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. J. C **76** (2016) 196.

[87] T. Sjöstrand et al., *A brief introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008) 852.

[88] T. Sjöstrand et al., *An introduction to PYTHIA 8.2*, Com. Phys. Commum. **191**

(2015) 159.

[89] T. Gleisberg et al., *Event generation with SHERPA 1.1*, JHEP 02 (2009) 007.

[90] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP 03 (2008) 038.

[91] F. Cascioli et al., *Scattering amplitudes with open loops*, Phys. Rev. Lett. **108** (2012) 111601.

[92] F. Buccioni et al., *OpenLoops 2*, Eur. Phys. J. C **79** (2019) no. 10, 1.

[93] A. Denner, S. Dittmaier, and L. Hofer, *Collier: a fortran-based complex one-loop library in extended regularizations*, Comput. Phys. Commun. **212** (2017) 220.

[94] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section in the lepton+jets channel at $\sqrt{s} = 13$ TeV with the ATLAS experiment*, Phys. Lett. B **810** (2020) 135797.

[95] ATLAS Collaboration, *Search for the Standard Model Higgs boson decaying into $b\bar{b}$ produced in association with top quarks decaying hadronically in pp collisions at $\sqrt{\mathbf{s}} = 8$ TeV with the ATLAS detector*, JHEP 05 (2016) 160.

[96] CMS Collaboration, *Search for the associated production of the Higgs boson with a top-quark pair*, JHEP 09 (2014) 87.

[97] G. Bevilacqua et al., *Assault on the NLO wishlist: $pp \to t\bar{t}b\bar{b}$*, JHEP **2009** (2009) 109.

[98] A. Bredenstein et al., *Next-To-Leading Order QCD Corrections to $pp \to t\bar{t}b\bar{b} + X$ at the LHC*, Phys. Rev. Lett. **103** (2009) 012002.

[99] A. Bredenstein et al., *NLO QCD corrections to top anti-top bottom anti-bottom production at the LHC: 2. Full hadronic results*, JHEP **3** (2010).

[100] M. V. Garzelli, A. Kardos, and Z. Trocsanyi, *Hadroproduction of $t\bar{t}b\bar{b}$ final states at LHC: predictions at NLO accuracy matched with Parton Shower*, JHEP **2015** (2015) no. 3, 83.

[101] M. V. Garzelli, A. Kardos, and Z. Trocsanyi, *$t\bar{t}b\bar{b}$ hadroproduction at NLO accuracy matched with parton shower*, PoS EPS-HEP2013 **253** (2013).

[102] F. Maltoni, G. Ridolfi, and M. Ubiali, *b-initiated processes at the LHC: a reappraisal*, JHEP **2012** (2012) 22.

[103] R. Frederix, E. Re, and P. Torrielli, *Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@ NLO*, JHEP **2012** (2012) no. 9, 130.

[104] S. Frixione, G. Ridolfi, and P. Nason, *A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, JHEP **2007** (2007) 126.

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

[106] The NNPDF Collaboration, *Parton distributions for the LHC Run II*, JHEP 04 (2015) 40.

[107] ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling for Top 2016*, ATL-PHYS-PUB-2016-020, (2016).

[108] A. Buckley, *ATLAS Pythia 8 tunes to 7 TeV data*, ATL-PHYS-PUB-2014-021, (2014).

[109] R. D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244.

[110] D. J. Lange, *The EvtGen particle decay simulation package*, Nucl. Instrum. Meth. A **462** (2001) 152.

[111] J. Bellm et al., *Herwig 7.1 Release Note*, arXiv:1705.06919 (2017).

[112] L. A. Harland-Lang et al., *Parton distributions in the LHC era: MMHT 2014 PDFs*, Eur. Phys. J. C **75** (2015) 1.

[113] ATLAS Collaboration, *Studies on top-quark Monte Carlo modelling with Sherpa and MG5_aMC@ NLO*, ATL-PHYS-PUB-2017-007, (2017).

[114] S. Höche et al., *A critical appraisal of NLO+PS matching methods*, JHEP **2012** (2012) no. 9, 49.

[115] S. Höche et al., *QCD matrix elements+ parton showers. The NLO case*, JHEP **2013** (2013) no. 4, 1.

[116] S. Catani et al., *QCD matrix elements+ parton showers*, JHEP **2001** (2002) no. 11, 063.

[117] S. Höche et al., *QCD matrix elements and truncated showers*, JHEP **2009** (2009) no. 05, 053.

[118] S. Alioli et al., *NLO single-top production matched with shower in POWHEG: s-and t-channel contributions*, JHEP **2009** (2009) 111.

[119] M. Cacciari et al., *Top-pair production at hadron colliders with next-to-next-to-leading logarithmic soft-gluon resummation*, Phys. Lett. B **710** (2012) 612.

[120] M. Czakon and A. Mitov, *NNLO corrections to top-pair production at hadron colliders: the all-fermionic scattering channels*, JHEP **2012** (2012) 54.

[121] M. Czakon and A. Mitov, *NNLO corrections to top pair production at hadron colliders: the quark-gluon reaction*, JHEP **2013** (2013) 1.

[122] M. Czakon, P. Fiedler, and A. Mitov, *Total top-quark pair-production cross section at hadron colliders through* $\mathcal{O}\left(\alpha_S^4\right)$, Phys. Rev. Lett. **110** (2013) no. 25, 252004.

[123] G. P. Salam, *Towards jetography*, Eur. Phys. J. C **67** (2010) 637.

[124] M. Cacciari, G. P. Salam, and G. Soyez, *The anti-k_t jet clustering algorithm*, JHEP 04 (2008) 063.

[125] M. Botje et al., *The PDF4LHC Working Group Interim Recommendations*, arXiv:1101.0538 (2011).

[126] J. Butterworth et al., *PDF4LHC recommendations for LHC run II*, J. Phys. G **43** (2016) no. 2, 023001.

[127] S. Höche, J. Krause, and F. Siegert, *Multijet merging in a variable flavor number scheme*, Phys. Rev. D **100** (2019) no. 1, 014011.

Erklärung nach §17(9) der Prüfungsordnung für den Bachelor-Studiengang Physik und den Master-Studiengang Physik an der Universität Göttingen:
Hiermit erkläre ich, dass ich diese Abschlussarbeit selbständig verfasst habe, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe und alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen wurden, als solche kenntlich gemacht habe.
Darüberhinaus erkläre ich, dass diese Abschlussarbeit nicht, auch nicht auszugsweise, im Rahmen einer nichtbestandenen Prüfung an dieser oder einer anderen Hochschule eingereicht wurde.

Göttingen, den 18. August 2021

(Stephen Eggebrecht)