GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Fakultät für
Physik

Master's Thesis

Messung des Wirkungsquerschnittes der Top-Antitop Paar Produktion in Kombination mit einem Charm-Anticharm Paar im dileptonischem Endzustand

Measurement of the associated production of top-antitop pair with charm jets using the DL1r b-tagging algorithm in the dileptonic final state

prepared by

Moritz Habbaba

from Ehringshausen

at the II. Physikalisches Institut

Thesis number:	II.Physik-UniGö-MSc-2022/02
Period:	1st October 2021 until 1st April 2022
First referee:	Prof. Dr. Arnulf Quadt
Second referee:	Prof. Dr. Stan Lai

Abstract

This thesis attempts to find a signal region for the measurement of $t\bar{t}+ \geq 1c$ production in the dileptonic channel using the b-tagging algorithm "DL1r" and evaluate the expected uncertainty of such measurement. The full proton-proton collision data at $\sqrt{s} = 13$ TeV recorded with the ATLAS detector at the LHC of run II (2015-2018) is used, which corresponds to an integrated luminosity $L_{int} = 139\,\text{fb}^{-1}$. Events are required to have exactly two leptons, at least three jets as well as at least three jets at 77% DL1r working point (WP) and at most two jets at 70% DL1r WP. As such, selected events are split into two signal regions, one requiring exactly three jets and the other requiring at least four jets. The selected variable is the pseudo-continuous DL1r discriminant of the third jet by p_T. A profile likelihood is used to investigate an interim expectation of the uncertainty of the signal strength with a limited number of systematics, as well as investigate the impact of the individual systematics on the uncertainty of the signal strength. In an Asimov fit, the expected signal strength uncertainty was found to be 41%, with the highest impact on the uncertainty coming from the uncertainty of the main backgrounds.

Contents

1. Introduction

A question humans have asked themselves for millennia is the nature of the inner workings of the world. Over the last century, a limited answer to this question was found in a number of elementary particles and their interactions. A number of theories were developed to explain and predict the behaviour of these elementary particles and their interactions, which resulted in the so-called Standard Model (SM) of particle physics. It combines the theory of the electro-weak interaction [1–6] with quantum chromodynamics (QCD)[7, 8], which describes the strong interaction. The predictions of the Standard Model could not be falsified time and time again by experiments such as the Large Hadron Collider (LHC) at CERN (Conseil européen pour la recherche nucléaire) in Geneva. One of the most prominent predictions of the Standard Model was the existence of the so-called Higgs boson [9–11], which was discovered at the ATLAS experiment at CERN in 2012 [12].

One particle described in the Standard Model of special interest is the top quark. It was discovered in 1995 by the DØ [13] and CDF [14] experiment at the Tevatron and has the highest mass among the particles described in the Standard Model. Due to its high mass, the top quark typically decays, before it can form bound states. This behaviour differentiates the top quark from all other quark flavours.

The top pair production with an additional $c\bar{c}$-pair can be used as a test for the gluon coupling of the top quark. Additionally, it can serve as a way to compare the coupling of the top quark to the charm quark with the coupling to the b-quark, since the theoretical modelling of the top pair production with an additional $c\bar{c}$-pair is similar to the top pair production with an additional $b\bar{b}$-pair. Additionally, the top pair production with an additional $c\bar{c}$-pair is an important background in several existing analysis, such as $t\bar{t}H(bb)$, $t\bar{t}t\bar{t}$, $H^+(tb)$ and can be seen in figure 1.1.

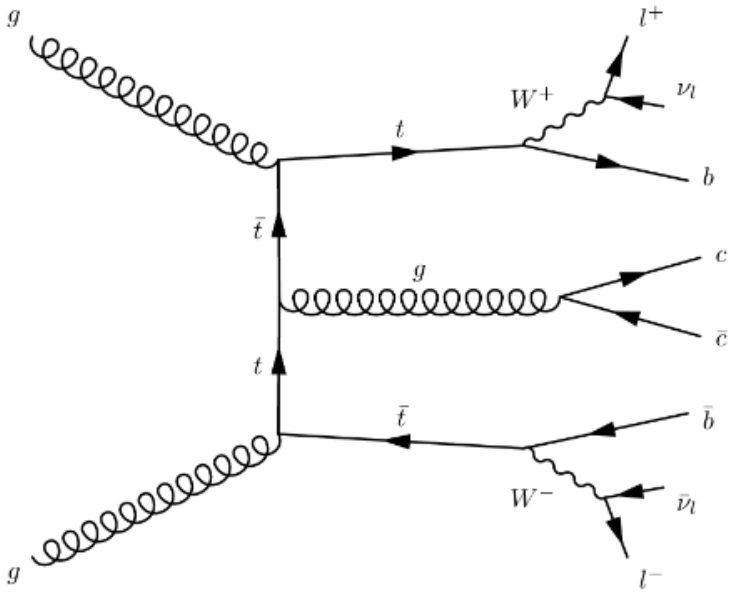

Figure 1.1.: Leading order Feynmann diagram of the production of a top-antitop pair in combination with a charm-anticharm pair.

The measurement of the top pair production with an additional $c\bar{c}$-pair meets theoretical and experimental difficulties. The theoretical difficulties are large uncertainties resulting from $t\bar{t}$+jets modelling problems due to the choice of factorisation and re-normalisation scales, since the high mass of the charm quark introduces an additional fixed energy scale. The experimental difficulties are in disentangling the $t\bar{t}$+$c\bar{c}$ events from the other $t\bar{t}$+jets events, especially from the large $t\bar{t}$+lights contribution.

While the top pair production with an additional $b\bar{b}$-pair has been measured by ATLAS and CMS at multiple occasions [15–21], the cross section for top pair production with an additional $c\bar{c}$-pair has only been measured once by CMS [22]. This can be explained by the challenge in distinguishing the additional charm jets from the large contribution of light flavour jets to the $t\bar{t}$+jets cross section. This distinction is easier for $t\bar{t}$+$b\bar{b}$ events due to the availability of b-tagging tools.

b-tagging algorithms have a long and stable history within the analysis tools of particle physics. Dedicated tagging algorithms for charm quarks on the other hand are not widely used as of this writing. This work investigates the possibilities of repurposing existing b-tagging tools for c-tagging. As such, the expected signal purity, that can be achieved by only using the DL1r b-tagging working points, is investigated.

This thesis presents first steps, that are taken with the goal to measure the production cross-section of the top pair production with an additional $c\bar{c}$-pair in the dileptonic final state at $\sqrt{s} = 13\,\text{TeV}$ using the full run II dataset of the ATLAS experiment. Using the DL1r b-tagging algorithm working points a signal region is determined. The precision measurement is limited due to a number of systematic uncertainties, which are taken into account using a profile likelihood fit.

This thesis first offers a limited overview of the theoretical background of the Standard Model in chapter 2, followed by an overview in chapter 3 about the experimental setup used to gather the dataset used in this analysis. General definitions and event selections are explained in chapter 4, while chapter 5 explains the b-tagging algorithm "DL1r". The samples used in the analysis are depicted in chapter 6 and chapter 7 depicts the strategy used in this analysis. Systematic uncertainties are investigated in chapter 8 and a few preliminary results are depicted in chapter 9. Finally chapter 10 gives an outlook over ways to continue this analysis.

2. The Standard Model

The Standard Model (SM) of particle physics [1–8, 23–27] was formulated in the 1970s and to this day produces predictions with remarkable accuracy. The beginnings of the Standard Model lie in separate theories describing the electromagnetic, weak and strong nuclear forces, from which a unified quantum field theory based on a local $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge symmetry was developed. This chapter will give a general overview over the most relevant parts of the Standard Model[1].

2.1. Elementary Particles

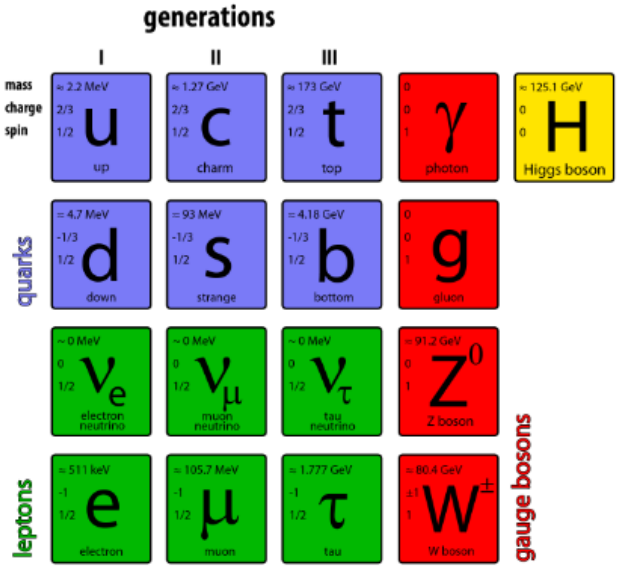

Figure 2.1.: Elementary particles in the SM. The mass, charge and spin values are taken from [29].

The SM contains 18 particles (and their corresponding anti-particles), which can be divided into fermions (half-integer spin) and bosons (integer spin). These 18 particles with their corresponding masses, charges and spins can be seen in figure 2.1.

[1]This chapter is based in part on [28]

Three of the four known fundamental forces can be explained using spin-1 gauge bosons as mediators of the respective forces. The electromagnetic force uses the massless photon γ as a mediator, which couples to electric charge. The massive $W^{\pm}$ and Z bosons are the mediators for the electroweak interaction and the massless gluons mediate the strong interaction, while coupling to colour charge. The fourth known fundamental interaction, gravity, is not included in the Standard Model, but can be neglected due to the low mass scale present in particle physics.

In addition, the SM also includes the Higgs boson, which corresponds to the Higgs mechanism [9–11] and is responsible for the masses of the other elementary particles via electroweak symmetry breaking [1] and the Yukawa coupling [2].

The fermions are subdivided into quarks and leptons, which, in contrast to the bosons, appear in three different generations, which only differ by their masses. Quarks differ from leptons by having colour charge and having a fractional electric charge. Leptons on the other hand do not interact via the strong force, are colour neutral and have integer charge (0 or 1). The higher generations of particles are unstable and decay into lower generations of particles. The particles in the lowest generation are assumed to be stable.

2.2. Local Gauge Invariance

The dynamics of free fermions can be described via the relativistic Dirac equation

$$i\gamma^{\mu}\partial_{\mu}\psi - m\psi = 0, \tag{2.1}$$

where ψ is a four-component spinor, describing the fermion, and γ^{μ} are the four-dimensional gamma matrices ($\mu = 0, 1, 2, 3$). Using the Euler-Lagrange equation the relativistic Dirac equation can be derived from the Lagrange density

$$\mathcal{L} = i\overline{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\overline{\psi}\psi, \tag{2.2}$$

where $\overline{\psi} = \psi^{\dagger}\gamma^{0}$ is the adjutant spinor. This Lagrange density consists of a momentum term and a mass term. The Standard Model uses quantum field theories to describe the interactions between the elementary particles. It also requires, that these quantum fields

are locally gauge invariant. This can be demonstrated with the example of quantum electrodynamics (QED), which describes the electromagnetic interaction. QED follows a U(1) local gauge symmetry, which means spinors following this symmetry transform via $\psi \mapsto \psi' = e^{-iq\lambda(x)}$. If the Lagrangian remains unchanged during such a transformation, the resulting theory is called locally gauge invariant under a U(1) transformation. The Lagrangian in equation 2.2 is not invariant under such transformation. To achieve that, an additional massless quantum vector field A_μ has to be introduced. This results in the Lagrangian

$$\mathcal{L}_{QED} = i\overline{\psi}\gamma^\mu D_\mu \psi - m\overline{\psi}\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \tag{2.3}$$

where $D_\mu = \partial_\mu + iqA_\mu$ is the covariant derivative and $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ describes the field strength tensor of A_μ. This is the Lagrangian, that describes the electromagnetic interaction between particles with charge q and massless photons A_μ. It can be directly seen, that the insistence on local gauge invariance results in the inclusion of a (massless) gauge boson. Including a massterm for the gauge field A_μ results in the Lagrangian no longer being invariant under local gauge transformation, which means, that the photon is required to be massless. Since the U(1) symmetry group is an Abelian symmetry group, there is no self-interaction between gauge bosons. This is not the case for theories using non-Abelian symmetries such as Quantum Chromodynamics (QCD), which uses an SU(3) symmetry. The resulting gauge bosons (the gluons) can therefore interact with each other.

2.3. Electroweak Unification

The electromagnetic and the weak interaction can be unified using a $\mathrm{SU(2)_L} \times \mathrm{U(1)_Y}$ gauge theory. This model was first described by Glashow, Weinberg and Salam and is therefore known as the Glashow-Weinberg-Salam (GWS) model [1–6]. Since W bosons only interact with left-handed particles (and right-handed anti-particles) the $\mathrm{SU(2)_L}$ term is included, where the L refers to left-handed particles. This results in parity violation of the electroweak interaction. The description of the electroweak interaction is done compactly, by arranging fermions in so-called left-handed weak isospin doublets χ_L. For this a new quantity called weak isospin $\vec{T}$ is introduced:

$$\begin{matrix} T_3 = +\frac{1}{2} \\ T_3 = -\frac{1}{2} \end{matrix} \quad \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L, \begin{pmatrix} u \\ d' \end{pmatrix}_L, \begin{pmatrix} c \\ s' \end{pmatrix}_L, \begin{pmatrix} t \\ b' \end{pmatrix}_L$$

This leaves the right-handed particles (and left-handed anti-particles), which are arranged into right-handed singlets χ_R. Since they do not interact with the W boson, there corresponding weak isospin values are zero. The downtype particles denote the weak eigenstates of the corresponding elementary particles and not their mass eigenstates and are as such denoted with a prime. The weak eigenstates are connected to the mass eigenstates via the CKM matrix, as shown in equation 2.4 [30, 31].

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & v_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.4}$$

As can be seen the downtype weak eigenstates consist of a mixing of the downtype mass eigenstates, which results in possible couplings between particle generations via the charged weak interaction.

Just as QED required local gauge symmetry, the GWS theory also requires invariance under a local gauge transformation. The difference is that GWS theory requires the transformation to be invariant under an $\mathrm{SU(2)_L} \times \mathrm{U(1)_Y}$ symmetry instead of a U(1) symmetry. Following the principles from section 2.2, this requires the introduction of four additional gauge fields, which correspond to the four generators of the symmetry groups. The $\mathrm{SU(2)_L}$ part of the symmetry requires the existence of the three gauge fields W_μ^i, that couple to the weak isospin T_i, while the $\mathrm{U(1)_Y}$ part of the symmetry requires an additional gauge field B_μ, that couples to the hypercharge $Y = 2(q - T_3)$. These four gauge bosons are the weak isospin eigenstates of the physical bosons measured in a detector, which correspond to the mass eigenstates of these gauge fields and result from a mixture of them. The $W^\pm$-boson can be described via $W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2)$. The remaining fields B_μ and W_μ^3 mix together via the Higgs mechanism, resulting in

$$A_\mu = B_\mu \cos\theta_W + W_\mu^3 \sin\theta_W, \tag{2.5}$$

$$Z_\mu = B_\mu \sin\theta_W + W_\mu^3 \cos\theta_W. \tag{2.6}$$

Here A_μ refers to the massless photon and Z_μ to the Z boson, while θ_W denotes the Weinberg angle.

2.4. Quantum Chromodynamics

In the same way, that the electroweak interaction is described by a gauge theory, the strong interaction can also be described via a gauge theory. In the case of the strong interaction this gauge theory is called quantum chromodynamics (QCD) and follows an $\mathrm{SU(3)_C}$ gauge theory, where the C refers to the colour charge. The three-dimensional special unitary group (SU(3)) has eight generators. As such, to ensure local gauge symmetry, eight gauge fields G^a_μ ($a = 1, 2, \ldots, 8$) have to be introduced, where each gauge fields corresponds to one of the eight generators in a process equivalent to the one described in section 2.2. In a similar fashion, these eight gauge fields correspond to eight different physical particles called gluons. By introducing a kinetic term for the gluons and by replacing the space derivative with the covariant derivative $D_\mu = \partial_\mu + \frac{ig_s}{2}\lambda_a G^a_\mu$, where g_s is the strong coupling constant and λ_a are the Gell-Mann matrices, the Lagrangian corresponding to the strong interaction can be derived:

$$\mathcal{L}_{QCD} = \sum_f \overline{q}_f (i\gamma^\mu D_\mu - m) q_f - \frac{1}{4} G^a_{\mu\nu} G^{\mu\nu,a}. \tag{2.7}$$

Where q_f denotes the spinor of a quark and $G^a_{\mu\nu} = \partial_\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f_{abc} G^b_\mu G^c_\nu$ are the structure constants of the $\mathrm{SU(3)_C}$ gauge symmetry, describing the non-commutativity of the generators. The index f runs over all quark flavours.

In the process of renormalisation, the strong coupling constant $\alpha = \frac{g_s}{4\pi}$ becomes dependent on the energy scale of the relevant process. Due to the non-commutativity of the generators, self-interactions of the gluons arise. This results in the dependency of the strong coupling constant on the energy scale Q^2 shown in equation 2.8 [32].

$$\alpha_s(Q^2) = \frac{\alpha_s(\mu^2)}{1 + \alpha_s(\mu^2)\frac{11N_c - 2N_f}{12\pi}} \ln \frac{Q^2}{\mu^2} \tag{2.8}$$

Here $N_{c/f}$ is the number of colours and the number of flavours respectively, while μ denotes an arbitrary energy scale, at which α_s is known. For $N_c = 3$ and $N_f = 6$ the coupling constant increases with increased particle distance. As such particles carrying colour charge can not exist freely in nature. In QCD this is known as colour confinement. If two particles carrying colour charge are separated, the energy in the connecting quantum field increases. This continues until the energy is sufficiently high to produce an additional

quark-antiquark pair. This process continues until a bunch of quarks remain in colourless bound states, formed in a process called hadronisation. As such, whenever particles carrying colour charge are produced, they can not be observed as free particles. Instead they are seen as a number of hadrons, the collection of which is called a jet.

2.5. The Top Quark

Discovered in 1995 by the DØ [13] and CDF [14] experiment at the Tevatron, the top quark is the particle with the highest mass in the standard model. In the Standard Model, the top quark is the up-type weak hyperspin partner of the b-quark with the corresponding spin and charge properties. Due to its large mass, the top quark has a stronger coupling to the higgs sector and a significantly smaller lifetime than other particles. Due to this short lifetime, the top quark decays faster then the time scale of hadronisation.

2.5.1. Production

The production of top quarks happens mainly at the LHC using proton-proton-collisions. Protons are not elementary particles. In a simple static quark model, protons consist of three quarks (uud). The bound state resulting of the combination of these three quarks is held together by the strong nuclear force mediated by the massless gluons. These gluons can split into virtual quark-antiquark pairs. As such, an interaction between two protons has to be seen as an interaction between the constituents of the protons called the partons. Following this picture, top pairs can be produced either via quark-antiquark annihilation or gluon-gluon fusion. In the first production channel, two quarks annihilate into a gluon, which then splits into a top-antitop pair. In the case of gg fusion, a top pair can be emitted either via a singular fused gluon, or via the propagation of a virtual top quark. The leading order Feynman diagrams are shown in figure 2.2.

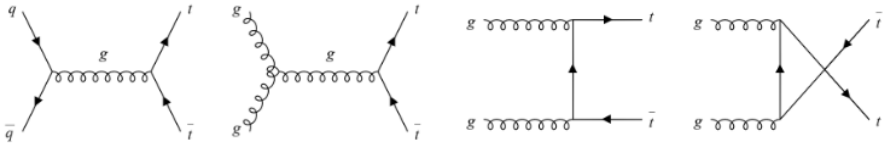

Figure 2.2.: The four leading Feynman diagrams of the top-antitop pairs in pp-collisions.

2.5.2. Decay

The top quark decays weakly via the emission of a W boson. In this interaction the top quark mainly decays to a b-quark and a W-boson. The cross sections of the other decay modes are so small, that they can be ignored for this analysis. Following the initial decay of the top quark, the W boson either decays into a charged lepton-neutrino pair, or a quark-antiquark pair. As such the signature of a single top quark in isolation is either a single jet, which can be tagged as a b-jet, accompanied by an isolated lepton and missing transverse energy, or three jets, of which one can be identified as originating from a b-quark.

$t\bar{t}$ pair signature

The experimental signature of a top pair production event can be inferred by the combination of these possible decay modes of singular isolated top quarks. This results in three possible decay modes for top pairs: The hadronic decay, in which both W bosons decay via quark jets, the semi-leptonic decay, in which one W boson decays hadronically, while the other decays leptonically and the dileptonic decay, in which both W bosons decay leptonically. This analysis will focus on the dileptonic decay mode, the schematic of which is shown in figure 2.3.

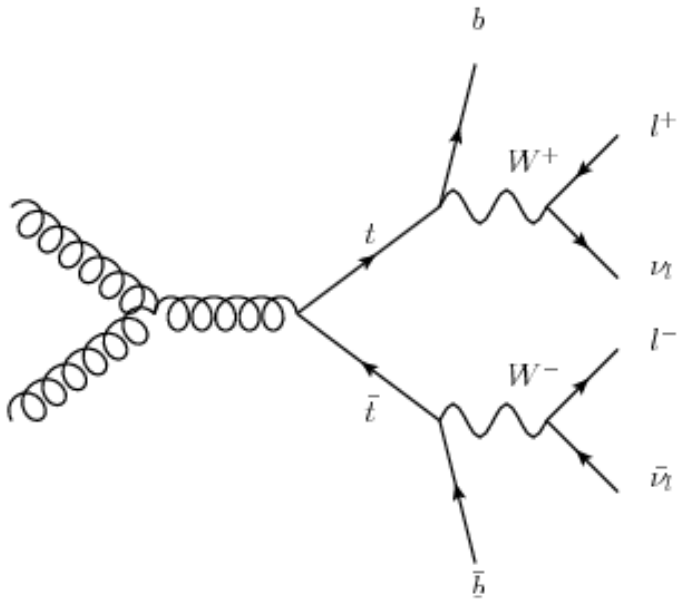

Figure 2.3.: Schematic of a dileptonic top pair production event.[2]

The process analysed in this work is $t\bar{t}+ \geq 1c$, which contains a top-quark pair and a charm-quark pair resulting mostly from gluon splitting. The event signature contains as

such at least one additional jet resulting from a charm quark.

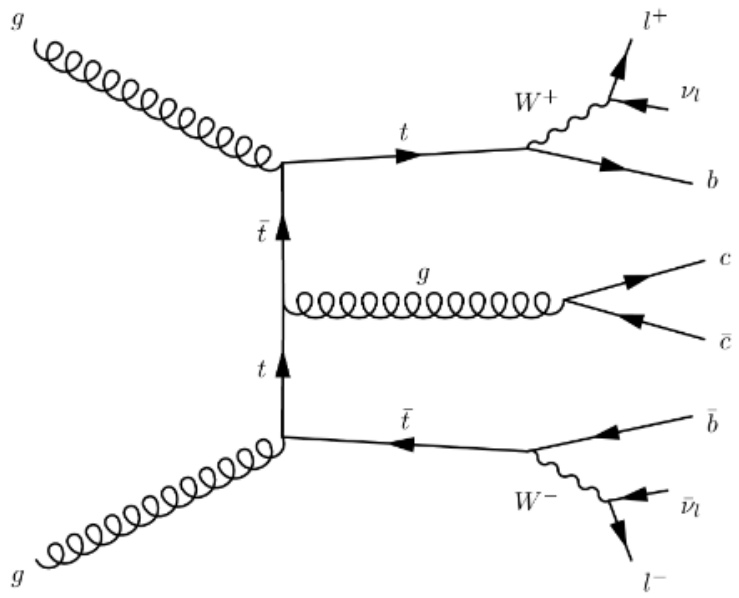

Figure 2.4.: Leading order Feynmann diagram of the production of a top-antitop pair in combination with a charm-anticharm pair.

2.6. The Charm Quark

Discovered in 1974 [33] the charm quark corresponds to the second generation up-type quark in the Standard Model. It is the weak hyperspin partner to the down-type strange quark with all the corresponding charge and spin properties. Due to its higher mass in comparison to up, down, and strange quark, it has a median free travel distance before hadronisation, that is higher than the light flavour quarks, and lower when compared to bottom quarks.

2.6.1. Production

The relevant production channel for this analysis is the production in a QCD process in combination with the production of a top-antitop pair. This specific process can be seen in figure 2.4.

Background

The signature of a $t\bar{t}+ \geq 1c$ event is not unique. The main background processes are other $t\bar{t}$+jet events, which can be categorised into $t\bar{t}+ \geq 1LF$ (Light Flavour) and $t\bar{t}+ \geq 1b$

events. Other background events are $t\bar{t}V$, where V describes a vector boson (either W or Z), $t\bar{t}H$, tW, Z+jets, W+jets and single top events. The modelling of the signal and the background as well as the data samples used is described in chapter 6.

Previous Cross-section measurements

The inclusive cross section of the associated production of a top-antitop pair with an additional $c\bar{c}$-pair was measured by CMS [22]. It used the dileptonic final states of $t\bar{t}$ events produced in proton-proton collisions at a centre-of-mass energy of 13 TeV. The data used in the analysis corresponds to a integrated luminosity of $41.5\,\mathrm{fb}^{-1}$. The distinction between the different flavours of $t\bar{t}$ + jets events is made via a neural network, that uses the output of a new charm jet identification algorithm.

3. Experimental Setup

The following chapter gives an overview over the experimental setup used[1]. The Large Hadron Collider (LHC) will be described in section 3.1, while the ATLAS detector located at the LHC, which was used to collect the data used in this analysis, will be described in section 3.2.

3.1. The Large Hadron Collider

The Large Hadron Collider (LHC) [34] is a superconducting particle accelerator located at CERN in Geneva. A general overview over the CERN accelerator complex is given in figure 3.1. Inside the LHC protons, are accelerated to high energies, before being brought to collision.

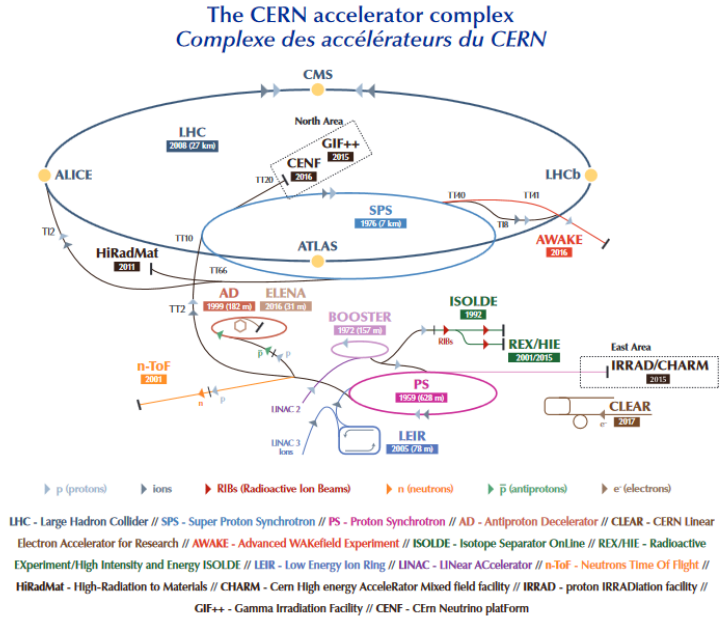

Figure 3.1.: Overview of the LHC and experiments. Protons pass several pre-accelerators before being injected into the main accelerator. © Cern.

[1]This chapter is based in part on [28].

The main part of the LHC consists of a synchotron with a circumference of 27km in a depth of approximately 100m. The required tunnel for this synchrotron was repurposed from the electron-positron collider LEP, that used it previously. Inside the tunnel two rings exist, that allow the acceleration of two proton beams in opposite direction. The protons pass a number of pre-accelerators before being injected into this main accelerator, as can be seen in figure 3.1. Protons are first accelerated to an energy of ca. 50 MeV using a linear accelerator (Linac2), during which the constant proton beam is split up into bunches. After the linear accelerator the protons are injected into the Proton Synchrotron Booster (PSB), that increases their energy to ca. 1.4 GeV. The next step in the pre-acceleration is the injection into the Proton Synchrotron (PS), which accelerates the proton bunches to an energy of 25 GeV and injects them into the Super Proton Synchrotron (SPS). After the protons are accelerated to an energy of ca. 450 GeV, they are finally injected into the LHC, where they reach their desired energy. During Run I of the LHC, this resulted in a center-of-mass energy of $\sqrt{s} = 7$ TeV (2011) and of $\sqrt{s} = 8$ TeV (2012). During Run II (2015-2018) the center-of-mass energy was increased to $\sqrt{s} = 13$ TeV.

The proton bunches inside the beampipes are guided along the pipe and focused using superconductive magnets, which are cooled using liquid helium at a temperature of $-271.3\,^{\circ}\mathrm{C}$. Guidance along the pipe path is done using 1232 dipole magnets, that are each 15 m long and keep the proton bunches on a circular orbit, while the focusing of the proton bunches is done using 392 quadrupol magnets, that are each 5-7 m long.

There are four main collision points. These correspond to the positions of the four main experiments at the LHC: ALICE, ATLAS, CMS and LHCb. The relative positions of which can be seen in figure 3.1.

3.2. The ATLAS detector

With a length of 44 m, a diameter of 25 m and a mass of 7000 t, the ATLAS detector [35] is the largest detector at the Large Hadron Collider (LHC) [36]. It consists of multiple layers of tracking detectors, calorimeters and muon chambers, as can be seen in figure 3.2. By utilising a number of trigger mechanisms and magnet systems, it is able to cope with the large number of particles originating in the proton-proton-collisions at the center of the detector.

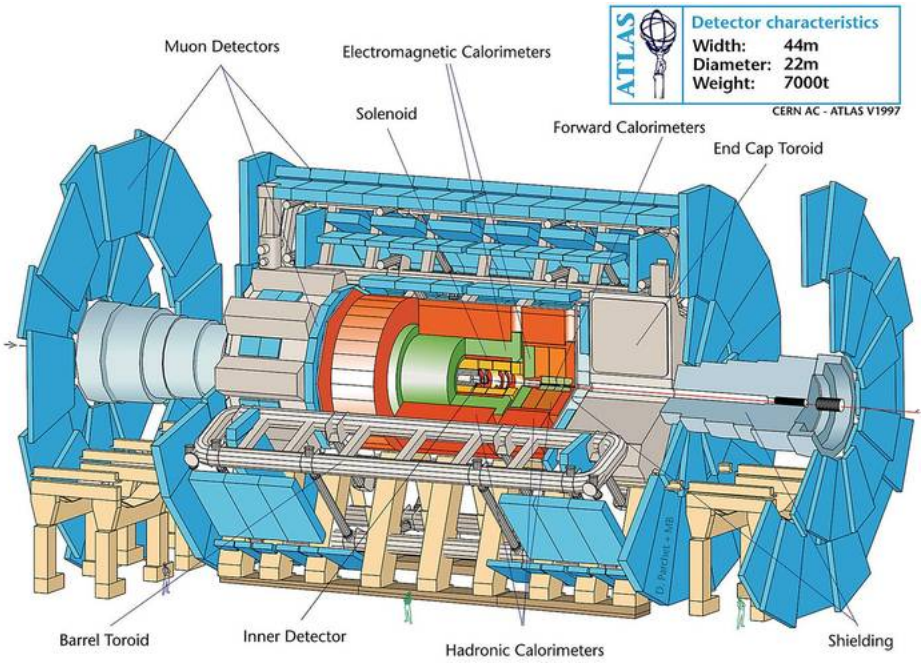

Figure 3.2.: Schematics of the ATLAS detector. © ATLAS Experiment 2022 CERN.

3.2.1. Tracking detector

The innermost layer of the ATLAS detector consists of a 6.2 m long cylinder with a diameter of 2.1 m. This cylinder hosts multiple types of tracking detectors and is enclosed by a 2 T strong magnetic field produced by the central solenoid. The highest challenge for the inner detector is the high particle flux, due to the proximity to the collision point. This requires a way to cope with the large particle flux, which is provided through the use of special granular detectors, that are used in this part and also enable the highest possible momentum and vertex resolution. Another challenge is the accompanying damages to the detector material and read-out mechanisms. The inner detector consists of three subdetectors, that are arranged cylindrically around the beam pipe. They have to fulfil contradicting objectives as they have to be as resilient as possible, while also consisting as little material as possible to not affect the energy measurement in the calorimeter.

The innermost subdetector is the ATLAS pixel detector required for track and vertex reconstruction. The spatial resolution of the pixel detector is 12 μm in the Rϕ-direction and 90 μm in the Z-direction. Each individual pixel is a semi-conductor detector, that produces a signal through the induction of electron-hole pairs by ionising radiation, which is in this case the particles originating from the collision point. Through the high spatial resolution of the pixel detector the trajectory of the ionisation inducing particles can

be reconstructed. This is also used for the reconstruction of the primary and secondary vertices and as such is very important for analysis processes such as b-tagging. To improve the b-tagging performance of the ATLAS pixel detector, an additional layer called the Insertable B-Layer (IBL) [37] was inserted after run I. The inclusion of this layer reduced the distance to the beam axis to 3.3 cm, by including a layer of $50 \times 250\,\mu\text{m}^2$ pixels.

The second subdetector of the innermost detector is the Semi-conductor tracker (SCT), which consists of multiple modules of two p-in-n silicon sensor layers, which are rotated 40 mrad with relation to each other. The principle in which this subdetector measures the position of an ionising particle is the same as for the ATLAS pixel detector. Through the rotated layering of the strips a spatial resolution of $17\,\mu$m in the Rϕ-direction is achieved.

The outermost part of the inner detector is the Transition Radiation Tracker (TRT), which consists of ca. 300000 scintillators embedded in passive material. The scintillators consist of a tube filled with a gas mixture based on Xe with small parts of CO_2 and O_2. The tubes have a diameter of 4 mm and contain a thin gold-coated tungsten wire in their center, which has a diameter of 0.03 mm. The gas is ionised by a traversing charged particle and the produced electrons undergo avalanche multiplication in the vicinity of the tungsten wire, which produces a readable signal. This signal is used to determine the position of the traversing charged particle. The TRT achieves a spatial resolution of $130\,\mu$m and can be additionally used for the identification of electrons, since it also registers transition radiation, which is produced by highly relativistic particles while transitioning between different mediums.

3.2.2. Calorimeter

The calorimeter measures the energy deposited inside of it. There are two types of calorimeters, the electromagnetic calorimeter and the hadronic calorimeter. The electromagnetic calorimeter measures the energy of particles, that mainly interact with their environment via the electromagnetic interaction. If on the other hand the particle possesses colour charge or consists of particles possessing colour charge, the hadronic calorimeter is used to measure the energy of the particle. In the ATLAS detector, both calorimeters are sampling calorimeters, consisting of active and passive material, which reduces the spatial dimensions of the detector. The active material of the calorimeter is used to measure the deposited energy of the primary particle. This is done by inducing the primary high-energy particle to shower into a cascade of secondary particles due to interactions with the medium. This process is directly proportional to the density of the medium,

which is why the active material is inter-spaced with high density passive material, that can not be used to measure the deposited energy, but more easily induces showering of the primary particle. To measure the complete energy of a primary particle, the shower has to be contained inside the calorimeter. Additionally, a sufficiently high energy resolution is required.

The electromagnetic calorimeter, which measures the energy of electrons and photons, is positioned in the pseudo-rapidity region $|\eta| < 3.2$. Since photons do not leave traces in the tracking detector, because they are electrically neutral, their trajectory has to be determined by their impact point in the electromagnetic calorimeter. The innermost layer of the electromagnetic calorimeter is finely segmented to allow for the precision measurement of their trajectory. The active material of the electromagnetic calorimeter is liquid argon while lead is used as the passive material.

The hadronic calorimeter on the other hand uses different active and passive material depending on the section of the detector. The Atlas Hadronic Tile Calorimeter is located in the pseudo-rapidity region $|\eta| < 1.7$ and uses steel as passive material and scintillating tiles as an active medium, in which the energy deposited by hadrons is emitted as light and measured by photomultipliers. For regions of high pseudo-rapidity ($1.5 < |\eta| < 3.2$) the active material is again liquid argon, while the passive material used as an absorber is copper. For regions of even higher pseudo-rapidity ($3.1 < |\eta| < 4.9$) the passive material is switched to tungsten.

3.2.3. Muon detector

The Muon spectrometer is the outermost part of the ATLAS detector and only muons and neutrinos are able to reach it. All other particles are stopped in the calorimeters. The trajectories of muons are diverted by the super-conductive torroid magnets, that are able to produce a magnetic field in strengths between 0.5 T and 3.5 T. The curvatures of these trajectories are used to determine the charge and momentum of the muons. The muon spectrometer consists of four different types of detectors. The Thin Gap Chambers (TGC) and Resistive Plate Chambers (RPC) are used for triggering and have to possess a high temporal resolution. The RPC possess a temporal resolution of 1.5 ns. They are mainly deployed in the central region of the detector within a pseudo-rapidity region of $|\eta| < 1.05$. They consist of finely segmented electrode plates, that encase a gas volume. Charged particles, that traverse the gas volume, ionize it and produce cascades of secondary particles in the gas, that are measured when they reach the electrodes. The

particle flux in high pseudo-rapidity regions is much higher, which results in the need for a different type of triggering detector. As such, TGCs are used in the pseudo-rapidity region of $1.05 < |\eta| < 2.7$, because they are able to handle such an enlarged particle flux. On the other hand they are only able to offer a temporal resolution of 4ns. Even if these two types of detectors are mainly used for triggering purposes, they still offer information on the track position of the muon. A more detailed measurement is used to supplement this first information. For this Monitored Drift Tubes (MDT), and Cathode Strip Chambers (CSC) are used. MDTs are used in the central region of the detector ($|\eta| < 2.0$). As the name suggests they consist of a number of layered drift tubes. These tubes are filled with an Ar-CO_2 mixture, which is ionised by traversing muons. The resulting electrons are collected at the electrodes to form a signal. Using this procedure, MDTs reach a spatial resolution of 25 μm. Similarly to the detectors used for triggering, the high pseudo-rapidity regions ($2.0 < |\eta| < 2.7$) require a detector, that can handle the higher particle flux. In the case of spatial precision measurements in the muon spectrometer CSCs are used. CSCs are multi-wire proportional chambers with their anodes split into strips, that are arranged perpendicular to the wires in the center of the chamber. Using this setup, CSCs reach a spatial resolution of 40 μm.

4. Reconstruction and selection

This chapter presents the definitions of the objects used in this analysis as well as the selection criteria used[1].

4.1. Object definitions

4.1.1. Leptons

Electrons

Electrons are reconstructed by matching tracks reconstructed in the Inner Detector to reconstructed energy deposits (clusters) in the electromagnetic calorimeter [39, 40]. Additionally they are required to meet $p_T > 10\,\text{GeV}$ and $|\eta| < 2.47$ selection cuts. To be classified as electrons, objects have to meet the medium identification criterion described in reference [40]. This criterion is based on a likelihood determinant, that combines a number of observables related to the shower shape in the calorimeter and the track matched to the clusters. Additionally the track of the candidate electron has to originate from the primary vertex of the event. The longitudinal and transverse impact parameters have to meet the requirements $|\text{IP}_z| < 0.5\,\text{mm}$ and $|\text{IP}_{r\phi}|/\sigma_{r\phi} < 5$, respectively. As described in section 4.2.2, further identification and isolation requirements are then applied after overlap removal, see section 4.1.3.

Muons

Muons are defined, by matching track segments in the Muon Spectrometer (MS) to tracks in the Inner Detector. This is done using the Loose identification criterion [41], after which tracks are re-fitted using information from both detector systems. Objects are required to meet $p_T > 10\,\text{GeV}$ and $|\eta| < 2.5$ selection cuts, to be defined as muons. Additionally, the track of the proposed muon has to originate from the primary vertex of the event. The longitudinal and transverse impact parameters have to meet the requirements $|\text{IP}_z| < 0.5\,\text{mm}$

[1] This chapter is in part based on [38].

and $|\mathrm{IP}_{r\phi}|/\sigma_{r\phi} < 3$, respectively. As described in section 4.2.2, further identification and isolation requirements are then applied after overlap removal, see section 4.1.3.

Hadronic taus

The method used to distinguish hadronically decaying τ leptons (τ_{had}) from jets is to use the track multiplicity and a multivariate discriminant based on the track collimation, further jet substructure, and kinematic information [42]. In addition to that, τ_{had} are required to meet $p_T > 25$ GeV and $|\eta| < 2.5$ selection cuts and pass the Medium τ-identification working point.

4.1.2. Jets

Small-R jets

Three-dimensional energy clusters [43] in the calorimeter are used to reconstruct jets using the anti-k_t algorithm [44] implemented in the FastJet package[45] with a radius parameter of 0.4. The electromagnetic scale response prior to jet reconstruction is used to calibrate each topological cluster. After reconstruction, each jet candidate is then calibrated to the jet energy scale, which is derived from simulations, as well as from situ corrections based on 13 TeV data [46]. After energy calibration, the prospective jets are required to meet $p_T > 25$ GeV and $|\eta| < 2.5$ selection cuts. Jets originating from non-collision sources and noise are identified and removed [47]. An additional requirement is made using an algorithm, that matches jets with $p_T < 120$ GeV and $|\eta| < 2.5$ to tracks with $p_T > 0.4$ GeV, to identify jets originating from the primary vertex. This is done to reduce pileup. The algorithm used is known as the jet vertex tagger (JVT) [48] and the default Medium working point is used.

b-tagging

This analysis uses the DL1r b-tagging algorithm to determine jet flavours. A detailed discussion of the methodology used can be found in chapter 5.

4.1.3. Overlap removal

Sometimes a single detector response is counted as multiple objects, such as leptons or small-R jets. To mitigate this effect overlap removal procedures are implemented to remove doubly counted objects. Leptons take precedence over jets, as such jets within the vicinity of $\Delta R_y = \sqrt{(\Delta y)^2 + (\Delta\phi)^2} < 0.2$ of a selected electron are removed. If a

jet survives such a selection and is within the vicinity of an electron of $\Delta R_y < 0.4$, the electron is removed instead. Muons in general are removed, if they are separated from a nearby jet by $\Delta R_y < 0.4$. This reduces the background from heavy-flavour decays inside jets. However, if the muon is near a jet with fewer than three associated tracks, the jet is removed instead. This avoids an inefficiency for high-energy muons undergoing significant energy loss in the calorimeter. In the case of any τ_{had} candidate, it is rejected, if it is in proximity of any electron or muon of $\Delta R_y < 0.2$.

4.1.4. Missing transverse energy

If one combines the transverse energy of all objects present in a respective event, with an additional term for present objects not defined above, it should combine to zero, with regard to the primary vertex. Missing transverse energy refers to the difference between the observed and expected sum of transverse energy and as such accounts for particles such as neutrinos, that are invisible to the detector.

4.1.5. Boosted objects

Large-R jets

For jets in the boosted category, an anti-k_t algorithm with a radius of $R = 1.0$ is used, which results in a collection of large-R jets. The uncertainties of the input small-R jets are directly propagated to the large-R jets without the need of additional recalibration. Such reclustered large-R jets are required to have a reconstructed invariant mass higher than 50 GeV, $p_T > 200$ GeV and at least two small-R jets as constituents.

4.2. Event selection

4.2.1. Trigger

Events are required to pass the lowest unprescaled single-lepton trigger. For this, the leptons have to either pass a low p_T threshold and an isolation requirement, or a higher p_T threshold with a looser identification criterion and no isolation requirement. The lowest p_T requirement is 20 (26) GeV for muons and 24 (26) GeV for electrons in the year 2015 (2016-2018).

The trigger menus used for the four years are listed in table 4.1.

Year	Single-electron trigger	Single-muon trigger
2015	HLT_e24_lhmedium_L1EM20VH HLT_e60_lhmedium HLT_e120_lhloose	HLT_mu20_iloose_L1MU15 HLT_mu50
2016-2018	HLT_e26_lhtight_nod0_ivarloose HLT_e60_lhmedium_nod0 HLT_e140_lhloose_nod0	HLT_mu26_ivarmedium HLT_mu50

Table 4.1.: Single-electron and single-muon trigger used for each year of data-taking.

4.2.2. Pre-Selection

Events are required to have exactly two electrons, two muons or one electron and one muon. The two leptons are required to have opposite charge. In the ee and $\mu\mu$ channel the invariant mass of the two leptons is required to be over 15 GeV and outside of the Z-boson mass window 83-99 GeV.

The number of jets is required to be atleast two jets, both of which are required to meet the 85% b-tagging working point.

5. The DL1r algorithm

The DL1r algorithm uses the output from multiple low level taggers as input to determine the flavour probability of a given jet. Is is widely used in ATLAS flavour tagging [49, 50] and uses SV1, JetFitter, ID3D, RNNIP and SMI as input. These are described below.

5.1. Low Level Taggers

Different algorithm strategies are used to exploit distinctive features of jets containing b- and c-mesons. The distinction is made between algorithms including secondary vertex reconstruction algorithms discussed in Section 5.1.1 and impact parameter based algorithms described in section 5.1.2. The additional information about muons contained in the jet originating from semi-leptonic decays of b- and c-mesons in the jet are included using a Soft Muon Tagger (SMT) described in section 5.1.3.

5.1.1. Vertex based taggers

For secondary-vertex reconstruction ATLAS uses two different algorithms, SV1 and JetFitter. SV1 reconstructs a single secondary vertex and JetFitter performs a topological decay reconstruction along the jet axis.

SV1 [50, 51]

The secondary-vertex finding algorithm SV1 [52] identifies a single secondary vertex in a jet. It does this by starting from all possible vertices with two tracks and then rejects all tracks that are compatible with the decay of long lived particles or have hadronic interaction with the detector material. From all the remaining tracks, an inclusive secondary vertex is reconstructed. SV1 provides eight discriminating variables as input for the high-level DL1r algorithm. These include the number of tracks associated with the SV1 vertex, the invariant mass of the secondary vertex as well as the energy fraction of the secondary vertex, which is derived by comparing the energy associated with the secondary vertex to the total energy of the jet, and the three-dimensional decay length significance.

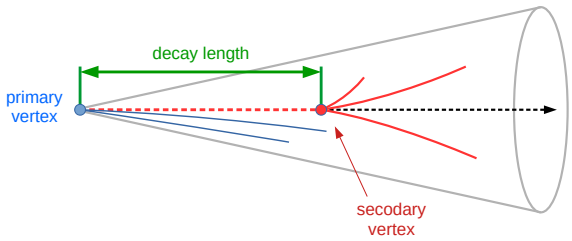

Figure 5.1.: Schematic description of the focus of the SV1 algorithm.

Figure 5.1 depicts on which part of the jet the SV1 algorithm focuses to provide input parameters to the DL1r algorithm, that can be used to infer the flavour of the investigated jet.

JetFitter [50, 51].

JetFitter reconstructs the topological vertex structure of heavy flavour decays inside the jet. It is based on a modified Kalman filter [53] and uses the intersections of particle tracks with the jet axis to reconstruct the topological vertex structure. JetFitter provides eight discriminating variables, that are used as input for the DL1r algorithm. These include the track multiplicity of the JetFitter vertex, the invariant mass of the secondary vertex as well as the energy fraction of the secondary vertex, which is derived by comparing the energy associated with the secondary vertex to the total energy of the jet, and the three-dimensional decay length significance. JetFitter additionally provides twelve variables associated with c-tagging as input for the DL1r algorithm. These exploit the fact, that charm hadron decays have only one secondary vertex with intermediate charged decay multiplicity, while still exhibit a decay length comparable to b-hadron decays. As such charm tagging variables can be included for the jets, that exhibit a singular secondary vertex and include the number, mass and energy of the particles associated with the singular secondary vertex. The smaller decay multiplicity of c-hadrons, when compared to b-hadrons, lead to higher energy decay products for the charm hadron decays, which result in higher rapidity.

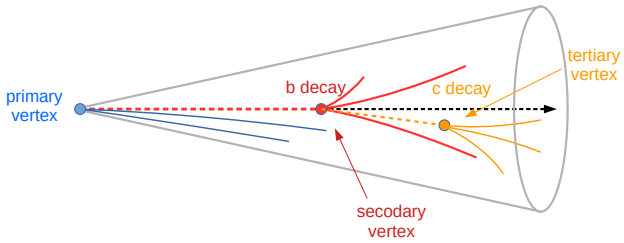

Figure 5.2.: Schematic description of the focus of the JetFitter algorithm.

Figure 5.2 depicts on which part of the jet the JetFitter algorithm focuses to provide input parameters to the DL1r algorithm, that can be used to infer the flavour of the investigated jet. Special focus is placed on the number of early track vertices when distinguishing between b- and c-flavour jets, by exploiting the specific jet topology of $b \mapsto c$-decays.

5.1.2. Impact parameter based taggers

ATLAS uses two different impact parameter based algorithms. One, IP3D, is used as a low level tagging algorithm providing input for both the DL1 and DL1r algorithm, while the RNNIP algorithm is only used as input for the DL1r algorithm to supplement the IP3D algorithm.

IP3D [50, 51]

IP3D starts with a track selection by requiring a transverse momentum $p_T > 1$ GeV, the transverse and longitudinal impact parameters to be $|d_0| < 1$ mm and $|z_0 \sin\theta| < 1.5$ mm and by requiring seven or more silicon hits with at most two silicon holes, at most one of which is in the pixel detector. Additionally the impact parameter is given a positive sign, if the position of crossing between the track and the jet axis is upstream of the primary vertex and negative otherwise. In this, d_0 describes the transverse impact parameter and $z_0 \sin\theta$ the longitudinal impact parameter. IP3D uses both the transverse and longitudinal impact parameter significance in a two-dimensional template to account for their correlation.

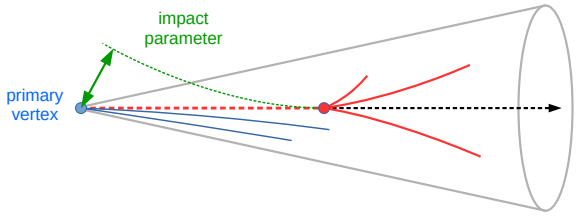

Figure 5.3.: Schematic description of the focus of the IP3D algorithm.

Figure 5.3 depicts on which part of the jet the IP3D algorithm focuses to provide input parameters to the DL1r algorithm, that can be used to infer the flavour of the investigated jet.

RNNIP [54]

IP3D assumes that the impact parameters of each track are uncorrelated, which is not the case, since multiple charged particles can emerge from a secondary or tertiary vertex with large impact parameters. The likelihood of a track with high impact parameters is higher the more tracks from the same vertex have high impact parameters. For light jets, no displaced decays occur, so no such relationship should exist.

Recurrent Neural Networks (RNN) are used to directly learn the connections between the impact parameters of different tracks in sequence and their output likelihood predictions for the different jet flavours.

5.1.3. Soft Muon Tagger

The Soft Muon Tagger (SMT)[50] is based on the reconstruction of muons coming from semileptonic decays of heavy-flavour hadrons. These muons have typically a high p_T^{rel} value relative towards the jet axis as well as a general high p_T value, when compared to other jet components. Regardless, when compared to leptons originating from the primary decay vertex, these values are low. As such the inter-jet muons are denoted as "soft" muons. The presence of muons in jets is enhanced for b-jets, when compared to charm- and light-jets due to the large branching ratios $BR(b \mapsto \mu\nu X) \approx 11\%$ and $BR(b \mapsto c \mapsto \mu\nu X) \approx 10\%$, where X denotes an unknown hadron.

The Soft Muon Tagger requires an angular separation between the muon and the jet axis of $\Delta R < 0.4$. Additionally the muons are required to have $p_T > 5\,\text{GeV}$, $d_0 < 4\,\text{mm}$ and $|\eta| < 2.5$. Light jets can meet these requirements via these three background processes:

- prompt muons from the nearby W boson can randomly be associated to light jets ($\sim 1\%$ contamination);
- muons coming from the decay in flight of light hadrons. These light hadrons are mostly pions and kaons ($\sim 1\%$ contamination);
- energetic hadrons, that travel through the calorimeter system and reach the MS and are as such mistakenly identified as muons ($< 0.1\%$ contamination).

To separate the muons coming from b- and c-hadron decays from these background processes contained in light jets, the three parameters ΔR, d_0 and p_T^{rel} are used. Since the p_T^{rel} is directly correlated to the mass of the parent particle, muons originating from b-hadron decays have in general higher p_T^{rel}.

Additional variables categorising the quality of the muon track are used. These variables are scattering neighbour significance ($\mathcal{S}$), the momentum imbalance significance ($\mathcal{M}$) and the ratio between the track curvature measured in the inner detector and the muon chamber ($\mathcal{R}$).

5.2. High Level Taggers

The DL1r algorithm is a high level tagger and takes the output of several low level taggers, described in section 5.1, as input. It is based on a Deep Learning Neural Network trained using Keras with the Therano [55] backend and the Adam optimizer [56]. It has a three dimensional output corresponding to the probability values for the three different flavour options, light, charm and b. The topology of the Neural Network consists of a mixture of fully connected hidden layers and maxout layers [57]. Information regarding batch normalisation can be found in Ref. [58].

5.2.1. DL1r score

The final discriminant of the DL1r algorithm is calculated by combining the three-dimensional output of the Neural Network:

$$\mathrm{DL1r}_{\mathrm{score}} \propto \frac{p_b}{p_c + p_{\mathrm{LF}}}$$

Pseudo-continuous (PC) b-tagging calibration is used. This allows for finer differentiation of jets and is referred to as the pseudo-continuous DL1r score in the following. This allows the differentiation into five classes according to the tightest WP with which the respective jet is tagged, as opposed to only two classes, tagged or not tagged, using a single WP. The different b-tagging working points calibrated using this discriminant are shown in table 5.1.

Working Point (WP)	DL1r score
85%	0.665
77%	2.195
70%	3.245
60%	4.565

Table 5.1.: Different Working Points of the DL1r algorithm corresponding to their respective DL1r scores.

These working points are used to define a number of working point flags. The exact definition can be seen in table 5.2.

DL1r WP flag	Definition
0	jets meeting no WP requirements
1	jets between 85% WP and 77% WP
2	jets between 77% WP and 70% WP
3	jets between 70% WP and 60% WP
4	jets at least at 60% WP

Table 5.2.: Definitions of the working point flags used in this analysis.

6. Samples

This section describes the samples, that are used in this analysis[1]. For the analysis, an estimation of all signal and background sources is needed. For this, a number of Monte-Carlo (MC) simulations based on the Standard Model are used. The signal and background contributions are simulated using a number of consecutive steps:

1. generate the events on parton level

2. model the parton showering and the hadronisation processes

3. simulate the response of the ATLAS detector to the produced particles

4. process data and MC-samples using the same reconstruction framework and apply object definitions and selection criteria as well as determine scale factors

Data samples are detailed in section 6.1, while Monte-Carlo samples are detailed in section 6.2. Object definitions can be found in section 4.1 and the event selection criteria can be found in section 4.2. For a more detailed overview of the data and Monte-Carlo samples, as well as a detailed listing of the exact DSIDs of the samples used, refer to appendix A.

6.1. Data samples

For this analysis, proton-proton collision data at $\sqrt{s} = 13$ TeV collected by the ATLAS detector from 2015 to 2018 is used. The selection criteria are detailed in section 4.2. Data were collected under stable LHC beam conditions with fully operational ATLAS detector. In table 6.1, all the datasets used in this analysis are listed together with their respective integrated luminosities[2]. The trigger menus used are listed in section 4.2.1.

[1] This chapter is in part based on [38].
[2] LuminosityForPhysics twiki page, https://twiki.cern.ch/twiki/bin/viewauth/Atlas/LuminosityForPhysics

Year	$\int L\mathrm{d}t$ [pb^{-1}]	dataset
2015	3219.6	data15_13TeV20170619/physics_25ns_21.0.19.xml
2016	32988.1	data16_13TeV20180129/physics_25ns_21.0.19.xml
2017	44307.4	data17_13TeV20180619/physics_25ns_Triggerno17e33prim.xml
2018	58450.1	data18_13TeV20190219/physics_25ns_Triggerno17e33prim.xml

Table 6.1.: Dataset used together with the corresponding total integrated luminosity[3].

The total integrated luminosity of the data sets used is 139.0 fb^{-1} with an uncertainty of 1.7 % [59]. The partial data sets used in this analysis are listed, with their uncertainties [60] in table 6.2.

Year	$\int L\mathrm{d}t$ [fb^{-1}]	uncertainty [%]
2015-2016	36.2	2.1
2017	44.3	2.4
2018	58.5	2.0
2015-2018	139.0	1.7

Table 6.2.: Partial data sets used in the analysis with their respective integrated luminosities [60].

6.2. Monte-Carlo samples

All nominal Monte-Carlo samples used in this analysis, where produced using the full ATLAS detector simulation [61] (FS) based on GEANT4 [62]. Other samples where produced using fast simulation (AFII) [63], where the full simulation is replaced by a parametrisation of the shower shapes [64]. The detector response of the ATLAS detector is based on the GEANT4 [62] framework for the nominal case, while additional samples are produced based on the fast parametrisation Atlfast-II (AFII) [63] of the detector response to estimate modelling uncertainties.

Different generators are used for different sample types. Heavy flavour decays in processes, that are not simulated using SHERPA [65], are simulated using EVTGEN v1.6.0 [66]. Additional interactions were produced using PYTHIA8.186 [67] and overlaid over the hard-scatter events to simulate the effects of pileup. To match the pileup conditions observed in data each of the additional interactions in the Monte Carlo samples is reweighted to data through the inclusion of an additional scale factor.

[3]Web directory for GoodRunsList at ATLAS group-data area, https://atlas-groupdata.web.cern.ch/atlas-groupdata/GoodRunsLists/

6.2.1. $t\bar{t}$+jet

The samples used to model the $t\bar{t}$+jet events include the semi-leptonic decay, to allow for the modelling of lepton fakes, and a number of b- or c-enhanced samples. The combined sample includes $t\bar{t}$+c signal events, as well as the $t\bar{t}$+b and $t\bar{t}$+light background events. The differentiation of these events is done via a specific flag for the event flavour set at truth level. All $t\bar{t}$+jet sample generators use a top mass of $m_t = 172.5\,\text{GeV}$ and the mass of b-quarks originating from the top decay is set to $m_b = 4.95\,\text{GeV}$.

$t\bar{t}$

The production of $t\bar{t}$ is modelled using the PowhegBox [68] at NLO in QCD in the five flavour scheme and by using the NNPDF3.0nlo PDF set [69]. Parton showering and hadronisation is handled using the Pythia8 [67] interface and the h_{damp} parameter, that controls the first gluon emission, is set to $1.5\,m_t$. This sample is used as a nominal $t\bar{t}+ \geq 1$c, $t\bar{t}+ \geq 1$b and $t\bar{t}+ \geq 1$ light prediction. The distinction between these three event types is made by searching for additional b- or c-quarks on truth level. If an additional b-quark is found the event is labeled as $t\bar{t}+ \geq 1$b. If no additional b-quark is found, and instead a c-quark is found, the event is labeled $t\bar{t}+ \geq 1$c. All other cases are labeled $t\bar{t}+$ light. Uncertainties related to initial state radiation (ISR) and final state radiation (FSR) are implemented by reweighing the original PowhegBox + Pythia (FS) sample using dedicated Monte Carlo weights. The details of this procedure are depicted in section 8.2.

6.2.2. Single top production

Single top t- and s-channel, as well as tW production is modelled using the PowhegBox v2 [70–72] generator. For t-channel events are produced in the four flavour scheme using the NNPDF3.0nlonf4 [69] PDF set. For s-channel and tW, the events are produced in the five flavour scheme using the NNPDF3.0nlo PDF set [69].

6.2.3. $t\bar{t}V$

The production of $t\bar{t}V$ ($t\bar{t}W$ or $t\bar{t}Z$) events is modelled using the MADGRAPH5 [73] generator, using the NNPDF3.0nlo PDF set [69], interfaced to Pythia8 [67].

6.2.4. V+jets

QCD V+jets (i.e W+jets or Z+jets) events are modelled using the Sherpa v.2.2.1 generator [65], while using the NNPDF3.0nnlo PDF set.

6.2.5. $t\bar{t}H$

Events with a top-quark pair in association with a Higgs boson ($t\bar{t}H$) are generated with Powheg-Box (v2) [68] using the NNPDF3.0nlo PDF set [69]. The showering and hadronisation are simulated with Pythia8 [67], and the assumed Higgs boson mass is 125 GeV.

All produced samples are reweighted with regard to several efficiencies to improve agreement with data. The weights used take into account the MC event weights as well as scale factors for the leptons, the pileup, the b-tagging (which include b- and c-tagging efficiencies, as well as light miss tagging rates) and the Jet Vertex Tagging (JVT) efficiencies. The samples are split according to the data-taking periods 2015+2016 (mc16a), 2017 (mc16d) and 2018 (mc16e), because of differing pileup conditions during these different data-taking periods. As such the distributions in each MC campaign are scaled to their respective integrated luminosities shown in table 6.2. The cross-sections used for this process where taken from the common XSection-MC16-13TeV.data file used by the Top working group [74].

For a more detailed overview over the specific DSIDs (Data Set Identification) used for all signal and background event types, as well as a listing of the cross-sections used for normalisation refer to appendix A.

7. Analysis strategy

This chapter discusses the general analysis strategy utilised[1]. Section 7.1 discusses the process, in which the relevant signal region was found, and section 7.2 presents the method used to fit the simulated samples.

7.1. Signal region selection

This section depicts the process of signal region selection starting from the pre-selection cuts shown in section 4.2.2. If the $t\bar{t}$+jets composition for events with different number of jets and only pre-selection cuts is compared, as depicted in figure 7.1, then for events with only two jets, a dominance of $t\bar{t}$+light events and a negligible number of $t\bar{t}$+c signal events is observed. As such, the events with only two jets can be safely excluded which results in a first addendum to the pre-selection cut by requiring events to have at least three jets.

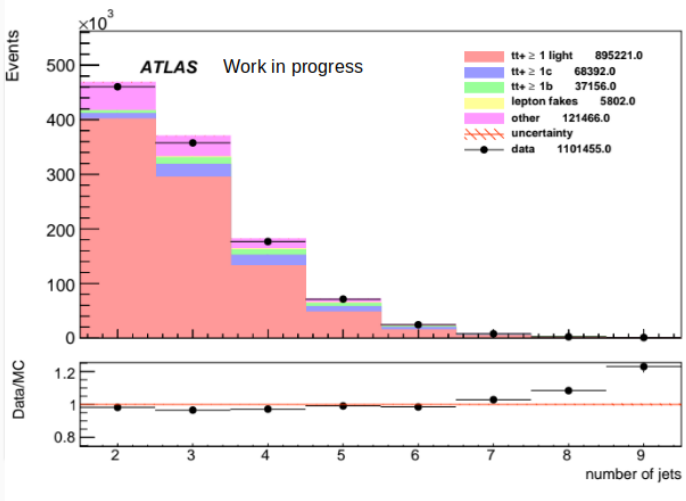

Figure 7.1.: Number of jets with only Pre-selection.

[1]This chapter is based in part on [28].

7.1.1. Working point investigation

After the implementation of this additional cut, this analysis investigated the signal to background ratio in dependence on the working point acceptance of the leading jets. The signal is $t\bar{t}+ \geq 1c$ events and the background consists of other $t\bar{t}$+jet and other non $t\bar{t}$+jet backgrounds, as listed in section 2.6.1.

As can be seen in table 7.1 the signal to background ratio declines with the tightness of the b-tagging. As such the highest signal to background ratio is found by using the 85% or the 77% working point without using additional conditions. The correspondent graphical depictions for the different working points can be found in figure 7.2 to figure 7.9 and show the distribution of the number of jets and the number of b-tagged jets for different working points.

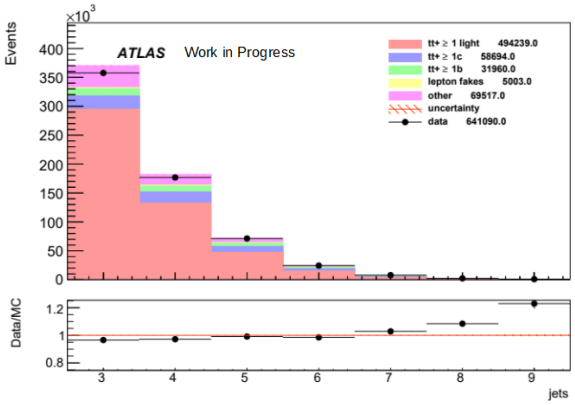

Figure 7.2.: Data-MC-comparison of the number of Jets for the two leading jets at 85% WP.

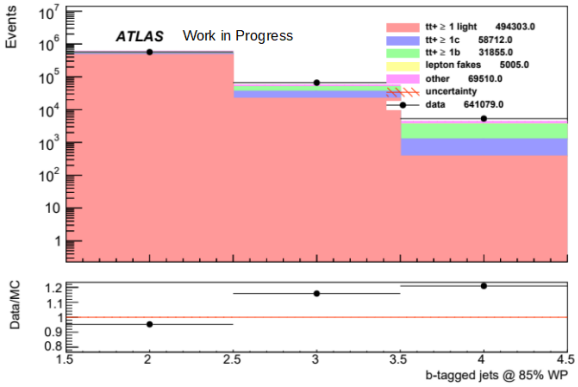

Figure 7.3.: Data-MC-comparison for the number of b-tagged jets at 85% WP.

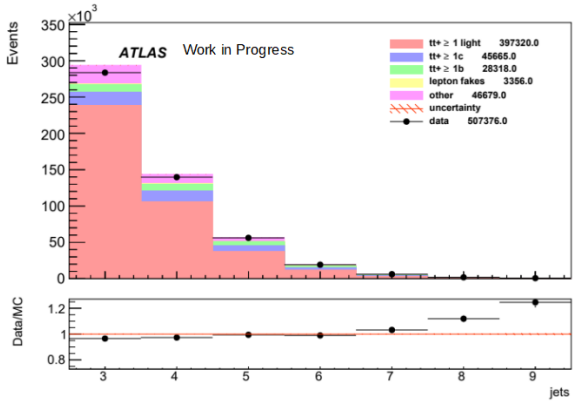

Figure 7.4.: Data-MC-comparison of the number of Jets for the two leading jets at 77% WP.

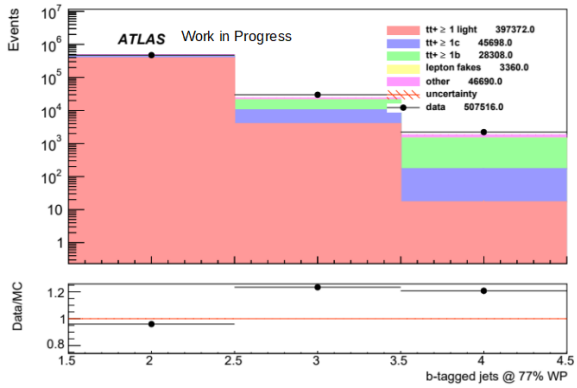

Figure 7.5.: Data-MC-comparison for the number of b-tagged jets at 77% WP.

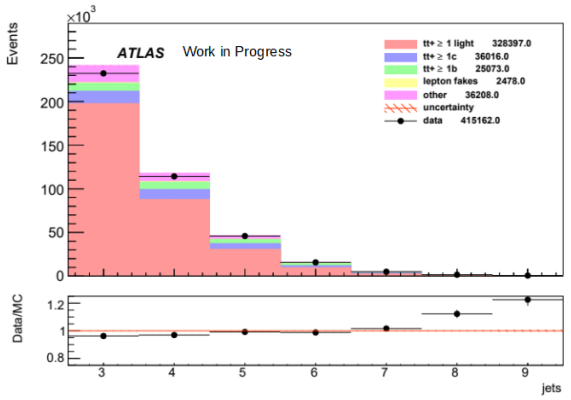

Figure 7.6.: Data-MC-comparison of the number of Jets for the two leading jets at 70% WP.

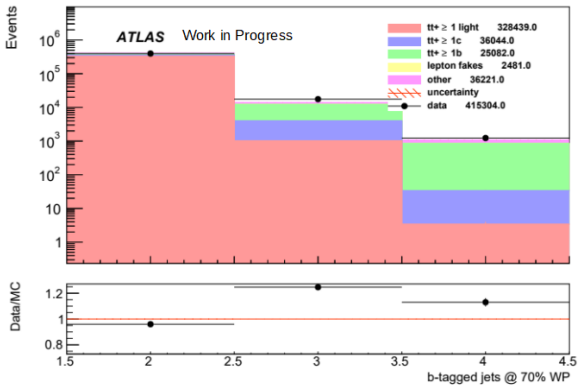

Figure 7.7.: Data-MC-comparison of the number for b-tagged jets at 70% WP.

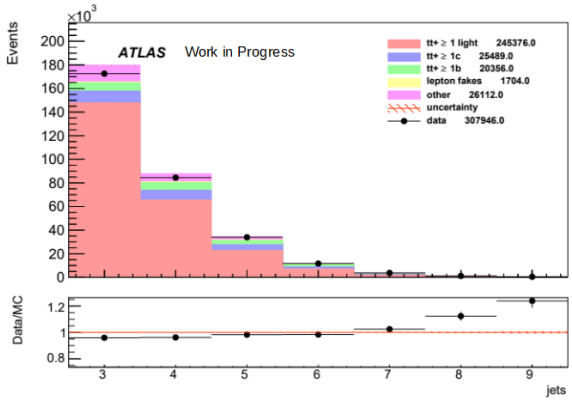

Figure 7.8.: Data-MC-comparison of the number of Jets for the two leading jets at 60% WP.

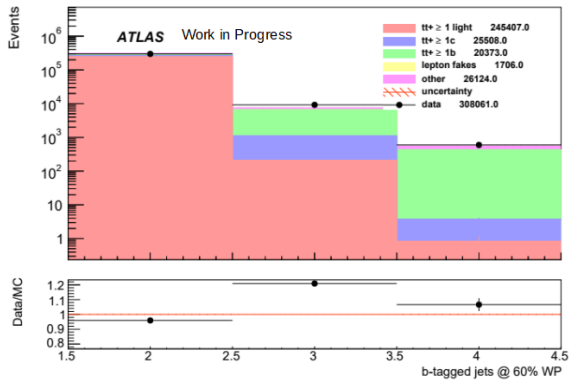

Figure 7.9.: Data-MC-comparison of the number for b-tagged jets at 60% WP.

	85% WP	77% WP
expected total yield	659400 ± 600	521400 ± 400
expected signal yield	58710 ± 60	45700 ± 50
expected background yield	600700 ± 500	475700 ± 400
expected S/B ratio	0.0977 ± 0.0001	0.0961 ± 0.0001
	70% WP	60% WP
expected total yield	428300 ± 300	319100 ± 300
expected signal yield	36040 ± 40	25510 ± 30
expected background yield	392200 ± 300	293600 ± 300
expected S/B ratio	0.0919 ± 0.0001	0.0869 ± 0.0001

Table 7.1.: Expected yield and expected signal to background ratio for different tightness requirements on the b-tagging.

7.1.2. DL1r score investigation for the third jet

The jets are sorted by their DL1r score. For this, every jet is assigned to one of the five bins corresponding to the b-tagging working point based on their DL1r score and then each bin is sorted by the p_T of the jet. Each bin is assigned a flag according to the definition in table 7.2.

DL1r WP flag	Definition
0	jets meeting no WP requirements
1	jets between 85% Wp and 77% WP
2	jets between 77% Wp and 70% WP
3	jets between 70% Wp and 60% WP
4	jets at least at 60% WP

Table 7.2.: Definition of the WP flags used in the following analysis.

If the signal to background ratio is plotted with respect to the working point, that the third jet meets, one observes a clear increase in the signal to background ratio, if the third jet meets the 77% working point, but misses the 70% working point. This is depicted in figure 7.12 and figure 7.13. This general pattern holds independent from the working point requirement on the leading two jets, as can be seen in figure 7.10 to figure 7.17.

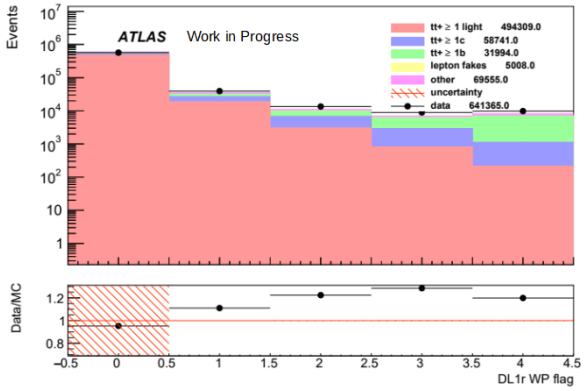

Figure 7.10.: Data-MC-comparison for the b-tagging score for the jet with third highest b-tagging score with the two leading jets at 85% WP.

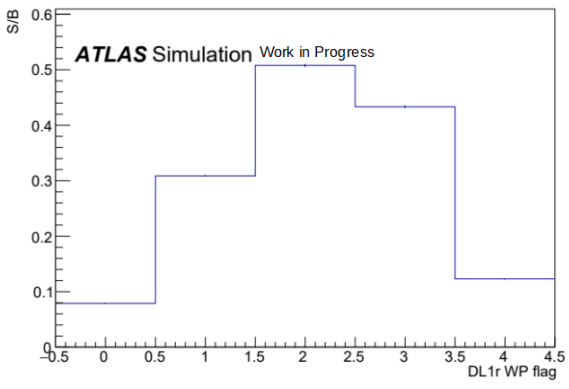

Figure 7.11.: Signal to Background ratio or the b-tagging score for the jet with third highest b-tagging score with the two leading jets at 85% WP.

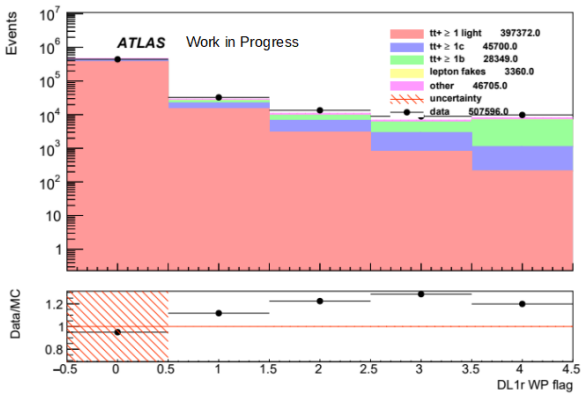

Figure 7.12.: Data-MC-comparison for the b-tagging score for the jet with third highest b-tagging score

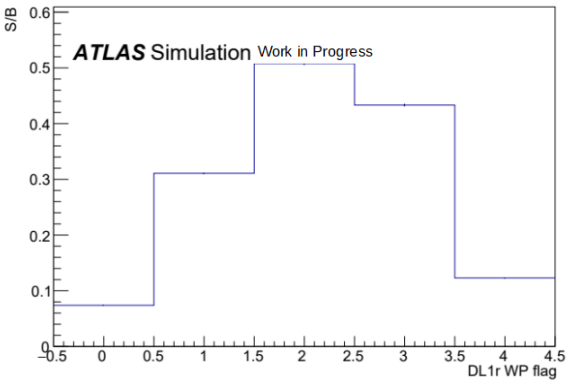

Figure 7.13.: Signal to Background ratio for the jet with third highest b-tagging score with the two lead-

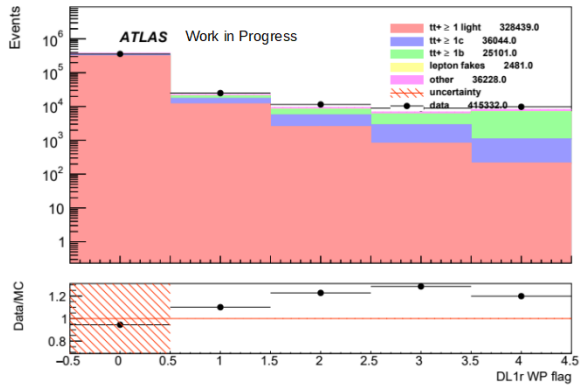

Figure 7.14.: Data-MC-comparison for the b-tagging score for the jet with third highest b-tagging score

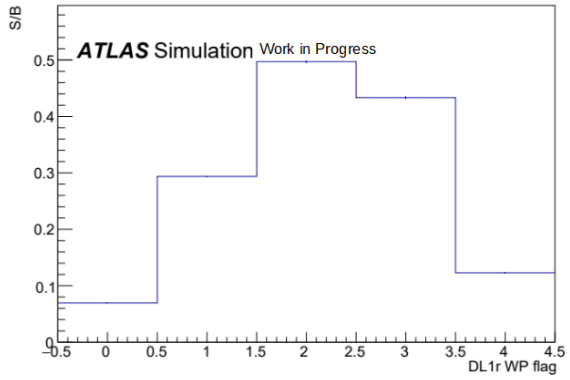

Figure 7.15.: Signal to Background ratio for the jet with third highest b-tagging score with the two lead-

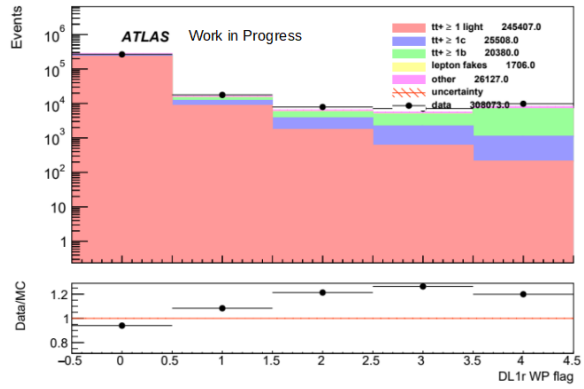

Figure 7.16.: Data-MC-comparison for the b-tagging score for the jet with third highest b-tagging score with the two leading jets at 60% WP.

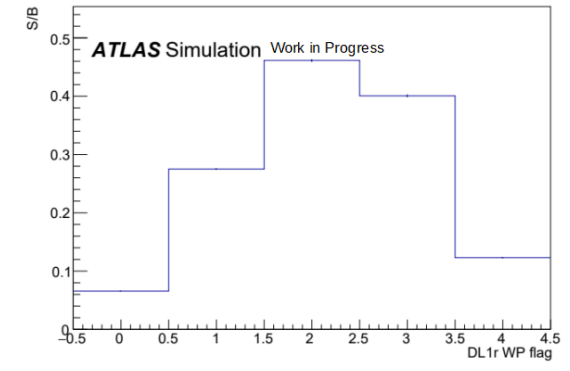

Figure 7.17.: Signal to Background ratio for the jet with third highest b-tagging score with the two leading jets at 60% WP.

As such the investigation focuses on this specific bin. The exact results in this bin for the differing working point requirements on the leading two jets are shown in table 7.3.

	85% WP	77% WP
expected total yield	11000 ± 30	11000 ± 30
expected signal yield	3700 ± 11	3700 ± 11
expected background yield	7290 ± 30	7290 ± 30
expected S/B ratio	0.508 ± 0.003	0.508 ± 0.003
	70% WP	60% WP
expected total yield	9390 ± 30	6520 ± 20
expected signal yield	3120 ± 10	2060 ± 10
expected background yield	6270 ± 30	4460 ± 20
expected S/B ratio	0.497 ± 0.003	0.462 ± 0.003

Table 7.3.: Expected yields and signal to background ratio for different tightness requirements on the b-tagging for the leading jets. The third jet passes the 77% WP and misses the 70% WP

As can be seen in table 7.3 the highest signal to background ratio can be found in events with leading jets at 85% or 77% working point. As such, the following analysis concentrates on events with leading jets at these working points.

Splitting the sample with regard to the number of jets

In the following, the samples are split between events with exactly three jets and at least four jets to investigate further possibilities to increase the signal to background ratio.

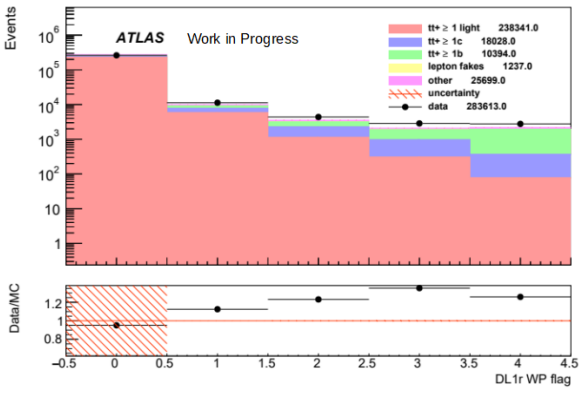

Figure 7.18.: Data-MC-comparison for the b-tagging score for the jet with 3rd highest b-tagging score with the two leading jets at 77% WP and exactly three jets.

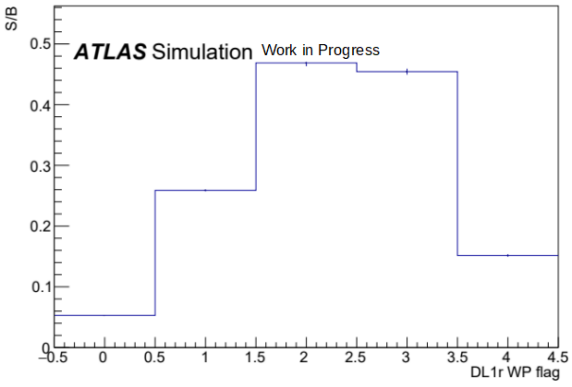

Figure 7.19.: Signal to Background ratio for the jet with 3rd highest b-tagging score with the two leading jets at 77% WP and exactly three jets.

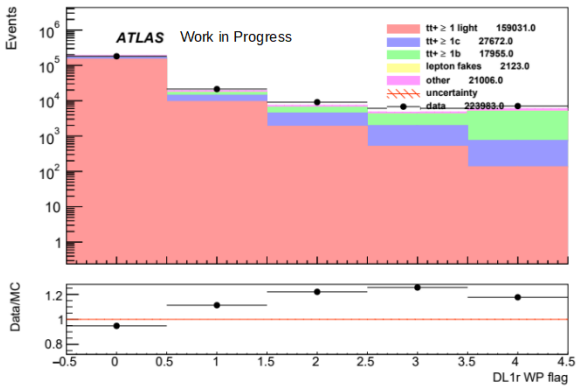

Figure 7.20.: Data-MC-comparison for the b-tagging score for the jet with 3rd highest b-tagging score

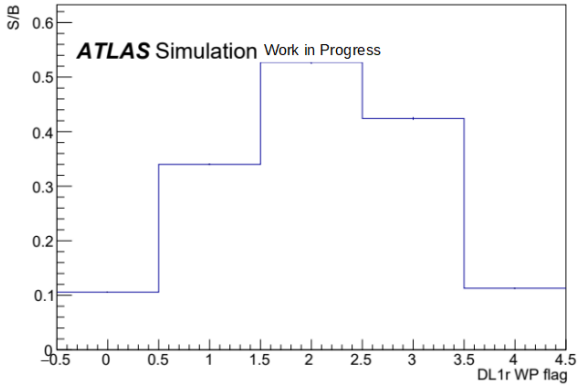

Figure 7.21.: Signal to Background ratio for the jet with 3rd highest b-tagging score with the two leading

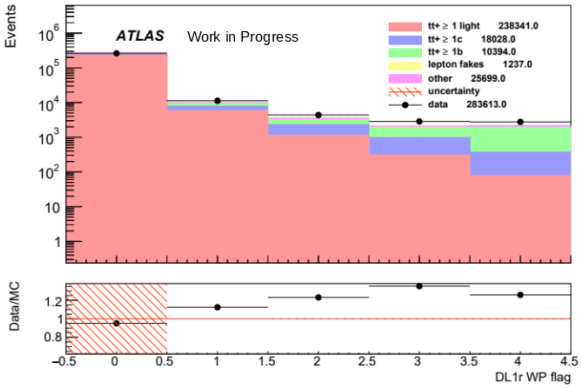

Figure 7.22.: Data-MC-comparison for the b-tagging score for the jet with 3rd highest b-tagging score

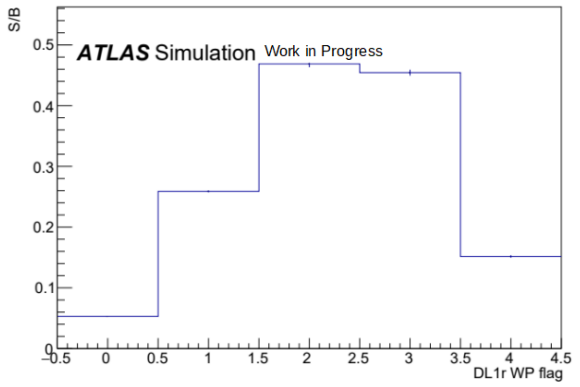

Figure 7.23.: Signal to Background ratio for the jet with 3rd highest b-tagging score with the two leading

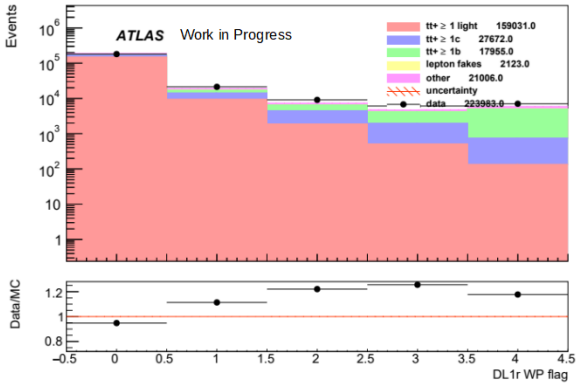

Figure 7.24.: Data-MC-comparison for the b-tagging score for the jet with 3rd highest b-tagging score with the two leading jets at 85% WP and at least four jets.

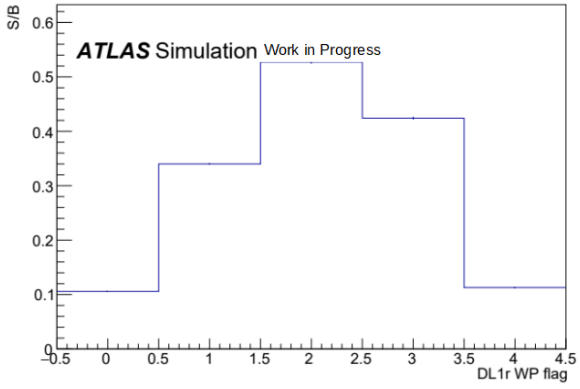

Figure 7.25.: Signal to Background ratio for the jet with 3rd highest b-tagging score with the two leading jets at 85% WP and at least four jets.

As can be seen in figure 7.18 to figure 7.21 the bins with highest signal to background ratio are consistent with the entire sample. The exact signal to background ratios are shown in table 7.4. Both 85% WP and 77% WP for the leading jets share the same yields, since the third highest jet passing the 77% WP implies, that the leading jets also pass the 77% WP regardless of the prior selection. This can also be seen by comparing the plotted distribution for both working point criteria, which can be seen in figure 7.18 to figure 7.25.

exactly 3 jets	85% WP	77% WP
expected total yield	3560 ± 30	3560 ± 30
expected signal yield	1140 ± 10	1140 ± 10
expected background yield	2420 ± 20	2420 ± 20
expected S/B ratio	0.469 ± 0.005	0.469 ± 0.005
4+ jets	85% WP	77% WP
expected total yield	7440 ± 20	7440 ± 20
expected signal yield	2570 ± 10	2570 ± 10
expected background yield	4870 ± 20	4870 ± 20
expected S/B ratio	0.527 ± 0.003	0.527 ± 0.003

Table 7.4.: Expected yield and expected signal to background ratio for different tightness requirements on the b-tagging and different jet requirements.

The highest signal to background ratio can be found for events with at least four jets, where the first two jets meet the 77% working point and the third jet meets the 77% WP but misses the 70% WP.

Selecting the signal region

In the previous section it was found that the signal region can be selected with a simple cut. By requiring at least three jets at 77% WP and at most two jets at 70% WP, the defined signal region can be achieved. This can be seen in figure 7.26 and figure 7.27, as the yields and signal to background ratio is identical to the results in table 7.3. This is shown in table 7.5.

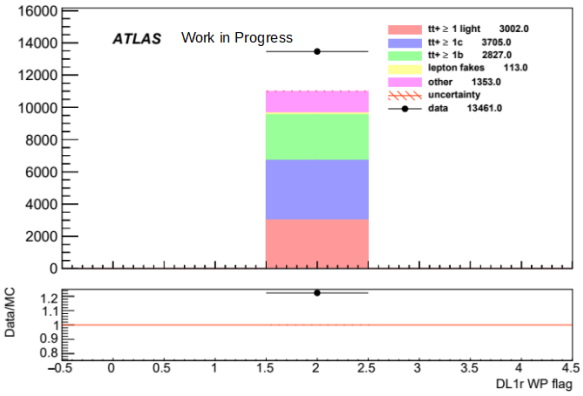

Figure 7.26.: Data-MC-comparison for the b-tagging score by requiring at least three jets at 77% WP and at most two jets at 70% WP

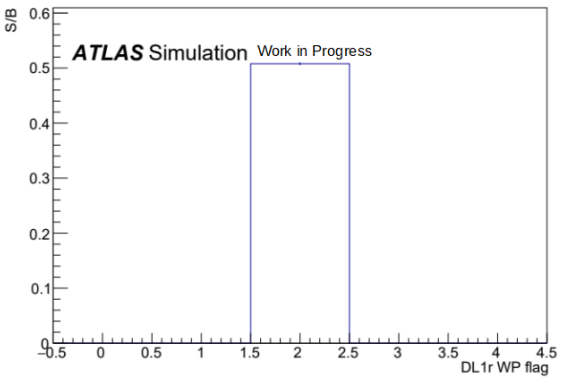

Figure 7.27.: Signal to Background ratio for requiring at least three jets at 77% WP and at most two jets at 70% WP

	DL1r sorted	Region selected
expected total yield	11000 ± 30	11000 ± 30
expected signal yield	3700 ± 11	3700 ± 11
expected background yield	7290 ± 30	7290 ± 30
expected S/B ratio	0.508 ± 0.003	0.508 ± 0.003

Table 7.5.: Comparison between the yields of the signal region achieved through the DL1r score sorting and the region cut.

7.1.3. DL1r score investigation of the fourth jet

If one selects the previously defined signal region and then investigates the DL1r WP flags of the jet with fourth highest DL1r score, one observes the distribution in figure 7.28 and the corresponding signal to background ratio in figure 7.29.

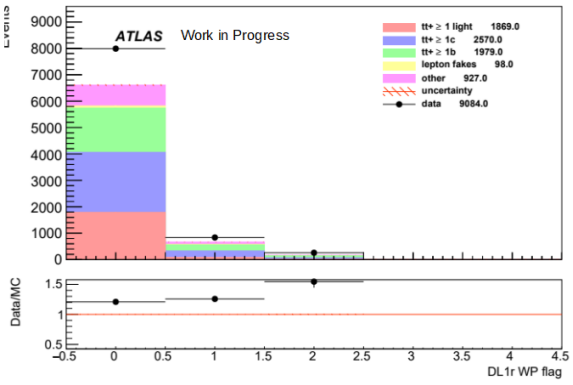

Figure 7.28.: Data-MC-comparison for the b-tagging score in the selected signal region for the fourth jet.

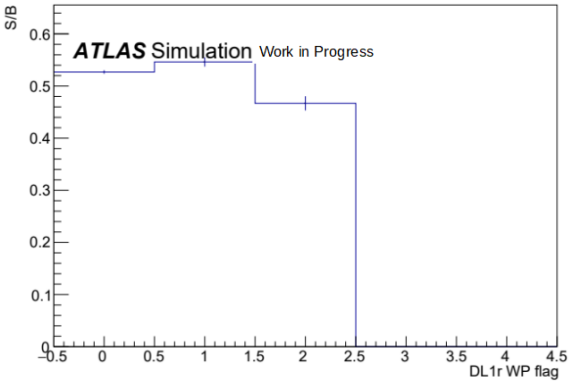

Figure 7.29.: Signal to Background ratio for previous plot.

As can be seen the highest expected signal to background ratio can be found if the fourth jet meets the 85% WP requirement but misses the 77% WP requirement. The exact values can be found in table 7.6.

expected total yield	666 ± 6
expected signal yield	235 ± 3
expected background yield	431 ± 5
expected S/B ratio	0.546 ± 0.009

Table 7.6.: Expected S/B-ratio for events with at most two jets at 70% WP, at most three jets at 77% WP and at least four jets at 85% WP.

The expected yields are much lower with this selection. As such, the expected gain in the S/B-ratio has to be settled with the reduced statistics. Since the bin with highest yield also has comparatively high S/B-ratio, one would expect that a combination of the first and second bin in figure 7.28 would also lead to an increase in S/B-ratio, without much sacrifices to the yield. The exact values for this combination can be seen in table 7.7.

expected total yield	7280 ± 30
expected signal yield	2516 ± 11
expected background yield	4760 ± 30
expected S/B ratio	0.529 ± 0.004

Table 7.7.: Expected S/B-ratio for events with at most two jets at 70% WP and at most three jets at 77% WP.

If the resulting signal to background ratio is compared to the result in table 7.4, no significant increase in S/B-ratio is observed, while reducing the expected yield.

7.1.4. Resulting signal regions

The cuts discussed in section 7.1.2 are used to define the signal regions used in this analysis. These cuts predict the highest signal to background ratio for events with the third jet b-tagged between the 77% and 70% working point of the DL1r b-tagging algorithm. This is consistent with tagging behaviour of the low level taggers discussed in section 5.1, the outputs of which are used as input in the DL1r b-tagging algorithm, in response to charm quarks. As can be seen in figure 7.26, the MC prediction deviates from the data. To compensate for the mismodelling as well as for other sources of uncertainties a profile likelihood fit to the data has to be performed.

7.2. Profile likelihood fit

A profile likelihood fit is performed using the signal strength ($\mu = \sigma_{t\bar{t}c\bar{c}}/\sigma^{\mathrm{SM}}_{t\bar{t}c\bar{c}}$) as the parameter of interest. The signal strength is defined as the ratio between the measured cross-section and the standard model prediction. This fit is performed using the TRExFitter package, which includes the HistFactory tools [75]. These tools can be used to fit multiple signal regions as well as control regions simultaneously. Multiple features, such as pruning, systematisation of systematic uncertainties, as well as smoothing of the histograms are used to perform a profile likelihood fit. The Likelihood is defined as

$$L = \prod_i \mathrm{Pois}(n_i|\mu \cdot s_i(\vec{\theta}) + b_i(\vec{\theta})) \times \prod_j C(\theta_j|\theta_j^0, \sigma_{\theta_j^0}). \tag{7.1}$$

Here i runs over the number of bins, while j runs over all nuisance parameters θ_j. n_i denotes the number of events in bin i, while s_i and b_i stand for the expected number of signal and background events, respectively. Nuisance parameters represent systematic uncertainties, where $\theta = 0$ denotes the nominal sample and $\theta = \pm 1$ corresponds to the $\pm 1\sigma$ variation of the respective uncertainty. In the likelihood function this is represented by the last term, which is called the constraint term for the systematics, where θ_j^0 is the nominal value of the nuisance parameter and the $\sigma_{\theta_j^0}$ is the prior estimate of the uncertainty. Here this is implemented via a Gaussian prior term with $\theta_j^0 = 0$ and $\sigma_{\theta_j^0} = 1$. The $\pm 1\sigma$ variations are interpolated to describe systematics continuously. Each systematic effect is split into a shape influence and a nominal influence. The shape influence is described using a linear

interpolation and the nominal influence is described using exponential interpolation. In combination with the Gaussian constraint term the exponential interpolation leads to a log-normal constraint of the normalisation component of the nuisance parameter θ. This prohibits negative event yields.

Additionally, the signal and background predictions are effected by statistical uncertainty. This is modelled by introducing an additional common nuisance parameter γ_i for all bins that scales the nominal Monte-Carlo prediction. The nominal value of this nuisance parameter is one, which distinguishes it from the other nuisance parameters. Additionally the constraint term used is a Poisson constraint term instead of a Gaussian constraint term. This constraint term takes into account the total statistical uncertainty within each bin [75].

Possible correlations between nuisance parameters ($\vec{\theta}$ and $\vec{\gamma}$) are not considered in the likelihood, but can be determined during the fit. The fit is performed by constructing the negative log-likelihood function and then minimising it using algorithms implemented in MINUIT [76]. All nuisance parameters as well as the parameter of interest (the signal strength μ) are allowed to vary freely during the fit. The difference between them is, that μ is being modelled as a free parameter, while the nuisance parameters are constrained by their respective constraint terms. If after the fit a nuisance parameter differs from its initial value, the data has strength to pull the respective parameter. If the post-fit uncertainty is smaller than $\pm 1\sigma$, the data is able to constrain the respective parameter statistically.

8. Systematic uncertainties

The nuisance parameters included in the fit correspond to specific statistical or systematic uncertainties of the model implemented in this analysis. In this chapter, all systematic uncertainties included are described[1]. Section 8.1 will describe the experimental uncertainties, the sources of which are the same for all simulated signal and background sample. Section 8.2 describes the theoretical systematic uncertainties. Special focus is placed on modelling uncertainties in regard to the $t\bar{t}$ samples as well as the uncertainties arising from the modelling of the main backgrounds. Not all systematic uncertainties effecting the parameter of interest are included in this analysis. For a full list of systematic uncertainties, that are to be included in this analysis at a later date, see chapter 10.

8.1. Experimental uncertainties

8.1.1. Luminosity and pile-up modelling

Luminosity

The uncertainty on the integrated luminosity for the full run-2 dataset given in section 6.1 is 1.7% [59], obtained using the LUCID-2 detector [77] for the primary luminosity measurements.

Pile-up modelling

The samples described in section 6.2 are reweighted to account for the uncertainty in the ratio of the predicted and measured inelastic cross-sections in the fiducial volume defined by $M_X > 13\,\text{GeV}$, where M_X is the mass of the hadronic system [78]. This is done by changing the nominal scale-factor applied to the pile-up distribution from the nominal 1.0/1.03 to 1.0/0.99 or 1.0/1.07 to derive the up and down systematic uncertainty.

[1] This chapter is in part based on [38] and [28].

8.1.2. Leptons

Uncertainties associated with leptons arise from the trigger, reconstruction, identification, and isolation, as well as the lepton momentum scale and resolution. Uncertainties arising from lepton momentum scale and resolution are discussed in chapter 10.

Trigger, reconstruction, identification, and isolation

The reconstruction, identification, and isolation efficiency of electrons and muons, as well as the efficiency of the trigger used to record the events, differ slightly between data and simulation, which is compensated for by dedicated scale factors (SFs). The efficiency scale factors are measured by using tag-and-probe techniques on $Z \mapsto l^+l^-$ data and simulated samples [39, 41]. The scale factors are then applied to the simulation to account for differences. This is done in the form of corrections to the MC event weight. In total, 4 components are included for electrons and 10 for muons.

8.1.3. Jets

Uncertainties associated with jets arise from the efficiency of pile-up rejection by the JVT, from the Jet Energy Scale (JES) and Resolution (JER), and from b-tagging. Uncertainties arising from JES and JER are discussed in chapter 10.

Jet Vertex Tagging

To account for differences in data and MC samples arising from JVT efficiencies, scale factors are applied. The efficiency scale factors are measured by using tag-and-probe techniques on $Z \mapsto \mu^+\mu^-$ data and simulated samples [48]. The effect of these SFs, as well as of their uncertainties, are propagated as corrections to the MC event weight.

B-tagging

B-tagging efficiencies in simulated samples are corrected to match efficiencies in data. As mentioned in Section 5.2.1, PC calibration is used in our analysis, therefore the 5-bins distribution of the DL1r score itself is corrected, instead of simply the efficiency for one given WP. For this correction, scale factors are derived from data [50, 79]. The b-tagging efficiencies, as well as the mistagging rates for c- and light-flavour jets are measured and the scale factors are calculated using the ratio between efficiencies in data and simulation. These scale factors are dependent on p_T and η and their uncertainties are propagated to

the analysis using 45 nuisance parameters for b-jets and 20 nuisance parameters for c-jets and light-jets respectively.

8.2. Theoretical uncertainties

8.2.1. Initial State Radiation

To estimate the impact of initial state radiation (ISR) modelling of the $t\bar{t}$ samples on the fit, three additional variations are defined. These variations are implemented as nuisance parameters in the fit. The three variations correspond to varying the strong coupling for QCD emission in the ISR by a factor of 0.549 (1.960) [80], which sets α_S^{ISR} to 0.140 (0.115) instead of the nominal 0.127. The other two variations correspond to varying the renormalisation scale and the factorisation using the parameters μ_R and μ_F, respectively. This is done by scaling these parameters independently by 2.0 (0.5) in respect to their nominal values to get their up (down) variations. The described strategy is also done for the single top samples.

8.2.2. Final State Radiation

To estimate the impact of final state radiation (FSR) modelling of the $t\bar{t}$ samples on the fit, one variation is defined, that is implemented as a nuisance parameter in the fit. It corresponds to varying the strong coupling for QCD emission in the FSR (α_S^{FSR}) by scaling the parameter μ_R^{FSR} by 2.0 (0.5) in respect to its nominal value to get its up and down variation. The described strategy is also done for the single top samples.

8.3. Pruning

The concept of pruning is the removal of systematics with negligible effects on the fit. This is done to reduce the complexity of the fit and the computational requirements. In this analysis systematics with an effect smaller than 0.5% on the normalisation are only considered for their impact on the shape, while systematics with an effect smaller than 0.5% on the shape are only considered for their impact on the normalisation. If both impacts are below their respective threshold, the systematic is excluded from the fit.

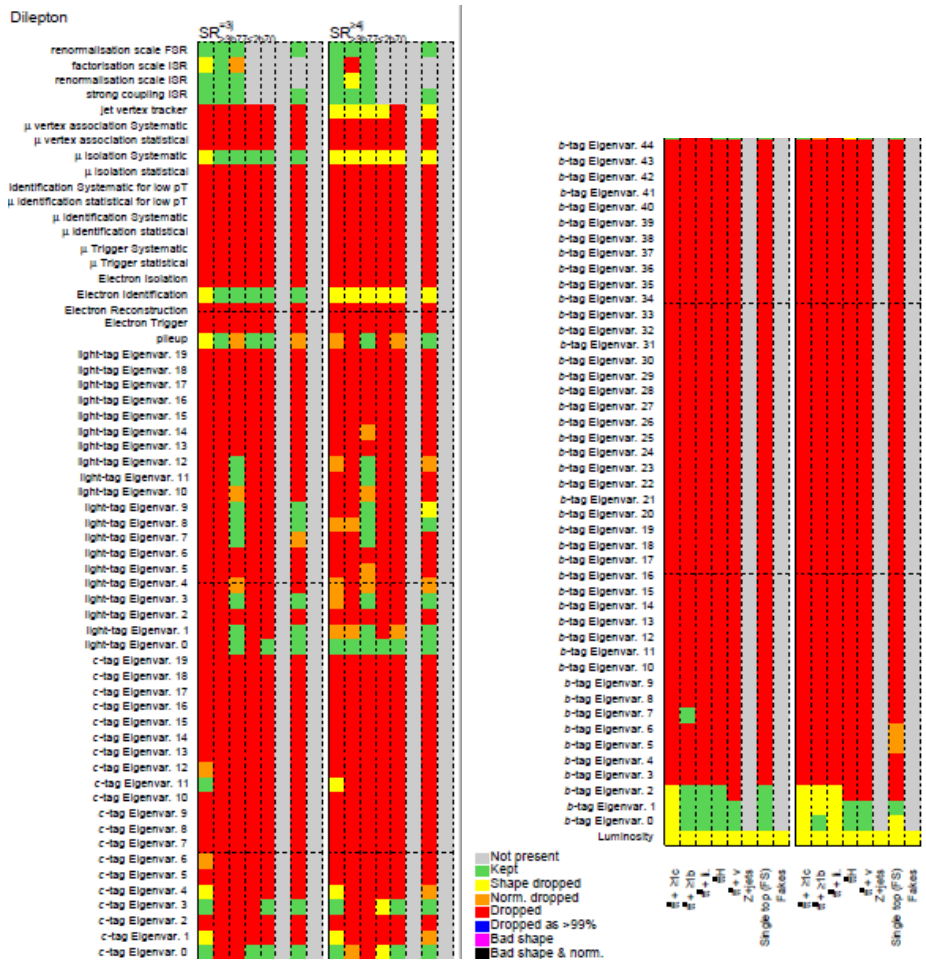

Figure 8.1.: Effect of pruning on the implemented systematics in both signal regions.

As can be seen in figure 8.1, the large majority of b-tagging systematics are pruned, while the modelling systematics related to initial and final state radiation are mostly kept. Systematics related to objects, such as electrons, muons and jets are mostly pruned, with notable exceptions corresponding to electron identification, muon isolation and pileup.

9. Results

This chapter presents first fit results. The signal regions described in section 7.1 are used and their pre-fit distributions can be seen in figure 9.1.

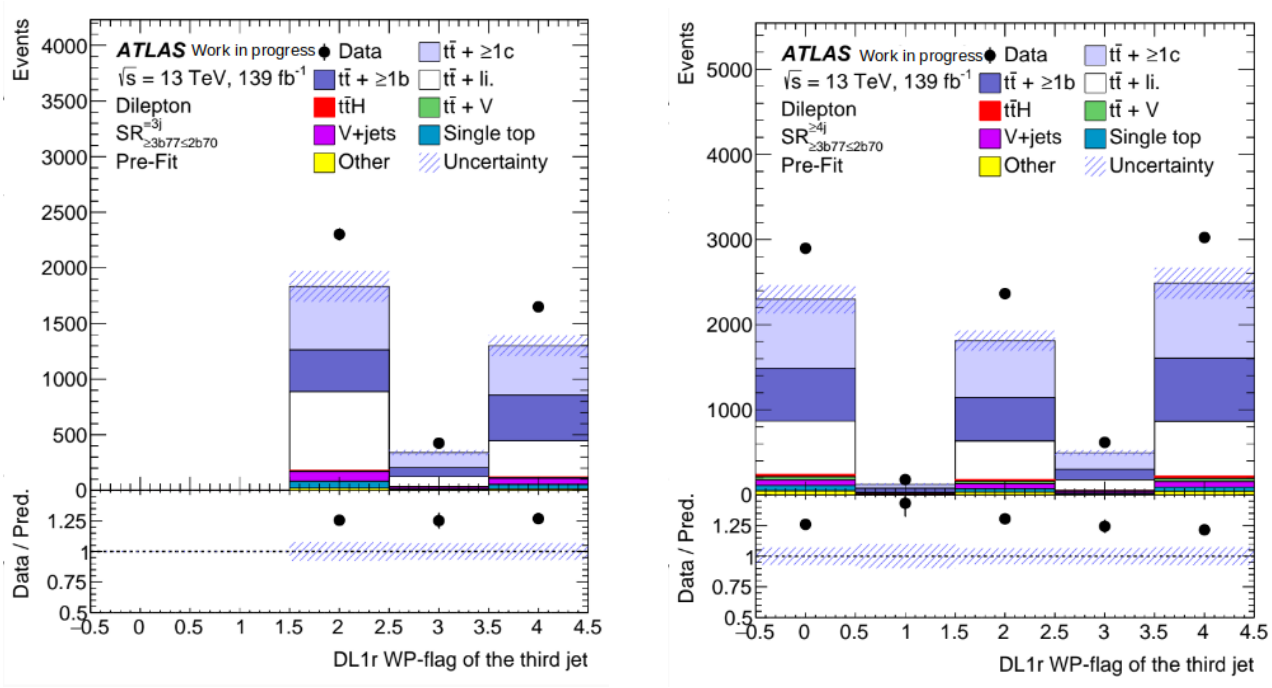

Figure 9.1.: Pre-fit distributions of the DL1r WP-flags for the third jet by p_T. Regions selected as described in section 7.1.2. On the left, region with exactly three jets. On the right, region with at least four jets. The total uncertainty (stat. + syst.) is shown and the first and last bin contain underflow and overflow events, respectively.

As can be seen, the total prediction does not agree with the data. This discrepancy is already present in the pre-selected sample, shown in figure 7.1, but it gets pronounced due to the low statistics in the selected region. However, the overall discrepancy is related almost only to the normalisation and not to the shape. The shown uncertainty bands take all systematics discussed in section 8.3 into account. Since not all relevant uncertainties are included, the uncertainty bands are undetermined.

9.1. Asimov fit

Before the MC prediction is fitted to data, the fit configuration has to be validated. A useful tool for this purpose is the Asimov fit. An Asimov fit fits the MC prediction to the nominal MC sample instead of data. This results in an assumed signal strength $\mu = 1.00$, while the nuisance parameters of the systematic uncertainties are set to zero.

9.1.1. Statistical only fit

A first estimation of the uncertainty of the fit in the selected signal regions can be obtained by investigating the statistical only fit by including only nuisance parameters corresponding to the statistical uncertainty of the signal and background samples. The signal strength can be extracted out of the fit and it results in $\mu = 1.00^{+0.33}_{-0.33}$. This corresponds to an uncertainty of 33%. This, as well as the values of the nuisance parameters corresponding to the main background samples can be seen in figure 9.2.

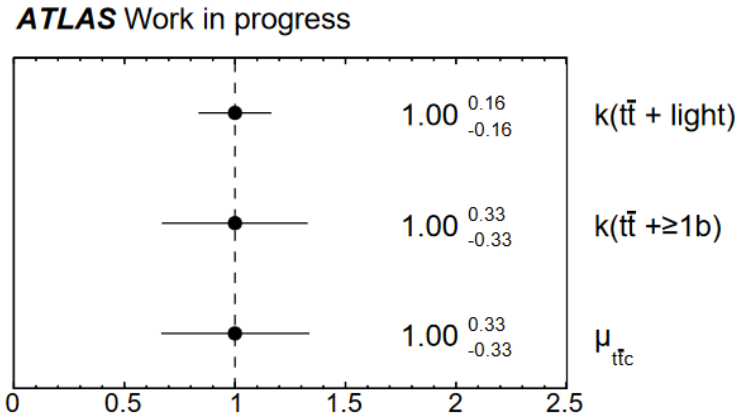

Figure 9.2.: Norm factors for the signal strength μ, as well as the statistical nuisance parameters of the main backgrounds in the statistical only fit.

By construction, the post-fit values of the signal strength and the statistical NPs of the main backgrounds are equal to one. The statistical uncertainties on the NPs are 16% for the $t\bar{t}$ + light sample and 33% for the $t\bar{t}+ \geq 1b$ sample. As such, it is expected, that the uncertainty on the $t\bar{t}+ \geq 1b$ sample will have a large impact on the uncertainty of the signal strength.

9.1.2. Inclusion of systematics

The NPs passing the pruning criteria discussed in section 8.3 are included into the fit. Since the MC prediction is fit to the pseudo-data consisting of the nominal MC samples,

the pull of the pseudo-data is expected to be zero. This can be seen in figure 9.3.

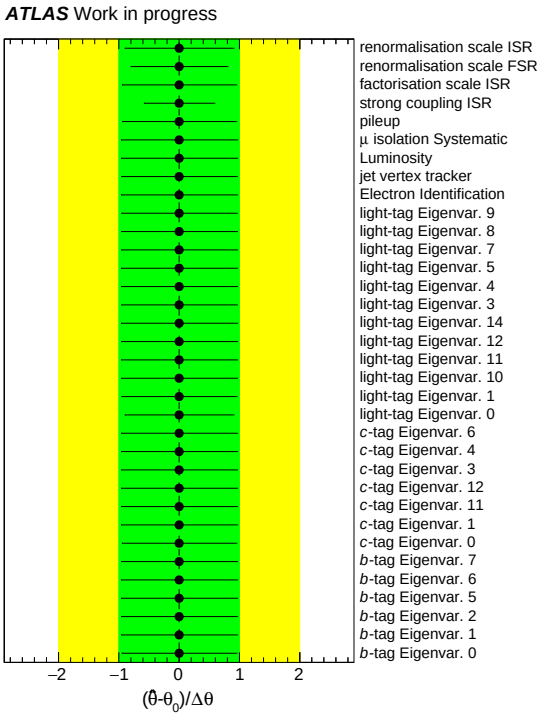

Figure 9.3.: Post-fit nuisance parameters in the Asimov fit. Shown are the 1σ and 2σ regions of the pre-fit uncertainty in green and yellow, respectively. The post-fit values and uncertainties of the NPs are shown in black.

As expected, all nuisance parameters take the post-fit value of zero and as can be seen, most of the NP uncertainties are not constrained by the pseudo-data. Notable exceptions are the modelling uncertainties corresponding to initial and final state radiation, with the NP corresponding to the uncertainty on the strong coupling in initial state radiation being constrained the strongest. The exact nature of these constraints is still to be investigated.

To investigate the impact of the individual nuisance parameters on the fitted signal strength, the nuisance parameters are ranked.

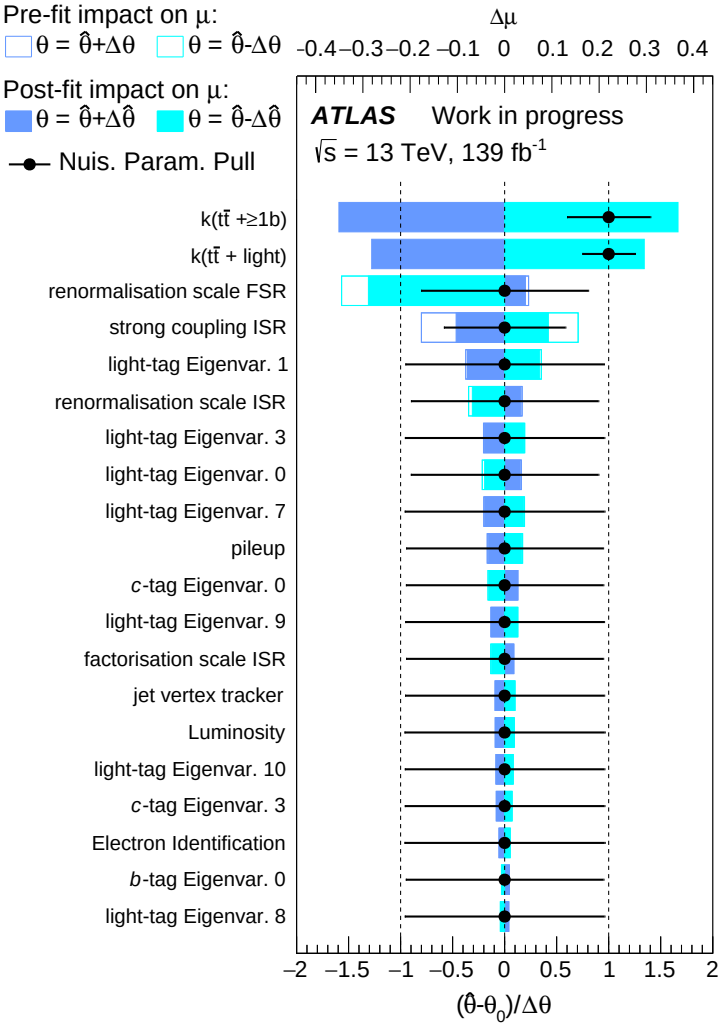

Figure 9.4.: Post-fit nuisance parameters in the Asimov fit ranked by their impact on the fitted signal strength. Constraints on the impact of uncertainties are shown with the fullness of the bars.

As shown in figure 9.4, the uncertainties on the main backgrounds have the highest impact on the signal strength. This is directly followed by modelling uncertainties related to initial and final state radiation. The next general group by impact on the signal strength are the light-mistag-rates, followed by c-tagging efficiencies and other instrumental uncertainties, such as pileup, luminosity, jet vertex tracking and with much smaller impact electron identification. The impact of b-tagging efficiencies can be safely said to be minimal. The only two NPs, that have their post-fit impact reduced are the two NPs, that are constrained by the pseudo-data. Additionally, the impact of the renormalisation scale in final state radiation is highly asymmetrical. An effect, that still needs to be investigated and possibly symmetrised. Since not all relevant systematic uncertainties are implemented, this ranking only represents an interim result and is subject to change.

The fit results can be plotted in the respective signal regions, which can be seen in figure 9.5.

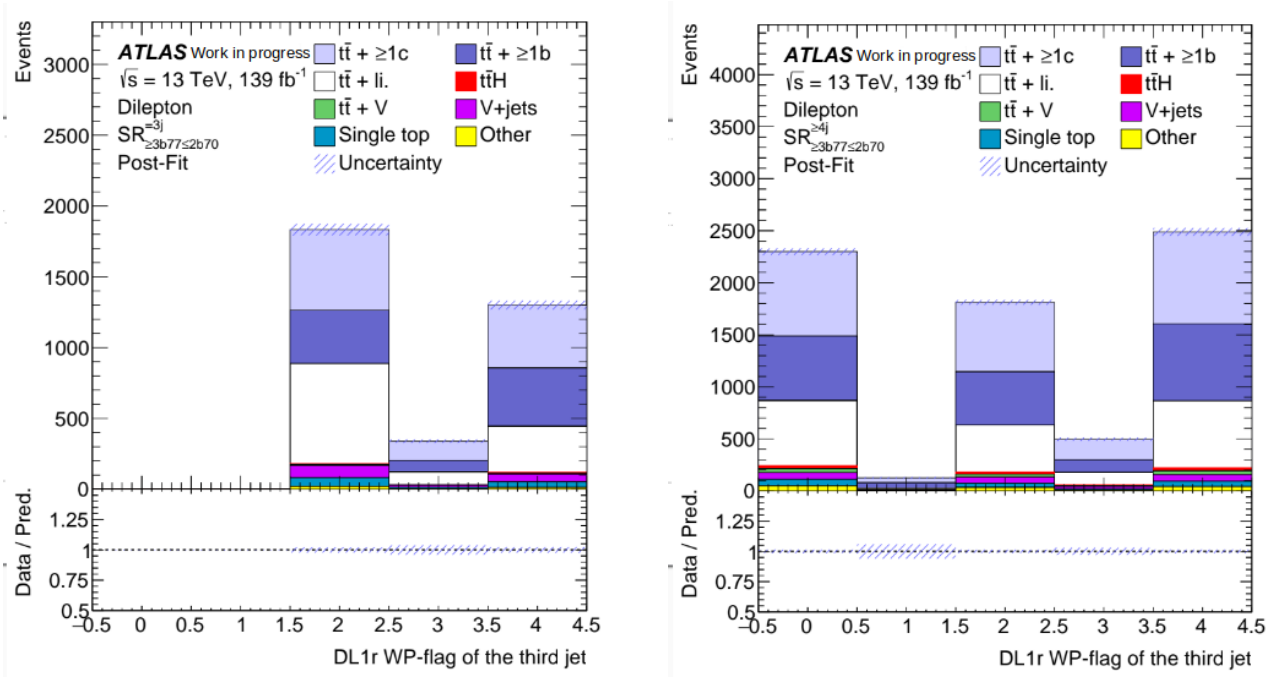

Figure 9.5.: Post-fit distributions of the DL1r WP-flags for the third jet by p_T. Regions selected as described in section 7.1.2. On the left, region with exactly three jets. On the right, region with at least four jets. The total uncertainty (stat. + syst.) is shown and the first and last bin contain underflow and overflow events, respectively.

When one compares the post-fit results with the pre-fit plot in figure 9.1, one finds a reduced uncertainty in the post-fit plot. This can be traced back to the constraints on the modelling uncertainties due to the pseudo-data.

An interim estimation of the uncertainty of the signal strength can be extracted from the fit. This results in a signal strength of $\mu = 1.00^{+0.41}_{-0.41}$, which corresponds to an uncertainty of 41% and can be seen in figure 9.6 along with the corresponding uncertainties on the main backgrounds.

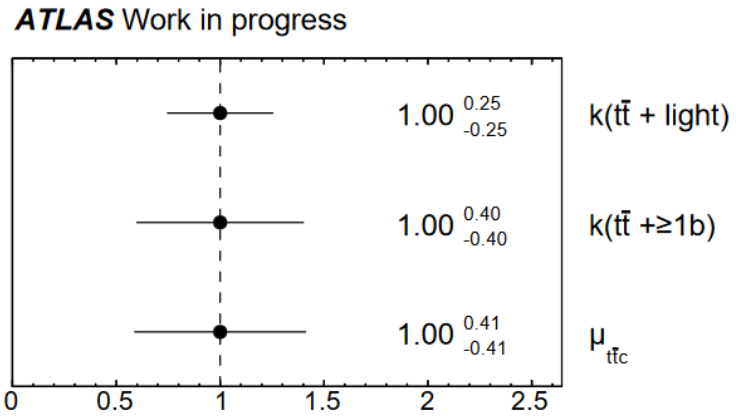

Figure 9.6.: Norm factors for the signal strength μ, as well as the interim nuisance parameters of the main backgrounds in the Asimov fit with limited systematic uncertainties.

By construction the post-fit values of the signal strength and the statistical NPs of the main backgrounds are equal to one. The interim uncertainties on the NPs are 25% for the $t\bar{t}$ + light sample and 40% for the $t\bar{t}+ \geq 1$b sample. Consistent with the ranking these uncertainties on the main backgrounds have a large impact on the uncertainty of the signal strength. By removing these NPs from the fit, the uncertainty of the signal strength gets lowered to approximately 17% ($\mu = 1.00^{+0.17}_{-0.16}$).

10. Summary and Outlook

10.1. Summary

The signal region found for $t\bar{t}+ \geq 1c$ events has been determined to be events with at least three jets at 77% WP and at most two jets at 70% WP. This results in a signal to background ratio of 0.508 ± 0.003 with a signal yield of 3700 ± 11. This signal region can then be split further between events with exactly three jets and events with at least four jets. Further requirements on the fourth jet do not result in a significant improvement of the expected signal region without much sacrifices to the yield.

Interim Asimov fits in the selected signal regions where conducted. The statistical only fit found an expected uncertainty of the signal strength μ of 33%. This result was found by including additional nuisance parameters for the main background types $t\bar{t}+ \geq 1b$ and $t\bar{t} + \text{light}$. An additional interim Asimov fit was conducted by including a number of systematic uncertainties corresponding to b-tagging efficiencies, object definitions and ISR and FSR modelling. From the corresponding fit the expected signal strength uncertainty was extracted, which was found to be 41%.

10.2. Outlook

This section attempts to give an overview over further steps, that can be taken in this analysis.

10.2.1. Full systematic implementation

The first step to continue this analysis, is to finish the implementation of the relevant systematic uncertainties in the fit. These uncertainties can be differentiated into experimental and theoretical uncertainties.

Experimental uncertainties

The number of uncertainties, yet to be implemented, in this category are limited. These uncertainties are related to energy scale and energy resolution for objects such as leptons and jets. These uncertainties originate from corrections made to the lepton and jet energy scales and resolution to the one observed in data. This is done via test beam data. The uncertainty due to energy scale is estimated via the $\pm 1\sigma$ variation, while the uncertainty due to energy resolution is done by smearing the lepton/jet momentum. The estimation of the lepton energy scale and resolution uncertainties is obtained via the reconstruction of the $Z \mapsto l^+l^-$ and $J/\Psi \mapsto l^+l^-$ processes by measuring the corresponding mass distribution. Additionally, the E/p ratio in $W \mapsto e\nu$ events is also investigated. The estimation of the uncertainty due to jet energy scale and resolution are obtained as described in Ref. [81].

Theoretical uncertainties

Parton showering and hadronisation The uncertainties related to parton showering and hadronisation can be estimated by comparing the nominal Pythia8 [67] sample with a sample produced with Herwig7 [82]. In both samples, the matrix element was generated using PowhegBox [68].

Top quark p_T reweighting Since only the NLO precision was used to model the top-quark p_T prediction, modelling uncertainties arise. This effect can be reduced by reweighting the samples to the NNLO QCD + NLO EW precision modelling. The effect due to this reweighting on the $t\bar{t}$ sample can be included into the fit as a nuisance parameter to estimate the uncertainty due to this mismodelling.

PDF uncertainties The PDFs used in the analysis are prone to uncertainties. These can be estimated using the PDF4LHC15 error set [83], which consist of 30 nuisance parameters. The uncertainties are estimated by reweighting the nominal prediction using the error set and propagating their impact to the variable used in the fit.

10.2.2. Pruning and symmetrization

Pruning

Due to the implementation of additional systematics and their possible influence on the already implemented uncertainties, as well as their in parts high expected impact, the prun-

ing procedure described in section 8.3 has to be revalidated, to check, if the used pruning cuts still do not significantly impact the expected uncertainty of the signal strength μ.

Symmetrization

A number of systematic uncertainties exhibit strongly one-sided behaviour when varying them by $\pm 1\sigma$. A good example for this behaviour is the impact of the renormalisation scale in the final state radiation. In these cases, a one-sided symmetrization could be implemented to better estimate the uncertainty due to this specific nuisance parameter. Such steps would have to be investigated and validated before their proper implementation.

10.2.3. Fit validation

Since not all systematic uncertainties are properly implemented in the presented Asimov fit in section 9.1, the fit validation as well as the investigation of the expected uncertainty due to the implemented nuisance parameters, as well as their ranking due to impact in an Asimov fit has to be redone. Especially a number of the theoretical uncertainties yet to be implemented, that are related to modelling such as generator uncertainties, are expected to have a significant impact on the uncertainty of the signal strength μ. Additionally, the observed restraints on the implemented uncertainties due to the pseudo-dataset have to be investigated and validated.

10.2.4. Fit to data

After the Asimov fit has been sufficiently validated and the possible constraints on the expected uncertainty of affected nuisance parameters are explained, the simulated samples will be fitted to data. With this the pull of actual data on the individual nuisance parameters can be investigated. Of special interest here are the pulls on the scale factors related to the strength of the main backgrounds as well as the signal strength. These directly indicate the observed discrepancy between data and the modelling of $\mathrm{t\bar{t}}$+jets events. Possible constraints on nuisance parameters have to be investigated again and to be compared to the results observed in the Asimov fit. The fit results can be validated by propagating the fitted nuisance parameters to variables not used in the fit for consistency checks. Additional consistency checks can then also be performed by splitting the used data set by production year and running the fit again. After a full validation of the observed results, the measured signal strength can then finally be converted into an observed cross-section, that can be compared with the results obtained by CMS [22].

A. Sample list

This appendix lists the full list of the DOAD containers used in this analysis. Section A.1 lists the data samples and section A.2 lists the Monte-Carlo samples.

A.1. Data samples

The DAOD data containers used in this analysis are shown in table A.1. They are skimmed using the selection criteria shown in section 4.2. The integrated luminosity is listed for each year.

Year	DAOD_TOPQ1 container	$\int L dt$ [pb^{-1}]
2015	data15_13TeV.AllYear.physics_Main.PhysCont.DAOD_TOPQ1.grp15_v01_p3794	3219.6
2016	data16_13TeV.AllYear.physics_Main.PhysCont.DAOD_TOPQ1.grp16_v01_p3794	32988.1
2017	data17_13TeV.AllYear.physics_Main.PhysCont.DAOD_TOPQ1.grp17_v01_p3794	44307.4
2018	data18_13TeV.AllYear.physics_Main.PhysCont.DAOD_TOPQ1.grp18_v01_p3794	58450.1

Table A.1.: Dataset used together with the corresponding total integrated luminosity.

A.2. Monte-Carlo samples

The MC TOPQ1 DAOD containers used in this analysis are shown in this section. The samples corresponding to each sub-campaign (mc16a, mc16d and mc16e) are listed for each DSID. The cross-section value used for normalisation is listed for each DSID and includes the branching ratio and/or filter efficiencies. The K-factor is also included and each cross-section is rounded to six significant digits. The cross-section values are stored in the common XSection-MC16-13TeV.data file used by the Top working group.

A.2.1. List of $t\bar{t}$ signal and background samples

The 5FS nominal $t\bar{t}$ samples are listed in this section. These samples are used to predict the nominal $t\bar{t}+ \geq 1c$ signal and the nominal $t\bar{t}+ \geq 1b$ and $t\bar{t}$ + light background. All of them are simulated using FS and are listed in table A.2.

Process/decay/filter		DAOD_TOPQ1 container	σ [pb]
$t\bar{t} \mapsto 1, 2l$	$t\bar{t}$	410470.PhPy8EG_A14_ttbar_hdamp258p75_nonallhad.deriv.DAOD_TOPQ1.e6337_s3126_r9364_p3832 410470.PhPy8EG_A14_ttbar_hdamp258p75_nonallhad.deriv.DAOD_TOPQ1.e6337_s3126_r10201_p3832 410470.PhPy8EG_A14_ttbar_hdamp258p75_nonallhad.deriv.DAOD_TOPQ1.e6337_s3126_r10724_p3832	452.352
$t\bar{t} \mapsto 1l$	$t\bar{t}$+BB	411073.PhPy8EG_A14_ttbar_hdamp258p75_ljets_BBFilt.deriv.DAOD_TOPQ1.e6798_s3126_r9364_p3832 411073.PhPy8EG_A14_ttbar_hdamp258p75_ljets_BBFilt.deriv.DAOD_TOPQ1.e6798_s3126_r10201_p3832 411073.PhPy8EG_A14_ttbar_hdamp258p75_ljets_BBFilt.deriv.DAOD_TOPQ1.e6798_s3126_r10724_p3832	3.34987
	$t\bar{t}$+B	411074.PhPy8EG_A14_ttbar_hdamp258p75_ljets_BFiltBBVeto.deriv.DAOD_TOPQ1.e6798_s3126_r9364_p3832 411074.PhPy8EG_A14_ttbar_hdamp258p75_ljets_BFiltBBVeto.deriv.DAOD_TOPQ1.e6798_s3126_r10201_p3832 411074.PhPy8EG_A14_ttbar_hdamp258p75_ljets_BFiltBBVeto.deriv.DAOD_TOPQ1.e6798_s3126_r10724_p3832	17.6669
	$t\bar{t}$+C	411075.PhPy8EG_A14_ttbar_hdamp258p75_ljets_CFiltBVeto.deriv.DAOD_TOPQ1.e6798_s3126_r9364_p3832 411075.PhPy8EG_A14_ttbar_hdamp258p75_ljets_CFiltBVeto.deriv.DAOD_TOPQ1.e6798_s3126_r10201_p3832 411075.PhPy8EG_A14_ttbar_hdamp258p75_ljets_CFiltBVeto.deriv.DAOD_TOPQ1.e6798_s3126_r10724_p3832	15.0340
$t\bar{t} \mapsto 2l$	$t\bar{t}$+BB	411076.PhPy8EG_A14_ttbar_hdamp258p75_dil_BBFilt.deriv.DAOD_TOPQ1.e6798_s3126_r9364_p3832 411076.PhPy8EG_A14_ttbar_hdamp258p75_dil_BBFilt.deriv.DAOD_TOPQ1.e6798_s3126_r10201_p3832 411076.PhPy8EG_A14_ttbar_hdamp258p75_dil_BBFilt.deriv.DAOD_TOPQ1.e6798_s3126_r10724_p3832	0.768225
	$t\bar{t}$+B	411077.PhPy8EG_A14_ttbar_hdamp258p75_dil_BFiltBBVeto.deriv.DAOD_TOPQ1.e6798_s3126_r9364_p3832 411077.PhPy8EG_A14_ttbar_hdamp258p75_dil_BFiltBBVeto.deriv.DAOD_TOPQ1.e6798_s3126_r10201_p3832 411077.PhPy8EG_A14_ttbar_hdamp258p75_dil_BFiltBBVeto.deriv.DAOD_TOPQ1.e6798_s3126_r10724_p3832	4.08276
	$t\bar{t}$+C	411078.PhPy8EG_A14_ttbar_hdamp258p75_dil_CFiltBVeto.deriv.DAOD_TOPQ1.e6798_s3126_r9364_p3832 411078.PhPy8EG_A14_ttbar_hdamp258p75_dil_CFiltBVeto.deriv.DAOD_TOPQ1.e6798_s3126_r10201_p3832 411078.PhPy8EG_A14_ttbar_hdamp258p75_dil_CFiltBVeto.deriv.DAOD_TOPQ1.e6798_s3126_r10724_p3832	3.45785

Table A.2.: List of all DAOD containers for PowhegBox+Pythia8 $t\bar{t}$ (5FS) MC samples used in this analysis. The prefix mc16_13TeV, present in all samples, is removed to reduce the table width. The samples are simulated in FS. The shown containers are used for the nominal $t\bar{t}+ \geq 1c$ signal and the nominal $t\bar{t}+ \geq 1b$ and $t\bar{t}$ + light background estimation.

A.2.2. List of smaller top background MC samples

Table **??**, A.4 and A.5 list the MC samples used to predict the single top background, while Table A.6 lists the MC samples used to predict the $t\bar{t}H$ background. All of the samples are simulated using PowhegBox+Pythia8 in FS. Table A.7 lists the MC samples used to predict the $t\bar{t}V$ background. The samples listed here are simulated using MadGraph5_aMC@NLO + Pythia8 in FS.

Process/decay/filter	DAOD_TOPQ1 container	σ [pb]
$tW^- \mapsto$ all	410646.PowhegPythia8EvtGen_A14_Wt_DR_inclusive_top.deriv.DAOD_TOPQ1.e6552_s3126_r9364_p3832 410646.PowhegPythia8EvtGen_A14_Wt_DR_inclusive_top.deriv.DAOD_TOPQ1.e6552_s3126_r10201_p3832 410646.PowhegPythia8EvtGen_A14_Wt_DR_inclusive_top.deriv.DAOD_TOPQ1.e6552_s3126_r10724_p3832	35.8495
$tW^+ \mapsto$ all	410647.PowhegPythia8EvtGen_A14_Wt_DR_inclusive_antitop.deriv.DAOD_TOPQ1.e6552_s3126_r9364_p3832 410647.PowhegPythia8EvtGen_A14_Wt_DR_inclusive_antitop.deriv.DAOD_TOPQ1.e6552_s3126_r10201_p3832 410647.PowhegPythia8EvtGen_A14_Wt_DR_inclusive_antitop.deriv.DAOD_TOPQ1.e6552_s3126_r10724_p3832	35.8591

Table A.3.: List of all DAOD containers for PowhegBox+Pythia8 tW MC samples used in this analysis. The prefix mc16_13TeV, present in all samples, is removed to reduce the table width. The samples are simulated in FS.

Process/decay/filter	DAOD_TOPQ1 container	σ [pb]
$t \mapsto 1l$	410658.PowhegPythia8EvtGen_A14_tchan_DR_inclusive_top.deriv.DAOD_TOPQ1.e6552_s3126_r9364_p3832 410658.PowhegPythia8EvtGen_A14_tchan_DR_inclusive_top.deriv.DAOD_TOPQ1.e6552_s3126_r10201_p3832 410658.PowhegPythia8EvtGen_A14_tchan_DR_inclusive_top.deriv.DAOD_TOPQ1.e6552_s3126_r10724_p3832	44.0587
$\bar{t} \mapsto 1l$	410659.PowhegPythia8EvtGen_A14_tchan_DR_inclusive_antitop.deriv.DAOD_TOPQ1.e6552_s3126_r9364_p3832 410659.PowhegPythia8EvtGen_A14_tchan_DR_inclusive_antitop.deriv.DAOD_TOPQ1.e6552_s3126_r10201_p3832 410659.PowhegPythia8EvtGen_A14_tchan_DR_inclusive_antitop.deriv.DAOD_TOPQ1.e6552_s3126_r10724_p3832	26.2330

Table A.4.: List of all DAOD containers for PowhegBox+Pythia8 t channel MC samples used in this analysis. The prefix mc16_13TeV, present in all samples, is removed to reduce the table width. The samples are simulated in FS.

Process/decay/filter	DAOD_TOPQ1 container	σ [pb]
$t \mapsto 1l$	410644.PowhegPythia8EvtGen_A14_schan_DR_inclusive_top.deriv.DAOD_TOPQ1.e6552_s3126_r9364_p3832 410644.PowhegPythia8EvtGen_A14_schan_DR_inclusive_top.deriv.DAOD_TOPQ1.e6552_s3126_r10201_p3832 410644.PowhegPythia8EvtGen_A14_schan_DR_inclusive_top.deriv.DAOD_TOPQ1.e6552_s3126_r10724_p3832	2.05720
$\bar{t} \mapsto 1l$	410645.PowhegPythia8EvtGen_A14_schan_DR_inclusive_antitop.deriv.DAOD_TOPQ1.e6552_s3126_r9364_p3832 410645.PowhegPythia8EvtGen_A14_schan_DR_inclusive_antitop.deriv.DAOD_TOPQ1.e6552_s3126_r10201_p3832 410645.PowhegPythia8EvtGen_A14_schan_DR_inclusive_antitop.deriv.DAOD_TOPQ1.e6552_s3126_r10724_p3832	1.28661

Table A.5.: List of all DAOD containers for PowhegBox+Pythia8 s channel MC samples used in this analysis. The prefix mc16_13TeV, present in all samples, is removed to reduce the table width. The samples are simulated in FS.

Process/decay/filter	DAOD_TOPQ1 container	σ [pb]
$t\bar{t} \mapsto 2l$	346345.PhPy8EG_A14NNPDF23_NNPDF30ME_ttH125_dilep.deriv.DAOD_TOPQ1.e7148_s3126_r9364_p3832 346345.PhPy8EG_A14NNPDF23_NNPDF30ME_ttH125_dilep.deriv.DAOD_TOPQ1.e7148_s3126_r10201_p3832 346345.PhPy8EG_A14NNPDF23_NNPDF30ME_ttH125_dilep.deriv.DAOD_TOPQ1.e7148_s3126_r10724_p3832	53.43
$t\bar{t} \mapsto 1l$	346344.PhPy8EG_A14NNPDF23_NNPDF30ME_ttH125_dilep.deriv.DAOD_TOPQ1.e7148_s3126_r9364_p3832 346344.PhPy8EG_A14NNPDF23_NNPDF30ME_ttH125_dilep.deriv.DAOD_TOPQ1.e7148_s3126_r10201_p3832 346344.PhPy8EG_A14NNPDF23_NNPDF30ME_ttH125_dilep.deriv.DAOD_TOPQ1.e7148_s3126_r10724_p3832	222.76
$t\bar{t} \mapsto 0l$	346343.PhPy8EG_A14NNPDF23_NNPDF30ME_ttH125_dilep.deriv.DAOD_TOPQ1.e7148_s3126_r9364_p3832 346343.PhPy8EG_A14NNPDF23_NNPDF30ME_ttH125_dilep.deriv.DAOD_TOPQ1.e7148_s3126_r10201_p3832 346343.PhPy8EG_A14NNPDF23_NNPDF30ME_ttH125_dilep.deriv.DAOD_TOPQ1.e7148_s3126_r10724_p3832	230.82

Table A.6.: List of all DAOD containers for PowhegBox+Pythia8 $t\bar{t}H$ MC samples used in this analysis. The prefix mc16_13TeV, present in all samples, is removed to reduce the table width. The samples are simulated in FS.

Process/decay/filter		DAOD_TOPQ1 container	σ [pb]
$t\bar{t}W$	$\mapsto$ all	410155.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttW.deriv.DAOD_TOPQ1.e5070_s3126_r9364_p3832 410155.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttW.deriv.DAOD_TOPQ1.e5070_s3126_r10201_p3832 410155.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttW.deriv.DAOD_TOPQ1.e5070_s3126_r10724_p3832	603.130
$t\bar{t}Z$	$Z \mapsto ee$	410218.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttee.deriv.DAOD_TOPQ1.e5070_s3126_r9364_p3832 410218.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttee.deriv.DAOD_TOPQ1.e5070_s3126_r10201_p3832 410218.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttee.deriv.DAOD_TOPQ1.e5070_s3126_r10724_p3832	41.315
	$Z \mapsto ee$ low m	410276.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttee_mll_1_5.deriv.DAOD_TOPQ1.e6087_s3126_r9364_p3832 410276.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttee_mll_1_5.deriv.DAOD_TOPQ1.e6087_s3126_r10201_p3832 410276.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttee_mll_1_5.deriv.DAOD_TOPQ1.e6087_s3126_r10724_p3832	18.4
	$Z \mapsto \mu\mu$	410219.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttmumu.deriv.DAOD_TOPQ1.e5070_s3126_r9364_p3832 410219.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttmumu.deriv.DAOD_TOPQ1.e5070_s3126_r10201_p3832 410219.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttmumu.deriv.DAOD_TOPQ1.e5070_s3126_r10724_p3832	41.3224
	$Z \mapsto \mu\mu$ low m	410277.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttmumu_mll_1_5.deriv.DAOD_TOPQ1.e6087_s3126_r9364_p3832 410277.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttmumu_mll_1_5.deriv.DAOD_TOPQ1.e6087_s3126_r10201_p3832 410277.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttmumu_mll_1_5.deriv.DAOD_TOPQ1.e6087_s3126_r10724_p3832	18.4
	$Z \mapsto \tau\tau$	410220.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_tttautau.deriv.DAOD_TOPQ1.e5070_s3126_r9364_p3832 410220.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_tttautau.deriv.DAOD_TOPQ1.e5070_s3126_r10201_p3832 410220.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_tttautau.deriv.DAOD_TOPQ1.e5070_s3126_r10724_p3832	40.9909
	$Z \mapsto \tau\tau$ low m	410278.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_tttautau_mll_1_5.deriv.DAOD_TOPQ1.e6087_s3126_r9364_p3832 410278.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_tttautau_mll_1_5.deriv.DAOD_TOPQ1.e6087_s3126_r10201_p3832 410278.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_tttautau_mll_1_5.deriv.DAOD_TOPQ1.e6087_s3126_r10724_p3832	1.97
	$Z \mapsto \nu\nu$	410156.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttZnunu.deriv.DAOD_TOPQ1.e5070_s3126_r9364_p3832 410156.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttZnunu.deriv.DAOD_TOPQ1.e5070_s3126_r10201_p3832 410156.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttZnunu.deriv.DAOD_TOPQ1.e5070_s3126_r10724_p3832	172.039
	$Z \mapsto qq$	410157.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttZqq.deriv.DAOD_TOPQ1.e5070_s3126_r9364_p3832 410157.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttZqq.deriv.DAOD_TOPQ1.e5070_s3126_r10201_p3832 410157.aMcAtNloPythia8EvtGen_MEN30NLO_A14N23LO_ttZqq.deriv.DAOD_TOPQ1.e5070_s3126_r10724_p3832	585.758

Table A.7.: List of all DAOD containers for MadGraph5_aMC@NLO +Pythia8 $t\bar{t}V$ MC samples used in this analysis. The prefix mc16_13TeV, present in all samples, is removed to reduce the table width. The samples are simulated in FS.

A.2.3. List of V + jets background MC samples

In this section the Sherpa samples, used to predict the V + jets background are listed. Table A.8, A.9 and A.10 list the samples corresponding to the $W \mapsto l\nu$ process, while table A.11, A.12 and A.13 list the samples corresponding to the $Z \mapsto ll$ process.

DAOD_TOPQ1 container	σ [pb]
364170.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364170.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	15299.7
364170.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364171.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364171.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	2418.36
364171.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364172.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364172.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	819.468
364172.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV0_70_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364173.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364173.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	611.539
364173.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364174.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364174.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	209.068
364174.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364175.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364175.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	94.8254
364175.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV70_140_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364176.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364176.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	196.791
364176.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364177.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364177.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	95.5097
364177.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364178.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364178.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	35.8940
364178.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV140_280_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364179.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364179.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	38.0738
364179.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364180.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364180.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	22.1657
364180.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364181.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364181.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	9.36888
364181.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV280_500_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364182.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV500_1000.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364182.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV500_1000.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	14.7703
364182.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV500_1000.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364183.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV1000_E_CMS.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364183.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV1000_E_CMS.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	1.19664
364183.Sherpa_221_NNPDF30NNLO_Wenu_MAXHTPTV1000_E_CMS.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	

Table A.8.: List of all DAOD containers for $W \mapsto e\nu$ Sherpa MC samples used in this analysis. The prefix mc16_13TeV, present in all samples, is removed to reduce the table width. The samples are simulated in FS. The samples are split in different filters, requiring the presents of specific hadrons, and in different slices of $\mathrm{Max}(H_T, p_T(V))$. The slice and filter used in each container is transparent in the container name.

DAOD_TOPQ1 container	σ [pb]
364156.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364156.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	15300.1
364156.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364157.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364157.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	2419.08
364157.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364158.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364158.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	819.041
364158.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV0_70_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364159.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364159.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	618.429
364159.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364160.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364160.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	213.411
364160.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364161.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364161.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	69.3299
364161.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV70_140_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364162.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364162.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	206.221
364162.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364163.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364163.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	95.5037
364163.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364164.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364164.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	35.8147
364164.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV140_280_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364165.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364165.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	38.2089
364165.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364166.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364166.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	22.2349
364166.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364167.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364167.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	9.32230
364167.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV280_500_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364168.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV500_1000.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364168.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV500_1000.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	14.5627
364168.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV500_1000.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364169.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV1000_E_CMS.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364169.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV1000_E_CMS.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	1.19761
364169.Sherpa_221_NNPDF30NNLO_Wmunu_MAXHTPTV1000_E_CMS.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	

Table A.9.: List of all DAOD containers for $W \mapsto \mu\nu$ Sherpa MC samples used in this analysis. The prefix mc16_13TeV, present in all samples, is removed to reduce the table width. The samples are simulated in FS. The samples are split in different filters, requiring the presents of specific hadrons, and in different slices of $\mathrm{Max}(H_T, p_T(V))$. The slice and filter used in each container is transparent in the container name.

DAOD_TOPQ1 container	σ [pb]
364184.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364184.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	15328.6
364184.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364185.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364185.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	2403.43
364185.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364186.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364186.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	829.089
364186.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV0_70_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364187.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364187.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	619.517
364187.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364188.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364188.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	204.113
364188.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364189.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364189.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	95.0974
364189.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV70_140_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364190.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364190.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	196.304
364190.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364191.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364191.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	95.6400
364191.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364192.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV140_280_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364192.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV140_280_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	38.8685
364192.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV140_280_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364193.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364193.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	38.1533
364193.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364194.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364194.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	22.1002
364194.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364195.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364195.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	9.38204
364195.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV280_500_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364196.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV500_1000.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364196.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV500_1000.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	14.5976
364196.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV500_1000.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364197.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV1000_E_CMS.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364197.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV1000_E_CMS.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	1.19713
364197.Sherpa_221_NNPDF30NNLO_Wtaunu_MAXHTPTV1000_E_CMS.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	

Table A.10.: List of all DAOD containers for $W \mapsto \tau\nu$ Sherpa MC samples used in this analysis. The prefix mc16_13TeV, present in all samples, is removed to reduce the table width. The samples are simulated in FS. The samples are split in different filters, requiring the presents of specific hadrons, and in different slices of $\mathrm{Max}(H_T, p_T(V))$. The slice and filter used in each container is transparent in the container name.

DAOD_TOPQ1 container	σ [pb]
364114.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364114.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	1586.66
364114.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364115.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364115.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	218.160
364115.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364116.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364116.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	123.302
364116.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364117.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364117.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	74.3928
364117.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364118.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364118.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	19.8296
364118.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364119.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364119.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	12.3085
364119.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364120.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364120.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	24.4068
364120.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364121.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364121.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	9.13863
364121.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364122.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364122.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	5.93117
364122.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364123.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364123.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	4.74799
364123.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364124.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364124.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	2.22321
364124.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364125.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364125.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	1.45716
364125.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364126.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV500_1000.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364126.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV500_1000.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	1.76308
364126.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV500_1000.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364127.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV1000_E_CMS.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364127.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV1000_E_CMS.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	0.144871
364127.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV1000_E_CMS.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364204.Sherpa_221_NN30NNLO_Zee_Mll10_40_MAXHTPTV0_70_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r9364_p3830	
364204.Sherpa_221_NN30NNLO_Zee_Mll10_40_MAXHTPTV0_70_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r10201_p3830	2273.18
364204.Sherpa_221_NN30NNLO_Zee_Mll10_40_MAXHTPTV0_70_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r10724_p3830	
364205.Sherpa_221_NN30NNLO_Zee_Mll10_40_MAXHTPTV0_70_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r9364_p3830	
364205.Sherpa_221_NN30NNLO_Zee_Mll10_40_MAXHTPTV0_70_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r10201_p3830	79.3319
364205.Sherpa_221_NN30NNLO_Zee_Mll10_40_MAXHTPTV0_70_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r10724_p3830	
364206.Sherpa_221_NN30NNLO_Zee_Mll10_40_MAXHTPTV70_280_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r9364_p3830	
364206.Sherpa_221_NN30NNLO_Zee_Mll10_40_MAXHTPTV70_280_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r10201_p3830	43.8516
364206.Sherpa_221_NN30NNLO_Zee_Mll10_40_MAXHTPTV70_280_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r10724_p3830	
364207.Sherpa_221_NN30NNLO_Zee_Mll10_40_MAXHTPTV70_280_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r9364_p3830	
364207.Sherpa_221_NN30NNLO_Zee_Mll10_40_MAXHTPTV70_280_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r10201_p3830	5.34493
364207.Sherpa_221_NN30NNLO_Zee_Mll10_40_MAXHTPTV70_280_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r10724_p3830	
364208.Sherpa_221_NN30NNLO_Zee_Mll10_40_MAXHTPTV280_E_CMS_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r9364_p3830	
364208.Sherpa_221_NN30NNLO_Zee_Mll10_40_MAXHTPTV280_E_CMS_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r10201_p3830	2.70825
364208.Sherpa_221_NN30NNLO_Zee_Mll10_40_MAXHTPTV280_E_CMS_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r10724_p3830	
364209.Sherpa_221_NN30NNLO_Zee_Mll10_40_MAXHTPTV280_E_CMS_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r9364_p3830	
364209.Sherpa_221_NN30NNLO_Zee_Mll10_40_MAXHTPTV280_E_CMS_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r10201_p3830	0.461307
364209.Sherpa_221_NN30NNLO_Zee_Mll10_40_MAXHTPTV280_E_CMS_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r10724_p3830	

Table A.11.: List of all DAOD containers for $Z \mapsto ee$ Sherpa MC samples used in this analysis. The prefix mc16_13TeV, present in all samples, is removed to reduce the table width. The samples are simulated in FS. The samples are split in different filters, requiring the presents of specific hadrons, and in different slices of $\mathrm{Max}(H_T, p_T(V))$. The slice and filter used in each container is transparent in the container name.

DAOD_TOPQ1 container	σ [pb]
364100.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364100.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	1589.63
364100.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364101.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364101.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	218.147
364101.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364102.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364102.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	124.013
364102.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364103.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364103.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	73.1486
364103.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364104.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364104.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	19.8411
364104.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364105.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364105.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	12.0800
364105.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364106.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364106.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	23.6806
364106.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364107.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364107.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	9.04446
364107.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364108.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364108.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	5.86387
364108.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364109.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364109.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	4.65413
364109.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364110.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364110.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	2.20916
364110.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364111.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364111.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	1.45419
364111.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364112.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV500_1000.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364112.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV500_1000.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	1.74358
364112.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV500_1000.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364113.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV1000_E_CMS.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364113.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV1000_E_CMS.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	0.144013
364113.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV1000_E_CMS.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364198.Sherpa_221_NN30NNLO_Zmumu_Mll10_40_MAXHTPTV0_70_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r9364_p3830	
364198.Sherpa_221_NN30NNLO_Zmumu_Mll10_40_MAXHTPTV0_70_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r10201_p3830	2272.17
364198.Sherpa_221_NN30NNLO_Zmumu_Mll10_40_MAXHTPTV0_70_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r10724_p3830	
364199.Sherpa_221_NN30NNLO_Zmumu_Mll10_40_MAXHTPTV0_70_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r9364_p3830	
364199.Sherpa_221_NN30NNLO_Zmumu_Mll10_40_MAXHTPTV0_70_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r10201_p3830	80.2086
364199.Sherpa_221_NN30NNLO_Zmumu_Mll10_40_MAXHTPTV0_70_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r10724_p3830	
364200.Sherpa_221_NN30NNLO_Zmumu_Mll10_40_MAXHTPTV70_280_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r9364_p3830	
364200.Sherpa_221_NN30NNLO_Zmumu_Mll10_40_MAXHTPTV70_280_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r10201_p3830	43.7616
364200.Sherpa_221_NN30NNLO_Zmumu_Mll10_40_MAXHTPTV70_280_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r10724_p3830	
364201.Sherpa_221_NN30NNLO_Zmumu_Mll10_40_MAXHTPTV70_280_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r9364_p3830	
364201.Sherpa_221_NN30NNLO_Zmumu_Mll10_40_MAXHTPTV70_280_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r10201_p3830	4.98763
364201.Sherpa_221_NN30NNLO_Zmumu_Mll10_40_MAXHTPTV70_280_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r10724_p3830	
364202.Sherpa_221_NN30NNLO_Zmumu_Mll10_40_MAXHTPTV280_E_CMS_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r9364_p3830	
364202.Sherpa_221_NN30NNLO_Zmumu_Mll10_40_MAXHTPTV280_E_CMS_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r10201_p3830	2.69107
364202.Sherpa_221_NN30NNLO_Zmumu_Mll10_40_MAXHTPTV280_E_CMS_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r10724_p3830	
364203.Sherpa_221_NN30NNLO_Zmumu_Mll10_40_MAXHTPTV280_E_CMS_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r9364_p3830	
364203.Sherpa_221_NN30NNLO_Zmumu_Mll10_40_MAXHTPTV280_E_CMS_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r10201_p3830	0.460399
364203.Sherpa_221_NN30NNLO_Zmumu_Mll10_40_MAXHTPTV280_E_CMS_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r10724_p3830	

Table A.12.: List of all DAOD containers for $Z \mapsto \mu\mu$ Sherpa MC samples used in this analysis. The prefix mc16_13TeV, present in all samples, is removed to reduce the table width. The samples are simulated in FS. The samples are split in different filters, requiring the presents of specific hadrons, and in different slices of $\mathrm{Max}(H_T, p_T(V))$. The slice and filter used in each container is transparent in the container name.

DAOD_TOPQ1 container	σ [pb]
364128.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364128.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	1587.20
364128.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364129.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364129.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	218.307
364129.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364130.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364130.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	124.552
364130.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364131.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364131.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	74.1331
364131.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364132.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364132.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	19.7090
364132.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364133.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364133.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	11.9878
364133.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364134.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364134.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	24.1858
364134.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364135.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364135.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	9.09596
364135.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364136.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364136.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	5.34266
364136.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364137.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364137.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	4.67189
364137.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364138.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364138.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	2.21896
364138.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364139.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364139.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	1.46542
364139.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_BBFilter.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364140.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV500_1000.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364140.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV500_1000.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	1.76454
364140.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV500_1000.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364141.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV1000_E_CMS.deriv.DAOD_TOPQ1.e5340_s3126_r9364_p3830	
364141.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV1000_E_CMS.deriv.DAOD_TOPQ1.e5340_s3126_r10201_p3830	0.144646
364141.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV1000_E_CMS.deriv.DAOD_TOPQ1.e5340_s3126_r10724_p3830	
364210.Sherpa_221_NN30NNLO_Ztautau_Mll10_40_MAXHTPTV0_70_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r9364_p3830	
364210.Sherpa_221_NN30NNLO_Ztautau_Mll10_40_MAXHTPTV0_70_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r10201_p3830	2275.81
364210.Sherpa_221_NN30NNLO_Ztautau_Mll10_40_MAXHTPTV0_70_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r10724_p3830	
364211.Sherpa_221_NN30NNLO_Ztautau_Mll10_40_MAXHTPTV0_70_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r9364_p3830	
364211.Sherpa_221_NN30NNLO_Ztautau_Mll10_40_MAXHTPTV0_70_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r10201_p3830	79.0832
364211.Sherpa_221_NN30NNLO_Ztautau_Mll10_40_MAXHTPTV0_70_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r10724_p3830	
364212.Sherpa_221_NN30NNLO_Ztautau_Mll10_40_MAXHTPTV70_280_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r9364_p3830	
364212.Sherpa_221_NN30NNLO_Ztautau_Mll10_40_MAXHTPTV70_280_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r10201_p3830	43.7204
364212.Sherpa_221_NN30NNLO_Ztautau_Mll10_40_MAXHTPTV70_280_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r10724_p3830	
364213.Sherpa_221_NN30NNLO_Ztautau_Mll10_40_MAXHTPTV70_280_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r9364_p3830	
364213.Sherpa_221_NN30NNLO_Ztautau_Mll10_40_MAXHTPTV70_280_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r10201_p3830	5.40297
364213.Sherpa_221_NN30NNLO_Ztautau_Mll10_40_MAXHTPTV70_280_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r10724_p3830	
364214.Sherpa_221_NN30NNLO_Ztautau_Mll10_40_MAXHTPTV280_E_CMS_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r9364_p3830	
364214.Sherpa_221_NN30NNLO_Ztautau_Mll10_40_MAXHTPTV280_E_CMS_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r10201_p3830	2.72399
364214.Sherpa_221_NN30NNLO_Ztautau_Mll10_40_MAXHTPTV280_E_CMS_BVeto.deriv.DAOD_TOPQ1.e5421_s3126_r10724_p3830	
364215.Sherpa_221_NN30NNLO_Ztautau_Mll10_40_MAXHTPTV280_E_CMS_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r9364_p3830	
364215.Sherpa_221_NN30NNLO_Ztautau_Mll10_40_MAXHTPTV280_E_CMS_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r10201_p3830	0.458025
364215.Sherpa_221_NN30NNLO_Ztautau_Mll10_40_MAXHTPTV280_E_CMS_BFilter.deriv.DAOD_TOPQ1.e5421_s3126_r10724_p3830	

Table A.13.: List of all DAOD containers for $Z \mapsto \tau\tau$ Sherpa MC samples used in this analysis. The prefix mc16_13TeV, present in all samples, is removed to reduce the table width. The samples are simulated in FS. The samples are split in different filters, requiring the presents of specific hadrons, and in different slices of $\mathrm{Max}(H_T, p_T(V))$. The slice and filter used in each container is transparent in the container name.

Bibliography

[1] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22**, 579 (1961)

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

[3] A. Salam, *Weak and Electromagnetic Interactions*, ed. Nobel Symposium No. 8 (Almqvist & Wiksell, Stockholm, 1968)

[4] A. Salam, *Weak and Electromagnetic Interactions*, Almqvist & Wiksell, Stockholm, nobel symposium 8 edition (1968)

[5] S. L. Glashow, J. Iliopoulos, L. Maiani, *Weak Interactions with Lepton-Hadron Symmetry*, Phys. Rev. D **2**, 1285 (1970)

[6] H. Georgi, S. L. Glashow, *Unified Weak and Electromagnetic Interactions without Neutral Currents*, Phys. Rev. Lett. **28**, 1494 (1972)

[7] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions*, Phys. Rev. Lett. **30**, 1346 (1973)

[8] H. D. Politzer, *Asymptotic Freedom: An Approach to Strong Interactions*, Phys. Rept. **14**, 129 (1974)

[9] P. W. Higgs, *Broken Symmetries, Massless Particles and Gauge Fields*, Phys. Lett. **12**, 132 (1964)

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

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

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

[13] DØ Collabrotation, *Observation of the Top Quark*, Phys. Rev. Lett. **74**, 2632 (1995)

[14] CDF Collaboration, *Observation of Top Quark Production in $\overline{p}p$ Collisions with the Collider Detector at Fermilab*, Phys. Rev. Lett. **74**, 2626 (1995)

[15] ATLAS Collaboration, *Study of heavy-flavor quarks produced in association with top-quark pairs at $\sqrt{s} = 7$ TeV using the ATLAS detector*, Phys. Rev. D **89**, 072012 (2014)

[16] ATLAS Collaboration, *Measurements of fiducial cross-sections for $t\bar{t}$ production with one or two additional b-jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, Eur. Phys. C **76(1)**, 11 (2016)

[17] 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 **2019(4)**, 46 (2019)

[18] CMS Collaboration, *Measurement of the cross section ratio $\sigma_{t\bar{t}b\bar{b}}/\sigma_{t\bar{t}jj}$ in pp collisions at $\sqrt{s} = 8$ TeV*, Phys. Lett. B **746**, 132 (2015)

[19] CMS Collaboration, *Measurements of $t\bar{t}$ cross sections in association with b jets and inclusive jets and their ratio using dilepton final states in pp collisions at $\sqrt{s} = 13$ TeV*, Phys. Lett. B **776**, 355 (2018)

[20] CMS Collaboration, *Measurement of the cross section for $\mathrm{t\bar{t}}$ production with additional jets and b jets in pp collisions at $\sqrt{s} = 13$ TeV*, JHEP **2020(7)**, 125 (2020)

[21] CMS Collaboration, *Measurement of the $t\bar{t}b\bar{b}$ production cross section in the all-jet final state in pp collisions at $\sqrt{s} = 13$ TeV*, Phys. Lett. B **803**, 135285 (2020)

[22] CMS Collaboration, *First measurement of the cross section for top quark pair production with additional charm jets using dileptonic final states in pp collisions at $\sqrt{s} = 13$ TeV*, Phys. Lett. B **820**, 136565 (2021)

[23] D. J. Gross, F. Wilczek, *Asymptotically Free Gauge Theories*, Phys. Rev. D **8**, 3633 (1973)

[24] S. Weinberg, *The Making of the Standard Model*, Eur. Phys. J. C **34**, 5 (2004), `hep-ph/0401010`

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

[26] G. 't Hooft, M. Veltmann, *Regularization And Renormalization Of Gauge Fields*, Nucl. Phys. B **44**, 189 (1972)

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

[28] K. Moor, *Measurement of the top quark pair production cross section in the lepton + jets channel at the ATLAS experiment at* $\sqrt{s} = 13$ TeV, II.Physik-UniGoe-MSc-2019/02 (2019)

[29] The Particle Data Group, P. Zyla, et al., *Review of Particle Physics*, Progress of Theoretical and Experimental Physics **2020** (2020)

[30] N. Cabbibo, *Unitary Symmetry and Leptonic Decays*, Phys. Rev. Lett. **10**, 531 (1963)

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

[32] M. Thomson, *Modern Particle Physics*, Cambridge University Press (2013)

[33] J. E. Augustin, et al, *Discovery of a Narrow Resonance in* e^+e^- *Annihilation*, Phys. Rev. Lett. **33**, 1406 (1974)

[34] L. R. Evans, *The Large Hadron Collider (LHC)*, Eur. Phys. J. C **34**, S11 (2004)

[35] *ATLAS Technical Proposal for a General-Purpose pp Experiment at the Large Hadron Collider at CERN*, CERN/LHCC/94-43, LHCC/P2, 1994 (unpublished)

[36] *The Large Hadron Collider LHC at CERN*, http://lhc-new-homepage. (2000)

[37] M. Capeans, et al (ATLAS Collaboration), *ATLAS Insertable B-Layer Technical Design Report*, Technical report (2010)

[38] A. Juste Rozas, et al, *Measurement of the* $t\bar{t}H$ *production cross-section in the* $H \to b\bar{b}$ *decay channel using 139* fb^{-1} *of pp collision data at* $\sqrt{s} = 13$ *TeV: Supporting note for ANA-HIGG-2018-11 and ANA-HIGG-2020-23*, Technical report, CERN, Geneva (2019)

[39] ATLAS Collaboration, *Electron reconstruction and identification efficiency measurements with the ATLAS detector using the 2015 and 2016 LHC proton-proton collision data*, Eur. Phys. C **79(8)** (2019)

[40] ATLAS Collaboration, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, Technical report, CERN, Geneva (2016)

[41] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton-proton collision data at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. C **76(5)** (2016)

[42] ATLAS Collaboration, *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*, Technical report, CERN, Geneva (2015)

[43] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. C **77(7)** (2017)

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

[45] M. Cacciari, G. P. Salam, G. Soyez, *FastJet user manual*, Eur. Phys. C **72(3)** (2012)

[46] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Rev. D **96(7)** (2017)

[47] *Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector*, Technical report, CERN, Geneva (2015)

[48] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in proton-proton collisions at* $\sqrt{s} = 8$ *TeV using the ATLAS detector*, Eur. Phys. C **76(11)** (2016)

[49] M. Cristinziani, et al, *Flavour tagging algorithms in 2018*, Technical report, CERN, Geneva (2018)

[50] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, Technical report, CERN, Geneva (2017)

[51] ATLAS Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, Technical report, CERN, Geneva (2016)

[52] ATLAS Collaboration, *Secondary vertex finding for jet flavour identification with the ATLAS detector*, Technical report, CERN, Geneva (2017)

[53] R. E. Kálmán, *A new approach to linear filtering and prediction problems" transaction of the asme journal of basic* (1960)

[54] ATLAS Collaboration, *Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector. Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector*, Eur. Phys. J. C **77**, 580. 49 p (2017), 1705.02211

[55] The Theano Development Team, *Theano: A Python framework for fast computation of mathematical expressions* (2016), arXiv:1605.02688

[56] D. P. Kingma, J. Ba, *Adam: A Method for Stochastic Optimization* (2017), arXiv:1412.6980

[57] I. J. Goodfellow, D. Warde-Farley, M. Mirza, A. Courville, Y. Bengio, *Maxout Networks* (2013), arXiv:1302.4389

[58] S. Ioffe, C. Szegedy, *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift* (2015), arXiv:1502.03167

[59] ATLAS Collaboration, *Luminosity determination in pp collisions at* $\sqrt{s} = 13$ *TeV using the ATLAS detector at the LHC*, Technical report, CERN, Geneva (2019)

[60] ATLAS Collaboration, *Luminosity determination in pp collisions at* $\sqrt{s} = 13$ *TeV using the ATLAS detector at the LHC*, Technical report, CERN, Geneva (2019)

[61] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. C **70(3)**, 823 (2010)

[62] S. Agostinelli, et al, *Geant4—a simulation toolkit*, NIM-A **506(3)**, 250 (2003)

[63] W. Lukas, *Fast Simulation for ATLAS: Atlfast-II and ISF*, Technical report, CERN, Geneva (2012), URL https://cds.cern.ch/record/1458503

[64] ATLAS Collaboration, *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, Technical report, CERN, Geneva (2010)

[65] T. Gleisberg, S. Hoeche, F. Krauss, A. Schälicke, S. Schumann, J. Winter, *SHERPA 1.α, a proof-of-concept version*, J. High Energy Phys. **0402**, 056 (2004), hep-ph/0311263

[66] D. J. Lange, *The EvtGen particle decay simulation package*, NIM-A **462(1)**, 152 (2001)

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

[68] S. Frixione, G. Ridolfi, P. Nason, *A positive-weight next-to-leading-order Monte Carlo for heavy flavour hadroproduction*, JHEP **2007(09)**, 126 (2007)

[69] The NNPDF collaboration, *Parton distributions for the LHC run II*, JHEP **2015(4)**, 40 (2015)

[70] R. Frederix, E. Re, P. Torrielli, *Single-top t-channel hadroproduction in the four-flavour scheme with POWHEG and aMC@NLO*, JHEP **2012(9)**, 130 (2012)

[71] S. Alioli, P. Nason, C. Oleari, E. Re, *NLO single-top production matched with shower in POWHEG: s- and t-channel contributions*, JHEP **2009(09)**, 111 (2009)

[72] E. Re, *Single-top Wt-channel production matched with parton showers using the POWHEG method*, Eur. Phys. C **71(2)**, 1547 (2011)

[73] Alwall, J. and et al, *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **2014(7)**, 79 (2014)

[74] ATLAS Top Reconstruction sub-group, *Common file for MC samples cross-sections and shower generators*, URL `https://atlas-groupdata.web.cern.ch/atlas-groupdata/dev/AnalysisTop/TopDataPreparation/XSection-MC16-13TeV.data`

[75] K. Cranmer, et al (ROOT Collaboration), *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, Technical report, New York U., New York (2012)

[76] F. James, *Minuit - a system for function minimization and analysis of the parameter errors and correlations*, Comput. Phys. Commun. **10**, 343 (1975)

[77] G. Avoni, et al, *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, J. Instrum. **13(07)**, P07017 (2018)

[78] ATLAS Collaboration, *Measurement of the Inelastic Proton-Proton Cross Section at* $\sqrt{s} = 13$ TeV *with the ATLAS Detector at the LHC*, Phys. Rev. Lett. **117**, 182002 (2016)

[79] ATLAS Collaboration, *Calibration of light-flavour b-jet mistagging rates using ATLAS proton-proton collision data at* $\sqrt{s} = 13$ *TeV*, Technical report, CERN, Geneva (2018)

[80] *ATLAS Pythia 8 tunes to 7 TeV data*, Technical report, CERN, Geneva (2014)

[81] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at* $\sqrt{s} = 13$ TeV *with the ATLAS detector*, Phys. Rev. D **96**, 072002 (2017)

[82] J. Bellm, et al, *Herwig 7.0/Herwig++ 3.0 release note*, Eur. Phys. C **76(4)**, 196 (2016)

[83] J. Butterworth, S. Carrazza, A. Cooper-Sarkar, A. D. Roeck, J. Feltesse, S. Forte, J. Gao, S. Glazov, J. Huston, Z. Kassabov, R. McNulty, A. Morsch, P. Nadolsky, V. Radescu, J. Rojo, R. Thorne, *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43(2)**, 023001 (2016)

Acknowledgements

I want to thank Prof. Dr. Arnulf Quadt for giving me the opportunity to work on this analysis, Dr. Jelena Jovicevic and Dr. Elizaveta Shabalina for guiding and supporting me in my understanding of the topic, Sreelakshmi Sindhu for her answering many questions on a daily basis, Steffen Korn for his help in setting up the analysis script, Andreas Kirchhoff for his help in resolving a weighting issue at the start of this project and Ishan Pokharel for his help in switching the analysis framework for the implementation of the systematics.