GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Fakultät für
Physik

Bachelor's Thesis

Klassifikation von in Monte-Carlo-Simulationen generierten $t\bar{t} + Jets$-Ereignissen

Classification of the $t\bar{t} + jets$ events in Monte Carlo simulations

prepared by

Titus Stanislaus Czajka

from Berlin

at the II. Physikalischen Institut

Thesis number: II.Physik-UniGö-BSc-2021/01

Thesis period: 5th April 2021 until 12th July 2021

First referee: Prof. Dr. Arnulf Quadt

Second referee: Prof. Dr. Ariane Frey

Abstract

Der *heavy flavour classification* Algorithmus (HFC) sortiert $t\bar{t}+Jets$-Ereignisse, generiert durch eine inklusive $t\bar{t}$ Monte-Carlo Simulation, in verschiedene Kategorien. Diese Klassifikation basiert auf der Identifikation von Bottom- und Charm-Hadronen auf *truth*-Level. Es wird festgestellt, dass sie zu großen Teilen im Einklang mit dem `jet_truthflav` (JTF) Algorithmus auf *reconstruction*-Level ist. Die Transversalimpulse der Jets und führenden *b*-Hadronen auf *truth*-Level werden als einflussreiche Parameter des HFC identifiziert. Es wird für beide Parameter festgestellt, dass eine Erhöhung dieser zu mehr Ereignissen in der $t\bar{t}+light$ Kategorie und weniger in allen anderen Kategorien führt. Mit steigendem Parameter nimmt auch der Anteil der Ereignisse zu, die von JTF und HFC als übereinstimmend gekennzeichnet werden. Gleichzeitig wächst jedoch der Anteil derjenigen Kategorien, in denen der JTF Algorithmus mehr *b*-Jets findet als der HFC Algorithmus. Ein kleines lokales Maximum wird bei einem Wert von 5 GeV als Parameter für das führende *b*-Hadron in einem Jet gefunden, übereinstimmend mit dem selben Auswahlkriterium des JTF Algorithmus. Im Allgemeinen wird aber festgestellt, dass eine Veränderung des Jet p_T Parameters einen größeren Einfluss besitzt als das p_T des führenden b-Hadrons. Es wird gezeigt, dass alle Generatoren bis auf kleine Schwankungen in ihren Vorhersagen übereinstimmen.

Stichwörter: Physik, Bachelorarbeit, Monte-Carlo Ereignis Generatoren, Top Quark, *heavy flavour*, *truth flavour* Klassifikation, Ereignisklassifikation, $t\bar{t}+H$ Untergrund Studien

Abstract

The algorithm for heavy flavour classification (HFC) sorts $t\bar{t}+jets$ events of inclusive $t\bar{t}$ Monte Carlo simulations into different categories based on the presence of heavy flavour (charm, bottom) hadrons at truth level. It is found to be in good agreement with a reconstruction level truth flavour tagger, the `jet_truthflav` (JTF) algorithm. The transverse momenta of truth jets and of the leading *b*-hadron in a jet are found to be influential on the HFC. It is observed that the events classified as $t\bar{t}+light$ increase with higher HFC cut parameters, while event numbers decrease in all other HFC categories. The number of consistent classifications between HFC and JTF is shown to rise with increasing cut value as well as the number of classifications where the JTF tags more *b*-jets than expected from the HFC. A small local maximum is observed for cuts on leading *b*-hadron p_T at 5 GeV, corresponding to the cut parameter configuration of the JTF. The dependence of the HFC is stronger for jet p_T cuts than for cuts on the leading *b*-hadron p_T. A good agreement between generators is observed throughout the analysis.

Keywords: physics, bachelor thesis, Monte Carlo event generators, top quark, heavy flavour, truth flavour labelling, event classification, $t\bar{t}+$ H background studies

Contents

1 Introduction **1**

2 The Standard Model **3**

2.1 General overview of the Standard Model 3

2.2 Weak hadronic interactions 5

2.3 The Higgs mechanism 6

2.4 Top quark physics 8

2.4.1 The $t\bar{t} + H$ process 9

2.5 Shortcomings of the Standard Model 10

3 The ATLAS detector at the LHC **11**

3.1 Setup of the ATLAS detector 11

3.1.1 Overview of the main detector components 13

3.2 The ATLAS trigger system 15

4 Simulation of particle collisions **17**

4.1 Monte Carlo generators 17

4.2 Detector emulation 19

4.3 Jets and their identification 19

4.4 Analysis setup 20

5 Heavy Flavour Classification (HFC) **23**

5.1 Working principle 23

5.1.1 The `jet_truthflav` algorithm (JTF) 24

5.2 Further event selection 25

5.3 Comparison of HFC and JTF 26

6 Modifications to the HFC **31**

6.1 Parameter analysis 31

6.1.1 Azimuthal angle 32

6.1.2 Pseudorapidity 33
6.1.3 Jet transverse momentum 33
6.1.4 Leading b-hadron transverse momentum 36
6.2 New parameter configurations 36
6.2.1 b-jet transverse momentum 37
6.2.2 Leading b-hadron transverse momentum 38
6.3 Influence of cut variations on the comparative performance of HFC and JTF 40

7 Conclusions **45**

A Tables comparing the HFC and JTF for different generators **49**

B Plots showing the parameter dependence of of all HFC categories **55**

C Plots showing the parameter dependence of the comparative performance of and JTF **59**

D Tables comparing HFC and JTF for different parameters of the HFC **63**

Nomenclature

Abbreviation	Explanation
CKM matrix	Cabbibo-Kobayashi-Maskawa matrix
DL	dilepton
EM	electromagnetic
FCNC	flavour changing neutral current
HF	heavy flavour (bottom and charm)
HFC	heavy flavour classification
HFSC	heavy flavour simple classification
HLT	high level trigger
JTF	`jet_truthflav` algorithm
L1	level 1 Trigger
LO	leading order
NLO	next to leading order
PDF	parton distribution function
QCD	quantum chromodynamics
QED	quantum electrodynamics
QFD	quantum flavour dynamics
SL	single lepton
SM	Standard Model
SUSY	supersymmetry

1 Introduction

About six hundred years ago, an Italian monk discovered a long forgotten ancient manuscript in a German monastic library. This manuscript was the only surviving copy of a didactic poem of over 7,400 lines giving a full description of Epicurus' natural philosophy. The poem, called *De rerum natura* was written by a Roman called Lucretius and its first book contains a comprehensive description of Epicurus' theory of the indivisible constituent parts of matter, the atoms.

The ideas put forth in the book were revolutionary even at the time of its rediscovery and very much at odds with the christian doctrine of the church. A prominent supporter of the ideas put forth in the book, Giordano Bruno, was one of the last people burnt at stake for heresy by the inquisition in 1600. It has been argued that the book provided the ideas for a large step towards the modern age [1]. And indeed, at about the same time that Bruno was burnt, another Italian scholar, Galileo Galilei, was to develop a predecessor of the scientific method still used and valued to this day.

Modern day particle accelerators unite these two century old concepts and are thus able to take high precision measurements at scales that were open only to speculation and logical reasoning in the days of Epicurus and Galilei. What is most remarkable is, however, that these machines were able to provide strong experimental evidence for the old atomistic ideas using the scientific method.

The simple idea of an atom, however, has been developed further into a complex mathematical model, the Standard Model of particle physics. This comes at a cost, namely the difficult computations necessary to arrive at an accurate prediction and a certain level of ambiguity in the results of such a prediction.

It is this interception of theory and experiment where the present thesis tries to make a contribution. In general terms, it is concerned with improving the matching between simulation output and detector response. The particular physics process for which this work tries to provide an improvement is the production of a Higgs boson in association with a top quark pair.

The $t\bar{t}$ associated Higgs production channel constitutes only about 1% of all Higgs boson productions, which is by itself a rare process. The subsequent decay of the Higgs boson into two bottom quarks happens in about 58% of the cases [2]. This process is called $t\bar{t} + H(H \to b\bar{b})$ in short and was first observed by ATLAS in 2018 [3]. It is well suited to measure the strongest Yukawa coupling of Standard Model between Higgs boson and top quark, as well as the second strongest Yukawa coupling between bottom quark and Higgs. A precise measurement of these couplings could potentially be sensitive to physics beyond the Standard Model (BSM), and further advance our knowledge at the foundation of physics.

To have a chance of measuring this process, the background must be filtered out most efficiently. The decay of the top quark into a bottom quark and a W-boson provides a clear signature in the detector and thus a good first step due to the possibility of a leptonic W-boson decay. Even after this selection, however, the process does not have a unique signal. The main remaining background is of the type $t\bar{t} + jets$ (e.g. $t\bar{t} + b\bar{b}$) [4]. Hence, an in-depth study of all processes involving a similar final state is necessary. To achieve this, it is useful to classify each event according to the number of heavy (bottom or charm) flavour jets that are present. To improve this classification of the simulated events is the aim of the work presented here.

This thesis begins with a description of the theoretical framework in Ch. 2 and the ATLAS particle detector at the LHC in Ch. 3, the experiment for which the analysis is performed. A short description of Monte Carlo event generation and the simulation framework used in this work is given in Ch. 4 before the heavy flavour classification (HFC) is explained and analysed in Ch. 5. The thesis concludes with an analysis of modifications to the HFC in Ch. 6 and a conclusion in Ch. 7.

2 The Standard Model

The Standard Model (SM) offers the most successful description of elementary particle physics to date. It describes the strong, weak, and electromagnetic interaction, but not the fourth fundamental force of Nature, gravity. Since its development in the 1960s and 70s, the SM has withstood numerous experimental tests, most notably the discovery of neutral weak currents in the Gargamelle bubble chamber at CERN in 1973 [5], the discovery of the top quark by CDF, and DØ at TEVATRON in 1995 [6, 7] and the discovery of the Higgs boson by ATLAS, and CMS at CERN in 2012 [8, 9].

2.1 General overview of the Standard Model

The Standard Model is a renormalisable [10–12], relativistic quantum field theory with a local $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge symmetry [13–17]. The symmetry group $SU(3)_C$ corresponds to quantum chromodynamics (QCD), the framework behind the strong interaction. $SU(2)_L \times U(1)_Y$ is the gauge symmetry of the electroweak interaction developed by S. Glashow, A. Salam and S. Weinberg [18–21], a unified theory of quantum flavourdynamics (QFD), describing the weak force, and quantum electrodynamics (QED).

The SM particles, shown in Fig. 2.1, can be divided into two categories: Fermions and Bosons. Fermions in the SM have a spin of 1/2 and are the constituent particles of matter. Bosons have an integer spin and come in two types: Vector and scalar bosons, a distinction depending on their associated coupling mechanism.

Fermions The SM encompasses twelve fermions: six quarks and six leptons. For each fermion except neutrinos there exists a corresponding antifermion with the same mass but opposite charge. Quarks are split into two categories, up-type quarks, where the third component of the weak isospin is $I_z = +1/2$, and down-type quarks, where $I_z = -1/2$. A pair of up-type and down-type quarks can be grouped to form either a left handed isospin doublet ($I = 1$) or a right handed isospin singlet ($I = 0$). Similarly, there are two

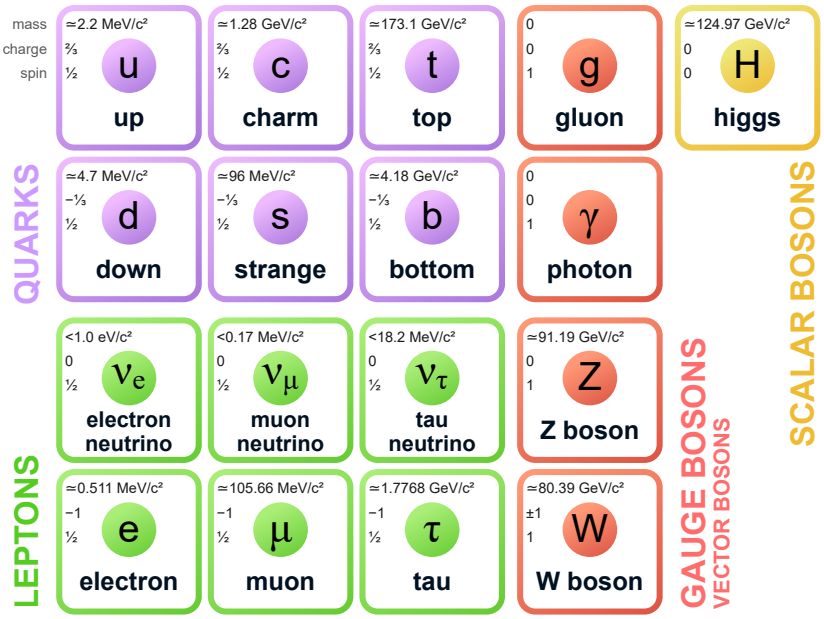


Figure 2.1: The Standard Model of particle physics. Fermions are depicted in green and purple, bosons in red and yellow.
Image credit: MissMJ under CC BY 3.0

different types of leptons, electrically charged leptons (electron, muon and tau-lepton) and electrically neutral leptons (neutrinos), where $I_z = -1/2$ and $I_z = +1/2$, respectively. The key difference between quarks and leptons is their coupling to the strong force: Quarks carry a colour charge and participate in the strong interaction. Leptons do not carry any colour charge and therefore do not interact via the strong force. All electrically charged fermions couple to the electromagnetic interaction. This excludes neutrinos, who couple only to the weak interaction, as do all other fermions.

Gauge bosons Gauge bosons are the force carriers within the framework of the SM. The photon (γ) mediates electromagnetic interactions, the gluon (g) strong interactions and the three massive gauge bosons (Z, $W^{\pm}$) mediate weak interactions.

The coupling strengths of the weak and electromagnetic interactions are much smaller than one and almost scale invariant, which means that they increase slowly with increasing energy. This enables precise perturbative calculations for most weak and electromagnetic processes. The strong force coupling constant α_S, on the other hand, decreases significantly with increasing energy, as shown in Fig. 2.2, a phenomenon called asymptotic freedom [23]. At high energy scales the coupling is well below unity, which allows perturbative QCD calculations. At the other end of the energy scale, the strong force coupling

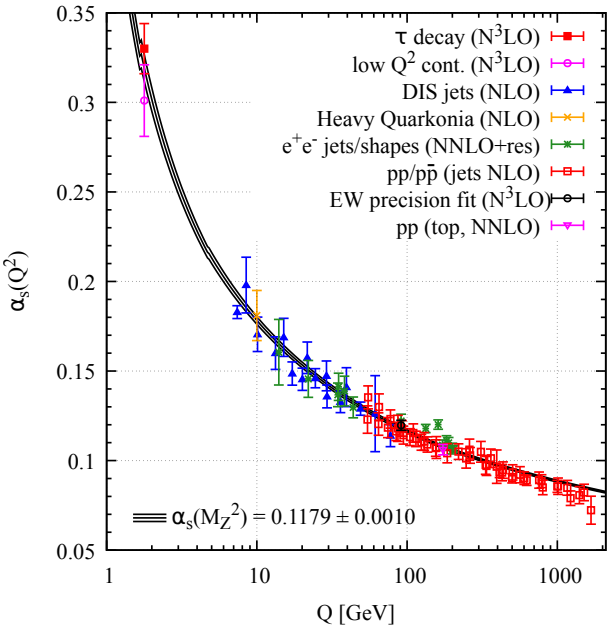


Figure 2.2: Different measurements of the running coupling of α_S as a function of the energy scale [22]. The highest order of the QCD perturbative expansion used to extract α_S is given in brackets. The coupling is very strong at low energy scales and decreases rapidly with increasing energy.

is much larger, rendering all perturbative methods useless (cf. Ch. 4).

The strong coupling at low energy scales is the reason why no quark can ever be directly observed on its own. If two quarks are pulled apart from each other, large amounts of energy are necessary and hadronisation eventually becomes energetically preferable to further separation [24]. This is called quark confinement and is a very important concept in high energy physics, since an accumulation of particles travelling in a similar direction, called a jet, can thus be identified as the remnant of a quark or gluon.

2.2 Weak hadronic interactions

Many observations indicate a conservation of flavour in SM interactions. However, this is not a priori obvious as it is not a conservation law originating from a gauge symmetry [25]. Charged weak interactions show that this is not always the case. Theoretically this can be explained by the difference between quark mass eigenstates and flavour eigenstates, as described by the Cabbibo-Kobayashi-Maskawa (CKM) matrix [26, 27]. Consequently, it

can be deduced that any weak charged current vertex involving quarks introduces a factor

$$-i\frac{g_W}{\sqrt{2}}\left(\bar{u},\bar{c},\bar{t}\right)\gamma^\mu\frac{1}{2}\left(1-\gamma^5\right)\begin{pmatrix}V_{ud} & V_{us} & V_{ub}\\ V_{cd} & V_{cs} & V_{cb}\\ V_{td} & V_{ts} & V_{tb}\end{pmatrix}\begin{pmatrix}d\\ s\\ b\end{pmatrix} \tag{2.2.1}$$

to the matrix element, where g_W is the coupling constant of the weak interaction and γ^i are the Dirac γ-matrices. d, s, b and $\bar{u}, \bar{c}, \bar{t}$ represent the corresponding quark spinors and adjoint spinors, respectively. The matrix elements V_{ij} define the relative strength of the interaction between quark i and quark j.

The entries V_{ij} can be parameterised by four independent, real valued numbers (e.g. Wolfenstein parametrisation [28]) that are free parameters of the SM. Using experimental neutrino scattering data, the entries of the CKM matrix were found to be [22]

$$\begin{pmatrix}|V_{ud}| & |V_{us}| & |V_{ub}|\\ |V_{cd}| & |V_{cs}| & |V_{cb}|\\ |V_{td}| & |V_{ts}| & |V_{tb}|\end{pmatrix}\approx\begin{pmatrix}0.974 & 0.227 & 0.004\\ 0.226 & 0.973 & 0.041\\ 0.009 & 0.040 & 0.999\end{pmatrix}. \tag{2.2.2}$$

This matrix thus reflects the experimental observation that flavour changing charged currents do happen. but much less often than flavour preserving currents. It also reflects the circumstance that flavour changes preserving the quark generation are favoured over generation changes. This effect is amplified in charged weak interactions involving heavy quarks.

Flavour changing neutral currents (FCNC) are predicted by the SM as well, but they are strongly suppressed since a single FCNC interaction vertex is not predicted (the SM requires multiple vertices for a FCNC). Leading order FCNCs are, however, predicted by many theories that go beyond the SM, which is why these processes are under close investigation in modern high energy physics experiments (e.g. at ATLAS [29]).

2.3 The Higgs mechanism

A locally gauge invariant theory of the weak interaction predicts massless gauge bosons. The discovery of the W-boson in 1983 by the UA1 [30] and UA2 [31] experiments at the SPS collider, however, showed that it was not massless. A model developed already in 1964 was able to solve the contradiction of theory and experiment. This approach by R. Brout, F. Englert, P. Higgs, G. Guralnik, R. Hagen, and T. Kibble introduces a

spontaneous symmetry breaking at some energy threshold which gives a nonzero rest mass to the heavy gauge bosons [13, 14, 16]. Gauge invariance demands the introduction of a new field and hence a new boson, the Higgs boson, which couples to all massive particles. The corresponding coupling relies on a spontaneous symmetry breaking of the complex scalar Higgs field ϕ with a potential of the form

$$V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2 \ . \tag{2.3.1}$$

From this and the requirement of massless photons, one can calculate the mass of the W-boson

$$m_W = \frac{1}{2} g_W v \tag{2.3.2}$$

and of the Z-boson

$$m_Z = \frac{1}{2}\frac{g_W}{\cos\theta_W} v \ , \tag{2.3.3}$$

where v is the vacuum expectation value of the Higgs field, g_W the weak interaction coupling strength, and θ_W the weak mixing angle. Together with the mass m_H of the Higgs boson, they are the free parameters needed to describe the electroweak interaction and can only be determined experimentally.

Yukawa Coupling The fermion masses are also generated by the Higgs mechanism through spontaneous symmetry breaking, similarly to the generation of the gauge boson masses [21]. The coupling strength g_f of the Higgs to a fermion f, another free parameter of the SM, is called Yukawa coupling and is proportional to the fermion mass, as described by the relation

$$g_f = \sqrt{2}\frac{m_f}{v} \ . \tag{2.3.4}$$

The value of g_f for all fermions (except neutrinos) is of order $\mathcal{O}(1)$, with the strongest Yukawa coupling being the one to the top quark ($g_t \approx 0.995$). This particular Yukawa coupling is believed to be sensible to physics beyond the Standard Model [32] and thus of special interest to the modern day particle physicist.

Production and Decay At the LHC, the dominant Higgs boson production channel is through gluon-gluon fusion, similar to the process shown in Fig. 2.3(a). Less frequent production channels are, among even rarer processes, vector boson fusion or Higgsstrahlung, where a vector boson radiates off a Higgs boson.

The Higgs boson decays in more than half of all cases into a $b\bar{b}$-pair, which is detectable

Table 2.1: Branching ratios of the most common Higgs boson decay channels, assuming a Higgs mass of $m_H = 125\,\text{GeV}$ [2].

Decay Channel	Branching Ratio
$H \to b\bar{b}$	58%
$H \to W^+W^-$	21%
$H \to \tau^+\tau^-$	6%
$H \to ZZ$	3%
$H \to c\bar{c}$	3%
$H \to \gamma\gamma$	0.2%

via good background studies and a decent b-tagging strategy [3]. The same methods could in principle be used to observe the rarer decay into a c-quark pair, but the low charm-tagging efficiency complicates this at the moment. A decay into a W-boson pair is the second most likely option and, together with the $H \to ZZ$ decay, one of the most easily detectable decay channels due to the possibility of an all-lepton decay and an asymmetric angular distribution. This advantage over the more probable, but less clear signature from hadronic decays was exploited in combination with the $H \to \gamma\gamma$ channel for the discovery of the Higgs by ATLAS [8] and CMS [9] in 2012. Tab. 2.1 gives an overview over the most common decay channels.

2.4 Top quark physics

Within the framework of the SM, top quark physics plays a special role. With a mass of about 173 GeV, it is the heaviest of all elementary particles in the SM and therefore it has an extremely short lifetime due to the large phase space. The lifetime of $\tau \approx 0.5 \times 10^{-24}\,\text{s}$ is in fact so small that the top quark decays before any bound states can form. This means that top quarks are not observed as jets, but through their decay via a weak interaction into a bottom quark and a W-boson in almost all of the cases, as shown in Fig. 2.3(a). In theory, the weak decay also allows for the production of strange and down quarks. This is, however, highly unlikely, since the entry corresponding to a decay into a bottom quark in the CKM-matrix is almost unity (see Eq. 2.2.2).

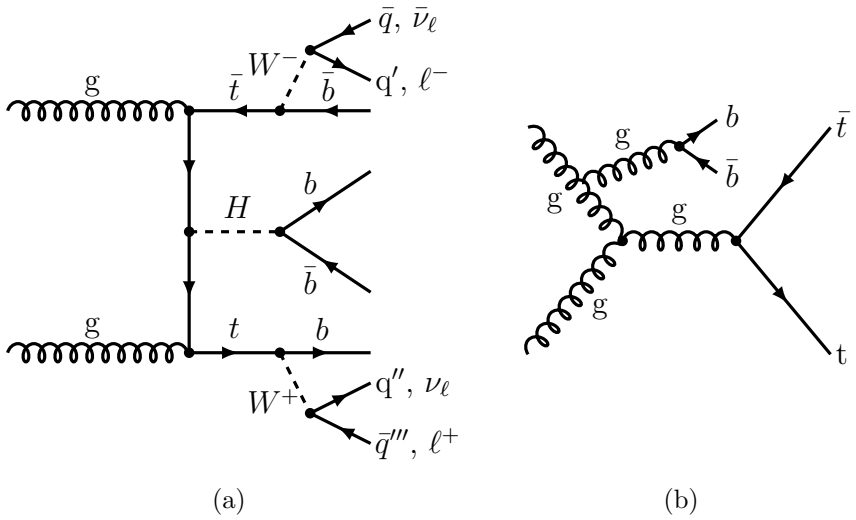


Figure 2.3: Figure (a) shows the leading order Feynman diagram for the Higgs production in association with a $t\bar{t}$-pair through gluon-gluon fusion. The three possible final states are shown, depending on the decay of the W-bosons: all hadron, single lepton and dilepton channels. Fig. (b) shows an example for a leading order background process to the $t\bar{t} + H$ ($H \to b\bar{b}$) event in Fig. (a).

2.4.1 The $t\bar{t} + H$ process

As already mentioned, the top quark Yukawa coupling is the strongest in the SM and thus of particular interest. The easiest way to analyse this interaction is the top quark associated Higgs boson production, as depicted in Fig. 2.3(a) [33–38]. This is a very rare process, contributing about 1% to the total Higgs boson production cross section [2]. The two top quarks have, however, a distinctive signature and therefore allow for a good detection sensitivity to this process. About 58% of the Higgs bosons produced in this way are predicted to decay into two b-quarks (cf. Tab. 2.1), which also opens the possibility of measuring the b-quark Yukawa coupling, the second largest predicted by the SM.

To further simplify the detection of this process, it is most convenient to select those events where either one or both W-bosons decay leptonically (electron or muon), which happens in about a quarter of all cases (about 11% per lepton) [22]. The produced high energy leptons give a clear trigger signature and also reduce the main $t\bar{t}+jets$ background, the separation of which is the largest challenge to an analysis of the $t\bar{t} + H$ ($H \to b\bar{b}$) process [4]. A Feynman diagram for an exemplary leading order background process is shown in Fig. 2.3(b).

2.5 Shortcomings of the Standard Model

Despite its experimental success, the Standard Model is not a complete description of nature. There are several issues, most notably the omission of gravity and its theoretical framework, general relativity. This is, however, not the only shortcoming of the framework, there are several others. Some of them are:

- Experiments show that neutrinos have a nonzero rest mass, but the current Standard Model predicts them to be massless [22].
- The top quark mass is much higher than the mass of all other quarks. The Standard Model provides neither an explanation for this behavior, nor for many other relations between parameters of the model. The strength difference between the four fundamental forces is another example.
- Astronomers have shown that the energy in the universe consists of only about 4.9% ordinary matter. The remaining 95% are attributed to dark matter (26.6%) and dark energy (68.5%) [39]. The Standard Model provides no explanation for these forms of energy.

There are many theories trying to explain some of the shortcomings. One famous approach is the introduction of additional particles to create a symmetry between fermions and bosons, grouped together under the name supersymmetric (SUSY) theories. Another approach is the introduction of extra dimensions, especially helpful in the incorporation of gravity. Even though particle detectors are on the lookout, experimental evidence for any of these theories is yet to be found.

3 The ATLAS detector at the LHC

The Large Hadron Collider (LHC) is currently the largest and highest energy particle accelerator in the world. It was built between 1994–2008 at a construction cost of about 6.5 billion Swiss francs. The collider is situated in a 27 km underground tunnel at the French–Swiss border next to Geneva, originally built for the LEP collider [40]. Since 2018, the LHC is in a long shutdown period to upgrade the accelerator and detectors in preparation for a third period of data taking (Run 3), which is predicted to begin in 2022. At the LHC, proton bunches travel in opposite directions along a circular path and collide at four interaction points (IP) with a centre of mass energy of up to $\sqrt{s} = 13\,\text{TeV}$. To reach such a high energy, more than one accelerator is necessary: The protons are initially boosted in a linear accelerator and gain further energy in two synchrotrons before injection into the LHC. Each one of the four detectors (ATLAS, CMS, ALICE and LHCb) is built around one interaction point. The embedding of the LHC and its detectors into the accelerator complex at the site of the European Organisation for Nuclear Research (CERN) is shown in Fig. 3.1.

To meet the challenges posed by the high luminosity of up to $2 \times 10^{34}\,\text{cm}^{-2}\text{s}^{-1}$, about 50 individual collisions per bunch crossing and bunch separation times of down to 25 ns [41], a sophisticated detector design is as necessary as a good trigger system to select and accurately record any interesting collision data. The following sections describe this using the example of the ATLAS detector.

3.1 Setup of the ATLAS detector

The ATLAS (**A** **T**oroidal **L**HC **A**pparatu**S**) detector is the larger of the two general purpose particle detectors at the LHC, the other being CMS. It was designed to enable a reconstruction of the proton-proton collision that is as complete as possible. Additionally, the detector incorporates a few special features that allow an accurate measurement of some particular processes. The discovery of the Higgs boson by accurate measurement of

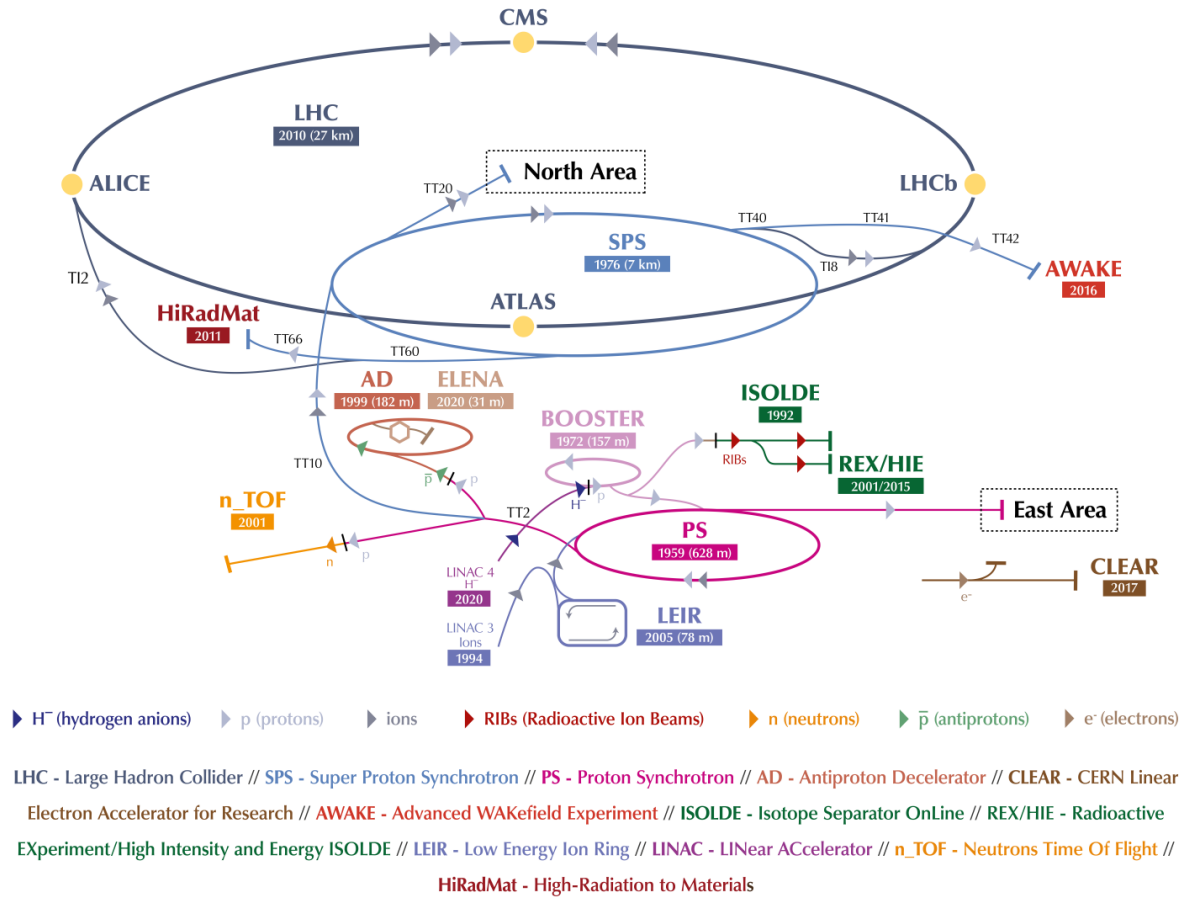


Figure 3.1: Embedding of ATLAS and the LHC into the CERN accelerator complex. The protons start at a hydrogen bottle and get accelerated in the linear accelerator LINAC 4 (previously LINAC 2) before injection into the PS first stage booster ring. From there the protons are accelerated further in the SPS booster ring before injection into the LHC. *©2018-2021 CERN*

lepton and photon energy is one example of these specific goals of the detector design [42]. It was famously achieved in 2012 through a combination of $H \to ZZ \to 4\ell$, $H \to WW$ and $H \to \gamma\gamma$ observations by both general purpose detectors at the LHC, ATLAS [8] and CMS [9].

As shown in Fig. 3.2 the detector is composed of several layers that are radially arranged, the tracking detector close to the interaction point, the electromagnetic and hadronic calorimeters in the middle and the muon detectors on the outside. The left and right end of the detector (end-caps) are constructed similarly, but the layers are now arranged along the z-axis, parallel to the beamline. Hence, the detector possesses a radial symmetry and a point symmetry in the z-direction with small deviations due to structural components [42]. To ensure an accurate measurement of particle momenta, four magnets

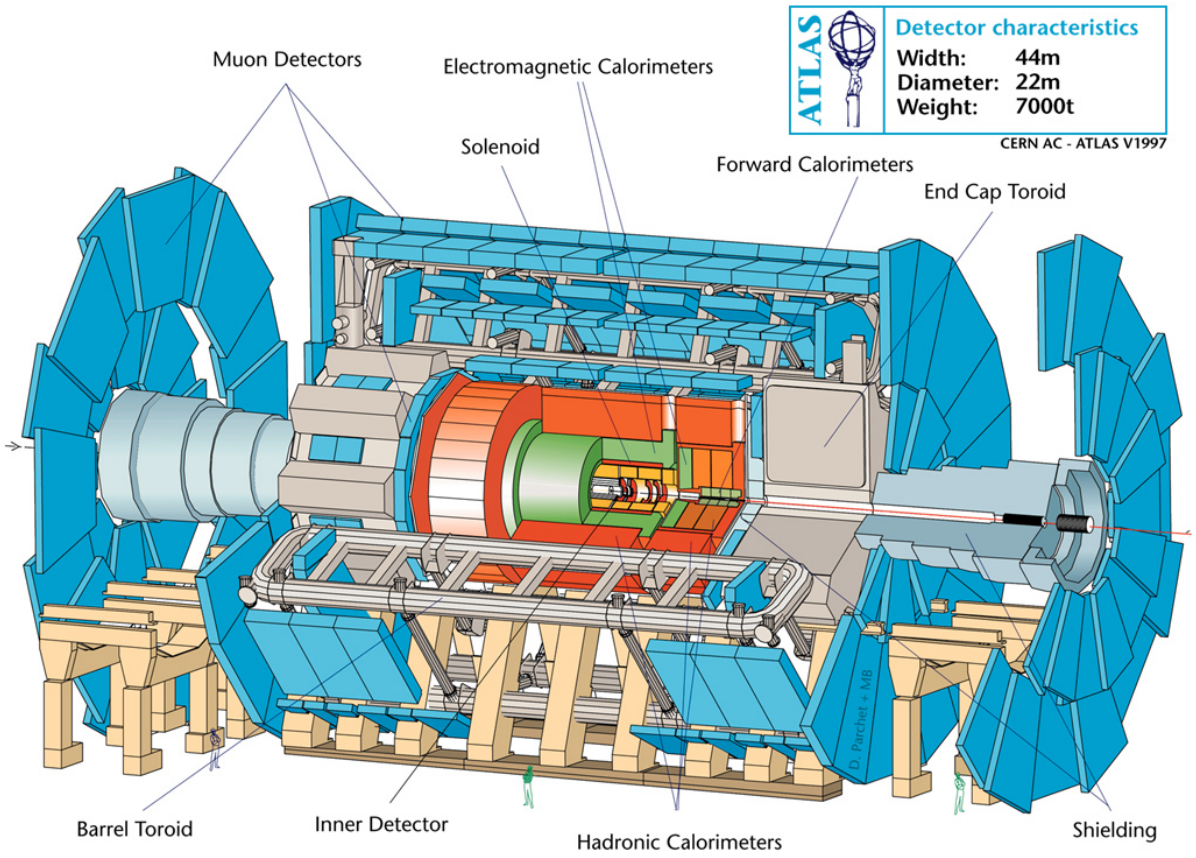


Figure 3.2: General setup of the ATLAS detector. *©1998-2021 CERN*

are embedded within the ATLAS detector: one axially aligned solenoid, which provides a field of about 2 T to the inner detector, two end-cap toroids (1 T in the muon detectors) and one barrel toroid (0.5 T), both comprising eight superconducting coils [42].

3.1.1 Overview of the main detector components

In the following paragraphs, pseudorapidity $\eta = -\ln\tan(\theta/2)$ will be used as a measure of the polar angle θ alongside the radius r and the azimuthal angle φ. With this definition, the angular distance as viewed from the interaction point can be written as[1]

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}\ . \tag{3.1.1}$$

This section is heavily based upon the description of the ATLAS detector in Ref. [42].

[1] η is not Lorenz invariant with respect to boosts along the z-axis, but differences in η are. The azimuthal angle φ is measured perpendicularly to the z-axis and thus invariant under Lorenz boosts along the z-axis by definition. Hence, the measure ΔR is independent of a particle's momentum in the z-axis. The concept is important in parton collisions, where the exact parton momentum is unknown.

Inner detector The main purpose of the inner detector, a cylinder of about two metres in diameter and six metres in length, is the measurement of particle tracks and thus particle momentum. To obtain maximum spatial accuracy, its components are semiconductor pixel modules close to the beam pipe (3–12 cm) and Silicon strip modules further away (30–51 cm). They achieve resolutions of at least $10 \times 115\,\mu\text{m}^2$ (r-ϕ and z) and $17\,\mu\text{m}^2$ (r-ϕ), respectively [43]. For Run 2, a new innermost pixel detector component, the insertable b-layer (IBL) with a resolution of $10 \times 60\,\mu\text{m}^2$, has been installed at $r \approx 3$ cm from the beampipe to enhance the track reconstruction performance [43]. The pixel (including IBL) and strip modules cover the range where $|\eta| < 2.5$. For $|\eta| < 2.0$ the transition radiation tracker (TRT), an array of drift tubes, covers a radial range of 55–108 cm and thus allows a more accurate track reconstruction, which improves the accuracy of momentum measurements. Additionally, it provides a good mechanism for electron identification.

Electromagnetic and hadronic calorimeter The ATLAS calorimeter system covers a range of $|\eta| < 4.9$ and is split into two parts, an inner electromagnetic (EM) and an outer hadronic calorimeter. The electromagnetic calorimeter is a liquid Argon/Lead sampling calorimeter with a very fine structure (up to $\Delta\eta = \Delta\phi = 0.025$) at a thickness of about 24 radiation lengths in the low-η regions. To achieve full ϕ-coverage and a uniform performance, the Lead-layers were folded to an accordion-like structure. The EM-calorimeter is designed to allow a particularly accurate measurement of electron and photon energy. The fractional energy resolution is about 1% for 100 GeV electrons and photons across the full detector range, except for the crack region between $\eta = 1.37$ and 1.52, due to the transition from barrel to end-cap calorimeters. There the energy resolution is worse by a factor of about ten. The hadronic calorimeter has a thickness of about ten interaction lengths and is made from an Iron/scintillator sampling calorimeter. It has a coarser granularity ($\Delta\eta = \Delta\phi \approx 0.1$) than the electromagnetic calorimeter, but it is still designed to provide accurate data on high energy jets. Both calorimeters together allow a good measurement of the total missing transverse energy E_T^{miss}, an important beacon on the path towards interesting physics.

Muon detector Because of their high mass, muons lose most of their energy through ionisation. Therefore they can penetrate matter much further than electrons with the same momentum, since the latter lose a more significant fraction due to Bremsstrahlung. To still achieve reasonable momentum resolution, the largest part of the ATLAS detector is covered by the muon system. It is designed to measure the momenta of muons by deflection in a magnetic field, similar to the inner detector. To get an optimal performance,

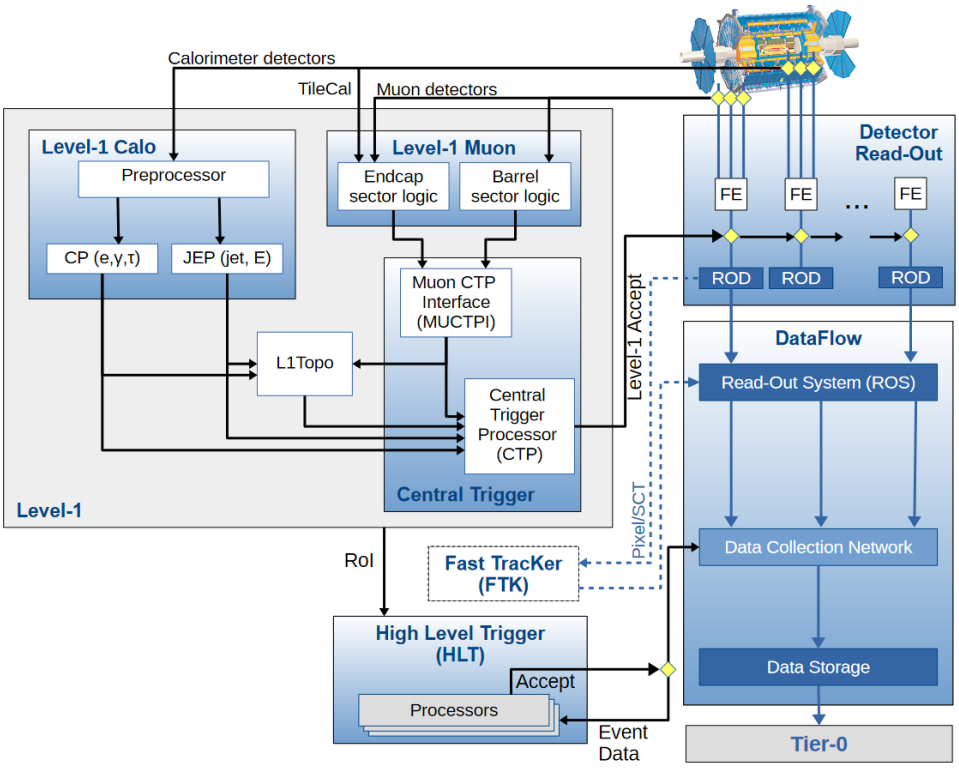


Figure 3.3: Trigger and detector read-out system of the ATLAS detector [44]. The hardware based Fast TracKer inner detector track reconstruction system was not used in the HLT during Run 2.

the muon system is made from different types of detectors. High precision ($\approx 40\,\mu$m) drift chambers and multiwire proportional chambers are installed to provide a good momentum measurement, and resistive plate chambers as well as thin gap chambers to obtain fast trigger signals.

3.2 The Atlas trigger system

With an interval of 25 ns between each bunch crossing, the rate at which events are generated is much higher than the rate at which the full detector data can be processed. Therefore it is necessary to discard physically uninteresting events as soon as possible, for example when two bunches miss each other and no collision occurs. Fig. 3.3 provides an overview over the ATLAS trigger system in Run 2. A detailed description can be found in Ref. [44].

The ATLAS trigger system comprises two stages: The hardware based first level trigger (L1) and a software based High-Level Trigger (HLT). The former reduces the bunch

crossing frequency of 40 MHz down to about 100 kHz with a latency of 2.5 μs. It receives coarsely grained data from the detector and decides upon acceptance of the event by identifying jets, calculating electron, muon and photon momenta. To have enough time for all necessary computations at its disposal, the L1 trigger relies on parallel processing and a pipeline structure whose clock is locked to the bunch crossing frequency. This system allows for a fast decision, however, it requires that the detector components store all their data until a decision has been made.

If the L1 trigger accepts the event, the detector signals are read out and stored temporarily. The HLT processes the data already obtained in the L1 trigger further by taking into account the complete data set of the event. It employs algorithms to reconstruct and analyse some structure within the data, for example track reconstruction or missing transverse momentum. Its output rate in Run 2 was on average 1.2 kHz, resulting in a data stream of about 1.2 GB/s to the Tier-0 permanent (*offline*) storage.

4 Simulation of particle collisions

A simulation of a collision event is necessary to compare the results of a proton-proton collision in a detector with the prediction made by some theory. Since the high complexity of such a process forbids an exact calculation, it is necessary to split it into many small. In this way an appropriate approximation can be used at each stage, as indicated by the colour-coding in Fig. 4.1, which illustrates the simulation of a $t\bar{t} + H$ event.

4.1 Monte Carlo generators

The start of any event generation is the calculation of the desired hard process, which in the case of this thesis is $t\bar{t} + H$ ($H \to b\bar{b}$). These interactions are highly energetic and can be accurately calculated with perturbative methods at leading or next to leading order (LO or NLO, respectively). A Monte Carlo algorithm serves as integrator in the calculation of the matrix elements, since it can efficiently handle integration over the large phase space of many-particle final states.

In a hadron collision, large momenta are transferred between different particles, which increases the likelihood of QCD Bremsstrahlung on top of the hard process. The radiated gluons can split themselves, due to their colour charge, thus producing a parton shower. Because an estimation of the number of particle splittings is not possible, a fixed order approximation does not yield good results. Instead, the evolution of the parton density functions are calculated using the DGLAP equations [46–48] that provide a relation between the parton densities at different energy scales. To avoid producing a particle twice by simulating the same parton at matrix level and in the parton shower (both red in Fig. 4.1), a good matching between those two stages is necessary. The two most common matching methods are POWHEG [49] and aMC@NLO [50] which are implemented in the corresponding matrix element generators. These algorithms play an important role, because programs specialised in matrix element calculations leave the parton shower and hadronisation calculations to different programs, such as PYTHIA 8 [51] and HERWIG 7

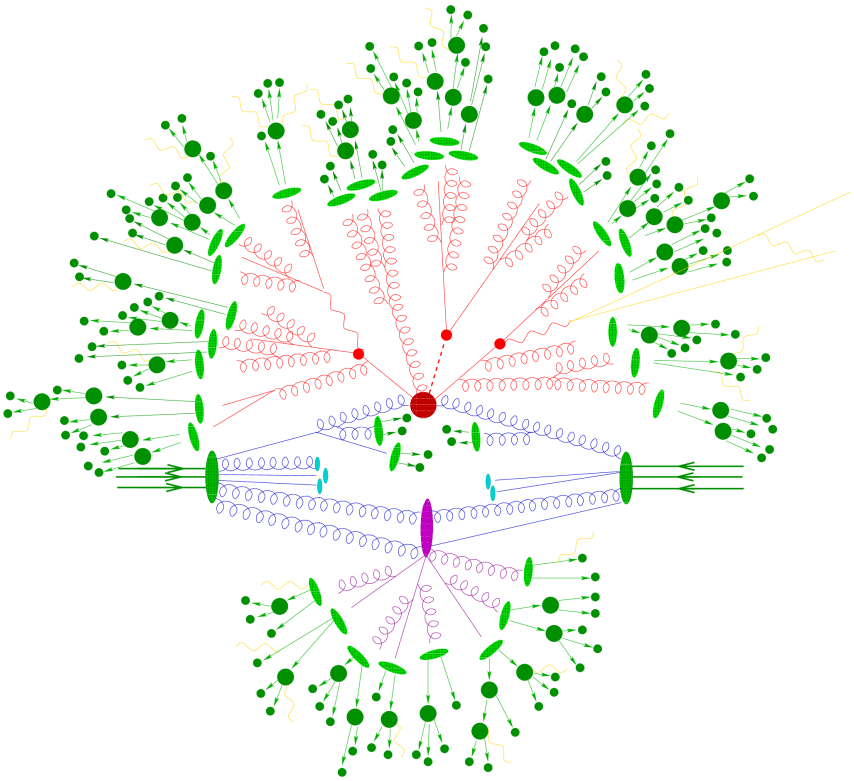


Figure 4.1: Simulation of a $t\bar{t}+H$ event [45]. The initial high energy parton collision is shown in blue, the parton shower in red, hadronisation and decay processes in green. Leptons and photons are depicted in yellow, underlying secondary events in purple. The $t\bar{t}+H$ production vertex is the big red circle at the centre of the figure. The two top-quarks decay into one W-boson and a b-quark. One of them decays hadronically, the other leptonically.

[52, 53] in the case of this analysis. The program Sherpa [54] is the exception to this rule, it performs all three stages of the event generation by itself.

Up to the parton shower, every process can be approximated by perturbative methods. For the hadronisation process, however, a different technique is necessary, due to the strength of the coupling constant α_S at low energy. This part of the simulation chain is shown in green in Fig. 4.1. Since a calculation from first principles is not yet possible, phenomenological methods that emulate the QCD processes during hadronisation based on experimental findings are deployed. Examples are the Lund string model, used in Pythia and the cluster model, used in Herwig and Sherpa. Finally, the hadron decay is simulated (dark green), to get an accurate representation of the stable final state particles.

Hadronic collisions require another approximation based on a phenomenological model: The partonic structure of any hadron needs to be simulated to get an estimate of underlying secondary events, happening due to the presence of other partons in the hadronic

collision (blue in Fig. 4.1). To achieve this, the kinematics of initial state particles is calculated from a parton distribution function (PDF) not only for the primary hard event, but also for all underlying activity, coloured purple in Fig. 4.1.

4.2 Detector emulation

The ATLAS detector is not an ideal machine, its calorimeters and tracking detectors are not able to tell what particles were produced, but only how much energy was deposited at some point in the detector. In order to compare experimentally observed and generated events, the response of the detector to the generated event must be computed [55].

In the case of ATLAS, this is done by a program called GEANT4 [56–58]. It emulates the response of the detector to the simulated event, based on the generator output and the experimentally observed behaviour of every component within the detector. This includes the limited spatial coverage (e.g. crack region) and the removal of all particles that cannot be observed, e.g. neutrinos or particles with a momentum below the detection threshold.

Truth and Reconstruction Level When improving analysis methods it is not enough to know the detector response. In order to study the efficiency of b-tagging algorithms, it is, for example, necessary to know exactly what type of particles are present in a jet. Hence, the event generator output (generated particles) is also combined into jets and other non-QCD particles and stored alongside the output of the reconstruction algorithm which runs on the digitised GEANT4 output. To distinguish both levels, the former is termed truth level, the latter reconstruction level. At truth level some cuts are usually introduced as well to assimilate this level to the response of the actual detector, but at a lower threshold than at reconstruction level.

4.3 Jets and their identification

Jets are formed when quark pairs with different momenta occur in particle collisions. Due to the running of α_S, hadronisation occurs and every quark appears in the detector as a cluster of particles, boosted in a similar direction. Top quarks are an exception to this rule, since they are so short lived that they decay before hadronisation can occur. If, however, the top quark is highly boosted, its decay products form a jet, called large-R jet which contains all particles from the decay, boosted in a similar direction.

The experimentalist's task is to collect all the particles belonging to a jet and then reconstruct the kinematics of the underlying particle. This is usually complicated by an overlap with neighbouring jets and the finite resolution of any detector. Jet algorithms perform this task of clustering particles in a reproducible way by comparing the distance between any two possible constituents of the jet based on a metric, for example the jet radius ΔR as defined in Eq. (3.1.1) [59]. Anti-k_t, the most frequently used jet reconstruction algorithm, is based on a sequential particle recombination that uses this measure multiplied by an energy dependent factor [60].

4.4 Analysis setup

The analysis in this thesis is based on eight inclusive $t\bar{t}$ samples at a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ for the Run 2 ATLAS integrated luminosity of $139\,\text{fb}^{-1}$. One combined single (SL) and dilepton (DL) sample is generated by a POWHEG + PYTHIA 8 setup, two each from POWHEG + HERWIG 7 and aMC@NLO + PYTHIA 8 (separate SL and DL) and three by SHERPA 2.2.10 (positive SL, negative SL and DL).

The POWHEG + PYTHIA 8 sample is the nominal sample for $t\bar{t}$ pair production at ATLAS and thus the reference sample in all generator comparisons made in this thesis. It will also be used in all cases where the clarity of an argument would suffer under the presence of data from all generators.

Each sample is processed in the same way. First, jets are reconstructed from stable truth particles (lifetime $\tau > 3 \times 10^{-11}\,\text{s}$) via anti-$k_t$ with a radius parameter $R = 0.4$. At truth level, jets with a transverse momentum below $p_T = 5\,\text{GeV}$ are discarded and at reconstruction level all jets with $p_T < 25\,\text{GeV}$ are as well. Additionally, all reconstruction level particles (including jets) are required to have a rapidity in the detectable region of $|\eta| < 2.5$, leptons must have $|\eta| < 2.47$ and $1.37 < |\eta| < 1.52$ due to the reduced sensitivity in the crack region. Events are chosen to be either in the single lepton[1] (SL) or dilepton channel, to select events with possible W-decays originating in a top quark. The criterion for the single lepton (SL) channel is the presence of exactly one lepton with $p_T > 27\,\text{GeV}$. In the dilepton (DL) channel, exactly one lepton with $p_T > 27\,\text{GeV}$ and one with $p_T \geq 10\,\text{GeV}$ are required. Additionally, all events used for the analysis must pass a further jet multiplicity cut, demanding at least two jets with $p_T > 25\,\text{GeV}$. No further event selections are applied in this analysis, except if explicitly mentioned in the

[1]The term *lepton* will from now on refer to electrons and muons only, unless anything else is stated alongside.

following chapters.

5 Heavy Flavour Classification (HFC)

After the cuts described in Sec. 4.4 have been applied, the main background to the $t\bar{t}+H$ signal is of the type $t\bar{t}+jets$ [3]. In order to enable a more event specific selection within this background, a non-overlapping selection based on jet multiplicity and jet flavour is made. In the case of a $t\bar{t}+H(H\rightarrow b\bar{b})$ measurement, for example, processes most closely related to the signal can be analysed in detail by selecting events that contain at least two b-tagged jets in addition to two b-jets originating in the top-quark decay.

The heavy flavour classification (HFC) provides a classification of the different types of events in the simulated data. It classifies events into categories based on truth-level information, as explained in the following section. Sec. 5.2 then provides an explanation of why the analysis focuses on particular regions of phase space. The final part of this chapter compares the HFC with a different b-jet classification algorithm, the `jet_truthflav` label.

5.1 Working principle

The classification of $t\bar{t}+jets$ events based on the jet flavour can be split into two parts. The heavy flavour simple classification (HFSC) and the more refined heavy flavour classification (HFC) [3, 61]. Both sort events based on the number of heavy flavour (HF) jets, discriminating between bottom- or charm-jets. Their general working principle is shown in Fig. 5.1 using the example of the former.

The HFSC takes into account all jets built from stable truth particles with $p_T > 15\,\text{GeV}$ and $|\eta| < 2.5$. It matches all hadrons to a truth jet and looks for those that were matched to at least one b-hadron with $p_T > 5\,\text{GeV}$ at a distance $\Delta R < 0.4$ from the respective jet axis. They are then called b-jets and any that do not originate in a top-quark or W-boson decay are called additional b-jets. An event is classified as $t\bar{t}+\geq 1b$ if the HFSC finds at least one additional b-jet. If this is not the case, but an additional c-jet is found based

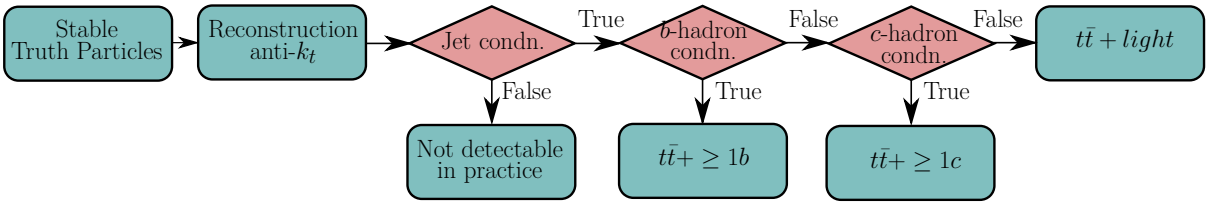


Figure 5.1: Flow chart of the heavy flavour simple classification. The jet condition selects jets with a transverse momentum above 15 GeV. The b- and c-hadron conditions check for hadrons of the corresponding flavour with $p_T > 5\,\text{GeV}$ that is matched to a jet. If at least one b- or c-jet is found using this method, the algorithm classifies the event as $t\bar{t}+\geq 1b$ and $t\bar{t}+\geq 1c$, respectively. If no jet can be matched to either type, the event is classified as $t\bar{t}+light$.

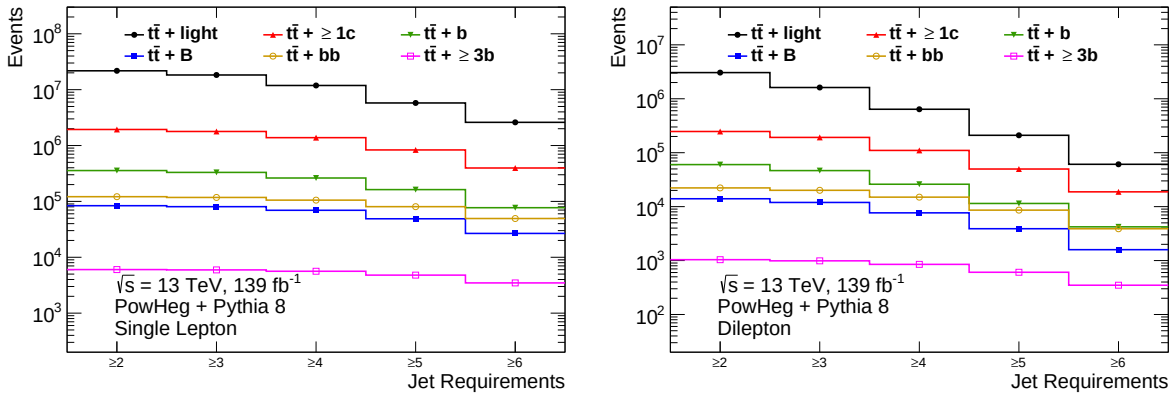
on the same tagging criteria but replacing b-hadron with c-hadron, the event is labelled $t\bar{t}+\geq 1c$. All events that are neither $t\bar{t}+\geq 1b$ nor $t\bar{t}+\geq 1c$ are classified as $t\bar{t}+light$, no matter whether this is due to a kinematic or flavour related criterion.

The HFC works in much the same way by classifying events based on the presence of additional b- or c-jets. However, a new distinction is used, based on the number of HF-hadrons found in a single jet: If only one b-hadron with $p_T > 5\,\text{GeV}$ is found within $\Delta R < 0.4$ of the jet axis, the jet is labelled b-jet. Similarly, it is called a B-jet, if there are two or more b-hadrons within that radius and the leading (most energetic) b-hadron has a p_T above 5 GeV. The HFC is discriminating based on the number and type of jets in the event. If one additional b-jet is found based on the method described above, the event is classified as $t\bar{t}+b$. Events with two b-jets are classified as $t\bar{t}+bb$, those with one B-jet as $t\bar{t}+B$. Any other event is classified as $t\bar{t}+\geq 3b$, including for example also those with only two B-jets.

For all events classified as $t\bar{t}+\geq 1c$, HFC works in an analogous way by classification into $t\bar{t}+c$, $t\bar{t}+C$, $t\bar{t}+cc$, and $t\bar{t}+\geq 3c$. If conflicts arise because both b- and c-hadrons are present in a jet, the jet is classified as b-jet. This level of detail within the $t\bar{t}+\geq 1c$ category is not very relevant for the analysis of the $t\bar{t}+H$ $(H \to b\bar{b})$ process and will thus not be examined in this thesis. It is, however, important to include in the HFC for future analyses, e.g. the $t\bar{t}+H$ $(H \to c\bar{c})$ process.

5.1.1 The `jet_truthflav` algorithm (JTF)

At reconstruction level there exists another algorithm that assigns a flavour tag to jets based on the type of hadron present. The `jet_truthflav` (JTF) algorithm works similarly



(a) (b)

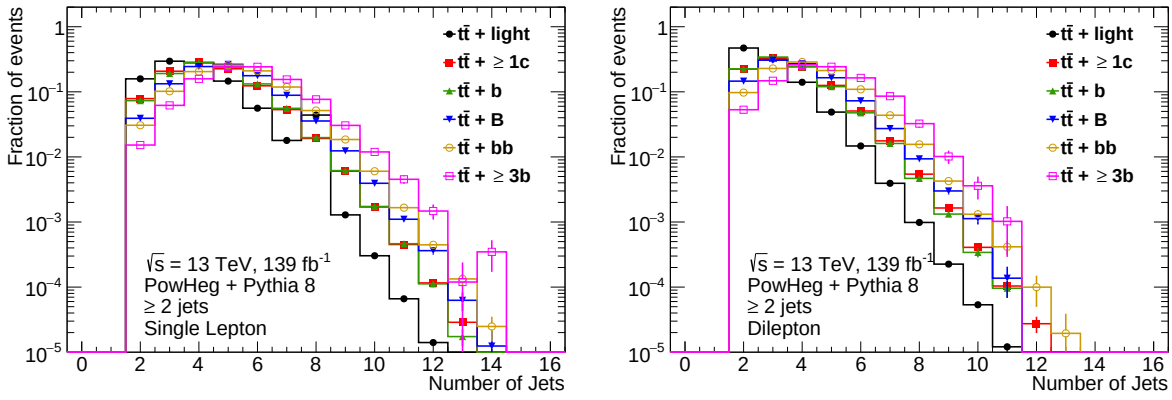
Figure 5.2: Comparison of different jet multiplicity requirements in the single (a) and dilepton channel (b) for the POWHEG + PYTHIA 8 sample. The total number of events decreases with increasing multiplicity requirement. This effect is more pronounced in the DL channel than in the SL channel. The influence on the relative composition of the different HFC categories is small.

to the first stage of the HFC: It starts with a list of all reconstructed jets and looks for weakly decaying hadrons with $p_T > 5\,\text{GeV}$ within $\Delta R < 0.3$ around each of the jet axes. If a hadron thus matched is found to be a b-hadron, the jet is called b-jet. The important difference between the JTF tag and the HFC apart from the slightly different radius parameter ($\Delta R_{HFC} = 0.4$) is that the latter requires the hadron to originate somewhere else than in the top decay. Another difference between the algorithms is due to the different object definitions at truth and reconstruction level, explained in Ch. 4.

Even though the two classifications, HFC and JTF, are different in their definition, they both have the goal of finding b-jets based on truth information. Comparing them will help to understand the different implementations and the consistency of both classification algorithms.

5.2 Further event selection

It should be noted that the requirement of at least two jets is only a loose selection criterion for the analysis of $t\bar{t} + H$ ($H \to b\bar{b}$) processes, since one would expect the signal to have at least four or six jets in the DL and SL channel, respectively. If, however, the task is background analysis, it is useful to perform a looser selection, allowing for an unconstrained study. This mainly affects the HFC categories requiring a lower jet



(a) (b)

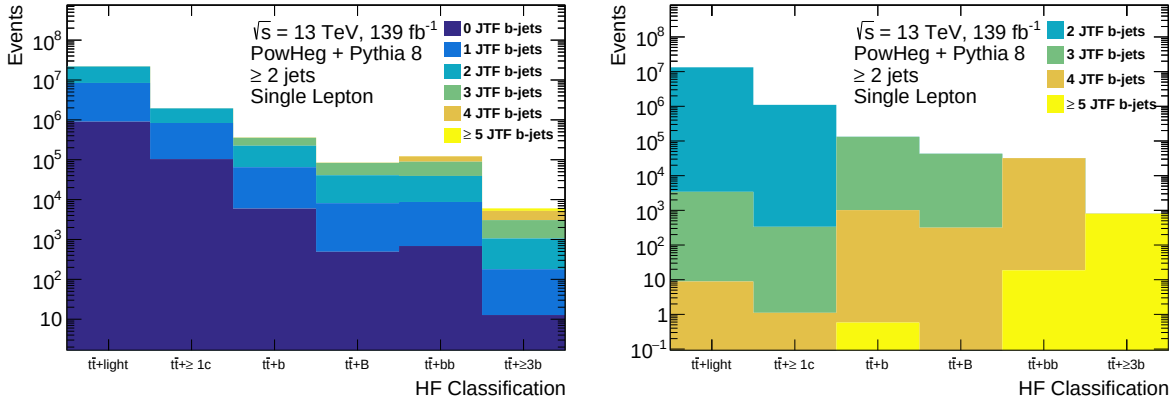
Figure 5.3: Comparison of the number of jets, broken down by number of true b-jets for the PowHeg + Pythia 8 sample. The plots clearly show the difference in single (a) and dilepton channel (b), resulting in two additional jets on average.

multiplicity, e.g. the $t\bar{t} + b$ classification which in principle requires only three jets. The jet requirement criterion has, however, as Figs. 5.2(a) and 5.2(b) show, only a small influence on the relative composition of the different HFC categories.

If one compares Figs. 5.2(a) and 5.2(b), it also becomes apparent that the DL channel is more sensitive to a jet requirement than the SL channel. This difference originates in the hadronic decay of one W-boson in the SL channel, leading to two additional jets in the final state. This hypothesis is supported by the comparison of the total number of jets in the SL and DL channel, shown in Fig. 5.3. From this, an important aspect of the subsequent analysis becomes apparent: It makes sense to study the SL and DL channel separately. In this thesis, the single lepton channel will be analysed more closely because it contains many more events. This is particularly important in the next chapter, where a smaller Monte Carlo sample will be used.

5.3 Comparison of HFC and JTF

Fig. 5.4 shows the differences between the HF classification and the JTF b-tag. Although the former categorises events while the latter tags jets, the comparison is possible under the assumption that almost all top quarks decay into a b-quark and a W-boson. From this, the number of additional b-jets in the HFC should be two above the number of JTF b-jets. From a HF classification of $t\bar{t} + b$ or $t\bar{t} + \geq 3b$ one would, for example, expect the



(a) (b)

Figure 5.4: The HFC is compared against the b-jet multiplicity determined by the JTF algorithm. Fig. (a) shows all b-jet multiplicities for the HFC categories, demonstrating that a large fraction of events is tagged with less b-jets than expected from the HFC (number of additional b-jets plus two). Fig. (b) shows only those events that are classified as expected (at the top of each bin) and those that have more b-jets according to the JTF than expected. The worst categories in this regard are $t\bar{t}+b$ and $t\bar{t}+B$.

JTF to tag three and five b-jets, respectively. Small deviations from this simple scheme are expected to occur due to W-bosons originating in a top quark and decaying into either b- or c-quarks. This effect has a larger impact on the $t\bar{t}+\geq 1c$ category, because the CKM matrix predicts a overwhelming rate of W-decays into *light*- and c-quarks and Higgs production is a very rare process. In such cases, the HFC would classify an event into a category containing less additional b-jets than inferred from the JTF tag.

However, the differences between the algorithms mentioned in Sec. 5.1.1 are the main reason for the inconsistent classification. Most of the events categorised differently by the two classifications have more additional b-jets than inferred from the JTF tag. Of the events classified as $t\bar{t}+b$, for example, the JTF tags less than the expected three b-jets per event in about 63% of all cases, outnumbering the consistently classified events, as Tab. 5.1 shows. In general, it can be concluded that on average the HFC classifies events in a category with more additional b-jets than expected from the JTF algorithm. This behaviour is expected to be related in part to the differences in jet-p_T cuts in HFC and at reconstruction level. These results are in accordance with a previous comparison in Ref. [62].

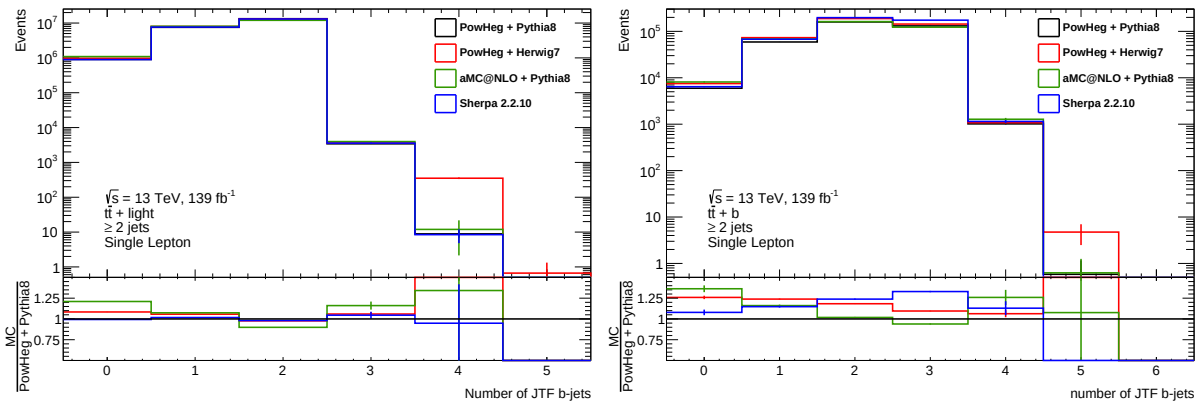
There are, however, also differences in classification that go the other way. Fig. 5.4(a) indicates that there are not many such events, but Fig. 5.4(b) proves that their number

Table 5.1: The table shows the percentage of events in each HFC-category contains 0, 1, 2, 3, 4, and ≥5 b-jets according to the JTF algorithm in the SL channel of the PowHeg + Pythia 8 sample. The expected number of b-jets for each HFC category, i.e. the number of additional b-jets plus the two b-quarks from top decay, is written in boldface. The numbers show what Fig. 5.4 illustrates, i.e. a large fraction of events in each category that contains less JTF b-jets and that the categories with the highest percentage of events with more b-jets than expected are $t\bar{t}+b$ and $t\bar{t}+B$.

HFC	Number of JTF b-tagged jets					
	0	1	2	3	4	≥ 5
$t\bar{t}+light$	4.1	34.6	**61.2**	2×10^{-2}	4×10^{-5}	0.0
$t\bar{t}+\geq 1c$	5.4	37.7	**56.9**	2×10^{-2}	6×10^{-5}	0.0
$t\bar{t}+b$	1.7	16.6	44.7	**36.8**	0.3	2×10^{-4}
$t\bar{t}+B$	0.6	9.1	39.1	**50.8**	0.4	0.0
$t\bar{t}+bb$	0.6	6.6	25.1	41.3	**26.4**	2×10^{-2}
$t\bar{t}+\geq 3b$	0.2	2.8	14.7	33.2	35.9	**13.3**

is not zero by showing only the consistently classified events and the ones with more JTF b-tagged jets than expected from the respective HFC category. The $t\bar{t}+b$ and $t\bar{t}+B$ classifications are the categories with the highest number of events that show this type of discrepancy. This happens for about 0.4% of all events in those categories. Most other categories have less events classified in this region, situated in the upper right of Tab. 5.1.

Fig. 5.5 shows that this observation can be made for all generators used in this analysis, using the example of the $t\bar{t}+light$ and $t\bar{t}+b$ category. Large discrepancies show up only for classifications with two more additional b-jets than expected from the number of JTF b-jets. In those regions, the PowHeg + Herwig 7 sample has about an order of magnitude more events than the other generators. A list of tables showing the event numbers per HFC category and JTF b-tag are provided in App. A together with the corresponding statistical uncertainties.



(a) (b)

Figure 5.5: The number of events per b-jet multiplicity according to JTF is compared for different generators. The agreement between the generators is high, for events classified as $t\bar{t} + light$ (a) and $t\bar{t} + b$ (b).

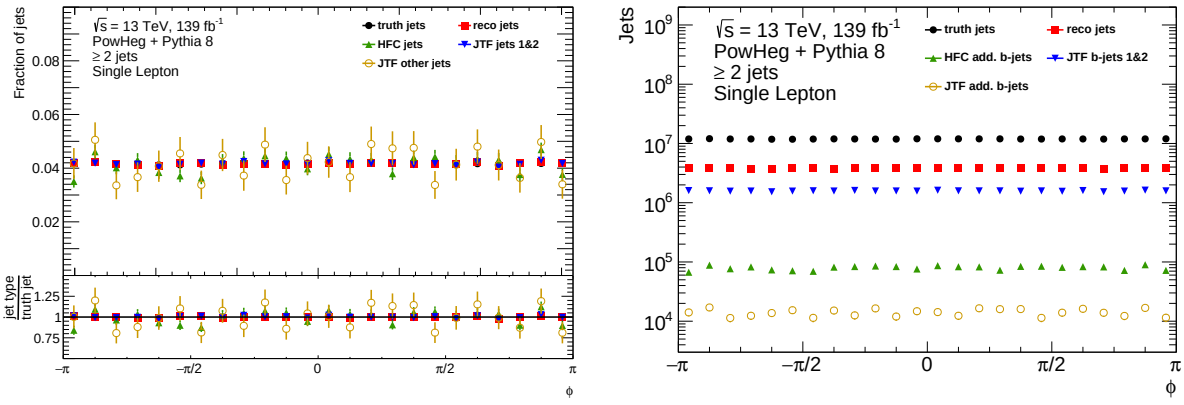
6 Modifications to the HFC

Although the previous chapter arrived at the conclusion that HFC works well compared to the JTF, there are still some incoherent classifications. Especially interesting are those cases where an event is classified by the HFC into a category containing less additional b-jets than expected from the b-jet multiplicity according to the JTF. The classifications that have more additional b-jets than expected are investigated as well, but it is expected that these are due to the stricter cuts on reconstruction level.

For this chapter a much smaller sample size is used for all generator setups, because the computational cost of the sample analysis is high. Another difference to the previous setup is the implementation of the HFC into the analysis software which was previously embedded into the Monte Carlo sample production. This enables a much more comprehensive analysis of modifications to the HFC in two ways: First, it reduces the computational cost for an analysis of a new HFC definition by some orders of magnitude, because sample production is even more costly than analysis. Second, it allows a study of the kinematic parameters of objects involved in the HFC (e.g. leading b-hadron p_T, cf. Sec. 6.1.4) which was not possible before.

6.1 Parameter analysis

First, a selection of kinematic parameters is analysed with the goal of finding those that could have a significant influence on the HFC if applied in a simple way, e.g. through a cut. The influence of samples from different generators will not be analysed, because a parameter should not be suitable for one generator only. The generator dependent behaviour will, however, be examined in Sec. 6.2 for those parameters deemed worth further consideration.



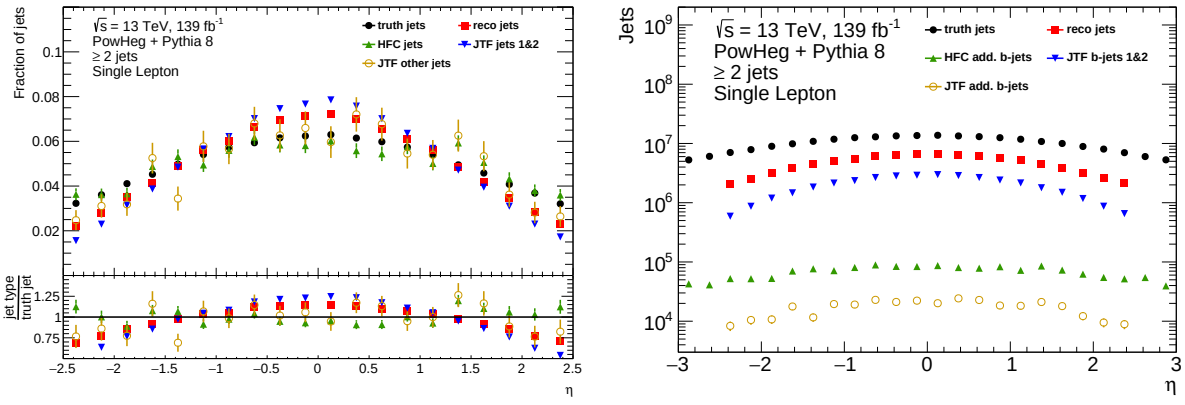
(a) (b)

Figure 6.1: Distribution of the azimuthal angle φ for different jet types. Fig. (a) shows no significant differences for different jet definitions (cf. Sec. 6.1.1). In Fig. (b) the absolute numbers of jets per definition are compared, showing that the additional b-jets contribute only with a small fraction to total number of jets.

6.1.1 Azimuthal angle

In principle, every parameter could possibly be optimised to yield a better classification performance. However, not all parameters are suitable, as the example of the azimuthal angle φ in Fig. 6.1 clearly proves. In the figure, the distribution of all jets at reconstruction and truth level are compared against distributions of the additional b-jets found by the HFC and the b-jets found by the JTF tag, both through matching to at least one b-hadron with $p_T > 5\,\text{GeV}$ (cf. Sec. 5.1). However, none of the HFC jet cuts are applied. Because the two b-jets originating in the top decay are on average expected to have a higher p_T than additional jets in a $t\bar{t}$ event, the two hardest b-jets as classified by the JTF tag are shown separately. They would not be classified as additional jets by the HFC and should thus not be used in a comparison. The remaining JTF b-jets will also be referred to as additional (JTF) b-jets.

The flat distribution across the complete range from $-\pi$ to π does not show any significant differences between different jet definitions. This behaviour is expected and almost obvious, since all objects observed in hadron collisions thus far have a flat φ distribution. This example serves, however, as a good reference for a variable not dependent on the type of jet to keep in mind when looking at the influence of other variables.



(a) (b)

Figure 6.2: Distribution of the pseudorapidity η for different jet types. Additional b-jets appear to have a slightly flatter distribution than jets defined by other definitions. The statistical uncertainties in Fig. (a) are, however, too large to draw any significant conclusions. Fig. (b) shows the absolute number of jets over a wider range of rapidities, thus making the $|\eta| < 2.5$ cut on reconstruction level jets visible.

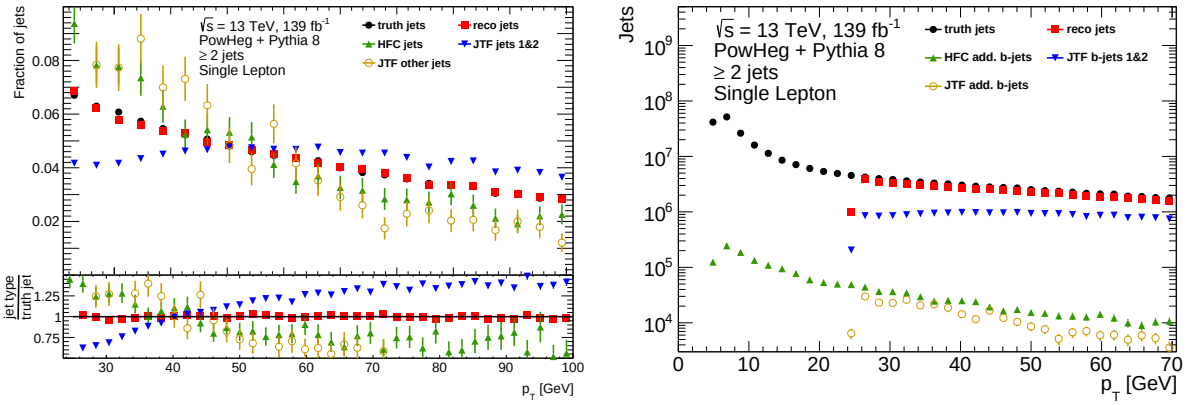
6.1.2 Pseudorapidity

A case less obvious is the pseudorapidity distribution. It does show differences in the η-distributions of the different jet types mentioned above, as Fig. 6.2 illustrates. It can be observed that the additional b-jets of both algorithms have a flatter distribution than the average jet. The two hardest b-jets from JTF have an even narrower distribution. Those differences are, however, only very small and not very significant. Although it may be possible to decrease the background with a narrower η cut, the small sample size used here is not able to resolve this at a scale fine enough.

The absolute numbers in Fig. 6.2(b) show, in addition to the $|\eta| < 2.5$ reconstruction level cut, that there are on average much less than two b-jets in a $t\bar{t}$ event. This is consistently true for both algorithms, leading to the conclusion that most additional jets are of light or charm flavour. This again highlights the importance of a good b-tagging algorithm, because it can significantly reduce the background.

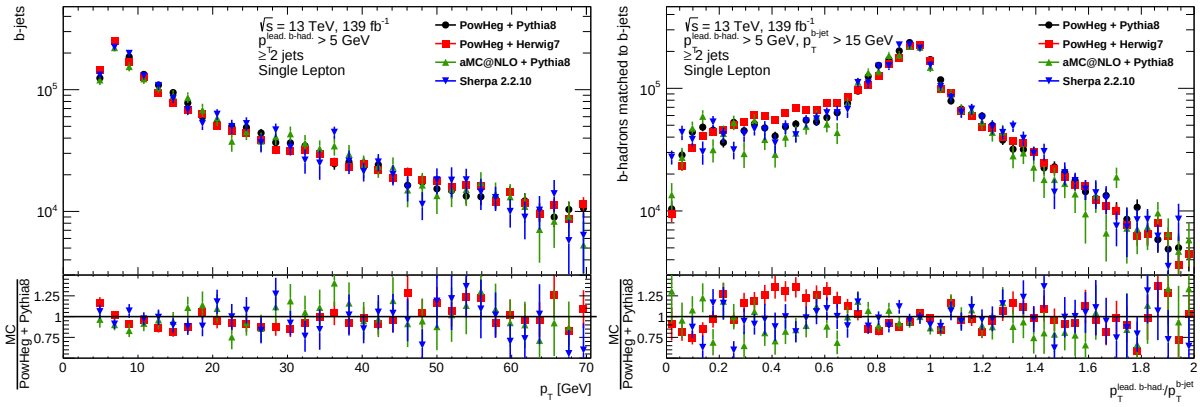
6.1.3 Jet transverse momentum

Of the two parameters investigated so far, neither is considered in the current HFC (except for a standard $|\eta| < 2.5$ cut), the reasons for which were analysed in the preceding sections.



(a) (b)

Figure 6.3: Distribution of the jet p_T for different jet types. The comparison of shapes in Fig. (a) suggests a strong dependence on the type of jet. Additional b-jets have on average a lower transverse momentum than the whole set of truth or reconstruction level jets and the two hardest JTF b-jets in particular. Fig. (b) shows the absolute p_T distribution of the different jet types. The reconstruction level cut at $p_T = 25\,\text{GeV}$ is also visible.



(a) (b)

Figure 6.4: p_T distribution of additional (HFC) b-jets matched to a b-hadron with $p_T > 5\,\text{GeV}$ (a) and ratio of matched b-hadron p_T with respect to the p_T of the corresponding additional b-jet (b) with $p_T > 15\,\text{GeV}$. A peak in the p_T distribution can be observed at a value corresponding to the 5 GeV cut on the leading b-hadron, taking into consideration that the leading b-hadron carries about 75% of the jet transverse momentum on average.

The jet p_T, however, is a selection criterion in the HFC. Only jets with a transverse momentum above 15 GeV pass the first stage of jet selection.

As the histograms in Fig 6.3 indicate, there is a strong dependence on p_T between the different jet definitions. The additional b-jet definitions employed by HFC and JTF have a very high fraction of events in the low energy range. The distributions of all jets on truth and reconstruction level are much less steep, although they too increase exponentially with decreasing energy. Only the two hardest b-jets as identified by JTF show a distribution with a maximum at about 40 GeV to 50 GeV. This could in part be due to the b-quarks from the top decay, as hypothesised earlier in Sec. 6.1.1.

A further point of interest is also the small maximum visible in Fig. 6.3 at about 8 GeV in the truth jet p_T distribution as well as in the HFC additional b-jet p_T distribution. Fig. 6.4 takes a closer look at this distribution for all additional b-jets according to the HFC. As just described, it falls off quickly if one excludes $p_T < 8$ GeV. For smaller values, the leading hadron requirement indirectly imposes a cut on the jet p_T. This influence of the b-hadron cut is visible above the actual boundary imposed by the cut, because on average, about 25% of the jet p_T do not originate in the matched b-jet, as Fig. 6.4(b) suggests. A leading hadron cut at 5 GeV thus translates into a jet p_T distribution peaked at about $5\,\text{GeV}/75\% = 6.6\,\text{GeV}$. Together with the exponential rise of the distribution, this gives a maximum some value slightly higher, just as observed.

It is a different circumstance that explains why a selection of highly energetic jets is reasonable. The reconstruction level jet p_T cut is at 25 GeV to prevent any soft b-jets from being considered in the subsequent analysis and because the $t\bar{t} + H$ ($H \to b\bar{b}$) signal is expected to have almost no jets below this threshold. The jet $p_T \geq 15$ GeV requirement has the same goal, only with a lower threshold, thus allowing for a p_T range where b-jets can be taken into account by the HFC but not the JTF. This discrepancy is one of the major differences between the two classifications and if the goal is to assimilate them, this cut will play a role, as Sec. 6.3 demonstrates. Even on the HFC on its own does this cut have a large influence on the number of classified events, as will be discussed in the next section.

In Figs. 6.4(a) and 6.4(b) a comparison between generators is also shown. They all agree well for the jet p_T distribution, although the uncertainties are very large due to the small sample size. The error bars are not much smaller for the p_T ratio cut in Fig. 6.4(b), but some differences between the Herwig 7 sample and the other generators are visible for p_T ratios below 0.8.

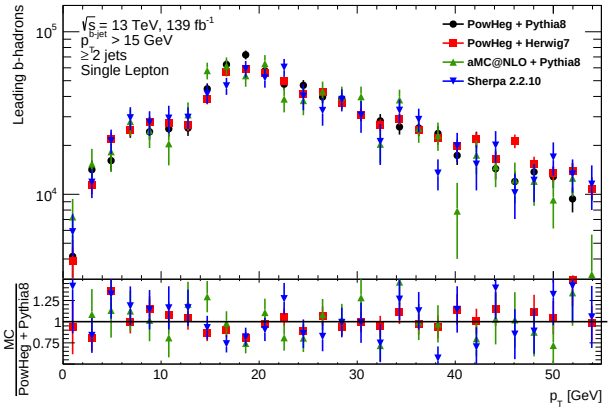


Figure 6.5: Distribution of leading b-hadrons matched to additional b-jets with $p_T > 15\,\text{GeV}$. A peak in the distribution is visible at about 20 GeV and a plateau around 10 GeV, a little lower than expected from the influence of the p_T cut applied to additional b-jets.

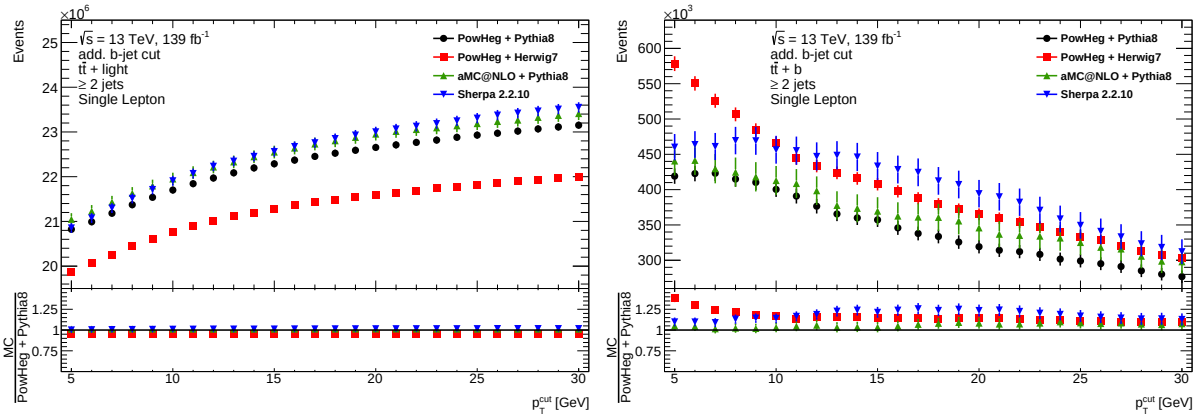
6.1.4 Leading b-hadron transverse momentum

Similarly to the jet p_T dependence, the number of leading b-hadrons decreases with an increasing cut on their p_T. This trend is clearly visible in Fig. 6.5, but it does not hold anymore for p_T lower than about 19 GeV. The distribution drops between 14 GeV and 20 GeV and rises slightly until it decreases again below about 6 GeV. This behaviour is not so easily explained by the 15 GeV cut on the additional b-jet p_T, from this one would expect a peak at about 11 GeV, i.e. at around 75% of the value of the jet cut, as indicated in Fig. 6.4(b). This result is, however, consistent with the same distribution presented in Ref. [61]. A comparison with the p_T of the b-hadrons used for matching by the JTF is not possible, because it is only implemented in the sample production algorithms and not available in the analysis files.

Unfortunately, the statistical uncertainties are too high to provide any meaningful information about the generator comparison. What can be deduced, however is that they all seem to agree well around the peak of the distribution. The deviation of the PowHeg + Herwig 7 sample observed in Ref. [61] at low p_T is not observed here.

6.2 New parameter configurations

From the preceding sections, it became clear that the angular parameters are not useful in the classification of additional HF-jets, at least not with the small sample size available.



(a) (b)

Figure 6.6: Influence of the cut on the p_T of additional b-jets on the HFC, using the example of the $t\bar{t} + light$ (a) and $t\bar{t} + b$ category (b), also comparing different generators. Although they agree well among each other, small differences in their qualitative behaviour are observed, in particular for the PowHeg + Herwig 7 sample.

The transverse momenta of (truth-level) jets and b-hadrons, however, were shown to provide significant grounds for a selection. This section determines the influence of both parameters on the HF classification before their influence on the relation between the HFC to the JTF is investigated in the next section.

6.2.1 b-jet transverse momentum

Fig. 6.6 shows the dependence of the HF classification on the cut of the truth jet p_T. As might be expected intuitively, the number of events identified as $t\bar{t} + b$ decreases with rising jet cut, while the number of $t\bar{t} + light$ events rises. This general trend is observed consistently throughout all generators used in this analysis and consistent to the results found in Ref. [61].

Interestingly, the shape of the distribution is very sensitive to the generator used. For the PowHeg + Herwig 7 sample, differences far above 25% with respect to the nominal PowHeg + Pythia 8 sample are observed in the $t\bar{t}+b$ category for jet cuts below 10 GeV. Another different response is shown by the Sherpa 2.2.10 sample, where the number of events classified into $t\bar{t} + b$ starts to decrease about 5 GeV higher than is the case for PowHeg + Pythia 8, the aMc@nlo + Pythia 8 and the PowHeg + Herwig 7 samples. Even though there are qualitative differences between the generators, their

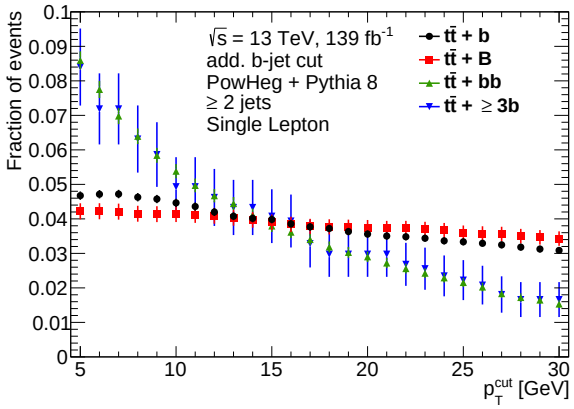


Figure 6.7: Comparison of the behaviour in the different $t\bar{t}+ \geq 1b$ subcategories for different additional b-jet p_T cuts using the PowHeg + Pythia 8 sample. Each subcategory is normalised to one over the range of the cuts. The figure indicates that the subcategories behave in a similar way, but with different slopes. $t\bar{t} + b\bar{b}$ and $t\bar{t}+ \geq 3b$ decrease much quicker than $t\bar{t} + b$ and $t\bar{t} + B$.

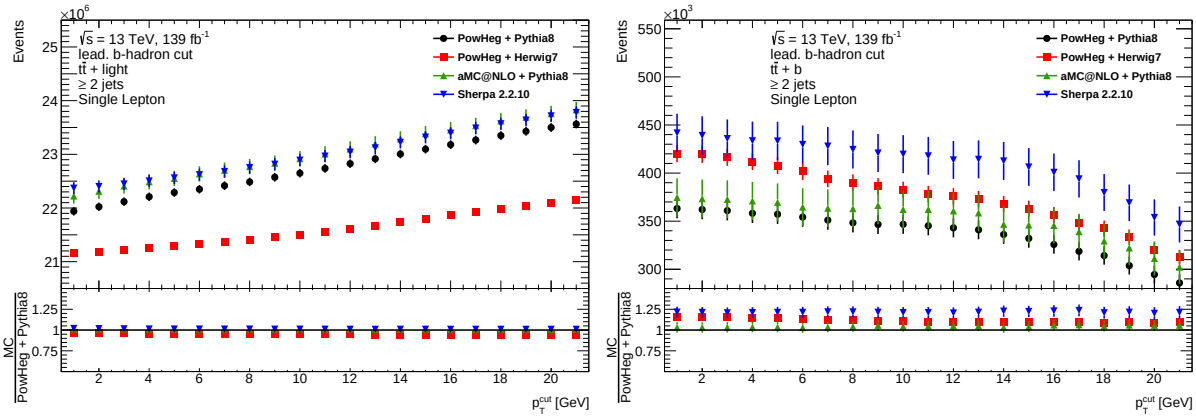
deviation is, however, in most cases still below 25%. The most accurate category is the $t\bar{t} + light$ bin, with differences of less than 10%.

An additional observation can be made in Fig. 6.7. As already stated, decreasing the jet cut increases the number of HF-classified events. This increase is, however, not uniform across all $t\bar{t}+ \geq 1b$ subcategories. The slope for categories with higher additional b-jet multiplicity is higher than for those with only one additional b-jet (or B-jet). This is consistent with the assumption of a large number of additional b-jets at low energy scales that falls off fast with increasing energy. The figure also shows that the $t\bar{t}+ \geq 1b$ subcategories behave in a similar way qualitatively, indicating that Fig. 6.6(b) can, to some extent, be taken as representative for all those subcategories. The exact behaviour of the different categories can be seen in the Fig. B.1 in App. B.

6.2.2 Leading b-hadron transverse momentum

The number of HF-classified events decreases when the p_T cut imposed on the leading HF-hadron is increased, in a similar way to the behaviour seen for different jet p_T cuts in the previous section. This, again, applies to all categories except $t\bar{t} + light$ where the relation is exactly the opposite. Fig. 6.8 shows this behaviour over a range of different b-hadron cuts, which is also consistent with the results in Ref. [61].

In terms of generator response, an interesting phenomenon is visible. In the $t\bar{t} + light$



(a) (b)

Figure 6.8: Influence of the cut on the p_T of leading b-hadrons on the HFC, using the example of the $t\bar{t} + light$ (a) and $t\bar{t} + b$ category (b), also comparing different generators. They generally agree very well with each other, with PowHeg + Herwig 7 and Sherpa 2.2.10 showing the most deviation from the other generators in the $t\bar{t} + light$ or $t\bar{t} + b$ category, respectively.

category, shown in Fig. 6.8(a), the PowHeg + Herwig 7 sample is, as seen before, the farthest away from the other generators which deviate from each other by less than 1%. This is different in the $t\bar{t} + b$ category, presented in Fig. 6.8(b), where the Sherpa 2.2.10 sample classifies the most events across the whole range. Apart from the offset, however, it performs similarly to the two Pythia 8 samples. The PowHeg + Herwig 7 sample shows a slightly steeper incline at low energy cuts.

The relation between different responses of the $t\bar{t}+ \geq 1b$ subcategories is shown in Fig. 6.9. All categories behave similarly for p_T cuts below 10 GeV, decreasing only slightly as the cut value increases. This changes for values above 10 GeV to 12 GeV, where the same phenomenon as in Fig. 6.7 can be observed, depending on the b-jet multiplicity of the subcategory. This is probably due to the circumstance that at b-hadron cuts that high, the limiting factor becomes the p_T cut on the matched jet, which translates to an average leading b-hadron p_T of about 11 GeV, as explained in Sec. 6.1.4. Fig. B.2 in App. B shows the dependence for all HFC categories and generators.

Comparing between the additional b-jet p_T cut and the leading b-hadron cut, one observes that the jet cut has the higher influence on the classification of events by the HFC.

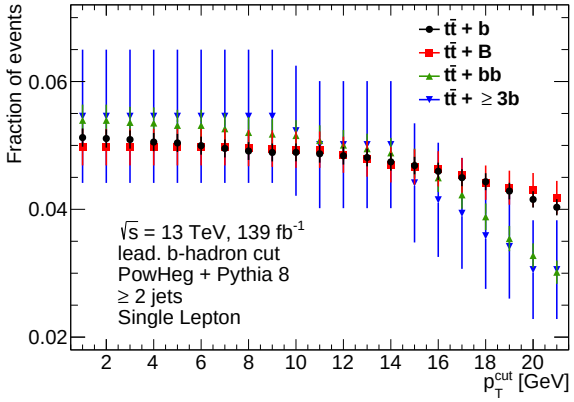


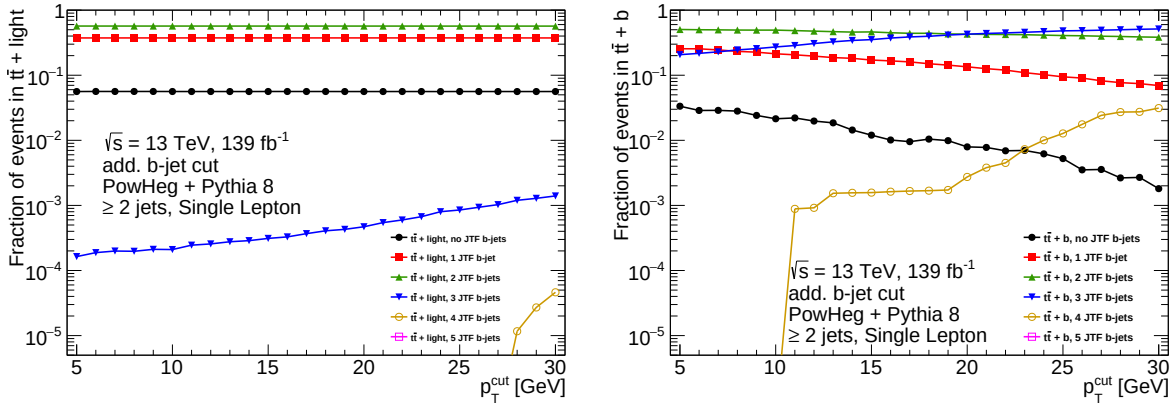
Figure 6.9: Comparison of the behaviour in the different $t\bar{t}+ \geq 1b$ subcategories for different p_T cuts on the leading b-hadron using the PowHeg + Pythia 8 sample. Each subcategory is normalised to one over the range of the cuts. Differently to the p_T cut on additional b-jets, the dependence is much weaker, only changing significantly above 10 GeV to 12 GeV depending on the category. Categories with higher b-jet multiplicity start to decrease earlier than those with low multiplicities.

6.3 Influence of cut variations on the comparative performance of HFC and JTF

It was established in the previous section that the HFC is sensitive to cuts of the additional b-jet p_T and of the leading b-hadron p_T, with the result that the jet cut is a little stronger. In this section, the influence of different HFC parameter configurations on the comparative performance of HFC and JTF will be investigated.

The influence of a change in the p_T of additional b-jets can be seen in Fig. 6.10, again using the example of the $t\bar{t}+light$ and $t\bar{t}+b$ categories. The first observation to be made is that as the cut gets higher, the fraction of events in the $t\bar{t}+light$ category that are classified as having a lesser or equal number of jets compared to the expectation according to the JTF (i.e. 0, 1, and 2 JTF b-jets) does not change significantly. The relative number of events with three JTF b-jets rises on the other hand by about an order of magnitude, but still contributing only up to about 0.1% at a jet cut of 30 GeV.

In the $t\bar{t}+b$ category, on the other hand, the fraction of events classified higher than expected, i.e. all events with less than three JTF b-jets, decreases significantly while the fraction of events classified as expected increases, becoming the largest contributor at a cut value around 20 GeV. At the same time, however, the fraction of events classified lower than the expectation (more than three JTF $b-$-jets) grows as well, contributing

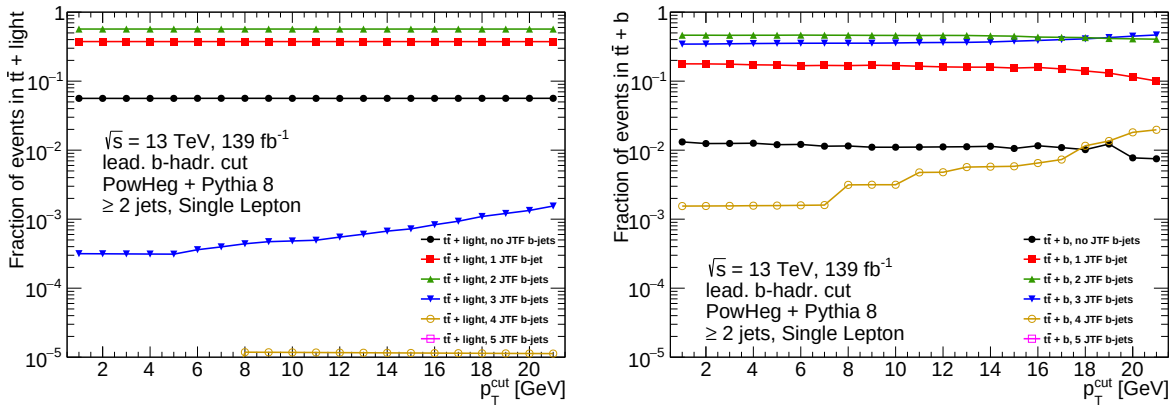


(a) (b)

Figure 6.10: The change of the relative composition of JTF b-jet multiplicities in the $t\bar{t}+light$ (a) and $t\bar{t}+b$ (b) HFC categories is shown for different p_T cuts on additional b-jets taken into account by the HFC. All other parameters are set to their current nominal value. In general, the fraction of events with multiplicities at or above the expected value (add. HFC b-jets plus two) rises with an increasing cut on the p_T of the additional b-jets in the HFC, while the fraction of events with less b-jets than expected decreases. This change is more pronounced in the $t\bar{t}+b$ category.

about 3% at the maximum cut shown in Fig. 6.10(b). That fraction drops abruptly at cut values below 11 GeV, although this is most likely due to the low statistics of the sample. Statistical uncertainties are, however, not drawn, since all categories are not at all independent from each other, but related by the total number of events in each category. The increasing p_T cuts are also not independent from each other, since those events classified at a higher cut value form a subset of all events classified at a lower cut. This is only possible, because the same sample is considered in all figures used in this section, if this would not be the case, the uncertainties would become necessary. For that case and further investigations, the figures for other HFC categories in App. C and absolute event numbers and their statistical uncertainties for a few representative cuts can be found in App. D.

The cut on the leading b-hadron has a very similar influence on the relative composition of events in the HFC categories. The main difference is that a significant reduction in the fraction of events with less JTF b-jets than expected in the $t\bar{t}+b$ category does not happen until a cut value of about 15 GeV is reached, the value of the cut imposed on the p_T of the additional b-jets. This is in accordance with the previous finding that the influence of cuts on the leading b-hadron is weaker than the one from cuts on the additional b-jet



(a) (b)

Figure 6.11: The change of the relative composition of JTF b-jet multiplicities in the $t\bar{t}+light$ (a) and $t\bar{t}+b$ (b) HFC categories is shown for different p_T cuts on leading b-hadrons taken into account by the HFC. All other parameters are set to their current nominal values. As for the cut on the p_T of additional b-jets, the fractions of events classified as expected and those where the JTF finds more b-jets rises with increasing cut value, while the fraction of events with less JTF b-jets than expected decreases. However, the change is much less rapid, as already observed for the influence of the cut on leading b-hadron p_T on the HFC as a whole. Fig. 6.12 examines the $t\bar{t}+b$ category closer in the range of 1 GeV to 10 GeV, finding a possible remnant of cuts done by the JTF algorithm.

p_T and not completely independent from it.

If one looks closely, an additional phenomenon in the classification of events in the $t\bar{t}+b$ channel can be observed, shown in Fig. 6.12. At about 5 GeV, a local maximum appears in the number of events classified as $t\bar{t}+b$ and containing three b-jets. Because of the dependence of the different cuts mentioned above, this is not a statistical fluctuation. The reason might instead be the 5 GeV cut on b-hadrons that is also deployed in the JTF algorithm, because the same peak does not appear in the histogram of different cut values containing the $t\bar{t}+b$ category as a whole.

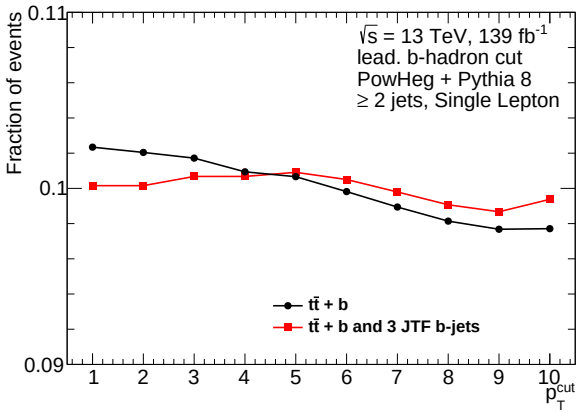


Figure 6.12: Excerpt from Fig. 6.11(b), showing a small rise in the number of events in the $t\bar{t}+b$ category and additionally tagged with three JTF b-jets. The same rise is not visible in the curve of the events classified as $t\bar{t}+b$ without the restriction. This indicates that the origin is not in the HFC, probably in the 5 GeV hadron cut of the JTF.

7 Conclusions

This thesis took a close look at the HF classification, that categorises generated $t\bar{t} + jets$ events. In Sec. 5.2, it was shown that it is justified to carry out the analysis in the single lepton channel with a jet multiplicity requirement of at least two jets, in order to be able to work with maximum statistics. Then the HFC was compared to the JTF algorithm that assigns a flavour to each jet on reconstruction level. It was found that both algorithms agree well in their present configuration, because for less than 0.5% of all events per HFC category the JTF finds more b-jets than expected from the HF classification. Due to the difference of algorithms and the much looser truth level jet p_T cut, the number of events tagged with less b-jets by the JTF than expected from HFC is very high, even outnumbering the consistently classified events in the $t\bar{t} + b$, $t\bar{t} + bb$, and $t\bar{t}+ \geq 3b$ categories. A good general agreement between the classifications for samples from different generators was also established.

In the following Ch. 6, different kinematic parameters were analysed in view of their influence on the HFC. The p_T of additional b-jets and of leading b-hadrons were found to be particularly well suited, also reflected by their use in the current HFC algorithm. It was found in Sec. 6.2 that an increased cut in both parameters leads to more events being classified as $t\bar{t} + light$, as one would expect, since a higher cut excludes more objects that are subsequently sorted into the $t\bar{t} + light$ category. It was also shown that the influence of a cut on the additional b-jet p_T has a larger influence on the HF classification than a cut on the p_T of the leading b-hadrons. This is probably due to a dependence of the variables upon each other. The peak in the ratio of leading b-hadron p_T and matched truth jet p_T supports this hypothesis.

Finally, it was demonstrated in Sec. 6.3 that an increased cut in both parameters leads to more events being classified consistently between JTF and HFC. Again, the p_T cut on additional b-jets was found to have a greater influence than a cut on the p_T of leading b-hadrons. A local maximum in the dependence of the consistently classified $t\bar{t} + b$ events on the leading b-hadron p_T cut that did not appear in the corresponding data for all $t\bar{t} + b$ events might be due to the 5 GeV cut on b-hadron p_T in the JTF algorithm, suggesting

that the algorithms can be assimilated to some extent by similar cut criteria.

In conclusion, the HFC performs well in comparison to the JTF tag. Taking into account the larger field of view of the HFC due to lower thresholds, it is expected that many more b-jets are found by the HFC than by the JTF flag. As demonstrated in Sec. 6.3, this can in part be compensated by assimilation of the cut parameters used in the HFC definition to those of the JTF. A complete removal of all discrepancies was, however, not possible, probably due to the differences between truth and reconstruction level.

Problematic are those discrepancies where the HF classification sorts events into categories that contain less additional b-jets than what is expected based on the information from JTF. This is a strange behaviour because the cuts imposed by the HFC are in principle much looser than those of the JTF tag. It has to be kept in mind, however, that this holds only under the assumption that the only bottom quarks originating in a top quark are just the ones from its decay. If one considers b-quarks from W-bosons, this behaviour could in part be explained, even though their effect will probably be much smaller than the observed discrepancies. As a first approximation, the numbers of wrongly classified events could be compared to the production cross sections of, for example, the $t\bar{t} + H$ $(H \to b\bar{b})$ process, to see how large the differences are. To further exclude any other origins of these classifications, a closer investigation into the nature of the JTF tag and its input variables is necessary to see where the JTF might tag an additional b-jet that the HF classification has not found. A close inspection of the algorithm that flags hadrons resulting from top-quark remnants could also be an interesting subject.

Since the comparative performance of the HFC against the JTF algorithm was found to be already close to an optimum, it is hard to predict an improvement that leads to significant changes in the classification. As was clearly found, a change in the p_T cut value on leading b-hadrons does not have a great influence. It might be worth the effort to increase the p_T cut on additional b-jets to a value in the range of 15 GeV to 20 GeV, since that would lead to a reduction of events classified higher than expected from the JTF and the curve describing the fraction of events that are classified lower than expected seems to flatten out at around 20 GeV. This might, however, be due to the small sample size. Due to the observed dependency between the p_T of leading b-hadrons and the matched jets it might also be possible to make improvements by lowering cuts on jet p_T and raising the cuts on the leading b-hadron.

It should also be noted that two kinematic parameters have not been considered in this thesis, the matching radius ΔR and the ratio between leading b-hadron and matched jet.

The latter was used for some part of the result discussion and was shown to have some sort of influence, but it is currently not considered in the HFC algorithm. The former is used and is a parameter of the JTF tag as well. Since both methods disagree on the value ($\Delta R < 0.4$ for HFC and $\Delta R < 0.3$ for JTF), it might also be a parameter worthwhile exploring.

Finally, the results obtained here should be verified with a larger sample size, to reduce the uncertainties that are currently in the way for some aspects of the analysis. Such an analysis, however, takes a lot of computational time. Apart from the additional aspects it would be able to shed light on, an investigation with higher statistics could also increase the certainty of the statements on observations made in this thesis.

A Tables comparing the HFC and JTF for different generators

Table A.1: Event numbers (a) and uncertainties (b) for the large POWHEG + PYTHIA 8 sample with HFC parameter settings 15 GeV for the p_T cut on jets and 5 GeV for the leading b-hadron p_T cut. The HFC category that is expected from a certain JTF b-jet multiplicity is written in boldface.

(a)

HFC	Number of JTF b-tagged jets						
	0	1	2	3	4	≥ 5	total
$t\bar{t} + light$	914995	7645350	**12665000**	3467	8	0	21228821
$t\bar{t} + \geq 1c$	105484	742500	**1118760**	337	1	0	1967082
$t\bar{t} + b$	6053	59961	161861	**133099**	1029	0	362005
$t\bar{t} + B$	506	7784	33250	**43169**	319	0	85029
$t\bar{t} + bb$	695	8160	30954	50897	**32549**	20	123277
$t\bar{t} + \geq 3b$	12	169	895	2036	2192	**805**	6112
total	1027746	8463925	14010720	233007	36100	805	23773918

(b)

HFC	Number of JTF b-tagged jets					
	0	1	2	3	4	≥ 5
$t\bar{t} + light$	707	2041	**2618**	42	2	0
$t\bar{t} + \geq 1c$	240	636	**778**	13	0	0
$t\bar{t} + b$	57	181	296	**267**	23	0
$t\bar{t} + B$	16	65	134	**152**	13	0
$t\bar{t} + bb$	19	67	130	166	**132**	3
$t\bar{t} + \geq 3b$	2	9	22	33	34	**26**

Table A.2: Event numbers (a) and uncertainties (b) for the large POWHEG + HERWIG 7 sample with HFC parameter settings 15 GeV for the p_T cut on jets and 5 GeV for the leading b-hadron p_T cut. The HFC category that is expected from a certain JTF b-jet multiplicity is written in boldface.

(a)

HFC	Number of JTF b-tagged jets						
	0	1	2	3	4	≥ 5	total
$t\bar{t} + light$	976951	7960030	**13017450**	3616	351	0	21958399
$t\bar{t} + \geq 1c$	87282	589518	**867200**	259	21	0	1544282
$t\bar{t} + b$	7495	73154	188624	**143590**	1078	4	413948
$t\bar{t} + B$	745	10982	46682	**61277**	412	3	120104
$t\bar{t} + bb$	604	8219	32553	53907	**34788**	21	130094
$t\bar{t} + \geq 3b$	11	273	1069	2408	2071	**613**	6447
total	1073089	8642179	14153580	265059	38724	613	24174519

(b)

HFC	Number of JTF b-tagged jets					
	0	1	2	3	4	≥ 5
$t\bar{t} + light$	1343	3879	**4989**	83	26	0
$t\bar{t} + \geq 1c$	405	1058	**1283**	22	6	0
$t\bar{t} + b$	117	371	600	**522**	45	2
$t\bar{t} + B$	37	144	299	**342**	27	2
$t\bar{t} + bb$	33	125	250	320	**256**	6
$t\bar{t} + \geq 3b$	4	22	45	67	62	**43**

Table A.3: Event numbers (a) and uncertainties (b) for the large aMc@nlo + Pythia 8 sample with HFC parameter settings 15 GeV for the p_T cut on jets and 5 GeV for the leading b-hadron p_T cut. The HFC category that is expected from a certain JTF b-jet multiplicity is written in boldface.

(a)

HFC	Number of JTF b-tagged jets						
	0	1	2	3	4	≥ 5	total
$t\bar{t} + light$	1090615	8085430	**11975780**	3962	11	0	21155799
$t\bar{t} + \geq 1c$	124173	772615	**1031996**	480	0	0	1929265
$t\bar{t} + b$	8115	68528	162306	**123103**	1278	0	363332
$t\bar{t} + B$	474	7270	27687	**32488**	453	0	68374
$t\bar{t} + bb$	809	9225	31690	48677	**27883**	18	118306
$t\bar{t} + \geq 3b$	2	175	794	1501	1719	**591**	4785
total	1224190	8943245	13230255	210215	31346	591	23641048

(b)

HFC	Number of JTF b-tagged jets					
	0	1	2	3	4	≥ 5
$t\bar{t} + light$	3025	8182	**9922**	195	12	0
$t\bar{t} + \geq 1c$	1061	2626	**3023**	62	0	0
$t\bar{t} + b$	275	807	1240	**1057**	108	0
$t\bar{t} + B$	73	262	499	**532**	58	0
$t\bar{t} + bb$	93	310	561	662	**483**	13
$t\bar{t} + \geq 3b$	9	40	92	123	125	**86**

Table A.4: Event numbers (a) and uncertainties (b) for the large SHERPA 2.2.10 sample with HFC parameter settings 15 GeV for the p_T cut on jets and 5 GeV for the leading b-hadron p_T cut. The HFC category that is expected from a certain JTF b-jet multiplicity is written in boldface.

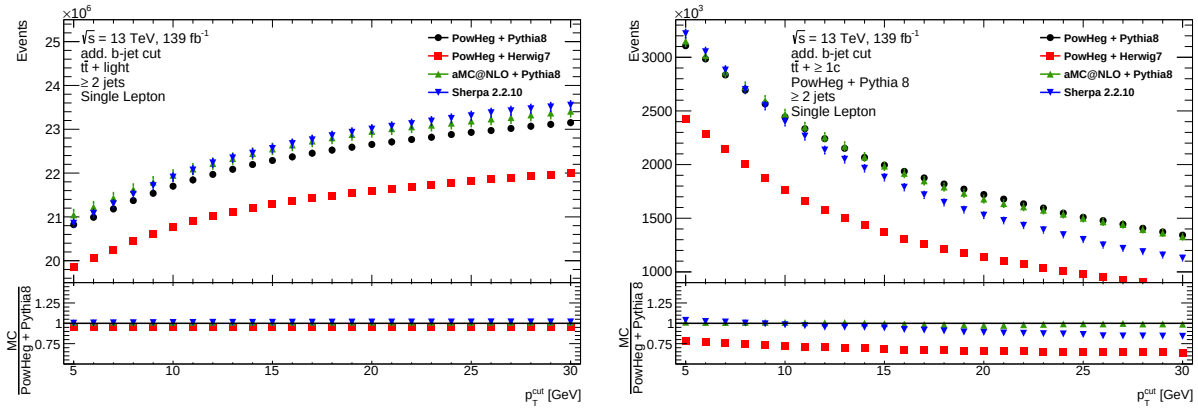
(a)

HFC	Number of JTF b-tagged jets						
	0	1	2	3	4	≥ 5	total
$t\bar{t} + light$	895417	7658130	**13066580**	3568	8	0	21623704
$t\bar{t} + \geq 1c$	94945	665887	**1025453**	384	0	0	1786669
$t\bar{t} + b$	6422	67764	197475	**174298**	1146	0	447106
$t\bar{t} + B$	551	8461	34969	**44847**	361	0	89192
$t\bar{t} + bb$	743	8733	34477	58994	**37182**	13	140144
$t\bar{t} + \geq 3b$	12	219	1087	2583	2430	**906**	7240
total	998092	8409195	14360043	284676	41129	906	24095833

(b)

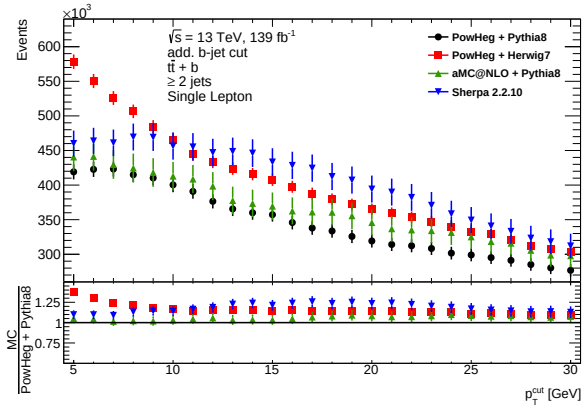
HFC	Number of JTF b-tagged jets					
	0	1	2	3	4	≥ 5
$t\bar{t} + light$	3123	8901	**11488**	221	6	0
$t\bar{t} + \geq 1c$	1144	3049	**3797**	72	0	0
$t\bar{t} + b$	302	986	1708	**1626**	131	0
$t\bar{t} + B$	93	351	736	**857**	76	0
$t\bar{t} + bb$	110	370	769	1030	**866**	19
$t\bar{t} + \geq 3b$	15	48	132	225	220	**176**

B Plots showing the parameter dependence of of all HFC categories

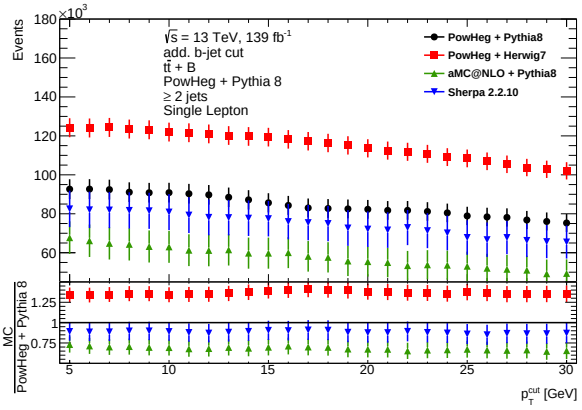


(a)

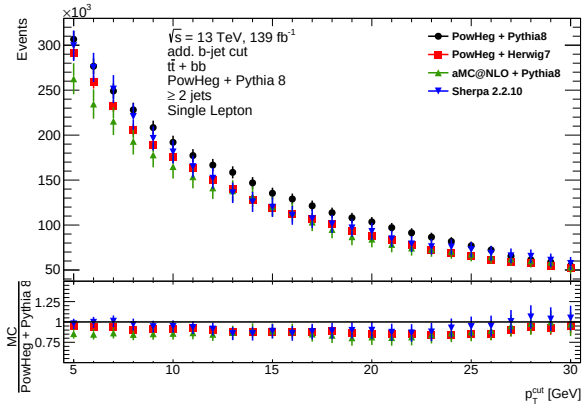
(b)



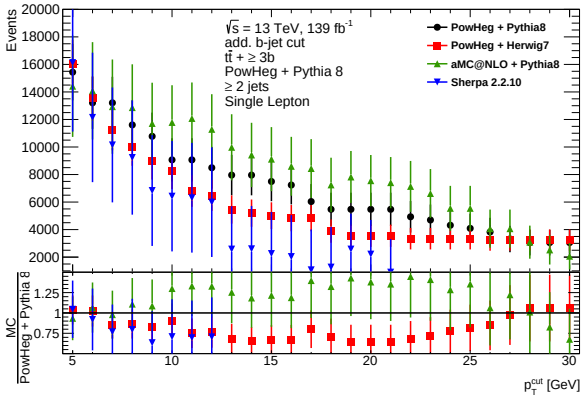
(c)



(d)



(e)



(f)

Figure B.1: Influence of the jet p_T cut on the number of events categorised in each HFC category. Additionally, the difference between generators is shown.

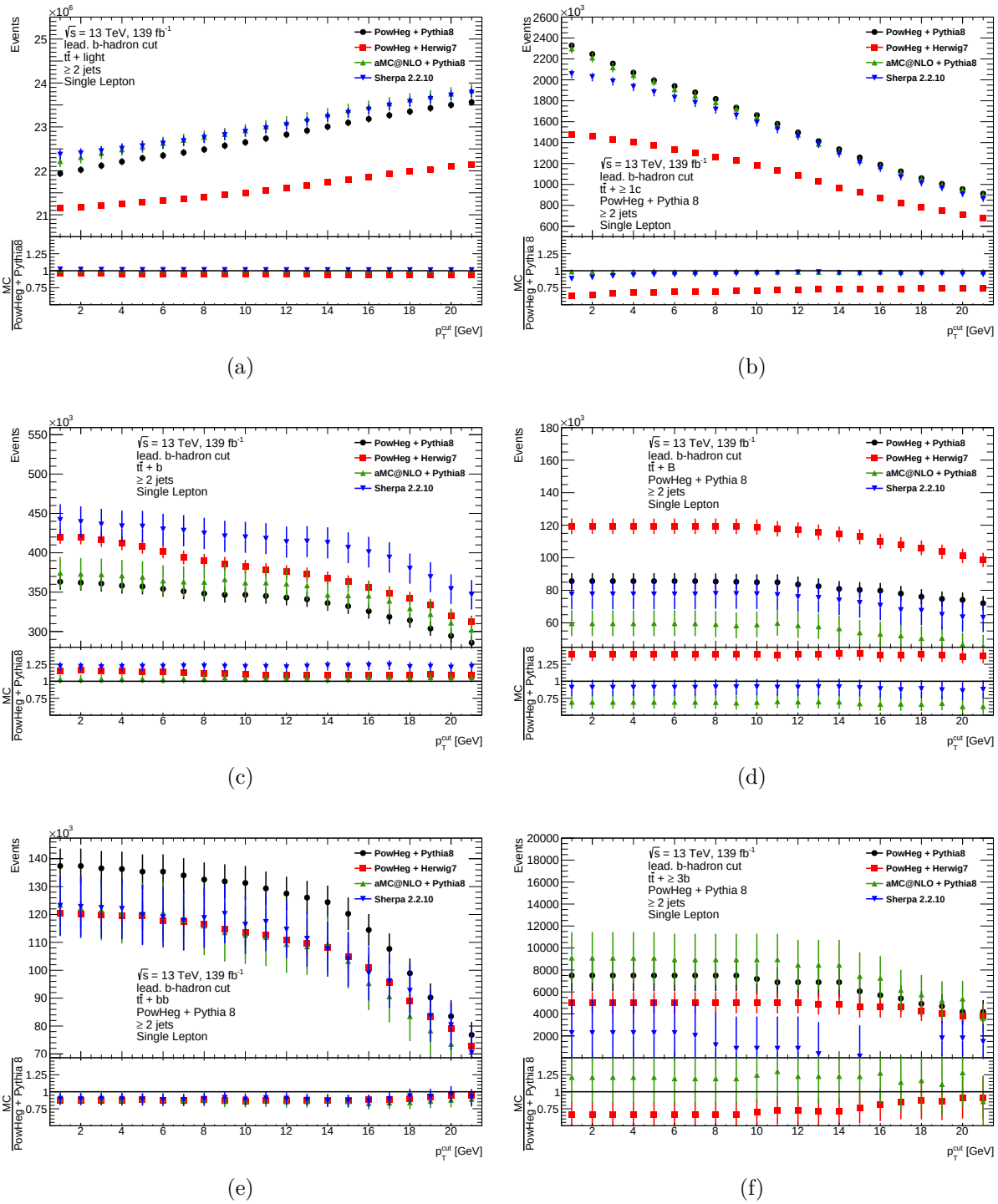


Figure B.2: Influence of the cut on the p_T of leading b-hadrons on the number of events categorised in each HFC category. Additionally, the difference between generators is shown.

C Plots showing the parameter dependence of the comparative performance of and JTF

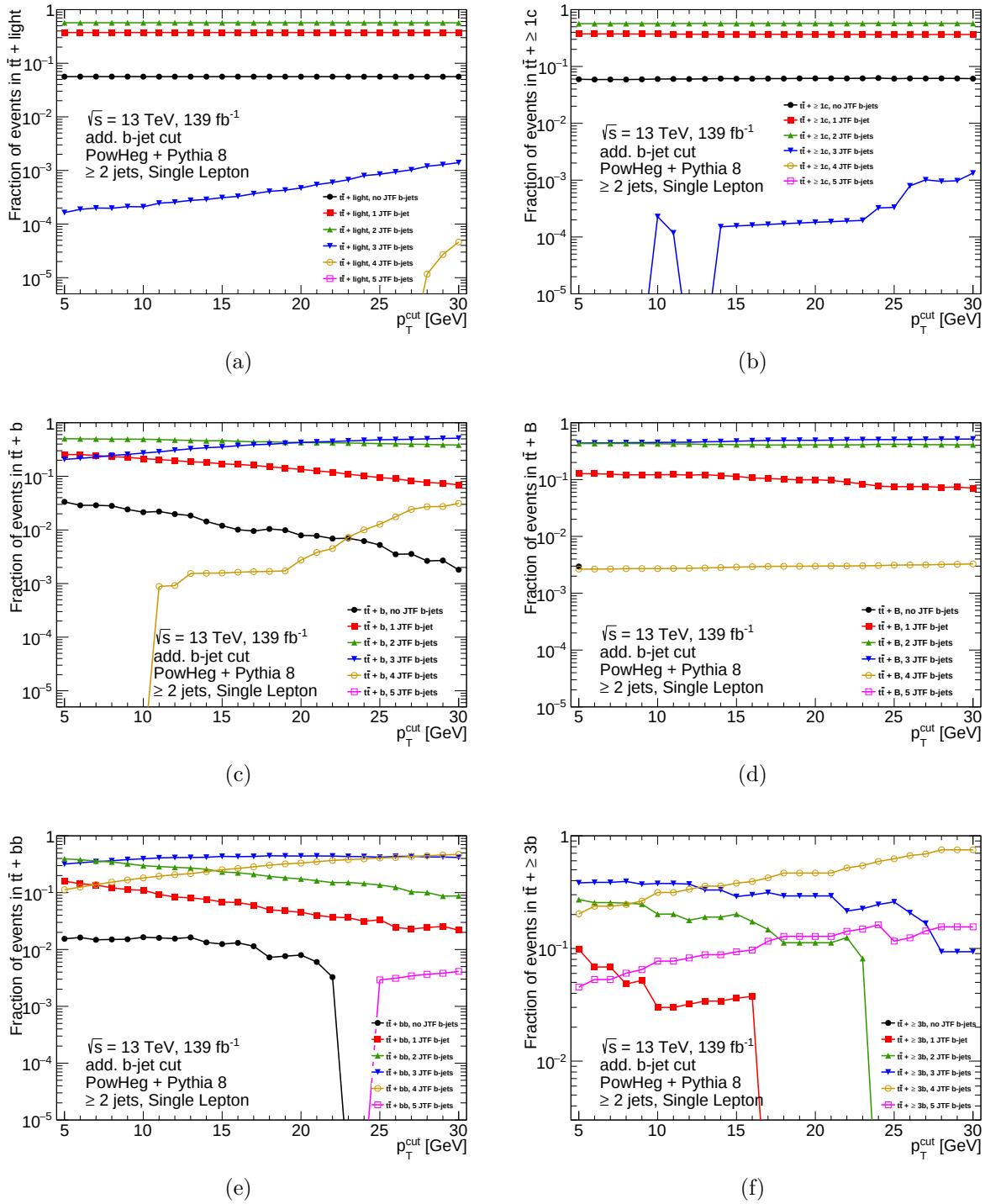


Figure C.1: Influence of the jet p_T cut on the fractional composition of each HFC category with regard to the different JTF b-jet multiplicities. All other parameters of the HFC are at their nominal values.

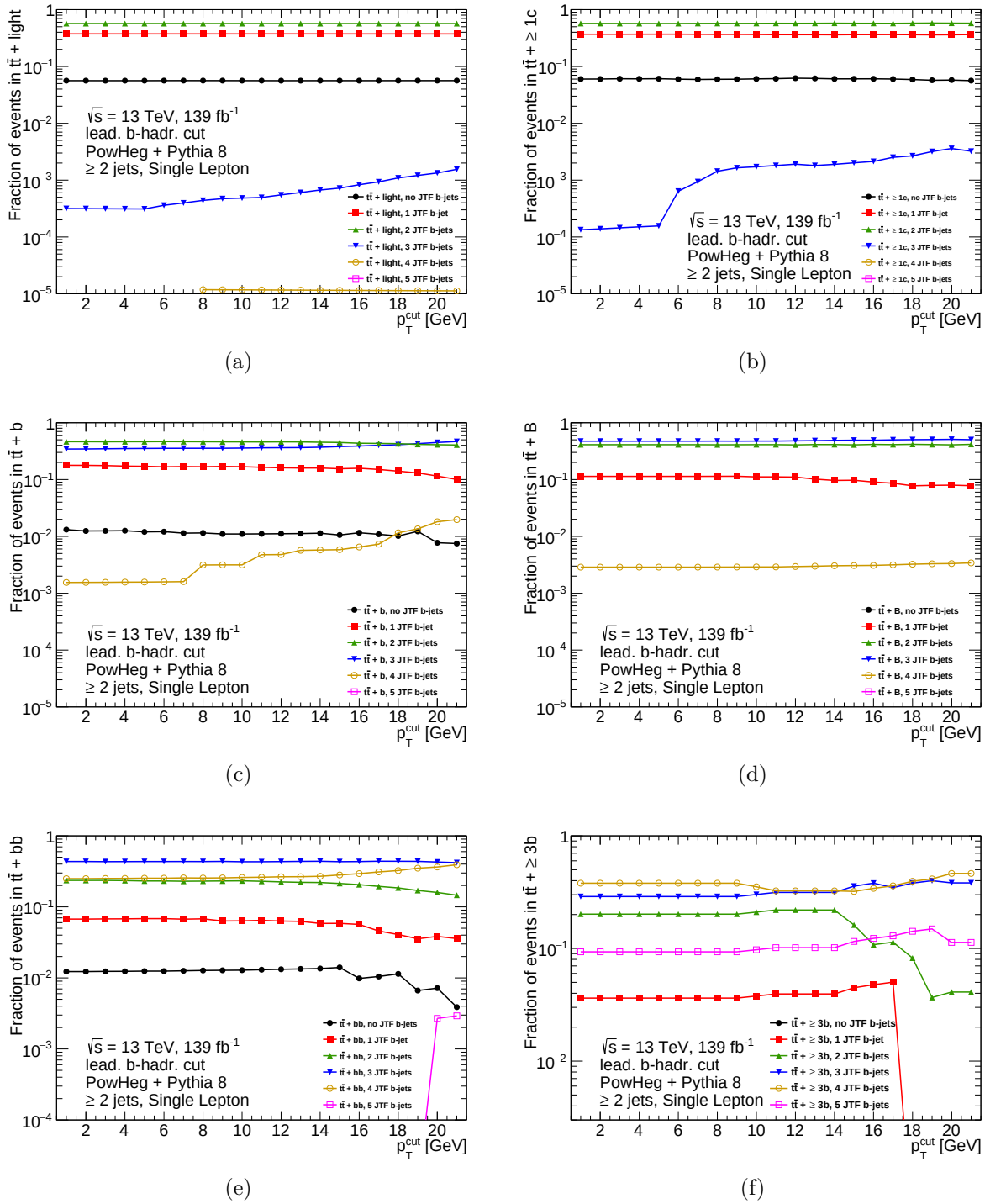


Figure C.2: Influence of cut on the p_T of leading b-hadrons on the fractional composition of each HFC category with regard to the different JTF b-jet multiplicities. All other parameters of the HFC are at their nominal values.

D Tables comparing HFC and JTF for different parameters of the HFC

Table D.1: Event numbers (a) and uncertainties (b) for the small PowHeg + Pythia 8 sample with HFC parameter settings 15 GeV for the p_T cut on jets and 5 GeV for the leading b-hadron p_T cut. The HFC category that is expected from a certain JTF b-jet multiplicity is written in boldface.

(a)

HFC	Number of JTF b-tagged jets						
	0	1	2	3	4	≥ 5	total
$t\bar{t} + light$	900789	7754070	**12683300**	6936	0	0	21345095
$t\bar{t} + \geq 1c$	101224	725993	**1135800**	312	0	0	1963329
$t\bar{t} + b$	3372	58206	165339	**125898**	562	0	353378
$t\bar{t} + B$	0	9430	35129	**40519**	246	0	85324
$t\bar{t} + bb$	1148	9001	31328	58807	**34295**	0	134582
$t\bar{t} + \geq 3b$	0	0	1511	2166	2846	**699**	7223
total	1006534	8556701	14052408	234639	37950	699	23890333

(b)

HFC	Number of JTF b-tagged jets					
	0	1	2	3	4	≥ 5
$t\bar{t} + light$	16035	46920	**59776**	1410	0	0
$t\bar{t} + \geq 1c$	5375	14410	**17937**	312	0	0
$t\bar{t} + b$	1015	4083	6858	**5915**	407	0
$t\bar{t} + B$	0	1635	3178	**3362**	246	0
$t\bar{t} + bb$	575	1600	2985	4121	**3074**	0
$t\bar{t} + \geq 3b$	0	0	634	766	907	**404**

Table D.2: Event numbers (a) and uncertainties (b) for the small PowHeg + Pythia 8 sample with HFC parameter settings 20 GeV for the p_T cut on jets and 5 GeV for the leading b-hadron p_T cut. The HFC category that is expected from a certain JTF b-jet multiplicity is written in boldface.

(a)

HFC	Number of JTF b-tagged jets						
	0	1	2	3	4	≥ 5	total
$t\bar{t} + light$	912867	7885510	**12890200**	10647	0	0	21699224
$t\bar{t} + \geq 1c$	91803	625291	**982117**	312	0	0	1699523
$t\bar{t} + b$	1947	40879	136435	**136383**	873	0	316518
$t\bar{t} + B$	0	7586	33565	**40305**	246	0	81704
$t\bar{t} + bb$	286	4441	18058	45647	**34274**	0	102709
$t\bar{t} + \geq 3b$	0	0	615	1601	2555	**699**	5472
total	1006904	8563709	14060991	234897	37950	699	23906551

(b)

HFC	Number of JTF b-tagged jets					
	0	1	2	3	4	≥ 5
$t\bar{t} + light$	16139	47326	**60276**	1713	0	0
$t\bar{t} + \geq 1c$	5133	13341	**16651**	312	0	0
$t\bar{t} + b$	779	3436	6211	**6176**	513	0
$t\bar{t} + B$	0	1474	3096	**3355**	246	0
$t\bar{t} + bb$	286	1138	2267	3626	**3072**	0
$t\bar{t} + \geq 3b$	0	0	357	654	859	**404**

Table D.3: Event numbers (a) and uncertainties (b) for the small PowHeg + Pythia 8 sample with HFC parameter settings 5 GeV for the p_T cut on jets and 5 GeV for the leading b-hadron p_T cut. The HFC category that is expected from a certain JTF b-jet multiplicity is written in boldface.

(a)

HFC	Number of JTF b-tagged jets						
	0	1	2	3	4	≥ 5	total
$t\bar{t} + light$	836707	7234270	**11861300**	3406	0	0	19935683
$t\bar{t} + \geq 1c$	149791	1124950	**1758790**	0	0	0	3033531
$t\bar{t} + b$	10506	102024	211408	**86743**	0	0	410681
$t\bar{t} + B$	273	11137	39850	**40540**	246	0	92049
$t\bar{t} + bb$	3323	46375	120748	97574	**34571**	0	302591
$t\bar{t} + \geq 3b$	0	1241	4189	5902	3132	**699**	15165
total	1000600	8519998	13996286	234167	37950	699	23791102

(b)

HFC	Number of JTF b-tagged jets					
	0	1	2	3	4	≥ 5
$t\bar{t} + light$	15475	45319	**57817**	1008	0	0
$t\bar{t} + \geq 1c$	6499	17931	**22269**	0	0	0
$t\bar{t} + b$	1742	5391	7718	**4915**	0	0
$t\bar{t} + B$	273	1759	3373	**3362**	246	0
$t\bar{t} + bb$	963	3595	5892	5268	**3088**	0
$t\bar{t} + \geq 3b$	0	573	1066	1264	951	**404**

Table D.4: Event numbers (a) and uncertainties (b) for the small PowHeg + Pythia 8 sample with HFC parameter settings 15 GeV for the p_T cut on jets and 2 GeV for the leading b-hadron p_T cut. The HFC category that is expected from a certain JTF b-jet multiplicity is written in boldface.

(a)

HFC	Number of JTF b-tagged jets						
	0	1	2	3	4	≥ 5	total
$t\bar{t} + light$	887176	7660800	**12529100**	6936	0	0	21084012
$t\bar{t} + \geq 1c$	114042	814021	**1282780**	312	0	0	2211155
$t\bar{t} + b$	3597	61474	167679	**124953**	562	0	358265
$t\bar{t} + B$	0	9430	35129	**40519**	246	0	85324
$t\bar{t} + bb$	1148	9001	32404	59752	**34295**	0	136603
$t\bar{t} + \geq 3b$	0	0	1511	2166	2846	**699**	7223
total	1005964	8554726	14048603	234639	37950	699	23883984

(b)

HFC	Number of JTF b-tagged jets					
	0	1	2	3	4	≥ 5
$t\bar{t} + light$	15916	46633	**59410**	1410	0	0
$t\bar{t} + \geq 1c$	5699	15265	**19065**	312	0	0
$t\bar{t} + b$	1039	4193	6903	**5889**	407	0
$t\bar{t} + B$	0	1635	3178	**3362**	246	0
$t\bar{t} + bb$	575	1600	3035	4157	**3074**	0
$t\bar{t} + \geq 3b$	0	0	634	766	907	**404**

Table D.5: Event numbers (a) and uncertainties (b) for the small PowHeg + Pythia 8 sample with HFC parameter settings 15 GeV for the p_T cut on jets and 12 GeV for the leading b-hadron p_T cut. The HFC category that is expected from a certain JTF b-jet multiplicity is written in boldface.

(a)

HFC	Number of JTF b-tagged jets						
	0	1	2	3	4	≥ 5	total
$t\bar{t} + light$	921328	7953550	**12980300**	12551	265	0	21867995
$t\bar{t} + \geq 1c$	80866	536938	**859043**	2859	0	0	1479707
$t\bar{t} + b$	3190	52558	157772	**124611**	1641	0	339774
$t\bar{t} + B$	0	8937	34027	**40073**	246	0	83285
$t\bar{t} + bb$	1148	7810	28620	55317	**33825**	0	126722
$t\bar{t} + \geq 3b$	0	0	1511	2166	2236	**699**	6613
total	1006534	8559795	14061274	237578	38216	699	23905498

(b)

HFC	Number of JTF b-tagged jets					
	0	1	2	3	4	≥ 5
$t\bar{t} + light$	16210	47528	**60484**	1878	265	0
$t\bar{t} + \geq 1c$	4825	12376	**15566**	1024	0	0
$t\bar{t} + b$	998	3888	6703	**5885**	681	0
$t\bar{t} + B$	0	1597	3129	**3346**	246	0
$t\bar{t} + bb$	575	1483	2857	4005	**3056**	0
$t\bar{t} + \geq 3b$	0	0	634	766	798	**404**

Bibliography

[1] S. Greenblatt, *The Swerve - How the Renaissance began*, Vintage Books, London (2012)

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

[3] ATLAS Collaboration, *Search for the standard model Higgs boson produced in association with top quarks and decaying into a $b\bar{b}$ pair in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **97**, 072016 (2018)

[4] ATLAS Collaboration, *Measurements of inclusive and differential fiducial cross-sections of $t\bar{t}$ production with additional heavy-flavour jets in proton-proton collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, JHEP **04**, 046 (2019)

[5] Gargamelle Collaboration, *Search for elastic muon-neutrino electron scattering*, Phys. Lett. B **46(1)**, 121 (1973)

[6] CDF Collaboration, *Observation of top quark production in $p\bar{p}$ collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74(14)**, 2626 (1995)

[7] DØ Collaboration, *Observation of the top quark*, Phys. Rev. Lett. **74(14)**, 2632 (1995)

[8] ATLAS Collaboration, *Combined search for the Standard Model Higgs boson using up to 4.9fb^{-1} of pp collision data at $\sqrt{s} = 7$ TeV with the ATLAS detector at the LHC*, Phys. Lett. B **710(1)**, 49 (2012)

[9] CMS Collaboration, *Combined results of searches for the standard model Higgs boson in pp collisions at $\sqrt{s} = 7$ TeV*, Phys. Lett. B **710(1)**, 26 (2012)

[10] G. 't Hooft, *Renormalizable Lagrangians for Massive Yang-Mills Fields*, Nucl. Phys. B **35**, 167 (1971)

[11] G. 't Hooft, M. J. G. Veltman, *Regularization and Renormalization of Gauge Fields*, Nucl. Phys. B **44**, 189 (1972)

[12] G. 't Hooft, M. Veltman, *Combinatorics of gauge fields*, Nucl. Phys. B **50(1)**, 318 (1972)

[13] F. Englert, R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13(9)**, 321 (1964)

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

[15] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. **145**, 1156 (1966)

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

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

[18] S. L. Glashow, *The renormalizability of vector meson interactions*, Nucl. Phys. **10**, 107 (1959)

[19] S. L. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. **22(4)**, 579 (1961)

[20] A. Salam, J. C. Ward, *Electromagnetic and weak interactions*, Phys. Lett. **13(2)**, 168 (1964)

[21] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19**, 1264 (1967)

[22] Particle Data Group, *Review of Particle Physics*, Prog. Theor. Exp. Phys. **2020**, 083C01 (2020)

[23] D. J. Gross, F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, Phys. Rev. Lett. **30**, 1343 (1973)

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

[25] E. Noether, *Invariante Variationsprobleme*, Kgl. Ges. d. Wiss. Nachrichten. Math.-phys. Klasse **1918(2)**, 235 (1918)

[26] N. Cabibbo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10(12)**, 531 (1963)

[27] M. Kobayashi, T. Maskawa, *CP-Violation in the Renormalizable Theory of Weak Interaction*, Progr. Theor. Phys. **49(2)**, 652 (1973)

[28] L. Wolfenstein, *Parametrization of the Kobayashi-Maskawa Matrix*, Phys. Rev. Lett. **51**, 1945 (1983)

[29] ATLAS Collaboration, *Search for the decay $B_s^0 \to \mu\mu$ with the ATLAS detector*, Phys. Lett. B **713**, 387 (2012)

[30] UA1 Collaboration (UA1), *Experimental Observation of Isolated Large Transverse Energy Electrons with Associated Missing Energy at $\sqrt{s} = 540$ GeV*, Phys. Lett. B **122**, 103 (1983)

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

[32] C. Englert, et al., *Precision measurements of Higgs couplings: implications for new physics scales*, J. Phys. G **41(11)**, 113001 (2014)

[33] J. N. Ng, P. Zakarauskas, *QCD-parton calculation of conjoined production of Higgs bosons and heavy flavors in $p\bar{p}$ collisions*, Phys. Rev. D **29**, 876 (1984)

[34] Z. Kunszt, *Associated production of heavy Higgs boson with top quarks*, Nucl. Phys. B **247(2)**, 339 (1984)

[35] W. Beenakker, et al., *Higgs Radiation Off Top Quarks at the Tevatron and the LHC*, Phys. Rev. Lett. **87**, 201805 (2001)

[36] S. Dawson, L. H. Orr, L. Reina, D. Wackeroth, *Next-to-leading order QCD corrections to $pp \to t\bar{t}h$ at the CERN Large Hadron Collider*, Phys. Rev. D **67**, 071503 (2003)

[37] CMS Collaboration, *Search for the associated production of the Higgs boson with a top-quark pair*, JHEP **09**, 087 (2014), [Erratum: JHEP 10, 106 (2014)]

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

[39] Planck Collaboration, *Planck 2018 results - VI. Cosmological parameters*, Astron. Astrophys. **641**, A6 (2020)

[40] L. Evans, P. Bryant, *LHC Machine*, JINST. **3(08)**, S08001 (2008)

[41] F. Pastore, *ATLAS Run-2 status and performance*, Nucl. Part. Phys. Proc. **270-272**, 3 (2016), 18th Montpellier International Conference on Quantum Chromodynamics (QCD 15)

[42] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST. **3(08)**, S08003 (2008)

[43] ATLAS Collaboration, *Alignment of the ATLAS Inner Detector in Run 2*, Eur. Phy. J. C **80**, 1194 (2020)

[44] ATLAS Collaboration, *Operation of the ATLAS trigger system in Run 2*, JINST. **15(10)**, P10004 (2020)

[45] 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*, World Scientific (2016)

[46] 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**, 641 (1977)

[47] V. N. Gribov, L. N. Lipatov, *e^+e^- pair annihilation and deep inelastic ep scattering in perturbation theory*, Sov. J. Nucl. Phys. **15**, 675 (1972)

[48] G. Altarelli, G. Parisi, *Asymptotic freedom in parton language*, Nucl. Phys. B **126(2)**, 298 (1977)

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

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

[51] T. Sjöstrand, et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191**, 159 (2015)

[52] M. Bahr, et al., *Herwig++ Physics and Manual*, Eur. Phys. J. C **58**, 639 (2008)

[53] J. Bellm, et al., *Herwig 7.0 / Herwig++ 3.0 Release Note*, Eur. Phys. J. C **76**, 196 (2016)

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

[55] ATLAS Collaboration (ATLAS), *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70**, 823 (2010)

[56] S. Agostinelli, et al. (GEANT4), *GEANT4–a simulation toolkit*, Nucl. Instrum. Meth. A **506**, 250 (2003)

[57] J. Allison, et al., *Geant4 developments and applications*, IEEE Trans. Nucl. Sci. **53**, 270 (2006)

[58] J. Allison, et al., *Recent developments in Geant4*, Nucl. Instrum. Meth. A **835**, 186 (2016)

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

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

[61] ATLAS Collaboration, *Search for the Standard Model Higgs boson produced in association with top quarks and decaying into $b\bar{b}$ in pp collisions at $\sqrt{s} = 13\,\text{TeV}$ with the ATLAS detector* (2018), ATL-COM-PHYS-2017-079

[62] ATLAS Collaboration, *Search for $t\bar{t}t\bar{t}$ Standard Model Production in the Single Lepton and Opposite Sign Dilepton Final State in proton-proton collisions at $\sqrt{s} = 13\,\text{TeV}$ using full Run 2 data set with the ATLAS Detector* (2021), ANA-TOPQ-2020-10-INT1

Acknowledgements

There are many people I have to thank for making this bachelor thesis possible. First of all, my thanks go to Professor Anulf Quadt for letting me write my thesis in his working group and for the good work ethic in his group. The strict adherence to deadlines and the demand for a bit of progress every weak meant hard work over the whole 14-week period. In the end I am, however, thankful that I knew my submission date right from the beginning and that I have a free summer ahead of me. Additionally, I want to thank him and Professor Ariane Frey for agreeing to referee my thesis.

Furthermore I want to thank Doctor Jelena Jovicevic and Doctor Elizaveta Shabalina for their comments on my work and the great patience they had with me, sometimes explaining the same thing many times again until I finally understood. They were always in reach for questions and I got quick and comprehensive answers every time. I am also very grateful to the comments Liza made on my thesis drafts, highlighting the logical flaws and missing pieces in my arguments.

Then there is Stephen Eggebbrecht. A lot of things would have cost me a lot more time if it would not have been for him. He was the go-to person for all technical problems, answering within minutes almost all the time. I also want to thank him for taking a lot of time, often giving long explanations to the more comprehensive questions I asked him, especially on the workings of event generators. He was also the one who produced the ntuples for me, which took a lot of his time especially before the HFC was available in the analysis code.

I also want to thank all those who supported me during the last two weeks of intensive analysis and writing, especially my little armada of proof readers, among them Liza and Stephen, but also Chrissi, Leo, Joel and Richard. Many thanks to all of them for their helpful comments on the thesis.

Finally, I want to thank my parents for always keeping up my spirit and sending encouragement, be it verbal or in edible form, especially during those intensive last weeks of the semester.

Erklärung nach §13(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 4. September 2021

(Titus Stanislaus Czajka)