Dissertation

**Search for Dark Matter in association with a top quark and a W-boson in $\sqrt{s} = 13$ TeV $pp$ collisions at the ATLAS detector**

Paul Moder

University of Freiburg
Faculty of Mathematics and Physics
Institute of Physic
August 03rd, 2022
Search for Dark Matter in association with a top quark and a $W$-boson in $\sqrt{s} = 13$ TeV $pp$ collisions at the ATLAS detector
Writing Period
05.01.2019 – 03.08.2022

Dean
Prof. Dr. Michael Thoss

Examiner
Prof. Dr. Beate Heinemann

Second Examiner
Prof. Dr. Markus Schumacher

Advisers
Prof. Dr. Beate Heinemann
Dr. Claudia Seitz

Date of Oral Examination
28.11.2022
Declaration

I hereby declare that I am the sole author and composer of my thesis and that no other sources or learning aids, other than those listed, have been used. Furthermore, I declare that I have acknowledged the work of others by providing detailed references of said work.

I hereby also declare that my Thesis has not been prepared for another examination or assignment, either wholly or excerpts thereof.

________________________  __________________________
Place, Date                      Signature
Abstract

The Standard Model (SM) is the only model in particle physics containing all observed elementary particles and their interactions. Over the years, its predictions were tested and observed in a number of experiments. However, there are still observations that can not be explained by the SM with one of the most prominent ones being the existence of Dark Matter (DM). While the existence of DM was first theorised through astronomical observations, extensions of the SM allow for a search of DM at the Large Hadron Collider (LHC) as well. Since DM can not be detected directly, final states analysing its existence at the LHC are always designed around high missing transverse energy.

This thesis presents a search for DM with data collected by the ATLAS detector in the years 2015-2018 corresponding to an integrated luminosity of 139 fb$^{-1}$ at a center of mass energy of 13 TeV. The analysis is based on an extended two Higgs doublet model (2HDM+a) where a pseudo-scalar mediator allows the production of DM in the final state. In addition to the DM, a top quark and a $W$-boson are produced in the final state. This thesis mainly focuses on a final state with zero leptons which results in both the $W$-boson and the top quark being required to decay hadronically. In addition to that, the $W$-boson can be expected to have a high momentum in the signal process. Therefore, a dedicated identification algorithm using large-radius jets is employed to select events with at least one hadronically decaying $W$-boson. This allows for a strong distinction against SM background events.

In this analysis, several different signal processes are considered, since different
parameters can affect the cross section and the distributions of the signal processes. These parameters include different values for the mass of the mediator \( a \), the mass of the heavy Higgs boson \( H \) and \( \tan \beta \). Since no significant excess was found in the signal regions when comparing SM prediction to data, upper exclusion limits on the cross section of the signal processes were calculated with a 95\% confidence level (CL). These limits are presented in three different model parameter planes. One such plane presents the upper limits for different values of the mass of the mediator, \( m_a \), and the mass of the heavy Higgs boson, \( m_H \), while keeping \( \tan \beta = 1 \). The second and third plane present the upper limits for different values of the mass of the heavy Higgs boson, \( m_H \), and \( \tan \beta \) while the mass of the mediator \( a \) stays constant at \( m_a = 250 \text{ GeV} \) and \( m_a = 150 \text{ GeV} \). In addition to the upper limits for signal processes with one top quark in the final state, this thesis also provides an upper exclusion limit where this signal process is combined with a process that includes two top quarks in the final state.
Kurzfassung


ein W-Boson, welches hadronisch zerfallen ist, in Ereignissen nachzuweisen. Dieser Algorithmus erlaubt für eine effiziente Unterscheidung zwischen Ereignissen, die aus Signal-Prozessen entstehen und Ereignissen, die zu Prozessen des SM gehören. Da der Signal-Prozess in dem unterliegenden Modell von verschiedenen Parametern abhängt, werden eine große Anzahl potentieller Signal-Prozesse zur gleichen Zeit analysiert. In diesen verschiedenen Prozessen werden vor allem die Parameter verändert, die die Masse des zusätzlichen pseudoskalaren Teilchens, \( m_a \), und die Masse des schweren Higgs-Bosons, \( m_H \), beschreiben. Darüber hinaus wird auch der Parameter \( \tan \beta \) variiert. Da bei Vergleich zwischen Daten und SM Vorhersage kein signifikanter Unterschied beobachtet wurde, werden die Ergebnisse in oberen Ausschlussgrenzen des Wirkungsquerschnitts der einzelnen Signal-Prozesse beschrieben, wobei diese in einem 95\%-igen Vertrauensbereich angegeben werden. Diese Ausschlussgrenzen werden dabei in drei Parameter-Bereichen präsentiert. Im ersten Bereich wird der Parameter \( \tan \beta = 1 \) konstant gehalten, während verschiedene Werte für die Massen des Mediators, \( m_a \), und des schweren Higgs-Bosons, \( m_H \), getestet werden. In den anderen beiden Bereichen wird die Masse des Mediators konstant auf \( m_a = 250 \text{ GeV} \) und \( m_a = 150 \text{ GeV} \) gehalten, während \( \tan \beta \) und die Masse des schweren Higgs-Bosons, \( m_H \), variiert werden. Neben den oberen Ausschlussgrenzen des Prozesses mit einem top-Quark im Finalzustand werden auch die oberen Ausschlussgrenzen angegeben, in welchem dieser Prozess mit einem Prozess kombiniert wird, in welchem zwei top-Quarks im Finalzustand vorkommen.
## Contents

1 Introduction 1

2 Theory 5
   2.1 The Standard Model of Particle Physics . . . . . . . . . . . . . . . . 5
      2.1.1 Elementary Particles and Interactions . . . . . . . . . . . . . . 5
      2.1.2 Electroweak Symmetry Breaking . . . . . . . . . . . . . . . . 10
      2.1.3 CP-symmetry . . . . . . . . . . . . . . . . . . . . . . . . . . . 16
   2.2 Dark Matter . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 17
   2.3 The 2HDM+a model . . . . . . . . . . . . . . . . . . . . . . . . . . . 22
      2.3.1 Basics of the 2HDM . . . . . . . . . . . . . . . . . . . . . . . 23
      2.3.2 The simplified 2HDM+a . . . . . . . . . . . . . . . . . . . . . 26

3 Statistical Framework 31
   3.1 Signal significance . . . . . . . . . . . . . . . . . . . . . . . . . . 31
   3.2 HistFitter . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 34
      3.2.1 Background-only fit . . . . . . . . . . . . . . . . . . . . . . . 35
      3.2.2 Background-only hypothesis test . . . . . . . . . . . . . . . . 37
      3.2.3 Upper limits and exclusion tests . . . . . . . . . . . . . . . . 38
      3.2.4 Expected statistical variation . . . . . . . . . . . . . . . . . 40

4 LHC and ATLAS 43
   4.1 Large Hadron Collider . . . . . . . . . . . . . . . . . . . . . . . . . 43
1 Introduction

With the technical advancements over the last century, it became possible to prove the existence of several particles. With that, the idea of finding the smallest building blocks and therefore a theory describing the universe was born. However, in order to understand these particles and their interactions, the understanding of the fundamental laws of physics had to be extended as well. One of these extensions came with the theory of special relativity by Einstein in 1905 [1] and later general relativity in 1916 [2]. A second extension was the theory of quantum electro dynamics (QED) by Schwinger, Feynman and Tomonaga in the 1940s [3, 4, 5]. With the help of these new theories, it was not only possible to understand already discovered particles, but also to build on these theories and predict more elementary particles and their interactions. One example is the theory by Glashow, Weinberg and Salam unifying the QED with the weak interaction [6, 7, 8], predicting the existence of the W- and Z-bosons in the 1960s which were later discovered by the UA1 experiment [9]. Another example is the independent postulation of the Higgs field by Higgs, Englert and Brout in the 1960s. This mechanism of electroweak symmetry breaking was introduced to explain the existence of masses for the elementary particles and as a result, it predicted the existence of another particle known as the Higgs boson [10, 11, 12]. However, it took over 40 years until the ATLAS and CMS experiments at the Large Hadron Collider (LHC) were both able to prove its existence in 2012 [13, 14].

Lastly, with the introduction of spark and bubble chambers in the 1950s, a large
variety of hadrons was discovered. It was postulated in 1963 by Gell-Mann that these hadrons were not elementary particles themselves, but that they were made up of three different flavour quarks [15]. However, it took until 1973, where Fritzsch, Leutwyler and Gell-Mann developed the theory of quantum chromo dynamics (QCD) which describes the color charge as the source of the strong field interaction [16]. Today, all of the discovered elementary particles as well as their interactions are summarized in the Standard Model of particle physics (SM) [17].

Over the years, the SM was tested by a variety of experiments and for most, its predictions were found to be accurate. However, there are still some observations that can not be explained by the SM. The most recent example features the measurement of the magnetic moment of the muon in 2021 which shows a significant deviation from the SM predictions [18, 19]. However, even previous to that, measurements, for example about the mass of the neutrinos showed that physics beyond the SM can be expected. Another prominent hint is the existence of Dark Matter (DM) which could include more elementary particles.

The existence of DM was already theorised in the early years of the 20th century. In 1932, Jan Oort among others studied galaxy rotation curves, which involve the velocity of the visible stars in dependence to their distance to the center of the galaxy. The distributions diverted significantly from the expected theory if only the mass of the visible objects was taken into account. However, by assuming the existence of non-visible massive particles, known as DM, the theory was able to explain the distributions [20]. Over the years, a lot of other experiments like experiments on the effect of Gravitational Lensing [21] or the measurement of the cosmic microwave background [22] supported the existence of DM as well. The combination of all these observations allowed the prediction that Dark Matter makes up around 85% of all matter in the universe.

The search for DM is also one large part of the data analysis at the ATLAS and CMS experiments at the LHC where a high number of proton-proton collisions with a center-of-mass energy of 13 TeV was measured over the years. However, for a possible
detection, a theory introducing weakly interacting massive particles (WIMPs) [23], that could be part of the DM, as well as an interaction between SM particles and these WIMPs is necessary. In some cases, this problem is solved by introducing a mediator which interacts with SM particles as well as the DM particles [24, 25]. However, since the measurements of the SM, for example the measurement of the Higgs couplings [26], heavily constrain such a model, an extension of the SM is needed. This extension needs to explain the observations made in SM measurements while also allowing for the production of DM particles in a similar way through the interaction with a mediator. One of these models is the two Higgs doublet model (2HDM) [27] which predicts the existence of a second Higgs doublet leading to a total of five Higgs bosons with one of these Higgs bosons being the one observed in the SM. Together with the mediator, this model is therefore called 2HDM+a. Final states with DM resulting from the interaction of the other four Higgs bosons are not as constrained and therefore are possibly observable at the LHC. Since DM does not interact directly with matter, processes with DM in the final state at the LHC are characterised by a high amount of missing transverse energy in association with different SM particles that can be measured directly by the detector. Final states that were mostly focused on in this model included Z- and Higgs-bosons as well as two top quarks [28]. A final state very similar to the one with two top quarks includes a single top quark and a W-boson. While the cross sections of these processes are usually lower than for processes with two top quarks, a recent analysis showed that the phase space can be extended by analysing this process in the context of the 2HDM+a as well [29].

The goal of this thesis is to present an extended search for the final state with DM in association with a single top quark and W-boson by introducing an additional analysis channel with zero leptons in the final state where the previous analysis focused on the channels with one and two leptons. Therefore, this analysis exploits the fact that the W-boson does often have a high transverse momentum which allows a W-tagging algorithm to select events with hadronically decaying W-bosons. This
can be a good discriminator between signal process and SM background events. This work is therefore structured as follows. Chapter 2 provides an overview over the theoretical background starting with a short presentation of the SM as well as a basis for the existence of DM. Afterwards, the 2HDM+a model is presented in more detail where the processes used in this analysis with a DM+$tW$ final state are also shown. Since this analysis requires a statistical interpretation of the data, the theory of this statistical framework is given in Chapter 3. In Chapter 4, an overview over the LHC and the ATLAS detector is presented and is followed by Chapter 5 which provides a more detailed explanation of the Track-Counting luminosity measurement at the ATLAS detector. Before presenting the analysis, Chapter 6 and Chapter 7 provide the necessary analysis information. Chapter 6 thereby presents the Monte Carlo samples used to predict the number of SM and signal events as well as an overview over the collected data. Chapter 7 provides the information for the object definitions for the measured particles used in this analysis. The strategy of the analysis is given in Chapter 8 with a detailed explanation of the variables that were used as well as their impact on the signal sensitivity. In Chapter 9, the overall estimation of the background is presented through the definition of regions used to model the major background processes with a data-driven method. Additionally, this chapter also provides an overview over important systematic uncertainties. Finally, the results of the analysis are presented in Chapter 10 and a short summary as well as outlook as to how the analysis could be improved in the future is presented in Chapter 11.
2 Theory

This chapter describes the theoretical foundation of this search for Dark Matter (DM) at the LHC with a top quark and a $W$-boson in the final state. Since the baseline model for this is an extension of the Standard Model itself, Section 2.1 introduces the Standard Model of particle physics (SM). Furthermore, because the analysis searches for DM, a short overview of the knowledge about DM is presented in Section 2.2. Finally, the model describing the interaction between the weakly interacting particles from the SM and the DM sector, the 2HDM+a model, is presented in more detail in Section 2.3.

2.1 The Standard Model of Particle Physics

2.1.1 Elementary Particles and Interactions

The Standard Model of particle physics [30] describes all the currently known elementary particles as well as three of the four fundamental interactions between them. The fundamental interactions are the electromagnetic, the weak, the strong and the gravitational interaction. In the SM, the first three of these interactions are described very precisely through a gauge group $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$. In this group, $SU(3)_C$ represents the strong interaction with the colour charge $C$ while the electromagnetic and weak interaction are represented by the $SU(2)_L \otimes U(1)_Y$ group which is therefore called the electroweak group. In this case, $L$ and $Y$ denote the
left chirality and hypercharge respectively which will be described in more detail in Section 2.1.2. Interactions between particles in the SM can only occur through the carrier of the interaction if both particles carry the respective charge. For the electromagnetic interaction, this is the electromagnetic charge. For the weak interaction, the charge is the weak hypercharge while it is the colour charge for the strong interaction.

All the currently known elementary particles of the SM are presented in Figure 1. Particles are divided into different groups depending on their properties. The first classification comes through the spin of the particle. Particles with a half-integer spin are called fermions while particles with an integer spin are called bosons. The bosons in the SM are elementary particles as well. Therefore, they are further specified as gauge bosons and can be interpreted as the carriers of the interactions. They are the mediators between particles that carry the charge of the corresponding interaction with a spin of 1. The main gauge bosons are the photon for the electromagnetic interaction, the $W^+$, $W^-$ and $Z$ bosons for the weak interaction and the gluon for the strong interaction. Examples for every interaction are shown in Figure 2.
While the photon and gluon are massless, the bosons of the weak interaction have comparatively high masses of 80.4 GeV and 91.2 GeV [32]. These particles decay into lighter particles with decay widths of 2.1 GeV and 2.5 GeV which results in a lifetime in the order of $10^{-25}$ s [32]. This is the reason why the weak interaction can only occur over short distances compared to the electromagnetic interaction which can occur over infinite distances. However, even though the gluon has no mass, the distance of the interaction between strongly interacting particles is small as well. This results from the fact that the strong force does not diminish with a higher distance. This means that the energy between two strongly interacting particles increases with distance. Therefore, if the distance between two particles reaches a certain point, a new pair of particles is produced reducing the interaction distance again. This phenomenon is called confinement [33]. While gauge bosons
can interact with the fermions that carry the corresponding charge, they are also able to interact with each other if they carry charges themselves. The $W$-bosons are carrying an electromagnetic charge and can therefore interact with a photon. Similarly, they are also carrying a weak charge and can therefore interact with each other. Another boson interaction is between the $W$- and $Z$-boson since both are carrying a weak hypercharge. For the strong interaction, only the gluon carries a colour charge, therefore it can only interact with other gluons but not other bosons. Examples of these interactions can be found in Figure 3 [17, 30, 34].

The second group of particles, the fermions, are further divided into six quarks and six leptons where a group of two is classified as a generation with the same flavour, creating three generations of quarks and leptons. Furthermore, the SM also predicts an anti-particle to each of these twelve fermions. These anti-particles are identical to the particle except that they have opposite charge. The different generations for leptons are electrons ($e$), muons ($\mu$) and taus ($\tau$) ordered by increasing mass together with their corresponding neutrino ($\nu$). All leptons carry a weak hypercharge and therefore interact through the weak interaction and with $W$- and $Z$-bosons. But since the neutrinos do not carry an electromagnetic charge, only the electron, muon and tau leptons, which all have an electromagnetic charge of -1, interact through the electromagnetic interaction and with the photon. However, none of the leptons carry a colour charge and therefore no lepton interacts through the strong interaction and
with the gluon. Furthermore, in the SM, neutrinos are massless particles. This was later disproven by different experiments, for example the measurement of the solar neutrino flux with the Homestake chlorine detector [35]. It observed a significant deficit in the flux of electron neutrinos $\nu_e$ compared to the prediction implying an oscillation to a different neutrino generation. A similar observation was made with the Super-Kamiokande experiment for atmospheric muon neutrinos and which was awarded with the Nobel Prize in the year 2015 [36]. The oscillations observed in these experiments are only possible if the mass difference between these neutrino flavours is different from 0 [37, 38]. However, since the masses of the neutrinos are very small and they only interact through the weak interaction, experiments to determine the mass like KATRIN [39] could only give upper limits with significant certainty. Compared to other particles of the SM, it is confirmed that the masses are smaller by several orders of magnitude. For the electron neutrino, the KATRIN experiment determined an upper limit of $m_{\nu_e} < 0.9$ eV with a 90% confidence level [39, 40].

For the six quarks in the SM, the generations always consist of an up-type quark combined with a down-type quark which are defined by their electromagnetic charge of 2/3 and -1/3 respectively. The combinations of quarks that form these generations are up (u) and down (d), charm (c) and strange (s) and top (t) and bottom (b). Since all quarks carry an electromagnetic charge, they interact electromagnetically. Like fermions, they are also interacting weakly and therefore with the $W$- and $Z$-bosons. In addition to that, they carry a colour charge which allows for strong interactions between quarks and with gluons. However, in comparison to the other two charges, the colour charge can not be measured since free particles can only have a neutral colour charge of 0. This results from the confinement that was described before which always leads to the production of a quark-antiquark pair if the distance between two strongly interacting particles becomes too large. This also means that in comparison to leptons, quarks are not free particles. They are always bound in particles that either have an even (mesons) or an odd number of quarks (baryons), where the most
common cases are particles with two or three quarks. A particle consisting of quarks can reach a neutral colour charge and therefore exist as a free particle in two ways. First, by combining the colour with its opposite colour, for example red and anti-red, a neutral colour charge is the result. Second, by combining all three colour charges, i.e. (anti-)red with (anti-)green and (anti-)blue, a colourless state can be achieved [17, 30, 34].

The last particle in the Standard Model is the Higgs boson. This is a particle that was very early predicted to exist due to the Higgs mechanism, described in detail in Section 2.1.2, but which could only be detected the first time in the year 2012 at the LHC with a mass of around $125 \text{ GeV}$ [12, 13]. The Higgs boson differs from the other gauge bosons as it has a spin of 0 instead of 1 and is not the carrier of one of the four interactions. It is still an important part of the SM as its existence is the proof for the explanation of how the particles in the SM acquire their masses which will be discussed in more detail in the next section.

### 2.1.2 Electroweak Symmetry Breaking

While the strong interaction is described in the SM through a $SU(3)_C$ gauge symmetry group for particles that carry the respective colour charge C, called Quantum Chromo Dynamic (QCD), the electromagnetic and weak interactions are described through a unification in the $SU(2)_L \otimes U(1)_Y$ group, which is called electroweak unification. The Lagrangian of this group can be defined as [41]

$$\mathcal{L}_{SU(2)_L \otimes U(1)_Y} = \mathcal{L}_f + \mathcal{L}_{\text{gauge}} + \mathcal{L}_{\text{Higgs}} + \mathcal{L}_{\text{Yukawa}}$$

(1)
where the different terms describe different parts of the SM. First, the Lagrangian for the gauge and fermion fields can be described as

$$\mathcal{L}_f = \sum_{\psi_L \in l_j, Q_j} \bar{\psi}_L i D_\mu \gamma^\mu \psi_L + \sum_{\psi_R \in l_j, u_j, d_j} \bar{\psi}_R i D_\mu \gamma^\mu \psi_R$$

$$\mathcal{L}_{gauge} = -\frac{1}{4} W^a_{\mu\nu} W^{a\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}$$

(2)

where $j = 1, 2, 3$ represents the three generations of leptons and $a = 1, 2, 3$ represents the three generators of the $SU(2)_L$. The field strength tensors

$$V^a_{\mu\nu} = \partial_\mu V^a_\nu - \partial_\nu V^a_\mu - g_2 \epsilon_{abc} V^b_\mu V^c_\nu \quad V^a_\mu \in W^a_\mu, B_\mu$$

(3)

correspond to the gauge fields $W^a_\mu$ and $B_\mu$ as well as their coupling constants $g_2$ and $g_1$ of the $SU(2)_L$ and $U(1)_Y$ respectively. $\epsilon_{abc}$ is the totally antisymmetric symbol. Since there is only one gauge field in the $U(1)_Y$ case, this means that the last term will always result in 0. These gauge fields are also part of the covariant derivative present in the Lagrangian of the fermion fields and defined as

$$D_\mu \psi = \left( \partial_\mu - ig_2 T_a W^a_\mu - ig_1 \frac{Y}{2} B_\mu \right)$$

(4)

where $Y$ is the weak hypercharge of the $U(1)_Y$ and $T_{1,2,3} = 1/2 \cdot \sigma_{1,2,3}$ are the three generators of the $SU(2)_L$ defined through the Pauli matrices $\sigma_{1,2,3}$ [41, 42].

In Equation (2), the fermion fields $\psi$ are divided into left-handed $\psi_L$ and right-handed $\psi_R$ components. The left-handed components are doublets from the different lepton and quark generations while the right-handed components are singlets. When
it can be noticed that there are no right-handed neutrinos in the SM. For all left-handed and right-handed fermions, a value for the hypercharge $Y$ and the weak isospin $T$ is defined. For left- and right-handed particles, the third component of the weak isospin $T_3$ is $1/2$ and $0$ respectively. The weak hypercharge $Y$ is then connected with this third component as well as the electromagnetic charge $Q$ through

$$Q = Y + T_3.$$  

With these two Lagrangian density terms, the interaction between electroweak bosons and fermions can be described through coupling constants. However, neither bosons nor fermions have any mass terms so far. This can be achieved through the other two terms as well as spontaneous symmetry breaking [41, 42].

The third term of Equation (1) describes a scalar field in the SM which is called the Higgs field

$$L = (D^\mu \Phi)^\dagger (D_\mu \Phi) - V(\Phi)$$  

with

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix}$$

and the potential

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2.$$
Fig. 1: An illustration of the Higgs potential (5) in the case that $\mu^2 < 0$, in which case the minimum is at $|\phi|^2 = -\frac{\mu^2}{2\lambda}$. Choosing any of the points at the bottom of the potential breaks spontaneously the rotational $U(1)$ symmetry.

The scalar particle corresponding to $\eta$ is massive with $m^2_\eta = -\mu^2 > 0$, whereas the scalar particle corresponding to $\xi$ is massless. This particle is a prototype of a (Nambu-)Goldstone boson. It is massless because there is a direction in field space, corresponding to changing the phase, in which the potential energy does not change. Its appearance is a general feature of models with spontaneously-broken global symmetries, as proven in [13]. The total number of such massless particles corresponds in general to the number of field directions in which the potential is flat. Nambu introduced this idea into particle physics in order to describe the (relatively light) pion of QCD [11], which he identified as a (pseudo-)Goldstone boson of chiral symmetry that would have no mass if the up and down quarks were exactly massless. The simple field-theoretical model is due to Goldstone [12].

We now discuss how this spontaneous symmetry breaking of symmetry manifests itself in the presence of a $U(1)$ gauge field [17, 19, 20]. In order to construct a theory that is invariant under local $U(1)$ phase transformations, i.e.

$$\phi \rightarrow e^{i\alpha} (x),$$

we introduce a gauge field $A_\mu$ that transforms under $U(1)$ as follows:

$$A'_\mu \rightarrow A_\mu + \frac{1}{q} \partial_\mu \alpha (x).$$

The space-time derivatives appearing in the kinetic term for the scalar field $\phi$ are replaced by covariant derivatives

$$D_\mu = \partial_\mu + iq A_\mu,$$

where $q$ is the conserved charge. Including kinetic terms for both the scalar field and the $A_\mu$ field:

$$\frac{1}{4} F_{\mu\nu} F^{\mu\nu},$$

where $F_{\mu\nu} \equiv \partial_\nu A_\mu - \partial_\mu A_\nu$, which is invariant under the $U(1)$ gauge transformation (11), we have the Lagrangian

$$L = [\left(\partial_\mu - iq A_\mu\right) \phi^*] [\left(\partial_\mu + iq A_\mu\right) \phi] - V(\phi^* \phi) - \frac{1}{4} F_{\mu\nu} F^{\mu\nu},$$

which we now analyze.

Figure 4: Higgs Potential in cases with $\mu^2 < 0$. [43]

The shape of this potential is dependent on the mass term $\mu^2$ and the Higgs field self-coupling $\lambda$. The self-coupling is set to be positive and real. Therefore, if $\mu^2$ is positive as well, the potential has exactly one minimum at $\Phi = 0$. However, if $\mu^2$ is negative, then the shape of the potential changes to what can be observed in Figure 4. This potential is symmetric around the point of origin where both the real and imaginary component of the potential are 0. However, at this point, the potential shows a local maximum and not minimum. Instead, the potential shows an infinite number of minima which represent the vacuum expectation value (VEV) where one of the infinite numbers is chosen to be

$$\langle \Phi \rangle_0 = \langle 0 | \Phi | 0 \rangle = \begin{pmatrix} 0 \\ v \end{pmatrix} \text{ with } v = \sqrt{-\frac{\mu^2}{\lambda}}.$$  \hspace{1cm} (10)

The symmetry of the $SU(2)_L$ is therefore spontaneously broken. However, it can also be noticed that the electromagnetically charged part of the potential is not broken for the VEV as it is 0. Therefore, the $SU(2)_L \otimes U(1)_Y$ group is spontaneously broken to the $U(1)_Q$ subgroup [41, 42, 43].

In order to quantise the potential around the VEV, the potential is written in terms
of four fields at first order

\[
\Phi = \begin{pmatrix}
\theta_1 + i\theta_2 \\
\frac{1}{\sqrt{2}} (v + H) - i\theta_3
\end{pmatrix}
\]

\[= e^{i\theta_3 \sigma^a / v} \begin{pmatrix}
0 \\
\frac{1}{\sqrt{2}} (v + H)
\end{pmatrix}
\]

(11)

where \(\theta_{1,2,3}(x)\) and \(H(x)\) are components of the additional fields and \(\sigma_{1,2,3}\) are Pauli matrices. Under a gauge transformation, this potential changes into

\[
\Phi \rightarrow e^{-i\theta_3 \sigma^a / v} \Phi = \frac{1}{\sqrt{2}} \begin{pmatrix}
0 \\
v + H
\end{pmatrix}
\]

(12)

This field is then included in Equation (7) which results in

\[
(D^\mu \Phi)^\dagger (D_\mu \Phi) = \frac{1}{2} (\partial_\mu H)^2 + \frac{1}{8} g_2^2 (v + H)^2 |W_{\mu}^1 + iW_{\mu}^2|^2 + \frac{1}{8} (v + H)^2 |g_2 W_{\mu}^3 - g_1 B_{\mu}|^2
\]

(13)

where the fields \(W_{\mu}^{1,2,3}\) and \(B_{\mu}\) are the gauge fields described in Equation (3). By defining new fields through mixing of these gauge fields

\[
W_{\mu}^\pm = \frac{1}{\sqrt{2}} \left( W_{\mu}^1 \mp iW_{\mu}^2 \right), \quad Z_{\mu} = \frac{g_2 W_{\mu}^3 - g_1 B_{\mu}}{\sqrt{g_1^2 + g_2^2}}, \quad A_{\mu} = \frac{g_2 W_{\mu}^3 + g_1 B_{\mu}}{\sqrt{g_1^2 + g_2^2}}
\]

(14)

the \(W^-\), \(Z\)-bosons as well as photons are defined. Through this addition of the field \(H\), the SM can now predict the masses of these bosons by including these definitions into Equation (13) and analysing the bilinear terms:

\[
m_W^2 W_\mu^+ W^-_\mu \rightarrow m_W = \frac{1}{2} v g_2 \\
\frac{1}{2} m_Z^2 Z_\mu Z^\mu \rightarrow m_Z = \frac{1}{2} v \sqrt{g_1^2 + g_2^2} \\
\frac{1}{2} m_A^2 A_\mu A^\mu \rightarrow m_A = 0
\]

(15)
which gives massive $W$- and $Z$-bosons as expected while the photon stays massless.

Furthermore, by plugging the VEV of the field in Equation (11) into the Higgs potential of Equation (9), the mass terms of the Higgs field can be predicted in a similar way

$$V(\Phi) = \lambda v^2 H^2 + \lambda v H^3 + \frac{\lambda}{4} H^4$$
with
$$\frac{1}{2} m_H^2 H^2 \rightarrow m_H = \sqrt{2\lambda v} = \sqrt{-2\mu^2} \quad (16)$$

which equates to the mass of the Higgs boson. The VEV of the SM was determined as well through the measurement of the Fermi constant $G_\mu$ in muon decays to be

$$m_W = \frac{1}{2} g_2 v = \left(\frac{\sqrt{2} g_2^2}{8 G_\mu}\right)^{1/2} \rightarrow v = \frac{1}{\left(\sqrt{2} G_\mu\right)^{1/2}} = 246 \text{ GeV.} \quad (17)$$

Finally, the masses of the fermions in the SM can also be predicted with the help of the Higgs field $H$ by analysing the final term of the Lagrangian, the Yukawa coupling term, which includes one term for each fermion field

$$L_{Yukawa} = -f_{\ell L} \bar{\ell}_L \ell_R \Phi_{\ell R} - f_{d_1} \bar{d}_1 \Phi_{d R} - f_{u_1} \bar{u}_1 \Phi_{u R} + \text{h.c.} \quad (18)$$

and which can be written through including the definition of the Higgs field in Equation (11) as

$$L_{Yukawa} = -\sum_{\psi_L \in \{\ell_L, Q_L\}, \psi_R \in \{\ell_R, d_R, u_R\}} \frac{f_\psi}{\sqrt{2}} (v + H) \bar{\psi}_L \psi_R + \text{h.c.} \quad (19)$$

In this case, the masses of the fermions can be retrieved as

$$m_{\ell, d, u} = \frac{f_{\ell, d, u} v}{\sqrt{2}} \quad (20)$$

and since a right-handed and left-handed component is needed to extract the mass this way, it results in a mass of 0 for all neutrinos in the SM. Furthermore, this term also describes the strength of the coupling between the Higgs boson and the fermions.
\[ g = \frac{m_{d_1} u_i}{v} \]  

(21)

which shows that the coupling is proportional to the mass of the fermion itself [41, 42, 43, 44, 45].

2.1.3 CP-symmetry

The charge conjugation parity symmetry (CP-symmetry) is the combination of the parity (P) and charge conjugation (C). The parity symmetry describes that a system is invariant under the flip of the spatial coordinates

\[ P\Psi(x, y, z) = \Psi(-x, -y, -z) \]  

(22)

where \( P \) describes the parity operator. From this, it results that applying the operator two times, the original state of the system is restored

\[ P^2\Psi(x, y, z) = P\Psi(-x, -y, -z) = \Psi(x, y, z) \]  

(23)

which means that under the assumption that \( \Psi(x, y, z) \) is an eigenstate of \( P \), the eigenvalues of the parity parameter are either +1 or -1

\[ P^2\Psi(x, y, z) = \lambda_P P\Psi(x, y, z) = (\lambda_P)^2 \Psi(x, y, z) = \Psi(x, y, z). \]  

(24)

The eigenvalue \( \lambda_P \) of a particle is defined as its intrinsic parity. Particles with an intrinsic parity of +1 are called “even” and particles with an intrinsic parity of -1 are called “odd”. It is defined that the intrinsic parity of fermions with a spin of \( 1/2 \) is +1 while the intrinsic parity of their anti-particles is -1. For bosons, they are classified according to their spin and intrinsic parity. A “scalar” and “pseudoscalar” boson is a boson with spin 0 and a parity of +1 and -1, respectively. For bosons with spin 1, they are called “vector” and “pseudovector” bosons if they have a parity of -1.
and +1, respectively [46].

The second part of the CP-symmetry is the charge conjugation describing the invariance when turning a particle into its anti-particle

\[ C |\psi\rangle = |\bar{\psi}\rangle \]  \hspace{1cm} (25)

where \( C \) is the conjugation operator. Similar to the parity, it can be concluded that \( C \) only has two possible eigenvalues, called the charge parity, with +1 and -1 since the anti-particle of the anti-particle returns the particle again

\[ C^2 |\psi\rangle = C |\bar{\psi}\rangle = |\psi\rangle . \]  \hspace{1cm} (26)

However, not all particles are eigenstates of \( C \). Only if all quantum numbers of a particle are 0, they are an eigenstate which for example is true for a photon, but not a proton [46].

The CP-symmetry is the combination of both symmetries. Therefore, the CP-operator is the product of \( P \) and \( C \). Similar to the parity, a particle is called CP-even and CP-odd if the product of \( C \) and \( P \) is +1 and -1 respectively [46]. For a time, it was assumed that while the SM violates the parity and charge conjugation individually, it does not violate the CP-symmetry. However, experiments, for example with decaying Kaons, showed a violation of CP-symmetry in the SM as well [47].

\[ \text{2.2 Dark Matter} \]

Dark Matter (DM) is one of the largest mysteries in the current theory of physics. DM thereby describes particles that are not part of the Standard Model and which are therefore not interacting directly through the three interactions described in Section 2.1. This makes the search for and proof of this type of matter very difficult. However, through gravity, the fourth of the fundamental interactions which is not
Figure 5: Comparison between theory and observations of galaxy rotation curves. The solid line shows the complete theory prediction combined from the three other lines that describe the contributions from gas (dotted), the visible disc components (dashed) and a potential halo made from Dark Matter (dashed-dotted). “NGC 6503” is the name for the observed galaxy of which more are listed in Ref. [48].

described in the SM, its existence is suggested through a variety of astrophysical measurements from which three examples will be presented here. One of the most known examples is the observation of galaxy rotation curves [48] which describe the orbital speed of stars or gas in relation to the radial distance from the galaxy’s center. In Newtonian dynamics, it would be expected that the velocity decreases with higher distance according to

$$v(r) = \sqrt{\frac{GM(r)}{r}} \quad \text{with} \quad M(r) = 4\pi \int \rho(r)r^2dr \quad (27)$$

where $\rho(r)$ is the mass density profile. However, observations show that the velocity in certain galaxies stays constant with the distance [48]. This difference can be explained by an additional halo of invisible matter that is defined as DM as it is shown in Figure 5.

Another example that strengthens the theory of DM in the universe is the gravitational lensing effect [49]. In General Relativity, the path of light is not straight, but
is affected by the mass of objects. The general idea of this can be described similarly to refraction of light where the speed of light is reduced in a gravitational field while the strength of this effect is dependent on the refraction index $n$:

$$n = 1 + \frac{2}{c^2}\Phi$$  \hspace{1cm} (28)

where $\Phi$ is the Newtonian potential and $c$ is the speed of light in a vacuum without disturbance through a massive object that distorts space time. Through this approach, the deflection angle of the light can be determined

$$\hat{\alpha} = \frac{4GM}{c^2b}$$  \hspace{1cm} (29)

which is only dependent on the mass $M$ of the object distorting space time and the impact parameter $b$ of the unperturbed light ray. $G$ in this equation is the gravitational constant [50]. An example of gravitational lensing can be found in Figure 6 which shows three different layers of galaxies. First, it shows stars in our own galaxy. Second, it shows members of the galaxy cluster “CL0024+17”. And finally, it also shows even more distant galaxies that lie behind this cluster. Through the gravitational lensing effect, these galaxies appear elongated and as radial arcs. By studying the effect of gravitational lensing, it can not only be determined that the total mass in a certain area is higher than the visible mass, but also what the distribution of the invisible mass is like [49].

The third example is the analysis of the cosmic microwave background (CMB) [22]. The CMB radiation is a remnant of the early universe. In the Big Bang theory, the universe started out small and hot. During this time, particle-antiparticle pairs were created and annihilated. Once the universe expanded and cooled down, this process was no longer possible and the energy density of the universe was dominated by photons. With further reduction of the temperature and therefore reduction of the energy density of photons, electrons and protons started to form the first hydrogen atoms and this period is therefore called recombination. With electrons and protons
no longer being free particles, the photons did not scatter with matter anymore and propagate freely building the radiation background that is known as the CMB. Through further expansion of the universe, this radiation was redshifted and is mostly isotropic with a temperature of around 2.726 K. However, while isotropic at the $10^{-5}$ level, small anisotropies were measured in the CMB in experiments like COBE (COsmic Background Explorer), WMAP (Wilkinson Microwave Anisotropy Probe) and Planck [22, 52, 53, 54, 55, 56]. These are shown in Figure 7. Anisotropies are the result of small fluctuations in the gravitational stability which would then later result in the large stars, galaxies and clusters of matter. Therefore, these anisotropies describe clusters of matter in the early universe. Photons with lower temperatures originated from a point in the universe with a higher concentration of matter. These
clusters of matter were created through gravitation and since Dark Matter does interact with matter through the gravitational force, it has an effect on the creation of these clusters as well. Therefore, by analysing these anisotropies, an evaluation of the existence and amount of Dark Matter can be done.

For this, anisotropies are described through an expansion in spherical harmonics $Y_{l,m}(\theta, \phi)$

$$\frac{\delta T}{T} = \sum_{l=2}^{\infty} \sum_{m=-l}^{l} a_{l,m} Y_{l,m}(\theta, \phi)$$

(30)

where $a_{l,m}$ can be described through the variance $C_l$

$$C_l = \left\langle |a_{l,m}|^2 \right\rangle \quad \text{and} \quad D_l = \frac{l(l+1)}{2\pi} C_l.$$

(31)

By assuming that the temperature fluctuations are Gaussian, they can be described through $D_l$ as a function of the multipole $l$. This was measured by the Planck project of the European Space Agency (ESA) and the results are shown in Figure 8. The data for the temperature power spectrum and therefore the temperature anisotropy of the CMB shows an oscillating pattern that is expected from the spherical harmonics description. Furthermore, the blue line represents the best-fit for the six parameters of the Lambda cold dark matter model ($\Lambda$CDM) which includes components for
Figure 8: Temperature power spectrum $D_{l}^{TT}$ from the Planck 2018 data. The determination $D_{l}^{TT}$ is done in two ways. In the first case ($2 \leq l \leq 29$), the red points represent the power spectrum estimations from the Commander component-separation algorithm [58]. In the second case ($l \geq 30$), the red points represent the maximum likelihood frequency-averaged temperature spectrum which is computed from the Plik cross-half-mission likelihood which is described in more detail in Ref. [59]. The nuisance parameters are determined through the best-fit results of the base-$\Lambda$CDM cosmology. The blue line represents the fitted base-$\Lambda$CDM theoretical spectrum that aligns best with the $D_{l}^{TT}$, $D_{l}^{TE}$, and $D_{l}^{EE}$ likelihoods [56].

dark energy, cold dark matter as well as ordinary matter. Therefore, the good agreement between data and the fitted model does not only function as another hint for dark matter, but it also allows an estimation for the composition between these components in the universe with around 68.3% Dark Energy, 26.8% Dark Matter and 4.9% ordinary matter [54, 56, 57].

2.3 The 2HDM+a model

While DM is a phenomenon that has only been observed in astrophysical data, searches for DM are also possible at particle colliders. The basis for these searches are theories beyond the SM which predict particles that can make up the DM in the universe like WIMPs [23] as well as the production of DM from SM particles. One of
these models is the 2HDM+a, a simplified model, which is an extension of the two Higgs doublet model where an additional mediator $a$ is included. Simplified models are defined with an effective Lagrangian and, in comparison to complete theories, are designed only to add a small amount of parameters and particles to the model. Simplified models for searches at the LHC have become quite popular, because the advantage is that it allows to study different possible models at the same time by designing the simplified models around the general final state of the search. In particular, for DM searches, this includes final states with a high amount of missing transverse energy since the proposed DM particles do not interact directly with the detector. However, these simplified models also come with the disadvantage that they can quickly become too simplified and therefore do not allow for an adequate search for DM at the LHC anymore. In the simplified model that is the basis for the analysis presented in this thesis, the two Higgs doublet model (2HDM) is extended with a single mediator $a$ that interacts with the 2HDM and the DM. This allows for processes with DM in the final state to be produced at the LHC. In Section 2.3.1, the general 2HDM will be shortly presented while Section 2.3.2 will focus on the simplified model and with that, the important final states for this analysis \cite{60, 61}.

\subsection*{2.3.1 Basics of the 2HDM}

The 2HDM is an extension of the SM. As it was explained in Section 2.1.2, one Higgs field is part of the SM that functions as the explanation for the elementary particles to acquire mass. In the 2HDM, a second Higgs field is predicted in addition which
leads to a more complex Higgs potential

\[
V(\Phi_1, \Phi_2) = m_{11}^2 \Phi_1^\dagger \Phi_1 + m_{22}^2 \Phi_2^\dagger \Phi_2 - m_{12}^2 (\Phi_1^\dagger \Phi_2 + \Phi_2^\dagger \Phi_1) \\
+ \frac{\lambda_1}{2} (\Phi_1^\dagger \Phi_1)^2 + \frac{\lambda_2}{2} (\Phi_2^\dagger \Phi_2)^2 + \lambda_3 (\Phi_1^\dagger \Phi_1) (\Phi_2^\dagger \Phi_2) \\
+ \lambda_4 (\Phi_1^\dagger \Phi_2) (\Phi_2^\dagger \Phi_1) + \frac{\lambda_5}{2} \left[ (\Phi_1^\dagger \Phi_2)^2 + (\Phi_2^\dagger \Phi_1)^2 \right] \\
+ \lambda_6 (\Phi_1^\dagger \Phi_1) (\Phi_1^\dagger \Phi_2 + \Phi_2^\dagger \Phi_1) + \lambda_7 (\Phi_2^\dagger \Phi_2) (\Phi_1^\dagger \Phi_2 + \Phi_2^\dagger \Phi_1)
\] (32)

where \( \Phi_1 \) and \( \Phi_2 \) are the two doublets similarly as they were described in Equation (9). Typically a \( Z_2 \) symmetry is imposed such that \( \Phi_1 \rightarrow -\Phi_1 \) and \( \Phi_2 \rightarrow \Phi_2 \) which forces \( \lambda_6 \) and \( \lambda_7 \) to be 0. From the remaining parameters, \( m_{12}^2 \) as well as \( \lambda_5 \) can be complex.

For both Higgs fields, a VEV can be found in the form

\[
\langle \Phi_1 \rangle_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_1 \end{pmatrix}, \quad \langle \Phi_2 \rangle_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v_2 \end{pmatrix}
\] (33)

where \( v_1 \) and \( v_2 \) are real and in order to be consistent with the observations of the SM also have to fulfill the condition \( v_1^2 + v_2^2 = (246 \text{ GeV})^2 \). With these two VEVs, an additional parameter is defined that combines both

\[
\tan \beta = \frac{v_1}{v_2}.
\] (34)

Furthermore, since there is an additional Higgs field, there are eight scalar degrees of freedom from which three are occupied by the Goldstone bosons that will later result in the \( W^\pm \) and \( Z \)-bosons. The other degrees of freedom result in a total of five Higgs particles: two CP-even scalar Higgs bosons which differ only in their mass (\( H \) and \( h \)), one CP-odd scalar Higgs boson (\( A \)) and two CP-even charged Higgs bosons (\( H^\pm \)). For the light CP-even Higgs boson, it is usually assumed to have the mass of the Higgs boson found in the SM with \( m_h = 125 \text{ GeV} \). Furthermore, the mixing angle \( \alpha \) describes the mixture between the massive and light CP-even scalar Higgs
boson eigenstates [62, 63, 64]

\[
\begin{pmatrix}
\rho_1 \\
\rho_2
\end{pmatrix} = \begin{pmatrix}
-\sin \alpha & \cos \alpha \\
\cos \alpha & \sin \alpha
\end{pmatrix} \begin{pmatrix} h \\ H \end{pmatrix}.
\] (35)

The Yukawa couplings in the 2HDM are given as

\[
\mathcal{L}_{\text{Yukawa}} = -\frac{\sqrt{2}}{v} \left[ \bar{Q}_L (M_d \Phi_1 + Y_d \Phi_2) d_R + \bar{Q}_L (M_u i\sigma_2 \Phi_1^* + Y_u i\sigma_2 \Phi_2^*) u_R \\
+ \bar{L}_L (M_l \Phi_1 + Y_l \Phi_2) l_R + \text{h.c.} \right]
\] (36)

where \( Q_L \) and \( L_L \) are the left-handed quark and lepton doublets and \( d_R, u_R \) and \( l_R \) are the singlets. The matrices \( M \) are non-diagonal fermion mass matrices while the matrices \( Y \) describe the Yukawa couplings to the scalar doublet with zero VEV. Through the \( Z_2 \) symmetry, the natural flavour conservation hypothesis is enforced making sure that the Lagrangian does not result in additional flavour-changing interactions. This means that at tree-level, only ever one of the Higgs doublets couples to fermions of a given charge. This can be achieved in four different ways resulting in four types of the 2HDM. In the “type I” model, only \( \Phi_2 \) couples to all types of fermions. In the “type II” model, \( \Phi_1 \) couples to down-type quarks and leptons while \( \Phi_2 \) couples to up-type quarks. In the “type X” model, \( \Phi_1 \) only couples to leptons while \( \Phi_2 \) couples to both types of quarks. And finally, in the “type Y” model, \( \Phi_1 \) couples to down-type quarks while \( \Phi_2 \) couples to up-type quarks and leptons [65, 66].

These types of the 2HDM can be further developed in the Yukawa alignment which guarantees that the matrices \( Y \) and \( M \) are proportional. This means that both types of matrices can be diagonalised at the same time resulting in

\[
Y_{d,l} = \zeta_{d,l} M_{d,l}, \quad Y_u = \zeta_u^* M_u
\] (37)
where the $\zeta_f (f = d, u, l)$ values are complex and are enough to describe all the new couplings in the different models. Furthermore, these couplings are proportional to either $\tan \beta$ or $\cot \beta$ depending on the type and which are summarised in Table 1.

In this analysis, “type II” is used [65].

### 2.3.2 The simplified 2HDM+a

A possible way to produce DM through SM particles is by mixing a SU(2) singlet CP-odd mediator $P$ with the CP-odd Higgs boson that is a result of the 2HDM described in Section 2.3.1. The resulting interaction term is

$$V_P = \frac{1}{2} m_P^2 P^2 + P \left( i b_P \Phi_1^1 \Phi_2 + h.c. \right) + P^2 \left( \lambda_{P_1} \Phi_1^1 \Phi_1 + \lambda_{P_2} \Phi_1^2 \Phi_2 \right)$$

where $m_P$ and $b_P$ are parameters with dimension of mass [66]. By assuming that $V_P$ does not break the CP-symmetry, the mediator $P$ does not develop a third VEV and remains a pure CP eigenstate. Furthermore, in this analysis, the couplings $\lambda_{P_1}$ and $\lambda_{P_2}$ do not affect the phenomenology and therefore are set to 3. Finally, the $b_P$ parameter describes the mixing between the two neutral CP-odd weak eigenstates. Similar to the mixture between the massive and light CP-even scalar Higgs boson eigenstates described in Section 2.3.1, this mixture can be described with a mixing angle called $\theta$ in this case. These two CP-odd eigenstates are denoted with $A$ which is the equivalent to the CP-odd Higgs boson in the 2HDM and $a$ which denotes the

<table>
<thead>
<tr>
<th>Type</th>
<th>$\zeta_d$</th>
<th>$\zeta_u$</th>
<th>$\zeta_l$</th>
</tr>
</thead>
<tbody>
<tr>
<td>“type I”</td>
<td>$\cot \beta$</td>
<td>$\cot \beta$</td>
<td>$\cot \beta$</td>
</tr>
<tr>
<td>“type II”</td>
<td>$-\tan \beta$</td>
<td>$\cot \beta$</td>
<td>$-\tan \beta$</td>
</tr>
<tr>
<td>“type X”</td>
<td>$\cot \beta$</td>
<td>$\cot \beta$</td>
<td>$-\tan \beta$</td>
</tr>
<tr>
<td>“type Y”</td>
<td>$-\tan \beta$</td>
<td>$\cot \beta$</td>
<td>$\cot \beta$</td>
</tr>
</tbody>
</table>

Table 1: Proportionality of the couplings to down-type quarks ($\zeta_d$), up-type quarks ($\zeta_u$) and leptons ($\zeta_l$) depending on the 2HDM type.
additional degree of freedom introduced in this section. Furthermore, by diagonalising the mass-squared matrices of the scalar states, relations between the parameters of $V(\Phi_1, \Phi_2)$ from Equation (32) and $V_P$ from Equation (38) can be introduced. These relations allow to trade the parameters $m_{11}, m_{12}, m_{22}, m_P$ as well as $b_P$ and $\lambda_1, \lambda_2, \lambda_4, \lambda_5$ for sines and cosines of the mixing angles as well as VEVs and masses of the physical Higgs bosons. This means that the physics of the 2HDM+a model can be fully described through the angles $\alpha, \beta$ and $\theta$ as well as $\lambda_3$ and the masses $m_h, m_H, m_A, m_{H^\pm}$ and $m_a$. To avoid constraints from electroweak precision measurements, the masses are required to satisfy $m_H = m_A = m_{H^\pm}$ in this analysis [66].

The DM in this model is introduced as Dirac fermion $\chi$ which couples to the newly introduced mediator $P$ as well through

$$\mathcal{L}_\chi = -iy_\chi P\gamma_5\chi$$

where $\gamma_5$ is the Dirac matrix and the coupling strength $y_\chi$ is set to be real in order to not violate the CP-symmetry. $y_\chi$ and the mass of the DM $m_\chi$ are two additional input parameters in this model. However, as long as the decays $A \to \chi\bar{\chi}$ and $a \to \chi\bar{\chi}$ are kinematically allowed, the mass parameter itself does not impact the phenomenology of the analysis which is why it is set to 10 GeV [66, 67].

In this model, a number of different processes can be analysed. These include processes like DM with a single Higgs boson or $Z$-boson as well as two top quarks in the final state which are shown in Figure 9 [66]. In this analysis, the focus
will lie on final states with DM in association with a single top quark. Three different types of contributions exist for the production of DM signatures: s-channel production, t-channel production and production in association with a W-boson. However, Ref. [68] showed, that the contributions of the s-channel production in simplified DM models are small compared to the other two due to its low cross section. From the other two contributions, two Feynman diagrams result respectively, one with a CP-even Higgs boson from the 2HDM as mediator and one without as Figure 10 shows. The two diagrams without the associated W-boson in the final state interfere destructively. This also applies to the two diagrams with the associated W-boson in the final state. This interference results in a higher total cross section for the process with an associated W-boson in the final state as Figure 11 shows. Furthermore, this figure shows two features of this process. The first one is the fact that the cross section decreases for increasing tan\(\beta\) at low tan\(\beta\), but then also starts to increase again after tan\(\beta = 10\). The second feature can be observed when comparing the two cross sections for the production in association with a W-boson. The inclusive process is combined from the two Feynman diagrams shown in Figure 10 which interfere destructively. The difference between the two diagrams is the intermediate production of a \(H^\pm\) boson during the process. Usually, the diagram with the intermediate CP-even Higgs boson is the dominant one as long as the process \(H^\pm \rightarrow W^\pm a\) is kinematically allowed. This can be observed in Figure 11(a) where

**Figure 10:** Feynman diagrams for the production of DM in association with a single top quark. (a) and (b) show the two diagrams without and (c) and (d) with an associated W-boson in the final state. The processes (a) and (b) as well as (c) and (d) interfere destructively [67].
Cross section for the production of DM from $pp$ collisions in association with a top quark at 14 TeV as a function of $\tan \beta$ with $\sin \theta = 1/\sqrt{2}$, $m_a = 150$ GeV and (a) $m_A = m_H = m_{H^\pm} = 500$ GeV and (b) $m_A = m_H = m_{H^\pm} = 1000$ GeV. The blue line represents the cross section for the $t$-channel production while the other two describe the cross section for the production in association with a $W$-boson. The full line describes the total cross section in this case while the pink dashed line describes the cross section for processes with an intermediate $H^\pm$ boson production [67].

$m_{H^\pm} = 500$ GeV and $m_a = 150$ GeV. However, as Figure 11(b) shows, the diagram without an intermediate $H^\pm$ boson becomes more dominant when increasing the mass of $H^\pm$ to 1000 GeV. The reason for this is that the probability for the production of a $H^\pm$ boson decreases with increasing mass due to the higher amount of energy required by the $b$-quark that produces the top quark and the $H^\pm$ boson. This analysis includes processes with $400$ GeV $\leq m_{H^\pm} \leq 2000$ GeV, $100$ GeV $\leq m_a \leq 450$ GeV and $0.3 \leq \tan \beta \leq 30$. [67].
3 Statistical Framework

For the design of the analysis presented in Chapter 8 as well as the results presented in Chapter 10, a general statistical framework is necessary. Since this analysis is designed by selecting data events with a higher sensitivity to the signal process shown in Section 2.3.2, a baseline for a good definition of a region needs to be presented. This is done using the recommended ATLAS signal significance variable described in Section 3.1. Furthermore, in order to statistically combine the results from the signal regions together with the background estimation and the systematic uncertainties, the framework “HistFitter” is used which is described in detail in Section 3.2 together with the different fits and their statistical background.

3.1 Signal significance

In order to define regions that maximise the sensitivity to the signal process, a variable is used that takes as inputs the prediction for the number of signal process events as well as the sum of the number of SM background events. This signal sensitivity significance $Z$ provides a variable that describes the expected significance of excluding the current background-only hypothesis, if the data is conform with $n = b + s$ where $n$ is the number of observed events and $b$ and $s$ are the number of predicted background and signal process events respectively. The probability of observing $n$ events is assumed to follow the Poisson distribution [69] in every bin of
the histogram

\[ P(n|b) = \frac{b^n}{n!} e^{-b} \]  \hspace{1cm} (40)

if we assume the background-only hypothesis to be true and

\[ P(n|s + b) = \frac{(b + s)^n}{n!} e^{-(b+s)} \]  \hspace{1cm} (41)

if we assume additional contributions through the signal process. The likelihood function is then the product of all Poisson distributions for every of the \( N \) bins in the histogram

\[ L(s) = \prod_{j=1}^{N} P(n_j|s_j + b_j) \]  \hspace{1cm} (42)

which reduces to \( L(s) = P(n|s + b) \) if we assume all events to be in a single bin histogram. However, this likelihood function does not take uncertainties on the background events into account. This can be achieved through the product with a secondary Poisson distribution \( P(m|\tau b) \) leading to

\[ L(s) = P(n|s + b) P(m|\tau b). \]  \hspace{1cm} (43)

This case represents the constrain on the background rate \( b \) through an auxiliary measurement \( m \) of this background rate where \( \tau \) describes the transfer factor relating the event rates in these two regions through \( \tau = b/\sigma^2 \) with \( \sigma \) being the uncertainty on the background rate \( b \) \cite{69}.

A maximum of the likelihood function can be observed and it is described by two maximum likelihood estimators \( \hat{s} = n - m/\tau \) and \( \hat{b} = m/\tau \). If the assumption of the background-only hypothesis is made which means that \( s = 0 \), an additional maximum likelihood estimator for \( b \) can be observed as \( \hat{b} = (n + m)/(1 + \tau) \). Furthermore, for every pair of event rates \( s \) and \( b \), a maximum likelihood ratio \( \lambda(s, b) \) can be calculated as

\[ \lambda(s, b) = \frac{L(s, b)}{L(\hat{s}, \hat{b})} \]  \hspace{1cm} (44)
which changes to

$$\lambda(0) = \frac{\mathcal{L}(0, \hat{b})}{\mathcal{L}(\hat{s}, \hat{b})} = \left(\frac{n+m}{1+\tau}\right)^{n+m} \frac{\tau^m}{n^mm^m}$$

(45)

in the case of the background-only hypothesis. With this, a test statistic \(q_0\) is defined as

$$q_0 = \begin{cases} -2\ln\lambda(0) & \hat{s} \geq 0 \\ +2\ln\lambda(0) & \hat{s} < 0 \end{cases}$$

(46)

While negative numbers of events are possible in certain scenarios, for example if the signal process interferes with a SM background process, this type of interference is not present for the signal process described in Section 2.3.2. Therefore, following here, only the case with \(\hat{s} \geq 0\) will be discussed. The reason for using this test statistic is “Wilk’s theorem” which describes that a test statistic like \(q_0\) which has one parameter of interest \((s)\) asymptotically approaches the \(\chi^2\) pdf for one degree of freedom if the number of events is sufficiently large enough [70]. It can then be shown [71] that the expected discovery significance is

$$Z = \sqrt{q_0} = \sqrt{-2\ln\lambda(0)}.$$  

(47)

If the previously introduced definition for the transfer factor \(\tau = b/\sigma^2\) is applied as well as \(m = \tau b\), then this results in

$$Z = \sqrt{2 \left(n \ln \left[\frac{n(b + \sigma^2)}{b^2 + n\sigma^2}\right] - \frac{b^2}{\sigma^2} \ln \left[1 + \frac{\sigma^2(n - b)}{b(b + \sigma^2)}\right]\right)}$$

(48)

which will be used to design signal regions with maximum signal sensitivity [69].

However, since the estimation of the total background uncertainty is a complex combination of different factors, a conservative assumption is made when designing the signal regions in Chapter 8 with the expected signal significance as baseline variable. In these cases, the uncertainty is set to 20% of the number of background events. For the results in Chapter 10, the uncertainties are calculated precisely.
3.2 HistFitter

For the final results of the analysis, a very similar strategy that was presented in Section 3.1 is used. However, there are two key differences compared to the expected signal significance. First, the likelihood that the analysis is based on includes the actual number of observed events instead of a number of expected events for either the background-only model or the model that includes the signal process as well. Furthermore, the systematic uncertainties will not be included via a Poisson distribution, but through a probability density function $C_{\text{syst}}(\theta^0, \theta)$ where $\theta$ is a set of nuisance parameters with $\theta^0$ being the central value of the auxiliary measurements around which $\theta$ can be varied, so that the likelihood function is maximised. In “HistFitter” [72], this variation has an upper and a lower limit, e.g. $\pm 1$ which describes a possible standard deviation of $1\sigma$. For independent systematic uncertainties, $C_{\text{syst}}$ is just the product of the probability density functions of the individual systematic uncertainties which can be described through a Gaussian distribution

$$C_{\text{syst}}(\theta^0, \theta) = \prod_{j \in S} G(\theta^0_j - \theta_j). \quad (49)$$

This leads to the following likelihood function

$$\mathcal{L}(N, \theta^0|\mu_{\text{sig}}, b, \theta) = \prod_{i \in \text{regions}} P\left(N_i|\mu_{\text{sig}}s_i + \sum_j \mu_j b_{i,j,\theta}, \theta\right) C_{\text{syst}}(\theta^0, \theta) \quad (50)$$

where $N_i$ is the observed number of events in region $i$, $s_i$ is the predicted number of signal events in region $i$, $\mu_{\text{sig}}$ is the signal strength parameter that scales the signal prediction and $b_{i,j,\theta}$ is the number of predicted events for background $j$ in region $i$ in dependence on the nuisance parameter set $\theta$ together with the normalisation parameter $\mu_j$ for that specific background. Similar to Section 3.1, a test statistic is
defined in the way

\[ q_{\mu_{\text{sig}}} = \begin{cases} 
-2 \ln \left( \frac{L(N, \hat{\mu}_{\text{sig}}, \hat{\theta})}{L(N, \mu_{\text{sig}}, \theta)} \right) & \mu_{\text{sig}} \geq 0 \\
-2 \ln \left( \frac{L(N, \hat{\mu}_{\text{sig}}, \hat{\theta})}{L(N, 0, \theta)} \right) & \mu_{\text{sig}} < 0 
\end{cases} \]  

(51)

where \( \hat{\mu}_{\text{sig}} \) and \( \hat{\theta} \) maximise the likelihood while \( \hat{\theta} \) maximises the likelihood for a specific value of \( \mu_{\text{sig}} \). Since no negative signals are assumed, the signal statistic uses the likelihood with \( \mu_{\text{sig}} = 0 \) as the estimator if the maximum likelihood estimator for the signal strength is negative.

The “HistFitter” framework now presents the results of the analysis in three different fits based on this test statistic that will be explained in more detail in the following sections.

3.2.1 Background-only fit

In the background-only fit, another type of regions, control regions, are considered in addition to the signal regions. Control regions are similar to the signal region definitions, but each control region aims at maximising the purity of one specific SM background while only including a negligible amount of signal events. Therefore, the signal contamination is an important part when defining control regions. Due to this assumption, \( \mu_{\text{sig}} \) is set to 0 which leaves the likelihood in Equation (50) for control regions only with the nuisance and the background normalisation parameters. This likelihood function contains the information of every control region and by maximising the likelihood function, a simultaneous fit of all background normalisation parameters is performed at the same time. These background normalisation parameters will not be treated as free parameters in the other fits of the “HistFitter” framework anymore and therefore do not contribute to the test statistic. However, it is necessary to extrapolate the results from the background-only fit to the signal and validation regions. This is schematically shown in Figure 12. For the predicted number of
background events in the signal/validation regions, this is done via transfer factors TF that are defined for a background process $p$ as

$$TF_p = \frac{MC_p(SR, \text{raw})}{MC_p(CR, \text{raw})}$$

(52)

where $MC_p(SR, \text{raw})$ and $MC_p(CR, \text{raw})$ describe the unnormalised events for the background $p$ in the signal and control region. Together with the number of observed data events corresponding to the background $p$ in the control region $N_p(CR, \text{obs.})$, an estimated normalised number of background events in the signal/validation region can be calculated

$$N_p(SR, \text{est.}) = TF_p \cdot N_p(CR, \text{obs.}) = \mu_p \cdot MC_p(SR, \text{raw})$$

(53)

with $\mu_p$ being the fitted normalisation parameter for the background $p$. The advantage of the transfer parameters is that the normalisation components can be absorbed into the ratio which reduces the uncertainty on the normalisation. However, the
nuisance parameters that are connected to both, the signal/validation regions and the control regions, have to be extrapolated as well. This is done through the typical error propagation method

\[
\sigma_{b,tot}^2 = \sum_i^n \left( \frac{\partial b}{\partial \eta_i} \right)^2 \sigma_{\eta_i}^2 + \sum_i^n \sum_{j \neq i} \rho_{i,j} \left( \frac{\partial b}{\partial \eta_i} \right) \left( \frac{\partial b}{\partial \eta_j} \right) \sigma_{\eta_i} \sigma_{\eta_j},
\]

(54)

where \( \sigma_{b,tot} \) is the total uncertainty on the extrapolated background prediction and \( \eta_i \) are floating parameters that are composed of the normalisation parameters \( \mu_k \) for the different backgrounds \( k \) and the nuisance parameters \( \theta_l \). \( \rho_{i,j} \) is the correlation coefficient between \( \eta_i \) and \( \eta_j \) that is evaluated from the background-only fit as well [72].

### 3.2.2 Background-only hypothesis test

Aside from the determination of the normalisation parameters, the background-only fit also provides a test for the background-only hypothesis. This test can be used to interpret if there is any potential new physics in the result by testing the validity of the background-only hypothesis against the measured data. For this, the test statistic from Equation (51) is applied with \( \mu_{\text{sig}} = 0 \) resulting in

\[
q_{\mu_{\text{sig}}} = \begin{cases} 
-2 \ln \left( \frac{\mathcal{L}(N,0,\hat{\theta})}{\mathcal{L}(N,\hat{\mu}_{\text{sig}},\hat{\theta})} \right) & \hat{\mu}_{\text{sig}} \geq 0 \\
-2 \ln \left( \frac{\mathcal{L}(N,0,\hat{\theta})}{\mathcal{L}(N,0,\bar{\theta})} \right) & \hat{\mu}_{\text{sig}} < 0
\end{cases}
\]

(55)

which means that the test statistic is 0 if the maximum likelihood estimator \( \hat{\mu}_{\text{sig}} \) is negative. The agreement or disagreement of the data with the background-only hypothesis is then quantified through the \( p \)-value which is defined as

\[
p_{\mu} = \int_{q_{\mu,\text{obs}}}^{\inf} f(q_{\mu}|\mu) dq_{\mu}
\]

(56)
where $q_{\mu}$ is the test statistic for the signal strength parameter $\mu$ while $q_{\mu, \text{obs}}$ is the value of the test statistic from the observed data and $f(q_{\mu} | \mu)$ is the PDF of $q_{\mu}$. For the background-only hypothesis, this results in

$$ p_0 = p_b = \int_{q_{0, \text{obs}}}^{\inf} f(q_0 | 0) dq_0. \quad (57) $$

The p-value describes the probability that the difference between data and hypothesis is observed due to uncertainty. Therefore, the lower the p-value, the less likely it is for the tested hypothesis to be true. Typically, a p-value of 0.05 is used to determine if a hypothesis is rejected since the interpretation is that the difference in data and hypothesis prediction is only a result of uncertainties in less than 5% of cases. This corresponds to the 95% confidence level [72, 73].

### 3.2.3 Upper limits and exclusion tests

If the background-only hypothesis could not be rejected in the background-only fit, two additional results are provided by “HistFitter” which are the model-independent and the model-dependent fit. In this case, upper limits or exclusion limits are calculated for signals. For this, the definition of test statistic from Equation (51) is slightly altered to

$$ q_{\mu_{\text{sig}}} = \begin{cases} -2 \ln \left( \frac{L(N, \hat{\mu}_{\text{sig}}, \hat{\theta})}{L(N, \hat{\mu}_{\text{sig}}, \hat{\theta})} \right) & \hat{\mu}_{\text{sig}} \leq \mu_{\text{sig}} \\ 0 & \hat{\mu}_{\text{sig}} \geq \mu_{\text{sig}} \end{cases} \quad (58) $$

The reason for setting the test statistic to 0 in the second case comes from the fact that the goal is to define upper limits or exclusion limits for $\mu_{\text{sig}}$ and therefore, only values that are larger than the one that maximises the likelihood function, i.e. being the best value for the agreement between data and prediction, make sense to consider. With this test statistic, a p-value as described in Equation (55) can be calculated for
different values of $\mu_{\text{sig}}$. However, the interpretation of this result changes depending on the model-dependent or model-independent fit.

For the model-dependent fit, the signal process with its prediction of the number of signal events is used. Two hypotheses are compared here: the background-only hypothesis, resulting in the $p$-value $p_b$ and the hypothesis where $\mu_{\text{sig}} = 1$, i.e. the hypothesis with background and signal resulting in a $p$-value $p_s$. From these two values, the $CL_s$ value is determined through

$$CL_s = \frac{p_s}{1 - p_b}$$  \hspace{1cm} (59)$$

which describes the ratio between the probability that the signal hypothesis is falsely excluded over the potential that the background-only hypothesis can be excluded. If this $CL_s$ value is smaller than 0.05, then the signal is excluded with a 95% confidence level. In order to also provide an upper limit on the signal strength parameter, several $p$-values for different values of $\mu_{\text{sig}}$ are calculated. The upper limit on $\mu_{\text{sig}}$ with a 95% confidence level is then defined as the value where the $p$-value falls under 0.05. Furthermore, the potential exclusion of a signal can be improved if statistically independent analysis channels are combined. In this case, the total likelihood function is the product of the likelihood functions of the different analysis channels $i$

$$\mathcal{L}(N, \mu_{\text{sig}}, \theta) = \prod_i \mathcal{L}_i(N_i, \mu_{\text{sig}}, \theta_i)$$  \hspace{1cm} (60)$$

where $\theta$ is the complete set of nuisance parameters combined from the individual sets of the channels.

In comparison, the model-independent fit provides upper limits for the number of events of a generic signal sample. Since there is no assumption about the distribution of the signal, a statistical combination cannot be done. Therefore, if signal regions are split into bins of a specific variable, these are combined into a one-bin signal region again. Otherwise, the fit is done similarly to the model-dependent case with the exception that a generic number of signal events has to be assumed for which
the upper limit on $\mu_{\text{sig}}$ is then calculated. Afterwards, this parameter is converted back into an upper limit on the generic signal events [72, 73].

### 3.2.4 Expected statistical variation

Besides the observed exclusion as well as upper limits, “HistFitter” also provides expected versions of these results together with a variation of $1\sigma$ and $2\sigma$. These results are achieved through an Asimov data set. This set does not consist of measured data, but produced “toy events”. In the following, the assumption is made that a data set follows the hypothesis with a specific signal strength parameter of $\mu_{\text{sig}}'$. Due to fluctuations in the data set, this parameter does not have to be the same as the maximum likelihood estimator $\mu_{\text{sig}}$. But by using Wald’s theorem [74], it can be found that

$$-2 \ln \lambda(\mu_{\text{sig}}) = \frac{(\mu_{\text{sig}} - \mu_{\text{sig}}')^2}{\sigma^2} + \mathcal{O}(1/\sqrt{N})$$

(61)

if there is only one single parameter of interest. In this case, $\mu_{\text{sig}}$ follows a Gaussian distribution with $\mu_{\text{sig}}'$ as the mean and $\sigma$ as the standard deviation while $N$ represents the sample size.

The Asimov data set is now defined such that when evaluating the maximum likelihood estimators, one obtains the true value $\mu_{\text{sig}}'$ that this data set follows. This means that the likelihood ratio for the Asimov data set follows

$$\lambda_A(\mu_{\text{sig}}) = \frac{\mathcal{L}_A(N, \mu_{\text{sig}}, \hat{\theta})}{\mathcal{L}_A(N, \mu_{\text{sig}}, \hat{\theta})} = \frac{\mathcal{L}_A(N, \mu_{\text{sig}}, \hat{\theta})}{\mathcal{L}_A(N, \mu_{\text{sig}}', \hat{\theta})}$$

(62)

while the standard derivation of that Gaussian distribution can be obtained through

$$-2 \ln \lambda_A(\mu_{\text{sig}}) \approx \frac{(\mu_{\text{sig}} - \mu_{\text{sig}}')^2}{\sigma^2}.$$ 

(63)
With this information, an Asimov data set can be created that follows $\mu'_{\text{sig}} = 0$ for the background-only hypothesis and therefore, an expected result can be obtained like described in Section 3.2.2. For the uncertainty band, the value of $\mu'_{\text{sig}}$ is substituted with $\mu'_{\text{sig}} \pm N\sigma$ and the process is repeated [72, 73, 74].
4 LHC and ATLAS

4.1 Large Hadron Collider

The Large Hadron Collider (LHC), part of the European Organisation for Nuclear Research (CERN) and located in the vicinity of Geneva, is the largest and most powerful hadron collider in the world. In a circular ring tunnel of 27 km circumference around 100 m underground, proton beams are accelerated up to an energy of 6.5 TeV and are then collided against each other, reaching a center-of-mass energy of up to 13 TeV, or against heavy-ion beams. Collisions are created at four interaction points around the LHC as shown in Figure 13 which represent the locations of the general purpose detectors ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid) as well as the ALICE (A Large Ion Collider Experiment) and LHCb (LHC-beauty) experiments. While the ALICE experiment mainly focuses on the collision of proton with lead ions and lead ions against each other, the LHCb experiment studies primarily b-hadron physics. In addition to these four main detectors, three smaller experiments are installed as well. The MoEDAL (Monopole and Exotics Detector at the LHC) experiment is located at the same interaction point as the LHCb and is searching for the existence of magnetic monopoles [76]. The TOTEM (TOTal Elastic and diffractive cross section Measurement) experiment is located around the CMS detector and measures the elastic scattering cross section of proton-proton collisions to a high accuracy [77]. The third small experiment is the LHCf (LHC forward) which is located around the ATLAS detector. Its goal is to
Figure 3: The LHC accelerator chain.

ability of a hard collision, the beam is squeezed as much as possible at the interaction point to a diameter of tens of microns. For these operating design conditions, it is expected that on average 20 additional $pp$ interactions will occur.

The LHC accelerator chain is shown in Fig. 3. Initially, 50 MeV protons are produced in the LINAC and accelerated to 1.4 GeV in the Booster. They are then injected in the Proton Synchrotron (PS) where they reach an energy of 26 GeV and are further accelerated to 450 GeV in the Super Proton Synchrotron (SPS). Finally, they are injected in the main ring where they reach a maximum energy of 7 TeV (the maximum to-date has been 3.5 TeV per beam).

The collisions at the LHC take place at the location of the four experiments, which are:
- Compact Muon Solenoid (CMS): One of the two large "general purpose" experiments.
- A Toroidal LHC Apparatus (ATLAS): The other of the two large "general purpose" experiments.

Figure 13: The CERN accelerator complex [75]

measure neutral particles in the forward direction of proton-proton collisions in order to provide more data for the simulation of high energy cosmic ray showers [78]. Figure 13 also shows the accelerator chain designed to reach energies up to 6.5 TeV per proton beam. In the beginning, the protons are produced from hydrogen ions and accelerated up to around 160 MeV in the linear accelerator LINAC [79]. Afterwards, the protons enter different accelerators, like the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS), where the protons reach energies of up to 50 GeV and 450 GeV respectively. From the SPS, the protons are injected into the LHC where two particle beams in opposite directions are accelerated in two adjacent beam lines until they reach the required energy. For this to be possible, the proton beams have to traverse the accelerator tubes without interactions. This is achieved through an ultra high vacuum with a pressure of only $10^{-11}$ to $10^{-10}$ mbar. The LHC alternates between eight straight sections where particles are accelerated through superconducting radiofrequency cavities and eight crossing points with four of these points used for collisions. In order to keep the beams on their trajectory, 1232
superconducting dipole magnets are used which need to be cooled down constantly to a temperature of -271.3°C with liquid helium. This allows the magnets to reach magnetic field strengths of up to 8.3 T. A second task is achieved through another type of magnets. Collisions in the LHC are done with proton bunches instead of individual protons. Each proton bunch consists of around $10^{11}$ protons. Over time, these bunches would disperse. Therefore, additional 392 superconducting quadrupole magnets are installed to focus both proton beams as well as focus the beams before the experiments to achieve higher luminosity. Up to 3564 proton bunches can circulate on both proton-beams at the same time. This means that collisions can occur every 25 ns which results in challenges for the experiments that will be discussed in the following section as well as Chapter 5 [75, 80, 81].

4.2 ATLAS detector

The ATLAS detector shown in Figure 14 is one of the four large scale experiments located at the LHC with a height of 25 m, a length of 44 m and a weight of 7000 t. Its primary focus is to measure particles produced in collisions between protons. In proton-proton collisions, a number of possible particles can be produced, each with their own mass, charge and interaction rate. However, due to the low life time of a lot of particles, these particles decay into more stable particles before they can reach the inner layer of the detector. For example, the top quark, the heaviest elementary particle in the Standard Model, has a lifetime of around $5 \cdot 10^{-25}$ s [83] which results in a decay length of less than $10^{-12}$ mm and therefore, it decays before reaching the innermost part of the detector. Only particles with a life time long enough to reach the detector can be measured. Because of this, the ATLAS detector is divided into several layers which are arranged cylindrical around the beam pipe with the center of the detector being the collision point of the protons. Each layer is optimised to measure a specific set of these particles that reach the detector where the sequence is determined by the interaction rate of the particles. Particles with a
higher interaction rate like electrons do not reach the outer parts of the detector as Figure 15 shows. Therefore, the innermost layers have to be optimised to measure these particles effectively. The first part of the detector is the Inner Detector which measures the tracks of charged particles. The second layer is the electromagnetic calorimeter which has the task of measuring the energy of particles interacting through electromagnetic processes, mainly electrons, positrons and photons. The hadronic calorimeter measures the energy of particles interacting mainly through strong interaction processes. These are hadrons which include particles like protons and neutrons as well as mesons. Finally, muons have a very low interaction rate with matter which is why they traverse through most of the detector. However, since they are charged particles, their momentum can be measured through a tracking detector, the muon spectrometer, designed specifically for muons. The illustration also shows that the neutrino is the only particle in the SM that can not be measured directly through the ATLAS detector due to its low interaction rate as well as its charge of 0. This particle can only be measured indirectly through missing transverse energy
using the conservation of energy in the transverse plane of the coordinate system of
the ATLAS detector [82].

For the coordinate system of the detector, the point of origin is set at the collision
point with the z-axis being defined by the beam axis of the proton direction. The
x-axis is defined as pointing to the center of the LHC while the y-axis points to
the surface. Since the protons move along the z-axis, the spatial coordinates of the
produced particles are usually divided into two directions. The first direction is the
longitudinal direction which is parallel to the z-axis. Observables with this spatial
orientation are labelled with the subscript “L”. The second direction is the transverse
direction which is parallel to the plane created by the x- and y-axis which is transverse
to the z-axis and with that the movement of the protons. Observables with this spatial
orientation get the subscript “T”. The transverse plane is also used for measuring
neutrinos indirectly. Since the protons move on the z-axis, their momentum in the
transverse plane is 0. Therefore, the vectorial sum of the momentum in transverse
direction, $p_T$, for all particles produced in the proton-proton collision has to be 0
as well. Discrepancies in this sum hint at neutral particles that do not have a high
interaction rate with matter which is true for neutrinos in the SM. This variable is either defined as missing transverse momentum ($\vec{p}_T^{\text{miss}}$) or missing transverse energy ($E_T^{\text{miss}}$).

Since the detector is arranged cylindrical around the proton-beam, the azimuthal angle $\phi$ which is the angle in the transverse plane and the polar angle $\theta$ describing the angle from the z-axis, are defined like in polar coordinates. Furthermore, the parity variable $y$ is defined which describes if a particle is moving more parallel or transverse to the beam axis. It is defined as

$$y = \frac{1}{2} \ln \left( \frac{E + p_L}{E - p_L} \right).$$  \hspace{1cm} (64)

and reaches values around 0 if the particle is moving close to transverse to the beam axis and values of $\pm \infty$ if the particle is moving close to parallel to the beam axis. However, this variable requires a good measurement of the energy and momentum of a particle. Therefore, an approximation can be done, assuming a mass of 0, that only requires the angle $\theta$ with

$$\eta = -\ln \left( \tan \frac{\theta}{2} \right).$$  \hspace{1cm} (65)

This variable is called pseudo-rapidity and for relativistic particles, it can be shown that $y \simeq \eta$ if $m = 0$ [85].

Lastly, the $\Delta R$ variable is used to measure the spatial distance between particles. For this, it uses the angular distance $\Delta \phi$ and the distance in pseudo-rapidity $\Delta \eta$ [81]

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

\hspace{1cm} (66)

4.2.1 The Inner Detector

The first layer of the ATLAS detector is the Inner Detector (ID) which is shown in Figure 16. The main task of the ID is the measurement of charged particle tracks. In
The layout of the Inner Detector (ID) is illustrated in figure 1.2 and detailed in chapter 4. Its basic parameters are summarised in table 1.2 (also see intrinsic accuracies in table 4.1). The ID is immersed in a 2 T magnetic field generated by the central solenoid, which extends over a length of 5.3 m with a diameter of 2.5 m. The precision tracking detectors (pixels and SCT) cover the region $|\eta| < 2.5$. In the barrel region, they are arranged on concentric cylinders around the beam axis while in the end-cap regions they are located on disks perpendicular to the beam axis. The highest granularity is achieved around the vertex region using silicon pixel detectors. The pixel layers are segmented in $R - \phi$ and $z$ with typically three pixel layers crossed by each track. All pixel sensors are identical and have a minimum pixel size in $R \times z$ of $50 \times 400 \mu m^2$. The intrinsic accuracies in the barrel are $10 \mu m$ ($R - \phi$) and $115 \mu m$ ($z$) and in the disks are $10 \mu m$ ($R - \phi$) and $115 \mu m$ ($R$).

The pixel detector has approximately 80.4 million readout channels. For the SCT, eight strip layers (four space points) are crossed by each track. In the barrel region, this detector uses small-angle (40 mrad) stereo strips to measure both coordinates, with one set of strips in each layer parallel to the beam direction, measuring $R - \phi$. They consist of two 6.4 cm long daisy-chained sensors with a strip pitch of $80 \mu m$. In the end-cap region, the detectors have a set of strips running radially and a set of stereo strips at an angle of 40 mrad. The mean pitch of the strips is also approximately $80 \mu m$. The intrinsic accuracies per module in the barrel are $17 \mu m$ ($R - \phi$) and $580 \mu m$ ($z$) and in the disks are $17 \mu m$ ($R - \phi$) and $580 \mu m$ ($R$). The total number of readout channels in the SCT is approximately 6.3 million.

A large number of hits (typically 36 per track) is provided by the 4 mm diameter straw tubes of the TRT, which enables track-following up to $|\eta| = 2.0$. The TRT only provides $R - \phi$ information, for which it has an intrinsic accuracy of $130 \mu m$ per straw. In the barrel region, the straws are parallel to the beam axis and are 144 cm long, with their wires divided into two halves, approximately at $\eta = 0$. In the end-cap region, the 37 cm long straws are arranged radially in wheels. The total number of TRT readout channels is approximately 351,000.

order to differentiate between different charges and masses of particles, the detector is surrounded by a magnetic field of 2 T. This magnetic field leads to a curved track of the charged particle which is dependent on its charge and mass. Since all of the measurable particles will cross the ID, the precision and granularity need to be high. This is achieved through a combination of a pixel detector, silicon microstrip trackers (SCT) and a Transition Radiation Tracker (TRT) which utilises straw tubes. With over 80 million individual readout channels over the region $|\eta| < 2.5$ in 1744 pixel sensors, the pixel detector is the part of the ID with the highest precision. The SCT in general is split into a barrel region where the detectors are concentrical to the z-axis and an end-cap region where they are arranged on perpendicular discs around the z-axis. The accuracy of these high precision detectors are $10 \mu m$ in the transverse plane and $115 \mu m$ on the z-axis. Similarly, the TRT is split into a barrel and an end-cap region with the straw tubes being arranged differently in the two regions. In the barrel region, the straw tubes are parallel to the z-axis with a length of 144 cm. In the end-cap region, the arrangement of the straw tubes is radially with 37 cm long
straws which results in 351,000 individual readout channels. In comparison to the SCT and pixel detector, the TRT only provides information in the transverse plane with an accuracy of 130 µm. While the three parts of the detector have similar tasks, their functionality differs. The pixel detector and the SCT work on a similar principle. They are both made out of silicon. Charged particles create electron-hole pairs in the semi-conducting material. Because the material is under a constant electrical field, these free charges then drift to the electrodes where they can be measured as a signal. Since the charged particle drifts to the closest pixel in the pixel detector or strip in the SCT, the track of the charged particle can be reconstructed by combining the pixels and strips that have measured a signal [82]. The TRT in comparison is filled with a gas mixture of xenon, carbon dioxide and oxygen. Besides tracking charged particles through the different straw tubes, the TRT has a secondary task which is the identification of electrons, especially against pions. This is achieved through transition radiation, which is produced when a charged particle with high momentum traverses a boundary between two media with different electric constants. The probability for transition radiation increases with the particle’s Lorentz γ-factor. For electrons and pions with an energy of up to around 100 GeV, the probability to produce transition radiation photons in the X-ray energy range is significantly higher for electrons than pions. Therefore, the number of photon hits in the X-ray energy range per track can be used to separate between electrons and pions [86, 87].

During the first shutdown of the LHC in the years 2013-2014, the ID was extended by the Insertable B-Layer which is shown in Figure 17. It was a new pixel detector layer consisting of silicon pixels between the already existing pixel detector and the beam pipe which had a reduced radius after the shutdown. The motivation for this additional layer was an improved resolution of the pixel detector closer to the interaction point. This is especially important for particles with low $p_T$ values where it improves the impact parameter resolution [88]. The impact parameters describe the closest distance to the collision point of the protons in the transverse plane, $d_0$, and longitudinal direction, $z_0$ [89].
Figure 17: The Insertable B-Layer of the ID added during the LHC shutdown in 2013-2014 [88].

4.2.2 The Calorimeter system

Around the ID is the calorimeter system of the ATLAS detector shown in Figure 18. The main task of this system is to measure the energy of particles. In order to measure the energy of different particles to a high precision, the calorimeter system is split into the electromagnetic calorimeter for measuring the energies of electrons and photons and the hadronic calorimeter measuring the energy of hadrons. In the electromagnetic calorimeter, the energy of high energy particles is measured in two steps. First, the absorber material creates electromagnetic showers. In the ATLAS detector, the absorber material is lead. If a high energy electron enters the calorimeter, high energy photons are produced via bremsstrahlung in the lead which in return produce electron-positron pairs through pair production. Both processes are repeated leading to a shower of particles. Once the energies of the particles are low enough, the detector can measure them in the second step. In the case of the ATLAS detector, this is done through liquid argon which is ionised by the electrons and photons creating a signal with its strength proportional to the energy of the particle similar to the ionised gas presented in the TRT in Section 4.2.1. The sum of
Figure 1.3: Cut-away view of the ATLAS calorimeter system.

Calorimeters must provide good containment for electromagnetic and hadronic showers, and must also limit punch-through into the muon system. Hence, calorimeter depth is an important design consideration. The total thickness of the EM calorimeter is $>22$ radiation lengths ($X_0$) in the barrel and $>24 X_0$ in the end-caps. The approximate 9.7 interaction lengths ($\lambda$) of active calorimeter in the barrel ($10 \lambda$ in the end-caps) are adequate to provide good resolution for high-energy jets (see table 1.1). The total thickness, including 1.3 $\lambda$ from the outer support, is 11 $\lambda$ at $\eta = 0$ and has been shown both by measurements and simulations to be sufficient to reduce punch-through well below the irreducible level of prompt or decay muons. Together with the large $\eta$-coverage, this thickness will also ensure a good $E_{\text{miss}} T$ measurement, which is important for many physics signatures and in particular for SUSY particle searches.

1.3.1 LAr electromagnetic calorimeter

The EM calorimeter is divided into a barrel part ($|\eta| < 1.475$) and two end-cap components ($1.375 < |\eta| < 3.2$), each housed in their own cryostat. The position of the central solenoid in front of the EM calorimeter demands optimisation of the material in order to achieve the desired calorimeter performance. As a consequence, the central solenoid and the LAr calorimeter share a common vacuum vessel, thereby eliminating two vacuum walls. The barrel calorimeter consists of two identical half-barrels, separated by a small gap (4 mm) at $z = 0$. Each end-cap calorimeter is mechanically divided into two coaxial wheels: an outer wheel covering the region $1.375 < |\eta| < 2.5$, and an inner wheel covering the region $2.5 < |\eta| < 3.2$. The EM calorimeter is a lead-LAr detector with accordion-shaped kapton electrodes and lead absorber plates over its full coverage. The accordion geometry provides complete $\phi$ symmetry without azimuthal cracks. The electromagnetic calorimeter in the ATLAS detector is accordion shaped. This provides a full $\phi$ coverage without any cracks while also allowing for a fast extraction of the signal. The electromagnetic calorimeter is further split into a barrel part providing the coverage for $|\eta| < 1.475$ and two end-cap parts providing coverage for $1.375 < |\eta| < 3.2$. These parts are then further divided. The barrel calorimeter consists of two half-barrels divided by a small gap of 4 mm at $z = 0$. Therefore, one half covers the positive and the other half the negative $\eta$ region. Each one has a length of 3.2 m with an outer and inner diameter of 2.8 m and 4 m weighing a total of 57 t. The end-cap components are divided as well into two coaxial wheels with the outer wheel covering the $1.375 < |\eta| < 2.5$ and the inner wheel covering the $2.5 < |\eta| < 3.2$ regions. These two wheels also differ in their precision. Because

all particle energies in the shower is then the energy of the original electron or photon. It is important to note that not only electrons and photons deposit their energy in the electromagnetic calorimeter. However, other particles like muons or hadrons do not deposit their whole energy and therefore can not be measured completely. The electromagnetic calorimeter in the ATLAS detector is accordion shaped. This provides a full $\phi$ coverage without any cracks while also allowing for a fast extraction of the signal. The electromagnetic calorimeter is further split into a barrel part providing the coverage for $|\eta| < 1.475$ and two end-cap parts providing coverage for $1.375 < |\eta| < 3.2$. These parts are then further divided. The barrel calorimeter consists of two half-barrels divided by a small gap of 4 mm at $z = 0$. Therefore, one half covers the positive and the other half the negative $\eta$ region. Each one has a length of 3.2 m with an outer and inner diameter of 2.8 m and 4 m weighing a total of 57 t. The end-cap components are divided as well into two coaxial wheels with the outer wheel covering the $1.375 < |\eta| < 2.5$ and the inner wheel covering the $2.5 < |\eta| < 3.2$ regions. These two wheels also differ in their precision. Because
The outer wheel is used for precision physics the calorimeter has a higher granularity compared to the inner wheel and is also segmented into three sections in depth while the inner wheel is segmented into two. [82, 90]

The hadronic calorimeter works on a similar principle as the electromagnetic calorimeter. By creating showers of particles, the high energy of one particle is split and the components can therefore be measured by the detector. And although these hadrons interact electromagnetically as well, their main interaction to deposit energy is through strong interactions. For this energy deposit, the high energy hadrons interact inelastically with the nuclei of the absorber material. The products of that reaction are hadrons with lower energy that interact with other nuclei again until they reach the pion threshold and pions are produced. These pions decay into a pair of photons which deposit their energy like it was described in the electromagnetic calorimeter. However, in comparison to the electromagnetic calorimeter, part of the energy is absorbed by the nuclei of the material itself leading to an excited state. This can either result in fission, creating electrons and photons through $\beta$- and $\gamma$-decay and therefore electromagnetic showers, de-excitation through spallation creating more nucleons which will interact with nuclei again, or neutron capture with followed fission. In the end, the sum of the energy of all electromagnetic showers represents the energy of the original hadron. However, this difference in shower creation does not only mean that the energy resolution is worse, but furthermore, that a hadronic calorimeter needs to have different properties than an electromagnetic calorimeter. For example, hadronic showers are longer than electromagnetic showers which is why a hadronic calorimeter needs more depth. Therefore, the inner and outer diameters of the 2.6 m long tile calorimeter are 2.28 m and 4.25 m respectively. The tile calorimeter is one of the three parts of the hadronic calorimeter at ATLAS covering the region up to $|\eta| < 1.7$. The other two parts of the hadronic calorimeter are the hadronic end-cap and the forward calorimeter covering the regions $1.5 < |\eta| < 3.2$ and $3.1 < |\eta| < 4.9$ respectively. Furthermore, because the hadronic calorimeter needs to create interactions between hadrons and nuclei, a different absorber material than
1.4 Muon system

The conceptual layout of the muon spectrometer is shown in figure 1.4 and the main parameters of the muon chambers are listed in table 1.4 (see also chapter 6). It is based on the magnetic deflection of muon tracks in the large superconducting air-core toroid magnets, instrumented with separate trigger and high-precision tracking chambers. Over the range $|\eta| < 1.4$, magnetic bending is provided by the large barrel toroid. For $1.6 < |\eta| < 2.7$, muon tracks are bent by two smaller end-cap magnets inserted into both ends of the barrel toroid. Over $1.4 < |\eta| < 1.6$, usually referred to as the transition region, magnetic deflection is provided by a combination of barrel and end-cap fields. This magnet configuration provides a field which is mostly orthogonal to the muon trajectories, while minimising the degradation of resolution due to multiple scattering. The anticipated high level of particle flux has had a major impact on the choice and design of the spectrometer instrumentation, affecting performance parameters such as rate capability, granularity, ageing properties, and radiation hardness.

In the barrel region, tracks are measured in chambers arranged in three cylindrical layers around the beam axis; in the transition and end-cap regions, the chambers are installed in planes perpendicular to the beam, also in three layers.

Figure 19: The muon system of the ATLAS detector [82]

lead is used. In the hadronic end-cap and forward calorimeter copper and tungsten are used while the active material measuring the energy of the electromagnetic shower particles is liquid argon like in the electromagnetic calorimeter. In the tile calorimeter, steel is used as the absorber material, while the electromagnetic shower is measured through scintillating tiles. In a scintillator, photons are produced through excitation of the material through ionisation of particles and the intensity of the produced light, which is increased by photomultiplier tubes, represents the energy [82, 91].

4.2.3 Muon Spectrometer

The muon spectrometer is the outmost part of the ATLAS detector and which is shown in Figure 19. In comparison to the other particles measured at the ATLAS detector, muons have very low cross sections when interacting with matter. This means that they traverse most of the detector without losing much of their energy.
However, this also results in the fact that the muon energy can not be measured by a calorimeter as they are not depositing all of their energy due to the low cross section. Therefore, the measurement of the muon energies is done through another tracking detector and since only muons are entering this tracking detector, the mass of the particles is known which allows for a more precise measurement. This is done by the first two modules of the muon spectrometer, the monitored drift tube chambers (MDT) and the cathode strip chambers (CSC). The MDT and CSC function similarly to the TRT of the ID presented in Section 4.2.1. They are filled with a gas mixture of argon and carbon dioxide which get ionised by the traversing muons. The electrons of the ionised particles are then drifting to the cathode because of an applied electromagnetic field which are either wires in the case of the MDT or strips in the case of the CSC. The main difference between the two detectors is the coverage and resolution. The CSC only covers the region $2.0 < |\eta| < 2.7$ while the MDT also covers areas with $|\eta| < 2.0$. However, the CSC is built to provide a better resolution for the muon tracks with $60 \, \mu\text{m}$ per CSC plane and $80 \, \mu\text{m}$ per MDT layer.

In order to measure the energy and momentum of these particles, a magnetic field orthogonal to the path of the charged particles through toroid magnets is applied with different magnetic field strengths depending on the region in the muon chamber [82].

The other two modules of the muon spectrometer, the resistive plate chamber (RPC) and the thin gap chambers (TGC), are part of the trigger system of the ATLAS detector which will be discussed in more detail in the Section 4.2.4.

### 4.2.4 Trigger System

As it was mentioned in Section 4.1, proton-proton collisions can occur every 25 ns which can lead to a rate of 40 MHz. With an event size of around 1.3 MB this means that the ATLAS detector would require a bandwidth of up to 52 TB/s. As this is not achievable, not all of the events can be written out. Therefore, a two level trigger
system at the ATLAS detector is installed which preselects the events that are written out. A schematic overview of the trigger system is shown in Figure 20. The trigger system has two levels, the Level-1 trigger (L1) and the High-Level Trigger (HLT). The L1 uses a limited amount of detector information to make a decision for discarding an event in around 2.5 μs reducing the rate of events to about 100 kHz. It mainly searches for electrons, muons, photons, jets and hadronically decaying τ-leptons with high $p_T$ values. In addition, it also selects events with high missing transverse energy and during the shutdown of the LHC, two additional systems were implemented to also use information of other variables depending on the L1 objects. If none of the particles or variables have required quantities in the event, it is discarded by the L1. For this decision, the trigger uses reduced granularity information from the different subsets of the detector. For muons, the object quantities are collected by the RPC and TGC (L1Muon) while it is the electromagnetic and hadronic calorimeter (L1Calo) that collects the object information for electrons, photons, hadrons and $E_T^{miss}$. The two additional systems are the L1 topological trigger (L1Topo) and the Central Trigger Processor (CTP). The L1Topo uses the particle information from the
L1 trigger to calculate further variables between the measured objects, for example the invariant mass, in less than 2 $\mu$s. To make this possible, around 15 algorithms are used which allow for a fast estimation. The CTP is the part of the trigger system which is making the decision of discarding an event or pass it to the HLT based on all these informations [92].

The HLT further reduces the rate of events that need to be written out to around 1 kHz. For this, it uses fast algorithms to either analyse the regions of interest (RoI) defined by the L1 or the full event information which is run on a unique PC farm and can process an event in around 0.2 s. Different event filters are applied at the same time in the HLT which can be used by different analyses to determine triggered events. In this analysis for example, a HLT trigger is used that only selects events with high missing transverse energy. Furthermore, for each of these triggers a prescale factor is calculated. If the number of events that need to be written out exceeds the bandwidth of the HLT, only every other event is written out and the prescale factor is later used to recalculate the real number of events. However, for physics analyses, triggers with a prescale factor should not be used if possible [42, 82, 92].
5 Luminosity measurement

The measurement of the delivered luminosity at the ALTAS detector is an important quantity for all analyses as it is necessary to calculate the number of expected events for different physics processes

\[ N_{exp} = \int_0^T L_{inst} dt \cdot \sigma = L_{int} \cdot \sigma \] (67)

where \( L_{int} \) is the integrated luminosity at the ATLAS detector over a certain time and \( \sigma \) is the cross section of the physics process. Furthermore, for high precision measurements, the uncertainty on the luminosity becomes one of its main limitations. Therefore, the precise measurement of the luminosity and with that reducing its uncertainty is one of the primary goals of the ATLAS luminosity program. This chapter will present studies to understand causes for the uncertainty in the track counting luminosity measurement. For this, Section 5.1 will first introduce important terms in the luminosity measurement. Section 5.2 presents the selection of tracks used in the track counting luminosity measurement. Then, Section 5.3 will show the motivation as well as the variables used for this analysis and finally, Section 5.4 presents the results of the analysis for the bunch structure dependent track counting luminosity measurement.
5.1 Luminosity theory

The luminosity delivered to the ATLAS detector is dependent on the number of proton-proton collisions delivered by the LHC. The protons are collided in bunches with around \(10^{11}\) protons per bunch and an average number of interactions of 20-50 protons per beam crossing. The luminosity per colliding bunch can be calculated as

\[
L_b = \frac{\mu \cdot f_r}{\sigma_{\text{inel}}} \tag{68}
\]

with the pile-up \(\mu\) being the average number of inelastic interactions per bunch crossing, the bunch revolution frequency \(f_r\) which is 11245 kHz at the LHC and the inelastic cross section of a proton-proton collision \(\sigma_{\text{inel}}\). When summing over all colliding bunches, the total instantaneous luminosity can be calculated as follows

\[
L_{\text{inst}} = \sum_{b=1}^{n_b} L_b = n_b \cdot \langle L_b \rangle = n_b \cdot \langle \mu \rangle \cdot \frac{f_r}{\sigma_{\text{inel}}} \tag{69}
\]

where \(n_b\) is the number of colliding bunches, \(\langle L_b \rangle\) is the mean bunch luminosity and \(\langle \mu \rangle\) is the average interaction rate. However, the luminosity delivered by the LHC is not necessarily the luminosity measured by the ATLAS detector which is also dependent on the detector efficiency \(\epsilon\). This changes Equation (68) to

\[
L_b = \frac{\mu_{\text{vis}} \cdot f_r}{\sigma_{\text{vis}}} \tag{70}
\]

with \(\mu_{\text{vis}} = \epsilon \cdot \mu\) and \(\sigma_{\text{vis}} = \epsilon \cdot \sigma_{\text{inel}}\) being the visible interaction rate per bunch crossing and the visible cross section respectively [93].

There are several ways to measure the visible interaction rate at the ATLAS detector which are used to reduce the overall uncertainty on the luminosity measurement. One of the methods is the track counting luminosity measurement where the luminosity is proportional to the average number of charged particle tracks measured by the ID.
Depending on the filling pattern, we differentiate the following terms:

1) Individual bunches (INDIVs)
   → One filled BCID between several non-filled BCIDs

2) Bunch trains
   → Series of consecutive filled BCIDs

3) Subtrains
   → If the BCID gap (number of non-filled BCIDs) between two bunch trains is smaller than 10, the 48b trains inside the large train are labeled subtrains.

The proportionality factor for a certain track selection has to be calibrated which is done at the ATLAS detector in specific beam configurations and with the luminosity measured by LUCID (LUminosity Cherenkov Integrating Detector) [94]. Beam configurations can be different depending on the “LHC fill” and are the subject of the analysis presented in Section 5.4. As it was mentioned in Section 4.1, proton bunches can be collided every 25 ns. Therefore, a specific LHC fill has 3564 positions which can be filled with proton bunches. These positions are called bunch crossing identifiers (BCID). However, not all of these positions are filled with proton bunches at all times. This defines different patterns in a proton beam depending on which of the positions are filled with proton bunches and which are left empty. The most important patterns that will be analysed are presented in Figure 21. The first pattern is defined by having one of the BCIDs being filled with a proton bunch while several of the surrounding BCIDs are empty. This pattern is called an “individual bunch pattern”. The second pattern includes several consecutively filled BCIDs and is therefore called a “bunch train”. The final pattern consists of several bunch trains where the number of empty BCIDs between these is less than 10. In this case, this pattern is considered as one large bunch train which consists of several smaller “subtrains”. Physics analyses are typically using LHC fills with bunch trains so that a lot of events can be measured in a short time. In the year 2018, the typical length of per event

\[ \mathcal{L}_{\text{inst}} \propto n_b \cdot \frac{N_{\text{trks}}}{N_{\text{evts}}} \cdot \frac{f_r}{\sigma_{\text{vis}}} \]  

(71)
a bunch train was either 48 consecutively filled BCIDs or bunch trains with a total length of 158 BCIDs which consist of three subtrains with 48 consecutively filled bunches and 7 empty BCIDs between them [93, 95].

5.2 Track selection

To measure the luminosity through the track counting method, a selection of charged particle tracks has to be defined. The goal of the analysis in Section 5.4 is to determine the effect of different track selections on the luminosity measurement with the track counting method. Therefore, three different track selections are defined in the ATLAS detector which are based on the Tight Primary selection of Ref. [96]. This selection includes that only tracks are counted which have a \( p_T > 400 \text{ MeV} \) and \(|\eta| < 2.5\). Furthermore, the selected tracks need to satisfy certain requirements in the pixel detector and the SCT while there is no requirement for the TRT. For a track to be selected it needs to have at least 9 and 11 hits combined in the pixel detector and SCT if \(|\eta| \leq 1.65 \) and \(|\eta| \geq 1.65 \) respectively. For the two innermost layers of the pixel detector the track needs to have at least one hit while there is no hole allowed in the SCT and pixel detector. A hole in the detector is thereby defined if the reconstructed particle track intersects a module but the module did not register a hit and which results in a “missing hit”. Lastly, the number of shared modules for a track is not allowed to exceed 1. A shared hit is defined as a hit that is used by more than one track. For the pixel detector, a module is then considered shared if it has at least one shared hit while a SCT module needs to have at least two shared hits to be considered a shared module.

The three track selections used in this analysis have even tighter requirements. Firstly, in addition to the Tight Primary selection, tracks are only counted if the transverse momentum is at least \( p_T > 900 \text{ MeV} \) and the impact parameter significance is \(|d_0/\sigma_{d_0}| < 7\). This combination of requirements defines the “2016 selection”. The other two selections have a tighter selection with \(|\eta| < 1.0\), but a looser requirement
Table 2: Differences in the track definitions for the three selections. All selections still apply the Tight Primary selection of Ref. [96] as well as $p_T > 900$ MeV and $|d_0/\sigma_{d_0}| < 7$.

for the pixel hole where at maximum one hole is allowed in the pixel detector. This defines the “2017 selection”. The third track selection is the “2017 + Si hit selection” where the only difference to the 2017 selection is that one combined additional hit is required in the pixel detector and SCT. An overview over the differences of the three selections is shown in Table 2 [97].

5.3 Motivation and analysis strategy

As it was mentioned in Section 5.1, the efficiencies for the track selections described in Section 5.2 need to be calibrated. This is done with luminosity measurements from LUCID in a special LHC fill which only includes individual bunches and an average interaction rate of $\langle \mu \rangle \sim 0.5$. However, these are not the beam configurations used in physics analyses. As described in Section 5.1, for physics analyses, bunch trains are used as patterns. Furthermore, the average interaction rate in these LHC fills is usually around 100 times higher at $\langle \mu \rangle \sim 50$. When estimating the luminosity in these LHC fills with the track counting method, it is therefore assumed that the calibration factor is constant when transferring it into the physics regime. But as Figure 22 shows, this is not true for all of the selections used in ATLAS. This figure shows the ratio between the luminosity measured with the track counting method for the three different track selections described in Section 5.2. The calibration factor is determined in the first fill of the figure. Therefore, the ratio between the
Figure 22: Ratio of the run-integrated track-counting luminosity calculated in the 2016 selection, the 2017 selection and the 2017+Si hit selection which are described in Section 5.2. The baseline selection is the 2017 selection. Four different LHC fills are compared with high and low average interaction rate as well as the pattern being either individual (isolated) bunches or bunch trains. [98]

luminosities is expected to be 1 since all of the track selections are calibrated with the same information from LUCID. The second fill changes the pattern to a bunch train pattern while keeping the low average interaction rate. It can be observed that the transfer of the calibration factor has the same effect in all three selections as the ratio between the measured luminosity is still 1. For the third fill, the pattern of individual bunches is kept while the average interaction rate is increased by around 100 times. It can be observed that the transfer of the calibration factor is now affected differently depending on the selection. This results in a slightly different luminosity depending on the track selection. Because the ratio between the 2016 selection and 2017 selection is larger than 1, the 2016 selection estimated a higher luminosity in this LHC fill than the 2017 selection while the opposite is the case for the comparison between the 2017 selection and the 2017+Si hit selection. However, the difference is still small compared to the last fill where both, the pattern and the average interaction rate are changed to what is expected in the physics regime. It can be observed that the transfer of the calibration factor has an effect of around 1-2% on
the measured luminosity for both the 2016 selection and the 2017 + Si hit selection when compared to the 2017 selection. A similar 1-2% effect was measured when determining the uncertainty of the calibration transfer by comparing the ratio of the luminosity measured in the tile calorimeter and with the track counting method

\[ R_{\text{TILE/trk}} = \frac{L_{\text{TILE}}}{L_{\text{trk}}} \]  

in fills with low and high average interaction rate. This is the reason why the calibration transfer is considered one of the largest contributors to the uncertainty of the luminosity measurement which in total is 1.7% [93] and therefore, an analysis understanding the reason for this behaviour is done [98, 99].

The analysis is done in several steps. First, the difference in behaviour between the three different track selections is investigated in two specific LHC fills which feature bunch trains, but which have different average interaction rates of \( \langle \mu \rangle \sim 0.3 \) and \( \langle \mu \rangle \sim 45 \). Both LHC fills are analysed in two ways. In the first part, the ratio of the \( \langle \mu \rangle \) values is calculated in every BCID individually between the different track selections. In the second part, the BCIDs are ordered according to their position inside a bunch train and the \( \langle \mu \rangle \) values are integrated over the whole run before the ratio between the \( \langle \mu \rangle \) values of the different track selections is calculated. Afterwards, this procedure is repeated for more LHC fills with different average interaction rates. The distribution of the ratio of the \( \langle \mu \rangle \) values in dependence on the bunch train position will then be fitted to a function and the parameters are compared in dependence on the average interaction rate. Because the analysis will use the ratio between the \( \langle \mu \rangle \) values of the different track selections, the uncertainty on the ratios will be correlated since some of the tracks in these track selections will be the same. If the correlation between these tracks is not taken into account then the uncertainties will be overestimated as Equation (73)
Figure 23: 2D histogram showing the amount of events with specific values of $\langle \mu \rangle$ for two different selections in a LHC fill with low average interaction rate. The x-axis represents the $\langle \mu \rangle$ values for the 2017 selection while the y-axis represents the $\langle \mu \rangle$ values for the (a) 2016 selection and (b) 2017 + Si hit selection.

\[
\frac{R_{BCID}^{BCID}}{\Delta R_{BCID}^{BCID}} = \sqrt{\left(\frac{\Delta \langle \mu \rangle_{BCID}^{Sel.1}}{\langle \mu \rangle_{BCID}^{Sel.1}}\right)^2 + \left(\frac{\Delta \langle \mu \rangle_{BCID}^{Sel.2}}{\langle \mu \rangle_{BCID}^{Sel.2}}\right)^2 - 2c \cdot \left(\frac{\Delta \langle \mu \rangle_{BCID}^{Sel.1}}{\langle \mu \rangle_{BCID}^{Sel.1}}\right) \left(\frac{\Delta \langle \mu \rangle_{BCID}^{Sel.2}}{\langle \mu \rangle_{BCID}^{Sel.2}}\right)}. \tag{73}
\]

Here, $R_{BCID}^{BCID}$ and $\Delta R_{BCID}^{BCID}$ are the ratio between the $\langle \mu \rangle$ values for a specific BCID and its uncertainty respectively. Similarly, $\langle \mu \rangle_{BCID}^{Sel.i}$ and $\Delta \langle \mu \rangle_{BCID}^{Sel.i}$ describe the $\langle \mu \rangle$ value and its uncertainty at a specific BCID and track selection, for example the 2017 selection. Lastly, $c$ is the correlation factor between the two $\langle \mu \rangle$ values of different track selections.

For this analysis, three versions for estimating the correlation are investigated. The first version uses the ROOT [100] integrated function for the correlation factor inside a 2D histogram. For this, the $\langle \mu \rangle$ values of every individual BCID are calculated for two track selections and the specific bin is filled with an event of weight 1.

The result for this is shown for the 2016 selection and 2017 + Si hit selection against the 2017 selection in Figure 23. The figure shows a strong correlation which is further confirmed by the calculated correlation factor of around 0.96 between the 2016 selection and 2017 selection and more than 0.99 between the 2017 +
Figure 24: Ratio between the \( \langle \mu \rangle \) values of different track selections using two different correlation factors. (a) and (b) use a correlation factor of 0.3 while (c) and (d) use a correlation factor of 0.7. (a) and (c) show the ratio between \( \langle \mu \rangle \) values of the 2016 selection and 2017 selection while (b) and (d) show the ratio between the \( \langle \mu \rangle \) values of the 2017 + Si hit selection and the 2017 selection.

A second test for the correlation is done using a \( \chi^2 \) method. For this, the ratio between the \( \langle \mu \rangle \) values in dependence on the bunch train position is calculated in a LHC fill with low average interaction rate while the uncertainties are estimated using different potential correlation factors. This is presented in Figure 24. Additionally, a constant function is then fitted onto these ratios and the value of \( \chi^2/ndof \) is
estimated where $\chi^2$ is calculated as

$$\chi^2 = \sum_{i=1}^{n} \frac{(x_i - e_i)^2}{\Delta x_i^2}$$

with $x_i$ being the actual value of the ratio between the $\langle \mu \rangle$ values and $e_i$ being the expected value which is 1. $\Delta x_i$ is the uncertainty on the ratio of $\langle \mu \rangle$ which depends on the applied correlation factor. $ndof$ is the number of degrees of freedom which is the difference between the number of points and the number of parameters of the fitted function. Since the fitted function is constant, the number of degrees of freedom is only 1 smaller than the number of points. It is expected that $\chi^2/ndof = 1$. Therefore, different correlation factors are applied to find the factor with the lowest difference of $\chi^2/ndof$ to 1. For the ratio between the 2016 selection and 2017 selection this was found to be between 0.9 and 0.95 while it was found to be higher than 0.99 for ratio between the 2017 + Si hit selection and the 2017 selection.

The last test for the correlation factor is similar to the $\chi^2$ test, but it uses the ratio of the $\langle \mu \rangle$ values of the individual BCIDs instead of the ones in dependence on the position in a bunch train. For every BCID, the square root of $\chi^2$ is calculated and filled into a 1D histogram. This results in a gaussian function with the mean at 0 and a $\sigma$ value dependent on the correlation factor. This is shown in Figure 25. Similarly to before, it is expected for $\sigma = 1$. For the ratios of $\langle \mu \rangle$ values between the 2016 selection and the 2017 selection this is reached for a correlation factor of around 0.95 while the correlation factor needs to be larger than 0.99 for the ratios between the $\langle \mu \rangle$ values of the 2017 + Si hit selection and 2017 selection.

With the three tests combined, it was decided to use a conservative correlation factor for both ratios. For the ratios between the $\langle \mu \rangle$ values of the 2016 selection and the 2017 selection a correlation factor of 0.9 was chosen, while a correlation factor of 0.99 was chosen for the ratios between the $\langle \mu \rangle$ values of the 2017 + Si hit selection and the 2017 selection.
Figure 25: 1D histogram of the square root of the $\chi^2$ values where one event represents the $\sqrt{\chi^2}$ value for the ratio of $\langle \mu \rangle$ values between two selections for an individual BCID. The mean value is always at around 0 while the $\sigma$ value of the gaussian function is dependent on the correlation factor. (a) and (b) show the 1D histograms for a correlation factor of 0.3 while (c) and (d) show the 1D histograms for a correlation factor of 0.7. (a) and (c) show the 1D histograms for ratios between $\langle \mu \rangle$ values of the 2016 selection and 2017 selection while (b) and (d) show the ratios between $\langle \mu \rangle$ values of the 2017+Si hit selection and the 2017 selection.

5.4 Analysis of the bunch structure dependence in the track counting luminosity measurement

The estimation of the luminosity with the calibration factor transferred from a LHC fill with low average interaction rate and individual bunches to the physics regime is investigated for the three different track selections introduced in Section 5.2.
This is done by calculating the ratio between the $\langle \mu \rangle$ values of two track selections respectively in different LHC fills. The first two chosen LHC fills both include bunch trains of 158 BCIDs with subtrains of 48 filled BCIDs and 7 empty BCIDs in between the subtrains. The difference between the two fills is the average interaction rate which is around 0.3 in one case and 45 in the other case. The results of this analysis are shown in Figure 26 which presents the dependence on the position along the bunch string and Figure 27 which presents the dependence on the bunch position inside the train. The first two distributions of both figures show the results for the LHC fill with a low average interaction rate. It can be observed that the ratio of $\langle \mu \rangle$ between the 2016 selection and 2017 selection is not dependent on either the position along the bunch string or the position inside a bunch train as the ratio of $\langle \mu \rangle$ between the selections is always around 1. This is further shown by comparing the ratios of $\langle \mu \rangle$ for BCIDs inside a bunch train to individual bunches in the same fill. The ratio for these individual bunches shows a similar behaviour being around 1. The same observation can be done when analysing the ratio between the $\langle \mu \rangle$ values of the 2017 + Si hit selection and 2017 selection.

However, this behaviour can be observed to change when analysing a LHC fill with a high average interaction rate as shown in the third and fourth distribution of both figures. When comparing the behaviour for the ratio of $\langle \mu \rangle$ between the 2016 selection and the 2017 selection, a pattern can be observed. For the positions along the bunch string, it is shown that the ratio of $\langle \mu \rangle$ between the selections for the first bunches inside the train is higher than the ratio of $\langle \mu \rangle$ for the bunches that are first in a subtrain. Furthermore, when comparing these ratios of $\langle \mu \rangle$ at the first position of bunch trains and subtrains to those in individual bunches, it can be observed that the ratio of $\langle \mu \rangle$ in bunches that are the first in a bunch train behave more like the ratio of $\langle \mu \rangle$ in individual bunches than the ratio of $\langle \mu \rangle$ in bunches that are first in a subtrain. A similar behaviour is observed when analysing the dependence of the bunches inside a train. It is shown in Figure 27(c) that the ratio of $\langle \mu \rangle$ at the beginning of the train is very close to the ratio of $\langle \mu \rangle$ in an individual bunch. However, with consecutively
Figure 26: Ratio of the run-averaged pile-up parameter $\langle \mu \rangle$ between the 2016 selection and 2017 selection ((a) and (c)) and the 2017 + Si hit selection and 2017 selection ((b) and (d)) in dependence on the position along the bunch string. (a) and (b) show the behaviour of the ratios of $\langle \mu \rangle$ for a LHC fill with low average interaction rate while (c) and (d) show the behaviour of the ratios of $\langle \mu \rangle$ for a LHC fill with high average interaction rate.

filled bunches, this ratio decreases until it reaches a plateau which is around 1% lower than the original value. After 7 empty BCIDs between the subtrains, the ratio of $\langle \mu \rangle$ for the first bunch in a subtrain is again increased, but not to the original value of the first bunch in the train. With following consecutive bunches, the ratio between $\langle \mu \rangle$ decreases again until it reaches the same plateau as before. The third subtrain follows the same behaviour as the second.

For the ratio between $\langle \mu \rangle$ values of the 2017 + Si hit selection and the 2017 selection, a slightly different behaviour can be observed. As Figure 26(d) shows, the ratio
Figure 27: Ratio of the run-averaged pile-up parameter $\langle \mu \rangle$ between the 2016 selection and 2017 selection ((a) and (c)) and the 2017 + Si hit selection and 2017 selection ((b) and (d)) in dependence on the bunch position inside the bunch train. (a) and (b) show the behaviour of the ratios of $\langle \mu \rangle$ for a LHC fill with low average interaction rate while (c) and (d) show the behaviour of the ratios of $\langle \mu \rangle$ for a LHC fill with high average interaction rate.

between the $\langle \mu \rangle$ values in bunches that are considered the first in train do not differ as strong from the ratio of $\langle \mu \rangle$ values in bunches that are considered first in a subtrain. However, when comparing these to ratios of $\langle \mu \rangle$ values in individual bunches, a stronger difference can be observed. Furthermore, when analysing the results of Figure 27(d), a different distribution for the ratio of $\langle \mu \rangle$ values can be observed as well. The ratio decreases very quickly by around 1% in the first three bunches of a train, before it starts increasing again until it reaches a plateau. However, after the 7 empty BCIDs, the ratio of $\langle \mu \rangle$ values in the first bunch of a subtrain is already
at a similar value compared to the ratio of $\langle \mu \rangle$ values in the first bunch of a train. Therefore, the distributions are very similar for every subtrain.

This behaviour is further analysed in other LHC fills with different average interaction rates. A very similar distribution for the ratio between $\langle \mu \rangle$ values of the track selections can be observed every time. Therefore, it is investigated if the observed effects are dependent on the average interaction rate. For this, fitting functions for both distributions are implemented. For the ratio of $\langle \mu \rangle$ values between the 2016 selection and the 2017 selection a combination of two functions is used. For the first part, where the ratio decreases, a linear function is included with

$$R(x) = a \cdot x + b$$

(75)

where $R(x)$ is the ratio at the bunch position $x$ inside the subtrain and $b$ being the value at $x = 0$. For a certain bunch position $x = p$ the distribution changes to a constant function with

$$R(x) = a \cdot p + b.$$  

(76)

The potential parameters of this fitting function are the variables $a$, $b$ and $p$. The values that will be compared between different average interaction rates are the number of bunches inside the subtrain it needs for the distribution to reach the plateau value, which is represented by $p$ and the difference between the first value in the subtrain and the value of the plateau represented by the constant value of Equation (76).

For the distribution of the ratio between $\langle \mu \rangle$ values of the 2017 $+ Si$ hit selection and the 2017 selection, a description through addition of two exponential functions is possible. However, for this analysis, two values are of interest which are independent of the fitted function. The first value to be compared between different average interaction rates is the number of bunches inside the subtrain it needs to reach the minimum value of the ratio between $\langle \mu \rangle$ values inside the subtrain. The second value is the difference between the ratio of the $\langle \mu \rangle$ values of the first bunch in a subtrain.
Figure 28: Distribution of the fitted values for the ratio between the $\langle \mu \rangle$ values of the 2016 selection and 2017 selection in dependence of the average interaction rate of the LHC fill. (a) shows the number of BCIDs it needs for the plateau to be reached while (b) shows the difference of the ratio between the $\langle \mu \rangle$ values at the first bunch in the subtrain and the ratio between the $\langle \mu \rangle$ values at the plateau.

and the minimum value inside the subtrain.

The results of the analysis are shown in Figure 28 and Figure 29. It can be observed that the number of bunches it needs to reach either the plateau or the minimum value is not dependent on the average interaction rate. It is constant with around 20-25 bunches in the ratio between the $\langle \mu \rangle$ values of the 2016 selection and 2017 selection and less than 10 for the ratio between the $\langle \mu \rangle$ values of the 2017 + Si hit selection and the 2017 selection. However, the difference in ratio between the $\langle \mu \rangle$ values for the first bunch in a subtrain to the plateau and to the minimum value is shown to be dependent on the average interaction rate where the difference increases with increasing average interaction rate. This is an expected behaviour as the ratio between $\langle \mu \rangle$ values was constant for low average interaction rates. For the ratio between the $\langle \mu \rangle$ values of the 2016 selection and the 2017 selection, a different behaviour can be observed for the first subtrain in a train compared to the second and third. This difference is not observed in the ratio between the $\langle \mu \rangle$ values of the 2017 + Si hit selection and the 2017 selection. However, this difference only originates from the fact that the ratio between the $\langle \mu \rangle$ values of the 2016 selection...
Figure 29: Distribution of the fitted values for the ratio between the $\langle \mu \rangle$ values of the 2017 + Si hit selection and 2017 selection in dependence of the average interaction rate of the LHC fill. (a) shows the number of BCIDs it needs for the minimum ratio in the subtrain to be reached while (b) shows the difference of the ratio between the $\langle \mu \rangle$ values at the first bunch in the subtrain and the ratio between the $\langle \mu \rangle$ values at the minimum.

and 2017 selection is already lower for the first bunch in the second and third subtrain compared to the first subtrain. This feature is not observed in the ratio between $\langle \mu \rangle$ values of the 2017 + Si hit selection and 2017 selection and therefore, no difference between the subtrains is expected.

The reason for this behaviour lies in the design of the pixel detector and the different track selection definitions. In the pixel detector, a pixel measures the charged particle if the signal reaches a certain threshold voltage. However, in order to measure another particle, the signal voltage needs to decrease again below the threshold. The time this takes is called “time over threshold” (ToT). If a second particle is traversing through the same pixel during the ToT, this particle will not be measured by the pixel. Especially for the 2016 selection, this has a significant consequence as this results in a pixel hole and therefore the track is not counted. This is visualised in Figure 30. This design is the reason why the measured average interaction rate for the 2016 selection is lower than the measured average interaction rate for the 2017 selection as the latter allows for one pixel hole in the selected charged particle.
Our Hypothesis: The ToT effect

- any given Pixel is not receptive for a new hit after it was hit for \( \Delta t = \text{ToT} \) (time-over-threshold)

- if occupancy is large: hit in Pixel X in BCID A, hit in same Pixel X in BCID B where \( \Delta t_{AB} < \text{ToT} \) → lose leading edge of hit in BCID B (Pixel shows a hit, but it has wrong BCID)

**Figure 30:** Visualisation of the time over threshold (ToT) in the ATLAS pixel detector. (a) shows the typical process. The incoming particle leads to an increase in voltage which is transferred into a signal with a leading edge (LE). (b) shows what happens if the second particle traverses the signal before the ToT is over. The voltage can not decrease below the threshold and therefore the second LE is missed and with that, the signal. [101]

tracks. Furthermore, this is the reason for the observations in Figure 27. For the first bunch in a bunch train, all pixels are susceptible and can measure the charged particle tracks in every pixel. Therefore, there is no noticeable difference between the 2016 selection and the 2017 selection. However, the second collision of proton bunches already happens 25 ns after the first collision. Since this is shorter than the ToT of the pixel detector, not all pixels are able to measure a signal resulting from this collision. Therefore, more charged particle tracks include a pixel hole and are therefore removed by the 2016 selection but not necessarily by the 2017 selection. This problem exaggerates with further bunch collisions in the bunch trains until it reaches a balance which is represented by the plateau. After the first subtrain, 7 empty BCIDs follow which allows the signal voltage in more pixels to decrease under the threshold again and is the reason for the increase in \( \langle \mu \rangle \) for the 2016 selection. However, it is not enough time for all pixels in the detector to be able to measure a new charged particle again. For this, around 20-30 empty BCIDs are needed which is the time between two large bunch trains. This is also the reason why the ratio of \( \langle \mu \rangle \) in the first bunch of a bunch train behaves similar to an individual bunch as all
pixels in the detector can measure the particle tracks. Furthermore, the reason why
this behaviour is not observed for LHC fills with a low average interaction rate, is
that the number of charged particle tracks is lower in these cases. Therefore, it is
more unlikely that a second charged particle in the consecutive bunches of a bunch
train traverses the same pixel of the detector, resulting in enough time between
measurements for the voltage to decrease under the threshold.

The same effect has a consequence on the ratio between $\langle \mu \rangle$ values of the 2017 +
$Si$ hit selection and the 2017 selection. However, since the track selections are far
more similar to one another with the 2017 + $Si$ hit selection only requiring one
additional hit in the SCT or pixel detector, the effect is not as significant. The
strength of this effect can be observed by comparing the ratio in the first proton
bunch of a train with the ratio in the plateau of Figure 27(d) which is a difference of
around 0.5% compared to the more than 1% difference observed in the comparison
between the 2016 selection and 2017 selection. However, a different effect from the
“01X” readout mode in the SCT can be observed in the comparison between the
2017 + $Si$ hit selection and the 2017 selection. This mode is designed so that a
channel is only read out if a signal was measured for the in-time bunch crossing, but
not for the preceding bunch crossing [102]. For fills with high average interaction
rate, it is more likely for a channel to not be read out since the preceding bunch
was also measuring a signal. Therefore, by requiring an additional hit, the number
of tracks in the 2017 + $Si$ hit selection is reduced after the initial bunch of the train
since more channels are not read out. In later bunches of the train, the previously
explained effect regarding pixel holes starts to affect both selections, reducing the
number of tracks and reducing the effect of the “01X” readout channel of the SCT.
[101].
6 Background samples and data sets

In order to find potential new physics, a well defined prediction of both the SM background as well as the signal process described in Section 2.3.2 is necessary for the statistical analysis described in Chapter 3. This prediction is provided by Monte Carlo (MC) generators [103] that simulate proton-proton collision events at the LHC. These MC events are generated for all of the potential SM backgrounds as well as the signal processes and the production of these events is discussed in this chapter. The simulation of these events is split into different parts which are illustrated in Figure 31 [104].

The first step of the MC event production is the generation of the hard scatter collision process. In this part, MC generators simulate the initial and final state particles based on the matrix element calculations. In the case where partons are part of the final state particles, the hadronisation of these is simulated by Parton Shower MC generators. Depending on the generator, this process can be part of the generator that also does the matrix element calculations. This is done in generators such as SHERPA [105] or PYTHIA8 [106]. However, generators like MadGraph5_aMC@NLO [107] or Powheg [108] only calculate the matrix element and the event information is then transferred to a second generator that simulates the parton shower. The matrix element and parton showers are created at a defined order for the underlying theory. The higher the order, the longer and more computationally intensive is the generation of events. Therefore, when generating particles, a balance between computing time and precision has to be chosen. Usually, events are generated at leading order (LO)
Figure 31: An illustration of the production of a MC event, in this case, the production of a SM Higgs boson in association with a $t\bar{t}$ pair. The different stages of the process are represented by the different colours. The incoming proton beams are represented by the green arrows while the partons of these protons are shown in blue. The red central circle is the hard interaction with the production of the Higgs boson and the $t\bar{t}$ pair which are represented by the smaller red circles. The decay of these particles as well as the hard QCD radiation is shown in red, while a secondary interaction of the proton partons, the underlying event, is represented in purple. The hadrons from the particle decays as well as radiation hadronise which is shown by the light green ellipses and finally decay themselves into more stable particles shown through the dark green colour. The yellow colour indicates photon radiation [104].

up to next-to next-to leading order terms (NNLO). New physics processes are thereby usually generated at LO and NLO while important SM background processes are generated at NNLO [42, 103, 109].

Since the modelling of the hadronisation processes of hadrons in the QCD theory is difficult, the generator relies on phenomenological models. For the partons, both from the colliding protons as well as parton shower calculations, parton density functions (PDFs) are used. The information on these PDFs are taken from PDF libraries and they describe the momentum fraction of the individual elementary particles inside the parton from quarks to gluons. These PDFs are determined in measurements of
specific processes, e.g. the production of a \( Z \)- and \( W \)-boson in association with jets in Ref. [110]. This also means that it is important what data the parton shower generator is tuned to. For example, the A14 tune of the PYTHIA8 shower generator was produced with the full Run 1 data of the LHC [111]. Similarly, proton-proton collisions at the LHC also feature scatter processes with lower energy. These are called underlying events and either result from other particles inside the proton with a lower energy fraction, additional interactions between the same particles as the hard scatter process but a reduced amount of energy or from particles from the radiation processes. Therefore, these are typically calculated in the parton shower generator through phenomenological models as well [42, 103, 109].

However, there is still an important difference between events generated in MC generators and events measured at the LHC which is the detector itself. Depending on the detector structure, the measured particle information can be slightly different to its real counterpart. The generated information is therefore referred to as the “TRUTH” information of the event. This TRUTH information is then further changed to represent the effect the detector response has on the measurement. This information is referred to as the reconstructed (“RECO”) information and represents a faithful simulation of how the events would be measured in the ATLAS detector.

The software including the ATLAS detector geometry is GEOMODEL [112] and it is then used in the GEANT4 framework [113]. However, not all MC events are generated with a full simulation of the detector in GEANT4 since it takes the most amount of time and resources during the modelling of events. Therefore, some samples, especially signal processes which usually have a lot of potentially varying parameters, are simulated with the ATLFAST-II simulation [114] which uses a faster but less precise package for the simulation of the effect that the calorimeter system has on the measured particles called FASTCALOsim [115]. The other parts of the detector like the inner detector are still modelled with the GEANT4 framework in this case [42].

The final difference between these generated MC events and the real data events
measured by the ATLAS detector is the pile-up during the proton-proton collisions at the LHC. To increase the number of collisions, protons are collected in bunches with more than $10^{11}$ protons per bunch. As it was explained in Chapter 5, the typical amount of proton-collisions per bunch crossing is 20-50. This means that other proton-proton collisions can have an influence on the data taken for a specific event. However, since the exact pile-up for every process would take up a lot of disc space, zero-bias events are either generated for example in Pythia8 or taken from measured data and then overlayed over the event generated by the MC generator [116]. Zero-bias events thereby describe events that are saved at random after a certain amount of time at the ATLAS detector without the necessity of any trigger in the trigger system firing during the event. This pile-up background is dependent on the LHC configuration and therefore differs between data taking periods. To account for this difference, the data as well as the MC generated events for the years from 2015-2018 which are analysed in this thesis, are split into three campaigns “MC16a”, “MC16d” and “MC16e” which represent the data taking periods from 2015-2016, 2017 and 2018 respectively and for which a different pile-up profile is overlayed [117]. With this, the data and the MC generated events are saved in the same way and therefore can be analysed and compared [42]. However, for data, not every event from the years 2015-2018 can be used. The ATLAS group provides a “GoodRunList” which describes what data from the whole dataset can be used in physics analysis [118]. The data used in this analysis corresponds to an integrated luminosity of $139.0 \pm 2.4 \, \text{fb}^{-1}$.

The MC generated events in this analysis are divided into background and signal samples, where the backgrounds describe the SM processes that can be expected while the signal samples describe processes presented in Section 2.3.2. For the background samples, a number of different MC generated events are used depending on the SM process and the information is summarised in Table 3. The table shows the generator used for the matrix element and parton shower. In addition to that, it also shows at what precision the cross section of these processes was calculated in the generator.
The SM background processes that are discussed in this analysis are the production of either a $W$- or a $Z$-boson in association with a number of jets ($V + \text{jets}$), the production of a top quark and an anti-top quark ($tt$), the production of a single top quark (Single top) as well as even rarer SM processes like the production of a $tt$ pair in association with either a $W$- or a $Z$-boson ($t\bar{t}V$), an individual top quark in association with a $Z$-boson ($tz$) and a single top quark in association with a $Z$- and $W$-boson ($tWZ$). Furthermore, the “diboson” and “triboson” samples describe SM backgrounds with a combination of two and three $W$- and $Z$-bosons respectively.

For the signal samples the MadGraph generator with version 2.6.2 is used for the generation of the matrix element, while it is interfaced with Pythia8 for the parton showering. The detector simulation is done with the Geant4 framework. The cross sections for these samples are calculated at leading order. Since the signal cross sections as well as the distribution of certain observables is dependent on parameters of the model, a number of signal samples is generated. The parameter for the mass of the DM particles as well as the parameter for $\sin \theta$ are always kept at $m_\chi = 10 \text{ GeV}$ and $1/\sqrt{2}$ respectively. In the first parameter plane, one other parameter is set to a specific value which is $\tan \beta = 1$. The other two parameters which are the mass of the massive Higgs boson $m_H = m_{H^\pm} = m_A$ and the mass of the mediator $m_a$ are varied between $400 \text{ GeV} \leq m_{H^\pm} \leq 2000 \text{ GeV}$ and $100 \text{ GeV} \leq m_a \leq 450 \text{ GeV}$. For the second and third parameter plane, instead of $\tan \beta$, the mass of the mediator $m_a$ is fixed to a value of $250 \text{ GeV}$ and $150 \text{ GeV}$ respectively. In this plane, the mass of

<table>
<thead>
<tr>
<th>Process</th>
<th>Matrix element generator</th>
<th>Parton Shower generator</th>
<th>Cross section calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V + \text{jets}$</td>
<td>Sherpa 2.2.1</td>
<td>Sherpa</td>
<td>NNLO</td>
</tr>
<tr>
<td>$tt$</td>
<td>Powheg-Box</td>
<td>Pythia8</td>
<td>NNLO</td>
</tr>
<tr>
<td>Single top</td>
<td>Powheg-Box</td>
<td>Pythia8</td>
<td>NNLO</td>
</tr>
<tr>
<td>$t\bar{t}V$</td>
<td>MadGraph5_aMC@NLO</td>
<td>Pythia8</td>
<td>NLO</td>
</tr>
<tr>
<td>$tz$</td>
<td>MadGraph5_aMC@NLO</td>
<td>Pythia8</td>
<td>LO</td>
</tr>
<tr>
<td>$tWZ$</td>
<td>MadGraph5_aMC@NLO</td>
<td>Pythia8</td>
<td>NLO</td>
</tr>
<tr>
<td>diboson, triboson</td>
<td>Sherpa 2.2.1-2.2.2</td>
<td>Sherpa</td>
<td>NLO</td>
</tr>
</tbody>
</table>

Table 3: Overview over the generators and precision used for the SM background samples.
the massive Higgs boson is again varied between the same values, but also the tan $\beta$ parameter is varied between $0.3 \leq \tan \beta \leq 30$.

Finally, since the final state of this signal process is very similar to final states from processes where DM is produced in association with a top quark and an anti-top quark, these processes are included in the results as well. For this, samples generated in a simplified model and for the DM+$t\bar{t}$ final state are rescaled to model the final state from the 2HDM+$a$ model described in Section 2.3.2. This is done by calculating the acceptances $A$ in the simplified model once for the mass of the mediator $a$ and once for the mass of the heavy Higgs boson $H$. The acceptance describes the ratio of the number of events after applying the definition of a selection and the number of events before any selection is applied. The acceptance of the DM+$t\bar{t}$ process in the 2HDM+$a$ model and the described selection can be calculated through

$$A_{\text{2HDM}+a}(m_H, m_a) = \frac{\sigma_H \cdot A_{\text{DMSimp}}(m_H) + \sigma_a \cdot A_{\text{DMSimp}}(m_a)}{\sigma_H + \sigma_a}$$

where $\sigma_H$ and $\sigma_a$ are the cross sections for the DM+$t\bar{t}$ process with the masses of $H$ and $a$ respectively. This combined acceptance can then be used to describe the amount of expected events from DM+$t\bar{t}$ processes in the 2HDM+$a$ model [119, 120].
In proton-proton collisions at the LHC, a high number of different particles can be produced such as leptons, quarks and bosons. However, most of these particles are unstable and therefore decay into more stable particles before reaching the inner detector of ATLAS. These stable particles are then measured. In order to reconstruct the processes during the proton-proton scattering necessary to find events with new physics, a multitude of algorithms is used to recreate the trajectories and energies of the decay particles. Therefore, the performance of these algorithms can impact the sensitivity of the analysis. This chapter is presenting the different algorithms used in the ATLAS experiment starting with simpler objects like tracks and calorimeter clusters in Section 7.1 which form the basis of more complex objects like leptons and jets presented in Section 7.2.

7.1 Low-Level Physics Objects

7.1.1 Track Reconstruction

The first step of track reconstruction in the pixel detector and SCT is to cluster the raw measurements. This is achieved through a connected component analysis [121]. These clusters create “space-points” which define the measured three-dimensional points in the detector the charged particles traverse through. For the pixel detector, one cluster represents one space-point. In the SCT, a space-point is created by
combining clusters from both sides of the strip layer. From these space-points, three are then used to create a seed which are further classified into different levels of purity. Seeds of the highest purity originate from only the pixel detector, then only the SCT and lastly through a combination of the two. For further maximisation of the purity, requirements on the momentum and impact parameters are placed. A Kalman filter [122] then creates the track candidates from the seeds by including information from additional space-points. These candidates are classified through “track scores” which are based on the fit quality including the $\chi^2$ of the track fit, the transverse momentum of the track where higher energetic tracks lead to a higher track score, while possible missing clusters along the estimated particle trajectory reduce the score. The track score defines the goodness of the track and therefore is used to define the order in which the tracks are processed. This is important when it comes to solving ambiguities in clusters with multiple track candidates which are identified by a neural network. In these shared clusters, tracks with a higher track score are reconstructed first and are therefore preferred. There are two rules for shared clusters: One cluster can not be shared by more than two tracks at the same time and one track can not have more than two shared clusters. If a new track candidate either leads to a cluster including more than two tracks or that an already accepted track includes more than two shared clusters, the candidate is re-evaluated. For this, the problematic clusters are removed, the track score is calculated again and the candidate is returned to the ordered list of tracks. To reduce the CPU time during this process only fitted tracks that fulfil minimum criteria are included in the ambiguity solver [123].

### 7.1.2 Primary Vertex Reconstruction

Interaction vertices are important to distinguish particles from different interaction points where the primary vertex describes the point where the inelastic proton-proton collision occurred. Primary vertex candidates are reconstructed in two steps: vertex
finding and vertex fitting. First, a seed position is defined where the coordinates in the transverse plane are the centre of the beam spot. For the \( z \)-coordinate, the closest point to the beam axis of the reconstructed tracks described in Section 7.1.1 is used. After the seed is determined, the vertex position is fitted through an iterative algorithm. For each track, a weight is defined depending on the compatibility with the vertex position. With these weights, the optimal vertex position is calculated through a \( \chi^2 \) minimisation and then the weights are recalculated and the process is repeated. After the final iteration, the tracks are re-evaluated and ones that are incompatible with the calculated vertex are removed. These tracks are then used to calculate another candidate for the primary vertex. However, only candidates with at least two associated tracks are considered valid primary vertex candidates [124]. In this analysis, the primary vertex is defined by the highest squared sum over all associated tracks where the tracks included in this sum need to have \( p_T > 0.5 \) GeV.

### 7.1.3 Calorimeter Clusters

For the measurement of the particle energies in the calorimeters, clusters of topologically connected cell signals are created on the basis of the cell signal significance \( \zeta_{\text{cell}}^{\text{EM}} \) which is defined as

\[
\zeta_{\text{cell}}^{\text{EM}} = \frac{E_{\text{cell}}^{\text{EM}}}{\sigma_{\text{noise,cell}}^{\text{EM}}}
\]

where \( E_{\text{cell}}^{\text{EM}} \) is the cell signal and \( \sigma_{\text{noise,cell}}^{\text{EM}} \) is the expected noise in the cell estimated in each run year. First, cluster seeds are defined by selecting cells with \( \zeta_{\text{cell}}^{\text{EM}} > 4 \) and they are ordered by decreasing \( \zeta_{\text{cell}}^{\text{EM}} \). Cells adjacent to these seeds and which satisfy \( \zeta_{\text{cell}}^{\text{EM}} > 0 \) are collected in the same proto-cluster. If one of the seed adjacent cells satisfies the tighter condition \( \zeta_{\text{cell}}^{\text{EM}} > 2 \), cells adjacent to this one with \( \zeta_{\text{cell}}^{\text{EM}} > 0 \) are added to the proto-cluster as well. If two seed cells are adjacent to one another or one cell with \( \zeta_{\text{cell}}^{\text{EM}} > 2 \) is adjacent to two seed cells at the same time, the proto-clusters are merged. This process is repeated until all adjacent cells with \( 0 < \zeta_{\text{cell}}^{\text{EM}} < 2 \) are
added to a proto-cluster. This leads to clusters with highly significant cores and less significant outer regions. This allows the inclusion of cells where the signal is close to the noise while also suppressing the noise effectively through the clustering algorithm [125].

7.2 High-Level Physics Objects

7.2.1 Electrons

Electrons deposit their energy mostly in the electromagnetic calorimeter described in Section 4.2.2 via electromagnetic showering. A smaller part of their energy is deposited through tracks in the ID described in Section 4.2.1. The reconstruction of electrons is therefore based on three measurements: localised clusters of energy in the electromagnetic calorimeter, charged particle tracks in the inner detector and a close matching of the tracks to the clusters in the $\eta - \phi$ space. First, track candidates are defined to the candidate calorimeter seed cluster if $-0.1 < \Delta \phi < 0.05$. If there are several tracks fulfilling the matching criteria, an algorithm is used to define the primary electron by taking into account the distance in $\eta$ and $\phi$ between the track and the cluster center measured in the second layer of the electromagnetic calorimeter, the number of hits in the silicon detector as well as the number of hits in the innermost silicon layer. The decision for an electron instead of a photon candidate is made if the associated track has at least four hits in the silicon layers and no association to a vertex from a photon conversion. If an electron candidate is found, another likelihood based identification is used to identify prompt electrons from background processes. Different likelihood thresholds correspond to different identification working points in ATLAS. Furthermore, isolation criteria are applied to differentiate prompt electrons from backgrounds where quark-pairs decay semileptonically or where hadrons are misidentified as leptons. These isolation criteria can use either calorimeter or track information or both simultaneously [126].
Aside from the identification and isolation, the electron energy measurement has to be calibrated in order to account for different detector effects, i.e. energy lost in the material upstream of the calorimeter and beyond the LAr calorimeter as well as energy deposited in neighbouring cluster cells. This is achieved in several steps which are described in more detail in Ref. [127]. The calibration is done with MC simulations and the calibration constants are determined in a multivariate algorithm. Later, the overall electron response is further calibrated to agree with simulated $Z \rightarrow ee$ events and is validated in $J/\psi \rightarrow ee$ events [127].

In this analysis, two types of electrons are defined. Baseline electrons must satisfy the $p_T > 4.5$ GeV and $|\eta| < 2.47$ requirements as well as the “LooseAndBLayer” likelihood identification operating point [126]. Furthermore, the longitudinal impact parameter $z_0$ relative to the primary vertex of the electron candidate needs to satisfy $|z_0 \sin \theta| < 0.5$ mm. The second type of electrons are signal electrons, which have the same requirements as baseline electrons, but they have to satisfy additional criteria. The likelihood operating point “Medium” must be satisfied which has a lower efficiency but also rejects background more effectively. Furthermore, the significance impact parameter for these electrons must be $|d_0/\sigma(d_0)| < 5$. Signal electrons also have additional isolation criteria. Electrons with $p_T < 200$ GeV have to satisfy the “Loose” isolation working point [128] while electrons with larger $p_T$ have to satisfy the “HighPtCaloOnly” working point.

### 7.2.2 Muons

Muons are reconstructed using the information from the inner detector as well as the muon spectrometer. The track reconstruction described in Section 7.1.1 is done for tracks in the inner detector. For the muon spectrometer a different method is used. Segments inside the MDT and CSC are formed by searching for hit patterns where they are reconstructed through a straight-line fit. Information from the RPC and TGC is further used to measure the track position orthogonal to the bending
plane. The segments in the middle layer of the muon spectrometer are then used as seeds for the algorithm which fits together the hits from segments in different detector layers to create muon candidates. For the building of a track, at least two matching segments are needed except in the barrel-endcap transition region where one high-quality segment is enough to form a track. While one segment can be used to build several tracks, an overlap removal in the algorithm is later used to either select the best assignment to a single track or assignments to at most two tracks at the same time. For the combination of muon candidates from the muon spectrometer and the inner detector, four types are defined where the ordering represents the preference when two muon types share the same ID track. The first reconstruction method is the combined muon where the track reconstructions are performed independently in the inner detector and the muon spectrometer and the muon candidate is then formed with an additional global refit that takes hits from both detectors into account. Typically, muons are reconstructed first in the muon spectrometer and then extrapolated to the inner detector. The second type is the segment-tagged muon where a track from the inner detector is classified as a muon if the extrapolated track is associated with at least one local track segment in the MDT or CSC chambers. This is done for muons that only cross one layer of the muon spectrometer usually because of a low $p_T$. The third category is the calorimeter-tagged muon where the muon is reconstructed if the track from the inner detector is matched to an energy deposit in the calorimeter which is compatible to a minimum-ionising particle. This type is used for muons in regions where the muon spectrometer is only partially instrumented. Finally, the extrapolated muon is only reconstructed with the muon spectrometer track together with a loose requirement that it is compatible with it originating from the interaction point. For this, the muon has to traverse at least two layers of the muon spectrometer chambers except in the forward region where it has to traverse three layers [129].

Similar to electrons, the energy of muons has to be calibrated. This is done by determining calibration constants through a binned maximum-likelihood fit between
data and MC simulated $J/\psi \rightarrow \mu\mu$ and $Z \rightarrow \mu\mu$ events. These calibration constants are then used to determine the corrected momentum $p_{T}^{\text{Cor,Det}}$ of the muon

$$
p_{T}^{\text{Cor,Det}} = \frac{p_{T}^{\text{MC,Det}} + \sum_{n=0}^{1} s_{n}^{\text{Det}}(\eta, \phi) \left( p_{T}^{\text{MC,Det}} \right)^{n}}{1 + \sum_{m=0}^{2} \Delta r_{m}^{\text{Det}}(\eta, \phi) \left( p_{T}^{\text{MC,Det}} \right)^{m-1} g_{m}}
$$

(79)

with $p_{T}^{\text{MC,Det}}$ being the uncorrected transverse momentum in the simulated MC events and $\Delta r_{m}^{\text{Det}}(\eta, \phi)$ and $s_{n}^{\text{Det}}(\eta, \phi)$ being the calibration constants dependent on $\phi$ and $\eta$. $g_{m}$ is a normally distributed number with mean 0 and width 1 [129].

Similarly to electrons, muons have identification criteria to suppress background while keeping a high efficiency with four different working points in ATLAS. However, compared to electrons, this identification is not likelihood based, but it mainly differentiates by using variables like $q/p$-significance, $\rho'$ and the normalised $\chi^{2}$ of the combined track fit. The $q/p$-significance is defined as the absolute value of the difference of the ratios between the charge of the muon and its momentum measured in the ID and the MS divided by the quadratic sum of the uncertainties for all variables. $\rho'$ describes the division of the absolute value of the difference between the measurements of the transverse momentum in the ID and MS and the $p_{T}$ of the combined track. Furthermore, isolation criteria are applied to muons as well to differentiate between muons from decays like $W$, $Z$ or Higgs and muons from semileptonic decays of hadrons and quarks. These isolation criteria are defined on a track and calorimeter basis [129].

In this analysis, two types of muons are defined. Baseline muons require $p_{T} > 4$ GeV, $|\eta| < 2.7$ and they have to satisfy the “Medium” identification criterion [130]. Similarly to electrons, the longitudinal impact parameter $z_{0}$ relative to the primary vertex has to satisfy $|z_{0}\sin \theta| < 0.5$ mm. The second type, signal muons, need to satisfy the same requirements with the addition of the transverse impact parameter significance $|d_{0}/\sigma(d_{0})| < 3$ and a “Loose” isolation criterion [130].
7.2.3 Jets

Jets are built from particle flow objects (PFO) where an algorithm is used to reconstruct jet objects. In the first step, the algorithm tries to match the tracks to a single topo-cluster described in Section 7.1.3. Then, for each combination of a track and a topo-cluster, the algorithm calculates the probability for the particle to deposit its energy in more than one topo-cluster and if necessary adds previously unmatched topo-clusters to the matched clusters. This is done by taking the topo-cluster position as well as the track momentum into account. However, in a last step, the algorithm also accounts for particle shower. By comparing the energy in the combined cluster with the expected deposited energy from the associated track as well as an energy density profile, cells are subtracted from the cluster until the remaining energy is consistent with the fluctuations in the shower of a single particle [131]. If this is the case, these modified clusters and the associated tracks are used as an input for the anti-\( k_t \) algorithm which reconstructs the jets from these clusters. Clusters created from the PFO objects are classified further into seeded objects if they have a \( p_T \) of at least 7 GeV. For these seeds \( i \), the distance to the surrounding clusters \( j \) is calculated as

\[
d_{ij} = \min \left( k_{ti}^{2p}, k_{tj}^{2p} \right) \frac{\Delta_{ij}^2}{R^2}
\]

where \( \Delta_{ij} = \sqrt{\Delta_{\eta}^2 + \Delta_{\phi}^2} \) is the angular distance between the two clusters, \( k_t \) is the transverse momentum of the clusters and the exponent \( p \) defines the general behaviour of the algorithm. For the anti-\( k_t \) algorithm, it is set to \( p = -1 \). The parameter \( R \) is the radius parameter. Additionally, a second distance is calculated

\[
d_{iB} = k_{ti}^{2p}
\]

which describes the distance of the seed cluster \( i \) to the beam. If \( d_{ij} < d_{iB} \) then the clusters are combined and the process is repeated with the newly combined cluster as seed cluster. If \( d_{iB} < d_{ij} \), then the cluster is defined as jet and is removed from
the collection. This process is repeated until no clusters remain [132].

In this analysis, two categories of jets are used. The first category includes jets that are reconstructed with a radius parameter $R = 0.4$ for the anti-$k_t$ algorithm. For these jets, jet energy scale (JES) and jet energy resolution (JER) corrections are applied to restore the jet energy to that of the jets reconstructed at particle level [133]. The jet energy scale calibration is done in bins of $p_T$ of the jet as well as its $\eta$ value and includes a number of steps. In the first stage, the $p_T$ and $\eta$ resolution of the jet is improved through the origin correction which recalculates the four-momentum of the jets so that it points to the calculated primary vertex. The next correction is done to account for the pile-up in two ways. The first step of this correction is a subtraction of the $p_T$ density while the second uses information from MC simulations. The fourth step uses truth information of the jets from dijet MC events to correct the four-momentum closer to the particle level. The last correction for jets from MC simulations is applied through the global sequential calibration which uses different variables from calorimeter and tracking detector information. For jets in data events, a final in situ calibration is applied which corrects the jets by using better measured reference objects like $Z$-bosons and photons in $Z +$ jets and $\gamma +$jets events [134]. The jet energy resolution is determined using a $p_T$ balance method. The $p_T$ asymmetry between the two leading jets in dijet events is determined

$$A = \frac{p_{T,1} - p_{T,2}}{p_{T,1} + p_{T,2}}$$  \hspace{1cm} (82)

and a Gaussian fit is applied to this distribution. From the width $\sigma(A)$ of this fit the jet energy resolution in dependence on the $p_T$, $\sigma(p_T)$, follows from [135]

$$\sigma(A) \approx \frac{1}{\sqrt{2}} \frac{\sigma(p_T)}{p_T}.$$  \hspace{1cm} (83)

This resolution is corrected in two ways. The first correction is the soft radiation correction in which the effect of a third jet on the asymmetry calculation is determined. In the second correction, the particle balance correction, the effect of the particle-level
balance (PB) is determined. The $p_T$ is thereby made of

$$p_{T,2}^{\text{calo}} - p_{T,1}^{\text{calo}} = (p_{T,2}^{\text{calo}} - p_{T,2}^{\text{part}}) - (p_{T,1}^{\text{calo}} - p_{T,1}^{\text{part}}) + (p_{T,2}^{\text{part}} - p_{T,1}^{\text{part}})$$

(84)

with the last term representing the PB. The PB is determined and then subtracted in quadrature from the in situ resolution to determine the effect and the necessary correction [135].

Furthermore, a jet-vertex tagger (JVT) algorithm is applied to these jets as well. For this, a two-dimensional likelihood function is used to calculate the probability in a plane of two variables: the corrected jet vertex fraction (corrJVF) and the scalar sum of the $p_T$ of the tracks that are associated with the jet and originate from the primary vertex divided by the $p_T$ of the fully calibrated jet ($R_{p_T}$). In this two-dimensional plane, the probability for a jet to either originate from a signal type event or a pile-up event are calculated using Monte Carlo dijet simulations [136].

In this analysis, for jets with $|\eta| < 2.4$ and $p_T < 60$ GeV, the “Tight” working point of the JVT needs to be satisfied which corresponds to a JVT score of at least 0.5. Jets with baseline quality have to further satisfy $|\eta| < 4.5$ and $p_T > 20$ GeV while jets with signal quality have to satisfy $|\eta| < 2.5$ and $p_T > 30$ GeV.

The second category includes jets that are reconstructed with a radius parameter $R = 1.0$ which is why they will be called large-$R$ jets from now on to distinguish them from the jets with $R = 0.4$. However, in comparison to the jets with $R = 0.4$, no track information is used for the reconstruction, but only information from the topoclusters of the calorimeter. Additionally, the local calibration weighting scheme (LCW) is applied to the cells of the topoclusters used for the reconstruction of the large-$R$ jets. This calibration is done to compensate for signal inefficiencies in the energy clusters. These inefficiencies come from the non-compensating nature of the ATLAS calorimeters which means that hadronic signals are smaller than electromagnetic signals from electrons and photons as well as inefficiencies due to noise suppression during the clustering and the energy loss in the inactive material of the calorimeter [137]. Furthermore, to reduce the contamination in these large-$R$ jets from pile-up, a
trimming method is applied. In this method, the constituents inside the large-\(R\) jet are reclustered in the \(k_t\) algorithm\(^1\) with a smaller radius parameter \(R_{\text{sub}} = 0.2\). The subjet \(i\) is then discarded if

\[ p_{T,i} < f_{\text{cut}} \cdot p_T^{R=1.0} \]  

\((85)\)

where \(p_{T,i}\) is the transverse momentum of the subjet, \(p_T^{R=1.0}\) is the transverse momentum of the large-\(R\) jet and \(f_{\text{cut}} = 0.05\) is the trimming fraction \([138]\). Similarly to jets with \(R = 0.4\), the JES and JER are corrected \([139]\). The \(p_T\) asymmetry for the JER is thereby determined in simulated events with two large-\(R\) jets. Furthermore, in addition to events where a large-\(R\) jet recoils against a \(Z\)-boson or a photon, events with a dijet system as recoil against the large-\(R\) jet are used as well to determine the JES correction. In addition to that, the jet mass scale (JMS) and jet mass resolution (JMR) are corrected. Most of the energy related properties of jets like the momentum are usually measured through the information from the calorimeter. However, they can also be determined by using information from the tracking detector. \(R_{\text{trk}}\) describes the average of the ratio between the measurements of the calorimeter and tracking detector

\[ R_{\text{trk}} = \left( \frac{m_{\text{calo}}}{m_{\text{track}}} \right) \]  

\((86)\)

and is proportional to the average of the ratio between calorimeter measurement and the truth information of the jet

\[ R_{\text{truth}} = \left( \frac{m_{\text{calo}}}{m_{\text{truth}}} \right) \]  

\((87)\)

Therefore, a ratio between these two variables where \(R_{\text{trk}}\) is determined from dijet data events and \(R_{\text{truth}}\) from MC simulated dijet events is expected to be 1 for well modelled variables. Differences from this value can be applied as a correction of the JMS \([140, 141]\). The JMR can similarly be improved by using the information from the tracking detector and the calorimeter together instead of just the calorimeter

\(^1\)For the \(k_t\) algorithm, the parameter \(p\) in Equation (80) is set to 1.
information. Following the definition of $R_{\text{truth}}$, the improved jet mass resolution can be determined from MC simulated events and through the JMS calibration from $R_{\text{trk}}$ in data events. By determining the median $\mathcal{M}$ and the 68% coverage of the inter-quantile range $iQR_{68}$ of the $R_{\text{truth}}$ distribution for a given $m^{\text{calo}}$ value, the mass resolution is defined as [142]

$$\sigma = \frac{iQR_{68}}{2 \cdot \mathcal{M}}. \quad (88)$$

In this analysis, large-$R$ jets are required to have a $p_T > 200$ GeV and $|\eta| < 2.0$.

### 7.2.4 $b$-tagging

Jets originating from a $b$-quark ($b$-jets) which contain $b$-hadrons as well as their decay products are important for signals with $b$-quarks in the final state. These $b$-hadrons usually differ from other hadrons in lifetime which, if they have a high enough $p_T$, results in a secondary vertex displaced from the primary vertex. This difference from the primary vertex together with other variables is exploited to tag $b$-jets with high efficiencies. First, low level algorithms are used for the reconstruction of the $b$-hadron decay and therefore locate the displaced vertex. Based on the results of these algorithms and other variables, high level algorithms are then deployed to separate $b$-jets from non-$b$-jets. The algorithm used in this analysis is a multivariate algorithm (DL1r) using a deep feed-forward neural network with 28 input variables in total to derive the probabilities for different jet flavours. The main discriminant is

$$D_{DL1r} = \ln \left( \frac{p_b}{f_c \cdot p_c + (1 - f_c) \cdot p_{\text{light}}} \right) \quad (89)$$

where $p_b$, $p_c$ and $p_{\text{light}}$ are the probabilities for the jet being a $b$-jet, a $c$-jet or a light-flavour jet respectively and $f_c$ being the effective $c$-fraction in the background training sample which is a hybrid of simulated $t\bar{t}$ and $Z'$ events. The algorithm is allowing different working points depending on the targeted $b$-jet efficiency [143].

In this analysis, the working point with a 77% efficiency in simulated $t\bar{t}$ events is
chosen. Furthermore, $b$-tagging is only performed on jets with a radius parameter $R = 0.4$, but not large-$R$ jets.

### 7.2.5 $W$-tagging

If a $W$-boson in the event is highly boosted with a $p_T > 200$ GeV and it decays hadronically then the decay products of both quarks can be reconstructed in one jet with a radius parameter $R = 1.0$. In this case, the large-$R$ jets described in Section 7.2.3 can be analysed in two different variables to differentiate the ones that originate from a hardonically decaying $W$-boson to those not originating from a $W$-boson. The first variable is the mass of the large-$R$ jet which needs to be in a $\pm 15$ GeV interval of the $W$-boson mass. The second variable is the $D_2$ variable which is defined as

$$D_2^{(\beta)} = \frac{e_3^{(\beta)}}{(e_2^{(\beta)})^3}$$

with

$$e_2^{(\beta)} = \frac{1}{p_{T,J}^2} \sum_{1\leq i<j \leq n_J} p_T^i p_T^j R_{ij}^2$$

$$e_3^{(\beta)} = \frac{1}{p_{T,J}^3} \sum_{1\leq i<j<k \leq n_J} p_T^i p_T^j p_T^k R_{ij}^\beta R_{ik}^\beta R_{jk}^\beta$$

The tagger allows for two working points. In this analysis, the working point with a tagging efficiency of 50% is chosen [145].

97
7.2.6 Missing Transverse Energy

Since the collision of proton-proton events in the center of the ATLAS detector is symmetric by design, the sum of all transverse momentum vectors $\vec{p}_T$ produced during this collision should be zero. However, due to some particles not interacting with the detector like neutrinos in the SM, an imbalance of transverse energy can occur with the direction being a result of the sum of all trajectories of invisible particles. While these energies and directions can not be measured directly, reconstructing the missing transverse energy $E_T^{\text{miss}}$ as well as its polar coordinates through the energies and momenta of the directly measurable particles, e.g. electrons, muons etc., can give information about these non interacting particles. The algorithm differentiates between a hard term and a soft term, where the hard term are the physics objects defined and used in the analysis. In this analysis, this includes the baseline definitions of jets, electrons and muons. In order to not double count objects, an overlap removal is applied to these objects which is defined in more detail in Section 7.2.7. Furthermore, large-$R$ jets are not included in the calculation of the missing transverse energy to avoid double counting with the jets of radius parameter $R = 0.4$. The soft term is calculated from tracks that originate from the primary vertex but are not matched to any of the electrons, muons or jets. This is typically the case because they have a low $p_T$ and are therefore not reconstructed as a physics object. This results in

$$E_T^{\text{miss}} = - \sum \vec{p}_T^e - \sum \vec{p}_T^\mu - \sum \vec{p}_T^\text{jet} - E_T^{\text{miss,soft}}$$  \hspace{1cm} (92)

for the missing transverse energy.

7.2.7 Overlap removal

In order to avoid double counting of energies in different objects, an overlap removal procedure between objects defined in Section 7.2 is applied. In this analysis, there
are four types of physics objects: Electrons, muons, jets with a radius parameter $R = 0.4$ and large-$R$ jets. An electron is removed with respect to another electron if they have a shared track in the ID. The electron with the higher $p_T$ is kept. An electron is rejected with respect to a muon if they have the same track in the ID except if the muon is a calorimeter-tagged muon. In this case, the muon is rejected. If the angular distance between an electron and jet is $\Delta R < 0.4$, then the electron is rejected against the jet. However, if the angular distance is $\Delta R < 0.2$ and the jet is either not a $b$-jet or has a $p_T > 100$ GeV then the jet is rejected against the electron.

If the angular distance between an electron and a large-$R$ jet is $\Delta R < 1.0$ then the large-$R$ jet is always rejected. The same is true if a muon and a large-$R$ jet have an angular distance of $\Delta R < 1.0$. If a muon and a jet have an angular distance of $\Delta R < 0.4$ then the muon is rejected unless the jet is not a $b$-jet, the number of tracks belonging to this jet is smaller than three and either $\Delta R < 0.2$ or the muon and jet are “ghost-associated”. Ghost-association describes a procedure where the tracks are treated as infinitesimally soft, low-$p_T$ particles with a $p_T$ of 1 eV and which are used when finding tracks connected to the subjets reconstructed in the calorimeter. The advantage of this procedure is that tracks can be matched to calorimeter subjets [146]. When a jet overlaps with a large-$R$ jet, both objects are kept. This is possible since the variables for the large-$R$ jets are only used for $W$-tagging.
8 Event selection

This chapter describes the general analysis strategy using the object and variable definitions of Chapter 7. Before optimising the event selection strategy, a general event cleaning has to be done in order to remove “bad events” from the analysis. This is described in Section 8.1. Afterwards, a short overview over the main observables is presented in Section 8.2 which will then be used in the optimisation of the analysis. A preselection is defined in Section 8.3 aiming at reducing the most dominant SM background processes by defining the general phase-space of the analysis. Afterwards, a more detailed selection is described for the tW_{0L} channel in Section 8.4 with a more thorough comparison between the signal and background processes in order to determine the variables able to increase the sensitivity of the analysis. After determining the variables, they are optimised by using a random-grid search algorithm [147] which optimises the selections for different variables at the same time providing the highest expected signal significance for the analysis. The selection requirements described in Section 8.4 already include the results from this algorithm. The tW_{0L} channel is later combined in Chapter 10 with two additional analysis channels. One channel includes one lepton in the final state (tW_{1L}) and the other includes two leptons in the final state (tW_{2L}). A short overview over the tW_{1L} channel is given in Section 8.5 while the tW_{2L} channel uses the same final state definitions as described in Ref. [29].
8.1 Basic Event selection

For simulated events from Monte Carlo simulations as well as data events, a general event selection is applied to remove what is considered “bad events”. The following events are removed from the collection:

- Events that contain liquid argon (LAr) noise burst
- Events with corrupted data
- Events with corrupted tile calorimeter measurements
- Incomplete events
- Events containing muons that do not originate from proton-proton collision but from cosmic rays
- Events with a jet that did not satisfy all jet cleaning requirements after the overlap removal
- Events that contain significant energy deposits in the hadronic endcap calorimeters or the electromagnetic calorimeters from instrumental effects, cosmic rays or non-collision particles
- Events with at least one muon with insufficient momentum resolution before the overlap removal:

\[
\sigma(q/p)/|q/p| > 0.4
\]

- Events without a primary vertex that has at least two associated tracks with \(p_T > 0.5\) GeV
- Events with electrons where the “DFCommonCrackVetoCleaning” label is set to false \(^1\)

\(^1\)This label resulted from a bug in the association of topoclusters to electrons and photons in the crack region (\(\eta \sim 3.2\)) which resulted in double counting of energies. This bug only appeared in events of the years 2015-2016 (“MC16a”).
8.2 Main discriminants

This section describes the variables that are calculated from the physics objects and which are used to define signal regions in Section 8.4 and Section 8.5 as well as control and validation regions in Section 9.1 and Section 9.2. First, there are different mass variables that will be used. The calculation of masses between two physics objects is done to reconstruct the mass of the mother particle these two objects originate from. Different types of mass variables are used in this analysis. The first type is the invariant mass $m_{12}$ between two physics objects which is defined in natural units as

$$m_{12}^2 = (E_1 + E_2)^2 - ((p_{x1} + p_{x2})^2 + (p_{y1} + p_{y2})^2 + (p_{z1} + p_{z2})^2)$$  \hspace{2cm} (94)

with $E_i$, $p_{xi}$, $p_{yi}$, $p_{zi}$ ($i = 1, 2$) being the energy and the $x$-, $y$- and $z$-coordinates of the momentum of objects 1 and 2. Similarly, the transverse mass $m_T$ is defined like the invariant mass of only the transverse coordinates, i.e. the $x$- and $y$-coordinates. If not specified, this variable is usually used to calculate the mass between a physics object and the $E_{miss}^T$ since there is no $z$-coordinate for the $E_{miss}^T$. In this case, this variable can be calculated through

$$m_T = \sqrt{2p_T E_{miss}^T (1 - \cos \Delta \phi (obj., \vec{E}_{miss}^T))}$$ \hspace{2cm} (95)

where $p_T$ is the transverse momentum of the physics object and $\Delta \phi (obj., \vec{E}_{miss}^T)$ is the difference in the $\phi$ coordinate between the physics object and transverse momentum vector $\vec{E}_{miss}^T$.

Aside from these two general calculations for masses, some more specific mass variables are used as well. The transverse mass $m_{T2}$ is a variable that takes two physics objects and the missing transverse energy as input. It is defined as

$$m_{T2}(\vec{p}_{T,1}, \vec{p}_{T,2}, E_{miss}^T) = \min_{\vec{q}_{T,1} + \vec{q}_{T,2} = E_{miss}^T} \left[ \max(m_T(\vec{p}_{T,1}, \vec{q}_{T,1}), m_T(\vec{p}_{T,2}, \vec{q}_{T,2})) \right]$$ \hspace{2cm} (96)
where $\vec{p}_{T,1}$ and $\vec{p}_{T,2}$ are the transverse momentum vectors of the two objects and $\vec{E}_{T}^{\text{miss}}$ is the missing transverse momentum vector [148]. This variable is unique since it is bounded from above by the mass of the mother particles if both mother particles decay into one visible and one invisible particle respectively, for example in events with leptonically decaying $t\bar{t}$ or $WW$ pairs. In these two cases, the distribution of $m_{T2}$ is bounded sharply from above by the mass of the $W$-boson when including the two leptons in the final state [149]. Furthermore, the asymmetric transverse mass variable $am_{T2}$ is a variation of $m_{T2}$ [150]. It uses one lepton and two jets with the highest $b$-tagging weight as input variables and is therefore defined as

$$am_{T2}(\vec{p}_{T,b1}, \vec{p}_{T,b2}, \vec{p}_{T,l}, \vec{E}_{T}^{\text{miss}}) = \min \left[ m_{T2}(\vec{p}_{T,b1} + \vec{p}_{T,l}, \vec{p}_{T,b2}, \vec{E}_{T}^{\text{miss}}), m_{T2}(\vec{p}_{T,b1}, \vec{p}_{T,b2} + \vec{p}_{T,l}, \vec{E}_{T}^{\text{miss}}) \right]$$

(97)

where $\vec{p}_{T,b1}$, $\vec{p}_{T,b2}$, $\vec{p}_{T,l}$ are the transverse momentum vectors of the jets with the highest and second highest $b$-tagging weight as well as the lepton respectively. It is the minimal value between two $m_{T2}$ calculations where the lepton is combined with one of the jets each time. It is a variable that can be used effectively to reduce SM $t\bar{t}$ background events where both top quarks decay leptonically but one of the leptons is not detected [150].

An additional mass variable is the $m_{W}^{\text{had}}$ variable. For this variable, an iterative clustering algorithm, described in detail in Ref. [151], is used to reconstruct hadronically decaying $W$-bosons. All signal jets with $R = 0.4$ in the event are reclustered again with the anti-$k_t$ algorithm, but with a larger radius parameter of $R = 3.0$. The radius of each of those large-radius jets is then iteratively reduced to the optimal radius $R(p_T) = 2 \cdot m_W / p_T$ with $p_T$ being the transverse momentum of the large-radius jet.

In the end, the mass of the jet with the mass closest to the $W$-boson mass is taken and defines the $m_{W}^{\text{had}}$ variable.

The $m_{bl}^{\text{min}}$ variable describes the minimum invariant mass between the $b$-jet with the
highest $p_T$ and either of two leptons:

$$m_{b_1l_1}^{\text{min}} = \min(m_{b_1l_1}, m_{b_1l_2}).$$  \hspace{1cm} (98)

The $m_{b_1l_1}^{i}$ variable on the other hand is an extension of the $m_{b_1l_1}^{\text{min}}$ variable which uses an additional jet. It is defined as

$$m_{b_1l_1}^{i} = \min \left[ \max(m_{l_1j_1}, m_{l_1j_2}), \max(m_{l_2j_1}, m_{l_2j_2}) \right]$$  \hspace{1cm} (99)

where $m_{l_nj_m}$ is the invariant mass of the lepton and jet. The two jets that are used for this variable are the jets with the highest $b$-tagging discriminator value. For $t\bar{t}$ and $t\bar{t}V$ events where both of the top quarks decay leptonically, this variable has a kinematic edge at around 160 GeV.

Another discriminant is the object-based missing transverse momentum significance $S_{E_T^{\text{miss}}}$. This variable discriminates between events where $E_T^{\text{miss}}$ arises from poorly reconstructed particles against events where $E_T^{\text{miss}}$ arises from invisible particles. It is a likelihood driven evaluation of the missing transverse momentum in an event. The higher $S_{E_T^{\text{miss}}}$, the higher the likelihood that the missing transverse energy arises from invisible particles. It is defined as

$$S_{E_T^{\text{miss}}} = \frac{|E_T^{\text{miss}}|}{\sqrt{\sigma^2_L(1 - \rho^2_{LT})}}$$  \hspace{1cm} (100)

where $E_T^{\text{miss}}$ is the vector of the missing transverse momentum. $\sigma_L$ is the total estimated longitudinal momentum resolution of all jets and leptons at a given $p_T$ and $|\eta|$. For jets, this resolution is defined as the maximum resolution between the ones measured in MC events and data. The total estimated momentum resolution is then the sum of the jets and leptons in the event projected on a basis parallel to $E_T^{\text{miss}}$. $\rho_{LT}$ is a correlation factor between the longitudinal and transverse momentum resolution of each object with respect to $E_T^{\text{miss}}$. Further details on the calculation of this variable can be found in Ref. [152].
Finally, two additional variables are \( \min[\Delta \phi(\text{jet}_{1-4}, E_{T}^{\text{miss}})] \) and \( \min[\Delta \phi(\text{jet}_{\text{all}}, E_{T}^{\text{miss}})] \). These describe the smallest azimuthal angular distance between the \( E_{T}^{\text{miss}} \) and either the four jets with the highest \( p_T \) or all signal jets. This is a variable used to discriminate against QCD events with two highly energetic jets in the final state. These events can have a high \( E_{T}^{\text{miss}} \) value if one of the jets is mismeasured. In this case, the direction of the \( E_{T}^{\text{miss}} \) is very close to the direction of one of the jets resulting in a low \( \min[\Delta \phi(\text{jet}_{1-4}, E_{T}^{\text{miss}})] \) and \( \min[\Delta \phi(\text{jet}_{\text{all}}, E_{T}^{\text{miss}})] \) value.

### 8.3 Preselection

As described in Section 2.3.2, the final state of the signal consists of DM particles, one top quark and one additional \( W \)-boson, which originates either from the \( H^\pm \) boson or the back to back production with the mediator \( a \). Since DM can not be detected, one of this final state’s most important variables is \( E_{T}^{\text{miss}} \). Due to the high mass of the mediator of at least 100 GeV, a high missing energy is assumed in this model. Therefore, a selection of \( E_{T}^{\text{miss}} > 250 \) GeV is applied. This defines the general phase-space. The value of 250 GeV is also chosen since it allows the usage of the missing transverse energy trigger in ATLAS with a 100% efficiency \[153\]. However, a phase-space in a high \( E_{T}^{\text{miss}} \) region is also often susceptible to mis-identifications for certain processes. Events including multiple jets usually do not have high \( E_{T}^{\text{miss}} \), but due to mismeasurements of one of the jets in the detector, these events can still appear to have a high \( E_{T}^{\text{miss}} \) value. In order to remove these events from the analysis, the selection \( \min[\Delta \phi(\text{jet}_{1-4}, E_{T}^{\text{miss}})] > 0.5 \) described in Section 8.2 is applied. This removes the described events since the direction of \( E_{T}^{\text{miss}} \) resulting from the mismeasurement is then parallel to one of the jets in the event.

Furthermore, one top quark is expected in the signal. Since it decays to a \( W \)-boson and a \( b \)-quark in more than 99% of cases \[154\], at least one \( b \)-tagged jet is required in the event. Since a \( b \)-jet originating from a top quark decay can be expected to have a high \( p_T \) due to the high mass of the top quark, the transverse momentum
of the $b$-jet with the highest $p_T$ needs to be at least 50 GeV. This selection mainly removes events where $b$-jets originate from different sources, for example secondary processes besides the hard scatter process. In a case with $b$-tagged jets in the event, the $t\bar{t}$ pair production usually becomes a dominant background in the analysis. In order to reduce this process, an upper limit on the transverse momentum of 50 GeV is placed on the $b$-tagged jet with the second highest $p_T$.

Since this analysis focuses on the analysis channel with zero leptons in the final state, another requirement for the event is that no baseline leptons defined in Chapter 7 are present. This means for the signal that both $W$-bosons decay hadronically. Together with the $b$-quark originating from the top quark decay, a total number of five quarks and therefore five jets can be expected in the signal final state. However, due to a jet possibly falling out of the detector acceptance, a slightly looser definition of at least four jets with high transverse momentum is required in the event. These jets can, but do not have to be $b$-tagged.

The $tW_{1L}$ channel follows a similar preselection strategy, but requiring exactly one signal lepton and no additional baseline leptons. Since this means that one of the $W$-bosons decays leptonically, the jet multiplicity is reduced to two and their $p_T$ requirements are adjusted as well.

A summary of the whole preselection for the $tW_{0L}$ and $tW_{1L}$ channels can be found in Table 4. Furthermore, Figure 32 shows four important kinematic variables after applying the $tW_{0L}$ preselection criteria. These include the distribution for the $E_T^{\text{miss}}$, the $p_T$ of the $b$-tagged jet with the highest momentum, the $p_T$ of the jet with the highest momentum and the $\min[\Delta\phi(\text{jet}_{1-4}, E_T^{\text{miss}})]$ variable. It can be observed that the dominant SM processes in this region are $t\bar{t}$ (54%), $Z+\text{jets}$ (23%) and $W+\text{jets}$ (14%).
<table>
<thead>
<tr>
<th>Variable</th>
<th>tW$_{0L}$</th>
<th>tW$_{1L}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trigger</td>
<td>$E_T^{\text{miss}}$</td>
<td>$E_T^{\text{miss}}$</td>
</tr>
<tr>
<td>$E_T^{\text{miss}}$ [GeV]</td>
<td>$\geq 250$</td>
<td>$\geq 250$</td>
</tr>
<tr>
<td>Number of baseline leptons</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Number of signal leptons</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$p_T^{\ell}$ [GeV]</td>
<td>$\geq 30$</td>
<td>$\geq 30$</td>
</tr>
<tr>
<td>Number of signal jets</td>
<td>$\geq 4$</td>
<td>$\geq 2$</td>
</tr>
<tr>
<td>$p_T^{j1}$ [GeV]</td>
<td>$\geq 100$</td>
<td>$\geq 30$</td>
</tr>
<tr>
<td>$p_T^{j2}$ [GeV]</td>
<td>$\geq 60$</td>
<td>$\geq 30$</td>
</tr>
<tr>
<td>$p_T^{j3}$ [GeV]</td>
<td>$\geq 60$</td>
<td>$-$</td>
</tr>
<tr>
<td>$p_T^{j4}$ [GeV]</td>
<td>$\geq 40$</td>
<td>$-$</td>
</tr>
<tr>
<td>Number of $b$-tagged jets</td>
<td>$\geq 1$</td>
<td>$\geq 1$</td>
</tr>
<tr>
<td>$p_T^{b1}$ [GeV]</td>
<td>$\geq 50$</td>
<td>$\geq 50$</td>
</tr>
<tr>
<td>$p_T^{b2}$ [GeV]</td>
<td>$\leq 50$</td>
<td>$\leq 50$</td>
</tr>
<tr>
<td>$\min[\Delta\phi(jet_{1-4}, E_T^{\text{miss}})]$</td>
<td>$\geq 0.5$</td>
<td>$\geq 0.5$</td>
</tr>
</tbody>
</table>

**Table 4:** The preselection for the tW$_{0L}$ and tW$_{1L}$ channels of the analysis.

### 8.4 The tW$_{0L}$ channel

This section outlines the selections used after the preselection defined in Section 8.3. This will be done in detail for the tW$_{0L}$ channel in this section while Section 8.5 will shortly present the selection for the tW$_{1L}$ channel. The variables, defined in Section 8.2, are explained in more detail by comparing the signal physics process described in Section 2.3.2 to the SM backgrounds which are the most dominant after the preselection.

As Section 8.3 described, the dominant backgrounds after applying the preselection requirements are $t\bar{t}$, $Z$ + jets and $W$ + jets with $t\bar{t}$ being the most dominant. Therefore, the first goal is to reduce this background while keeping most of the signal events. The preselection already defines that only events without any leptons and high $E_T^{\text{miss}}$ enter the selection. However, when analysing the SM $t\bar{t}$ process, it can be noticed that both of these requirements can only very rarely be fulfilled. Figure 33 shows all three possible decay channels for the SM $t\bar{t}$ process. The first one is the hadronic
Therefore, a high \( E_T \) measured in the detector assuming a perfect measurement and therefore, the event if a lepton in one of these final states is either an electron or muon, it would be the third decay channel include a neutrino allowing for a higher \( E_T^{\text{miss}} \) value. However, if a lepton in one of these final states is either an electron or muon, it would be measured in the detector assuming a perfect measurement and therefore, the event decay, where no lepton can be found in the final state with a branching ratio of 44%, the second one is the lepton+jets channel, where either one of the top quarks decays leptonically with a branching ratio of 45% and the third one is the dileptonic \( tt \) channel with two leptons in the final state and a branching ratio of 11% [32]. In the first case, it can be observed that all decay products are measurable in the detector. Therefore, a high \( E_T^{\text{miss}} \) in the final state is very unlikely. Only the second and third decay channel include a neutrino allowing for a higher \( E_T^{\text{miss}} \) value. However,
Figure 33: The three decay channels for the SM $t\bar{t}$ background: (a) the hadronic $t\bar{t}$ decay channel, (b) the lepton+jets $t\bar{t}$ decay channel (c) the dileptonic $t\bar{t}$ decay channel.

Figure 34: The two decay channels of the $\tau$ lepton: (a) the hadronic $\tau$ decay channel, (b) the leptonic $\tau$ decay channel. The $W$-boson in these cases is produced off-shell$^2$.

would not enter the zero lepton selection. The only decay channel which fulfils both requirements is one where at least one of the top quarks decays into a $\tau$ lepton which itself decays before reaching the detector. Like the top quark, the $\tau$ lepton can either decay leptonically or hadronically as shown in Figure 34. Therefore, if the $\tau$ lepton originating from the top quark decays hadronically, there is no measurable lepton in the final state as well as $E_T^{\text{miss}}$ through the neutrino production. The branching ratio for a top quark decaying into a $\tau$ lepton is around 11% and the

$^2$An off-shell particle is a virtual particle in a Feynman-Diagram that does not satisfy the energy momentum relation $E^2 - |\vec{p}|^2 = m^2$ where $m$ is the rest mass of the particle.
branching ratio for a hadronic $\tau$ lepton decay around 66% [154], which results in
a branching ratio for one top quark decaying into a $\tau$ lepton which then decays
hadronically of 7%. The branching ratio for both top quarks decaying this way is
0.5%. However, when comparing the amount of $t\bar{t}$ events before and after applying
the zero lepton requirement, it can be observed that the theoretical acceptance of
this process can not account alone for the amount of $t\bar{t}$ events present after applying
the preselection requirements. Without zero lepton requirement and adjusting the
jet number requirement to at least two jets, resulting from the two $b$-tagged jets,
around 105,500 $t\bar{t}$ events are expected. After applying the lepton requirement, there
are still around 17,500 $t\bar{t}$ events expected in the selection which amounts to 16%,
around double the amount of events that would be expected with this branching
ratio. The reason for this difference in the number of events can be found in detector
inefficiencies. The energies of the particles could be measured incorrectly, leading to
the existence of $E_T^{\text{miss}}$ in the final state, but it is more likely that certain particles
are outside of the detector acceptance which was described in Section 4.2.

This assumption can be confirmed by analysing the “TRUTH” information of the
SM background MC samples. As described in Chapter 6, the particles for the SM
background are simulated first and possible detector effects and inefficiencies are
added in a later step. Therefore, the “TRUTH” information allows for a categorisation
of the events depending on the simulated particle content. This is done in two ways
in Figure 35. Both figures show the $t\bar{t}$ background split into different categories. The
first figure shows the dominance of different $t\bar{t}$ decay channels with the “unknown”
category representing events where the TRUTH information could not be retrieved
properly in the derivation of the samples used. It can be observed that three channels
are the most dominant. The first one is a decay channel, where one top quark decays
into a $\tau$ lepton which then decays hadronically while the second top quark decays
hadronically. This is the SM $t\bar{t}$ process that is still expected after the preselection
requirements if there were no detector inefficiencies. The reason for the process with
both top quarks decaying to hadronically decaying $\tau$ leptons being less dominant is
due to the very low branching ratio of 0.5%. However, the other two dominant decay channels both include a lepton in the final state, either resulting from a $\tau$ lepton decaying leptonically or directly originating from the leptonic top quark decay. These events should not enter the selection, but they are due to detector inefficiencies. That the lepton is not measured accurately in the detector can result from different reasons. For example, a lepton could be misidentified as a jet or it is outside the detector acceptance. It could be expected that the direct leptonic decay would be more dominant due to the overall higher branching ratio. However, the process including a leptonically decaying $\tau$ lepton provides an additional neutrino in the final state which makes it more likely for this event to fulfill the high $E_T^{\text{miss}}$ preselection requirement. Therefore, this figure shows that a significant amount of the $t\bar{t}$ background can be expected to result from the lepton+jets channel which can be utilised in the analysis. The right hand figure in Figure 35 analyses the origin of the $b$-jet with the highest $p_T$ in the event. It shows if the $b$-jet either originates from the hadronically or leptonically decaying top quark. In the cases where a top quark decays into a $\tau$ lepton, the $b$-jet is assigned to the leptonically decaying top quark without further analysing the decay channel of the $\tau$ lepton. The “unknown” category consists of
events where the TRUTH information was either not retrieved properly like before or if it was not possible to assign the $b$-jet with the highest $p_T$ to a $b$-jet on TRUTH level. It can be observed that these two categories show a significant split in the $m_T(b_1, E_T^{miss})$ variable. For low values of $m_T(b_1, E_T^{miss})$, in the majority of events, the $b$-jet with the highest $p_T$ originates from the leptonically decaying top quark peaking at around 160 GeV. The reason for this is that the $E_T^{miss}$ and $b$-jet originate from the same top quark in this case and therefore, the $m_T(b_1, E_T^{miss})$ variable is bounded from above by the top quark mass. Therefore, the origin is found to be of the hadronically decaying top quark for high values of $m_T(b_1, E_T^{miss})$. This behaviour is utilised in the following selection.

Following this observation, the $m_T(b_1, E_T^{miss})$ variable can be used to reduce a significant amount of SM $t\bar{t}$ background by selecting only events with $m_T(b_1, E_T^{miss}) \geq 180$ GeV since the majority of $t\bar{t}$ events where the $b$-jet originates from the leptonically decaying top quark can be removed with this selection. However, an analysis of the other SM backgrounds and the signal has to be done to make sure that this selection does not remove the majority of the signal processes as well. Figure 36 shows the $m_T(b_1, E_T^{miss})$ variable for all SM backgrounds and four signals after applying the preselection requirements. The second panel in this figure represents the signal significance defined in Section 3.1 integrating event yields towards the right hand side. This panel can give a good representation of where to set the lower requirement for the selection. The peak from Figure 35 can also be observed for the sum of all SM backgrounds. Furthermore, this figure also shows that this peak mostly consists of $t\bar{t}$ background while the other backgrounds as well as the signal peak at higher values of $m_T(b_1, E_T^{miss})$. The second peak in this distribution is a result of the kinematic selections applied on the $E_T^{miss}$ and $p_T^{b_1}$ variables and has no connection to a specific physics process. Therefore, the selection $m_T(b_1, E_T^{miss}) \geq 180$ GeV reduces the $t\bar{t}$ background effectively while keeping most of the signal events.

While the $m_T(b_1, E_T^{miss})$ variable was able to effectively reduce $t\bar{t}$ events with mismeasured leptons, other SM backgrounds, like $W +$ jets and single top backgrounds,
can suffer from mismeasured particles as well especially in the phase space where events with high $E_T^{\text{miss}}$ values are selected. Since the only particles not detectable in the SM are neutrinos, there are only very few events in the SM with a $E_T^{\text{miss}}$ value as high as 250 GeV. However, due to mismeasurements in the detector, more events will be present in the selection. Comparatively, the signal includes DM particles in the final state which can not be detected and therefore, a high $E_T^{\text{miss}}$ value can occur without mismeasurements. In order to differentiate events with $E_T^{\text{miss}}$ resulting from undetectable particles from events where $E_T^{\text{miss}}$ results from mismeasurement, the variable $S_{E_T^{\text{miss}}}$ described in Section 8.2 is used. For higher values of $S_{E_T^{\text{miss}}}$, it is more likely for the $E_T^{\text{miss}}$ in the event to result from undetectable particles instead of mismeasurements. This is represented in Figure 37 which shows the $S_{E_T^{\text{miss}}}$ distribution after applying the preselection requirements and the $m_T(b_1, E_T^{\text{miss}}) \geq 180$ GeV selection. It can be observed that the signal processes show peaks at higher values for $S_{E_T^{\text{miss}}}$ compared to the SM background. To reduce the majority of the SM background while keeping most of the signal events, the selection $S_{E_T^{\text{miss}}} \geq 14$ is applied. While it can be observed in the significance panel
that a selection with a tighter requirement might be advantageous for the signal sensitivity of certain signal processes, this requirement was chosen so that the signal sensitivity for all signal processes is maximised first. The difference in distributions between the signal processes will be taken into account after the general selection definitions.

One of the variables already used in the preselection is the $\min[\Delta\phi(j_{1-4}, E_T^{\text{miss}})]$ variable. This variable was mainly used to reduce SM backgrounds with only jets in the final state. However, when analysing the distributions of SM backgrounds and signal processes in Figure 38 after applying preselection requirements as well as the previous two presented selections, it can be observed that the signal sensitivity can be increased by choosing a tighter selection requirement for this variable. It can be observed that especially the $t\bar{t}$ background has a higher contribution for low $\min[\Delta\phi(j_{1-4}, E_T^{\text{miss}})]$ values compared to other backgrounds and the signal processes. Therefore, applying the selection $\min[\Delta\phi(j_{1-4}, E_T^{\text{miss}})] \geq 0.9$ results in an increased significance for the signals.

One of the main differences between the signal and SM backgrounds are the DM
Figure 38: The min(Δφ(jet_1−4, E_T^{miss})) variable after applying the preselection requirements and m_T(b_1, E_T^{miss}) ≥ 180 GeV and S_{E_T^{miss}} ≥ 14 for all SM backgrounds as well as four signal processes where the number of events is multiplied by a factor of 10. The second panel shows the significance defined in Section 3.1 if the number of events would be added up for all bins with a higher value.

particles in the final state. However, since these are assumed to be stable and are not detected directly, they can only be measured indirectly through missing transverse energy as explained in Section 8.3. In contrast, other particles form the 2HDM+a described in Chapter 2 are not stable and therefore, their decay products can be measured. In the case of this analysis, the heavy Higgs boson H± has masses of at least 400 GeV up to 2000 GeV while the mediator a has masses from 100 GeV to 450 GeV. Therefore, in the cases with a H±-boson in the process, the p_T of the decay products which are the mediator a and the W-boson can be expected to be high. Similarly, W-bosons with a high p_T can also be expected in the case without a H± boson. In this case, the W-boson is produced back to back with the mediator a. If the mass of the mediator is high, the p_T of the W-boson can also be expected to be high. Furthermore, when a particle with a high p_T decays, the angle between its decay products is small. In the case of hadronically decaying W-bosons, this can lead to overlapping jets which can be reconstructed as one large-R jet. This allows tagging of hadronically decaying W-bosons as described in Section 7.2.5. In
The number of tagged $W$-bosons after applying the preselection requirements and $m_T(b_1, E^{\text{miss}}_T) \geq 180$ GeV, $S_{E^{\text{miss}}_T} \geq 14$ and $\min[\Delta \phi(\text{jet}_{1-4},E^{\text{miss}}_T)] \geq 0.9$ for all SM backgrounds as well as four signal processes. The second panel shows the significance defined in Section 3.1 if the number of events would be added up for all bins with a higher value.

Figure 39 also shows that the dominant backgrounds in this selections are $t\bar{t}$ and $Z + \text{jets}$. With the requirement of having a tagged $W$-boson in the event, more variables including this object can be analysed. The $m_T(b_1, E^{\text{miss}}_T)$ requirement defined previously mainly reduced the $t\bar{t}$ background where the $b$-jet with the highest $p_T$ in the event originates from the leptonically decaying top quark. Therefore, in most of the remaining $t\bar{t}$ events, the $b$-jet with the highest $p_T$ originates from
the hadronically decaying top quark. In the lepton+jets $t\bar{t}$ channel, this means, it originates from the same top quark which the tagged $W$-boson originates from as well. Comparatively, in the signal processes, the tagged $W$-boson usually originates from either the $H^\pm$ decay due to its high mass or the back to back production with the mediator, while the $b$-jet originates from the top quark decay. Therefore, the reconstructed mass from these two objects can be expected to be around the top quark mass for the $t\bar{t}$ background while the distribution of this variable can be expected to be more uniform for the signal processes.

However, before analysing this variable, some events need to be removed first. It is possible that a large-$R$ jet including the $b$-jet with the highest $p_T$ is tagged as a $W$-boson. In most cases, this is a result from the $b$-jet and other jets having a low angular distance which means they are reconstructed into a large-$R$ jet and then erroneously tagged as a $W$-boson. In order to remove these events from the selection, the distribution for the angular distance between the tagged $W$-boson and the $b$-jet with the highest $p_T$ is analysed in Figure 40. Since the radius for the large-$R$ jets is 1.0 in this analysis, the $b$-jet is part of the tagged $W$-boson if $\Delta R_{W\text{Tagged},b} < 1.0$. Therefore, to remove these events, the requirement $\Delta R_{W\text{Tagged},b} \geq 1.0$ is applied to the selection.

The $\Delta R_{W\text{Tagged},b}$ and $m_{W\text{Tagged},b}$ variables are constructed using the same final state objects. Therefore, they are expected to be correlated. This can be observed in Figure 41 showing the distribution of the $m_{W\text{Tagged},b}$ variable before and after the $\Delta R_{W\text{Tagged},b} \geq 1.0$ requirement is applied. The first peak in the distribution is removed by the selection since it is correlated to events with the $b$-jet being a part of the tagged $W$-boson. With these events excluded from the selection, a second peak in the distribution remains at around the top quark mass, and events in this peak mainly consist of $t\bar{t}$ events which is expected from the definition of this variable. Furthermore, only few events of the signal processes fall into this part of the distribution. Therefore, a selection requirement of $m_{W\text{Tagged},b} \geq 220$ GeV is applied.

A summary of all selection requirements for the tW_0L channel is presentend in
Figure 40: The angular distance between the tagged W-boson and the b-jet with the highest $p_T$ after applying the preselection requirements and $m_T(b_1, E_T^{miss}) \geq 180$ GeV, $S_{E_T^{miss}} \geq 14$, $\min[\Delta \phi(jet_{1-4}, E_T^{miss})] \geq 0.9$ and $N_{W-\text{tagged}} \geq 1$ for all SM backgrounds as well as four signal processes. The second panel shows the significance defined in Section 3.1 if the number of events would be added up for all bins with a higher value.

Figure 41: The reconstructed mass of the tagged W-boson and the b-jet with the highest $p_T$ after applying (a) the preselection requirements and $m_T(b_1, E_T^{miss}) \geq 180$ GeV, $S_{E_T^{miss}} \geq 14$, $\min[\Delta \phi(jet_{1-4}, E_T^{miss})] \geq 0.9$ and $N_{W-\text{tagged}} \geq 1$ (b) and $\Delta R_{W\text{Tagged}, b1} \geq 1.0$ for all SM backgrounds as well as four signal processes. The second panel shows the significance defined in Section 3.1 if the number of events would be added up for all bins with a higher value.
Table 5: The selection requirements for the $tW_{0L}$ channel. The preselection requirements of Table 4 for the $tW_{0L}$ channel are also applied.

Table 5. This selection is still defined so that it maximises the sensitivity of all signal processes. As it can be observed in Figure 37, the distributions of the signal processes can behave differently. Differences of the distributions mostly result from differences in $H^\pm$ masses and therefore can be primarily observed in observables like $S_{T,miss}$, but also $E_{T,miss}^\text{miss}$ itself. Therefore, the signal region defined in Table 5 is further split into a total of five regions with different values of $E_{T,miss}^\text{miss}$. These $E_{T,miss}^\text{miss}$ bins are defined as [250, 330, 400, 500, 600] GeV with the last bin being $E_{T,miss}^\text{miss} > 600$ GeV. All bins as well as the background composition together with four signal processes are shown in Figure 42. The yields of all SM backgrounds as well as one signal process are presented in Table 6 together with its percentage compared to the total expected background yields. It can be observed from this table that around 34% of the background consists of $Z + \text{jets}$ events and the second and third most dominant backgrounds are $W + \text{jets}$ and $t\bar{t}$ making up around 20% of the total background respectively.

8.5 The $tW_{1L}$ channel

The $tW_{1L}$ channel will be combined with the $tW_{0L}$ channel defined in Section 8.4. Therefore, this channel will be described shortly in this section. The main difference
Figure 42: The distribution for the $E_T^{\text{miss}}$ variable after applying the preselection and all selection requirements defined in Table 5 for all SM backgrounds as well as four signal processes. The binning in this distribution is done according to the binned signal region definitions of $[250,330,400,500,600]$ GeV where the last bin consists of events with $E_T^{\text{miss}} > 600$ GeV. The second panel shows the significance defined in Section 3.1 if the number of events would be added up for all bins with a higher value.

<table>
<thead>
<tr>
<th>Process</th>
<th>$E_T^{\text{miss}} \geq 250$ GeV</th>
<th>$tW_{0L}$ yields</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$E_T^{\text{miss}}$</td>
<td>Bin 1</td>
</tr>
<tr>
<td>$m_a = 250$ GeV, $m_H = 800$ GeV, $\tan \beta = 1$, $\sin \theta = 0.7$</td>
<td>77.9 ± 4.5</td>
<td>18.9 ± 2.2</td>
</tr>
<tr>
<td>$Z +$ jets</td>
<td>52.8 ± 2.8 (34%)</td>
<td>21.9 ± 2.4</td>
</tr>
<tr>
<td>$t\bar{t}$</td>
<td>28.5 ± 0.9 (18%)</td>
<td>14.3 ± 0.7</td>
</tr>
<tr>
<td>$W +$ jets</td>
<td>29.0 ± 2.1 (19%)</td>
<td>13.2 ± 1.6</td>
</tr>
<tr>
<td>Single $t$</td>
<td>19.6 ± 1.0 (13%)</td>
<td>7.8 ± 0.7</td>
</tr>
<tr>
<td>Diboson</td>
<td>13.1 ± 1.0 (8%)</td>
<td>5.6 ± 0.7</td>
</tr>
<tr>
<td>$tZ$</td>
<td>8.8 ± 0.2 (6%)</td>
<td>3.1 ± 0.1</td>
</tr>
<tr>
<td>Others</td>
<td>3.3 ± 0.1 (2%)</td>
<td>1.3 ± 0.1</td>
</tr>
<tr>
<td>Total bkg.</td>
<td>155.1 ± 3.9</td>
<td>67.2 ± 3.2</td>
</tr>
</tbody>
</table>

Table 6: Expected yields before fitting the results for all SM backgrounds as well as one specific signal process in the $tW_{0L}$ channel bins which are divided through $E_T^{\text{miss}}$ in $[250,330,400,500,600]$ GeV with the last bin including events with $E_T^{\text{miss}} \geq 600$ GeV. The uncertainties are only statistical uncertainties from the MC generated events. The percentages in the second column describe the ratio of the specific SM background in relation to the total background yields.
between the $tW_{0L}$ channel and the $tW_{1L}$ channel is the existence of at least one lepton in the final state. This means that the $tW_{1L}$ channel is statistically independent from the $tW_{0L}$ channel. Similar to the $tW_{0L}$ channel, a tagging of the $W$-boson is more effective if it originates from the $H^\pm$ due to its high mass or the back to back production with the mediator. However, a consequence of the one lepton requirement is that one of the $W$-bosons has to decay leptonically instead of both decaying hadronically. This can either be the $W$-boson originating from the top quark or the $W$-boson originating from the $H^\pm$ decay or the back to back production. This means that applying a $N^f_{W-\text{tagged}} = 1.0$ requirement for the whole $tW_{1L}$ channel will remove a significant amount of signal events in the case where the $W$-boson originating from the top quark decays hadronically. Therefore, this channel is split into two channels by using the $m_{\text{bjet},\text{ljet}}$ variable, which is the invariant mass between the $b$-jet with the highest $p_T$ and the jet with the highest $p_T$ that is not a $b$-jet. The $b$-jet with the highest $p_T$ in the event typically originates from the top quark. Comparatively, the jet with the highest $p_T$ and which is not tagged as a $b$-jet typically originates from the hadronically decaying $W$-boson. This can be utilised as shown in Figure 43. This figure shows that the distribution of the mass between the lepton and $b$-jet as well as the mass between the $b$-jet and non-$b$-jet is dependent on which $W$-boson decays hadronically. In the lower panel which describes the $m_{\text{bjet},\text{ljet}}$ variable, an edge around the top quark mass can be observed in the distribution only if the $W$-boson originating from the top quark decays hadronically. In comparison, if the $W$-boson originating from the $H^\pm$ decays hadronically, the distribution appears to be more uniform [67]. Therefore, the $tW_{1L}$ channel is split into a channel with a hadronically decaying top quark (SR$_{tW_{1L}}^{\text{had, top}}$) with $m_{\text{bjet},\text{ljet}} \leq 200 \text{ GeV}$ and a channel with a leptonically decaying top quark (SR$_{tW_{1L}}^{\text{lep, top}}$) with $m_{\text{bjet},\text{ljet}} \geq 200 \text{ GeV}$. In the latter case, the tagging of a $W$-boson is possible due to the high $p_T$ of the hadronically decaying $W$-boson either resulting from the high mass of the $H^\pm$ boson or the recoil against a high mass mediator $a$. Comparatively, the mass of the hadronically decaying $W$-boson has to be reconstructed by utilising the $m_W^{\text{had}}$ variable described
Figure 43: The distribution of the mass between the lepton and the $b$-jet with the highest $p_T$ (upper panel) and the distribution of the mass between the $b$-jet with the highest $p_T$ and the jet with the highest $p_T$ that is not a $b$-jet (lower panel) for the cases that the $W$-boson originating from the top quark decays hadronically (orange) and leptonically (blue) [67].

in Section 8.2 for the $S_{tW1L}^{\text{had}}$ channel. Since this variable uses jets with $R = 0.4$ to reconstruct jets with a higher radius parameter than $R = 1.0$, it is not as effective in tagging events with a $W$-boson. Therefore, a value of $m_W^{\text{had}} > 60$ GeV is used to keep most of the signal events.

Furthermore, the combination of the $m_{T\text{lep,MET}}$ and $am_{T2}$ variable is used to reduce the $t\bar{t}$, single top and $W+$ jets background. The $m_{T\text{lep,MET}}$ variable reduces events with one lepton in the final state originating from a $W$-boson. This is done because the lepton and $E_T^{\text{miss}}$ in these cases result from the decay of the same $W$-boson and therefore, the variable is bounded by the $W$-mass from above. Similarly, the $am_{T2}$ variable is bounded from above by the top quark mass in $t\bar{t}$ events with two leptonically decaying top quarks but one of the leptons is missed in the detector. The missed lepton contributes to the $E_T^{\text{miss}}$ value and therefore, by combining the lepton measured by the detector with each of the jets that have the highest $b$-tagging weight, the calculated $m_{T2}$ variable is bounded from above by the mass of the
Figure 44: The (a) $m_{T,\text{lep,MET}}$ and (b) $am_{T2}$ variables in the $tW_{1L}$ channel. The distribution for the $m_{T,\text{lep,MET}}$ variable is shown after all preselection requirements for the $tW_{1L}$ channel from Table 4 are applied while the $am_{T2}$ variable is shown after all preselection requirements and $m_{T,\text{lep,MET}} \geq 120$ GeV is applied reducing the $t\bar{t}$, single top and $W+\text{jets}$ backgrounds with one lepton in the final state effectively. The second panel shows the significance defined in Section 3.1 if the number of events would be added up for all bins with a higher value.

<table>
<thead>
<tr>
<th>Variable</th>
<th>SR$<em>{\text{lep,top}}^{tW</em>{1L}}$</th>
<th>SR$<em>{\text{had,top}}^{tW</em>{1L}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m_{\text{bjet1,ijet1}}$ [GeV]</td>
<td>$\geq 200$</td>
<td>$\leq 200$</td>
</tr>
<tr>
<td>Number of signal jets</td>
<td>$\geq 2$</td>
<td>$\geq 3$</td>
</tr>
<tr>
<td>$S_{E_{\text{miss}}}$</td>
<td>$\geq 15$</td>
<td>-</td>
</tr>
<tr>
<td>$m_{T,\text{lep,MET}}$ [GeV]</td>
<td>$\geq 130$</td>
<td>$\geq 200$</td>
</tr>
<tr>
<td>$am_{T2}$ [GeV]</td>
<td>$\geq 180$</td>
<td>$\geq 180$</td>
</tr>
<tr>
<td>$N_{W-\text{tagged}}$</td>
<td>$\geq 1$</td>
<td>-</td>
</tr>
<tr>
<td>$m_{\text{had}}^{W}$ [GeV]</td>
<td>-</td>
<td>$\geq 60$</td>
</tr>
</tbody>
</table>

Table 7: The selection requirements for the $tW_{1L}$ channel split into the SR$_{\text{lep,top}}^{tW_{1L}}$ and SR$_{\text{had,top}}^{tW_{1L}}$ channel. The preselection requirements of Table 4 for the $tW_{1L}$ channel are also applied.

The complete selection for both the SR$_{\text{lep,top}}^{tW_{1L}}$ and SR$_{\text{had,top}}^{tW_{1L}}$ channel are summarised...
in Table 7. Since there is a $N_{W}^{J:R=1,0}$ requirement in the \( \text{SR}_W^{\text{had, top}} \) region compared to the \( \text{SR}_W^{\text{lep, top}} \) region, a higher number of events can be expected in the \( \text{SR}_W^{\text{had, top}} \) region. Therefore, a similar strategy is used as described in Section 8.4 where the distribution of the $E_T^{\text{miss}}$ variable is split into five individual regions to increase the sensitivity for all signal processes. The bins of this region are defined as $[250, 300, 350, 400, 450]$ GeV with the final bin being defined as $E_T^{\text{miss}} > 450$ GeV. Due to the low number of events in the \( \text{SR}_W^{\text{lep, top}} \) channel, this region is not further split. An overview over the binned \( \text{SR}_W^{\text{had, top}} \) region as well as an overview over the \( \text{SR}_W^{\text{lep, top}} \) region can be found in Figure 45. The yields for both channels divided into the different SM backgrounds are presented in Table 8. It can be observed that $t\bar{t}$ and $t\bar{t}Z$ are the most dominant backgrounds in the \( \text{SR}_{tW1L}^{\text{lep, top}} \) channel making up around 36% and 21% of the total background respectively. Similarly, it can also be observed that nearly 50% of the SM background in the \( \text{SR}_{tW1L}^{\text{had, top}} \) channel consists of $t\bar{t}$ events. The second most dominant background are events from $W+$ jets processes with around 21% of the total background events.
Table 8: Yields for all SM backgrounds as well as one specific signal process in the SR$_{lep,top}$$^{tW_{1L}}$ channel and the SR$_{had,top}$$^{tW_{1L}}$ bins which are divided in $E_T^{miss}$ through $[250,300,350,400,450]$ GeV with the last bin including events with $E_T^{miss}$ $\geq$ 450 GeV. The uncertainties are only statistical uncertainties from the MC generated events. The percentages in the second and last column describe the ratio of the specific SM background in relation to the total background yields.
9 Background estimation

While the previous chapter described the definition of the signal regions for the analysis, this chapter will focus on the background estimation strategy and the necessary systematic uncertainties which are used as inputs in the “HistFitter” framework described in Section 3.2. For the background estimation, control regions are defined for the most dominant backgrounds in order to retrieve normalisation factors from the background-only fit. Furthermore, in order to validate these normalisations for the signal regions, validation regions which are kinematically closer to the signal regions are defined as well where the difference between data and MC predictions is compared. All of these regions are designed so that they are statistically independent from the signal regions defined in Chapter 8 as well as statistically independent to all the control regions defined in Section 9.1. An overview over the definitions, that are explained in more detail in Section 9.1 for control regions and Section 9.2 for validation regions, can be found in Figure 46. Finally, the systematic uncertainties will be shortly described in Section 9.3.

9.1 Control Regions

In order to estimate the dominant backgrounds in the signal region, control regions that are kinematically similar to the signal regions, but which maximise the purity in a specific background are defined. In order to estimate every one of these backgrounds with a good precision while also keeping the amount of normalisation factors as simple
as possible, a single control region (CR) is defined and a single normalisation factor for the fit is retrieved for every dominant background. For the tW_{0L} channel, the three major backgrounds in the signal regions are processes with a Z-boson decaying into a pair of neutrinos and which is produced in association with jets (Z+jets), a production of a top and anti-top quark (tt̅) as well as the production of a W-boson in association with jets (W+jets). For the first two background processes, a CR is defined similar to the tW_{0L} signal regions. However, in the case of the W+jets process, a definition for a region with a high purity for the W+jets process with zero leptons in the final state is not achievable. Therefore, the same normalisation factor is used that is later defined in the W+jets CR of the tW_{1L} channel.

For the tt̅ process, it was noticed in Section 8.4 that a large contribution are tt̅ events with one lepton in the final state, but this lepton is then mismeasured in the detector. Therefore, the CR for the tt̅ process aims to maximise this contribution by requiring one lepton in the final state instead of zero leptons as in the signal region. This is the main difference for the CR. Furthermore, in order to increase the purity of the process, the selection on the object based $E_{T}^{miss}$ significance is not used for this CR where it was set to be at least 14 in the signal regions. Finally, to increase the purity further and in order to stay statistically independent to the regions of the tW_{1L} channel, two additional selections $m_{T}^{lep,MET} < 130 \text{ GeV}$ and $a_{MT2} < 180 \text{ GeV}$.
are applied which replace the original $m_T(b_1, E_T^{\text{miss}})$ requirement in the $tW_{0L}$ signal region. This results in a CR with around 195 events and a $t\bar{t}$ purity of 83%.

For the $Z + \text{jets}$ CR, a different strategy is applied. Since it is challenging to design a region which maximises events for the process $Z \to \nu\nu$, a CR with two leptons in the final state is defined. The decay of the $Z$-boson to two leptons is kinematically very similar to the decay to two neutrinos. Therefore, a region maximising $Z \to ll$ events where the leptons are treated as part of the missing transverse momentum can be used to emulate a zero lepton region. This means that new variables have to be defined. For the missing transverse energy $E_T^{\text{miss}}$ is used which is adding the two leptons as invisible particles to the calculation of the missing transverse momentum. Furthermore, other variables that include the missing transverse momentum as a variable will use $E_T^{\text{miss},ll}$ instead of $E_T^{\text{miss}}$. Examples for these variables are $S_{E_T^{\text{miss},ll}}$, $\min[\Delta\phi(\text{jet}_1, E_T^{\text{miss},ll})]$ as well as $m_T(b_1, E_T^{\text{miss},ll})$ which are the equivalent variables to $S_{E_T^{\text{miss}}}$, $\min[\Delta\phi(\text{jet}_1, E_T^{\text{miss}})]$ and $m_T(b_1, E_T^{\text{miss}})$. Furthermore, to increase the purity of this CR, the two leptons have to originate from a $Z$-boson decay. This is done by requiring that these have opposite charges, but are of the same flavour (SF-OS) and that their invariant mass $m_{ll}$ lies in the interval of 81 GeV and 101 GeV. Finally, the variables surrounding the tagging of a $W$-boson are not applied to ensure that the number of events in the CR is high enough for a stable fit. This results in a CR with around 733 events and a $Z + \text{jets}$ purity of 85%. A summary of both CRs can be found in Table 9 while the $E_T^{\text{miss}}$ and $E_T^{\text{miss},ll}$ distributions are presented in Figure 47.

Additionally, CRs for the $tW_{1L}$ channel are defined as well and since the normalisation parameters for all these regions will be fitted simultaneously, they are shortly presented here. Like the $tW_{0L}$ channel, two SM processes contribute mainly to the total background. The $t\bar{t}$ process is a dominant background in the $tW_{1L}$ channel as well. However, in comparison to the $tW_{0L}$ channel, the $t\bar{t}$ events mainly consist of those where both top quarks decay leptonically and one lepton is mismeasured in the detector. Therefore, the strategy for the $t\bar{t}$ CR is different from the one in
<table>
<thead>
<tr>
<th>Variable</th>
<th>CR$_{tW0L}$ ($t\bar{t}$)</th>
<th>CR$_{tW0L}$ ($Z + \text{jets}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_{T}^{\text{miss}}$ [GeV]</td>
<td>$\geq 250$</td>
<td>$\leq 120$</td>
</tr>
<tr>
<td>$E_{T,\text{ll}}^{\text{miss}}$ [GeV]</td>
<td>–</td>
<td>$\geq 250$</td>
</tr>
<tr>
<td>$S_{E_{T}^{\text{miss}}}$</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>$S_{E_{T,\text{ll}}^{\text{miss}}}$</td>
<td>–</td>
<td>$\geq 14$</td>
</tr>
<tr>
<td>$\min[\Delta \phi(\text{jet}<em>{1-4}, E</em>{T}^{\text{miss}})]$</td>
<td>$\geq 0.5$</td>
<td>–</td>
</tr>
<tr>
<td>$\min[\Delta \phi(\text{jet}<em>{1-4}, E</em>{T,\text{ll}}^{\text{miss}})]$</td>
<td>–</td>
<td>$\geq 0.5$</td>
</tr>
<tr>
<td>Number of baseline leptons</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Number of signal leptons</td>
<td>1</td>
<td>2 (SF-OS)</td>
</tr>
<tr>
<td>$p_{T}^{l_1}$ [GeV]</td>
<td>$\geq 30$</td>
<td>$\geq 30$</td>
</tr>
<tr>
<td>$p_{T}^{l_2}$ [GeV]</td>
<td>–</td>
<td>$\geq 20$</td>
</tr>
<tr>
<td>Number of signal jets</td>
<td>$\geq 4$</td>
<td>$\geq 4$</td>
</tr>
<tr>
<td>$p_{T}^{j_1}$ [GeV]</td>
<td>$\geq 100$</td>
<td>$\geq 100$</td>
</tr>
<tr>
<td>$p_{T}^{j_2}$ [GeV]</td>
<td>$\geq 60$</td>
<td>$\geq 60$</td>
</tr>
<tr>
<td>$p_{T}^{j_3}$ [GeV]</td>
<td>$\geq 60$</td>
<td>$\geq 60$</td>
</tr>
<tr>
<td>$p_{T}^{j_4}$ [GeV]</td>
<td>$\geq 40$</td>
<td>$\geq 40$</td>
</tr>
<tr>
<td>Number of $b$-tagged jets</td>
<td>$\geq 1$</td>
<td>$\geq 1$</td>
</tr>
<tr>
<td>$p_{T}^{b_1}$ [GeV]</td>
<td>$\geq 50$</td>
<td>$\geq 50$</td>
</tr>
<tr>
<td>$p_{T}^{b_2}$ [GeV]</td>
<td>$\leq 50$</td>
<td>$\leq 50$</td>
</tr>
<tr>
<td>$N_{W-\text{tagged}}$</td>
<td>$\geq 1$</td>
<td>$\geq 0$</td>
</tr>
<tr>
<td>$\Delta R_{W\text{Tagged},b_1}$</td>
<td>$\geq 1.0$</td>
<td>–</td>
</tr>
<tr>
<td>$m_{W\text{Tagged},b_1}$ [GeV]</td>
<td>$\geq 220$</td>
<td>–</td>
</tr>
<tr>
<td>$m_{ll}$ [GeV]</td>
<td>–</td>
<td>$\in [81, 101]$</td>
</tr>
<tr>
<td>$m_{T}(b_1, E_{T}^{\text{miss}})$ [GeV]</td>
<td>–</td>
<td>$\geq 180$</td>
</tr>
<tr>
<td>$m_{T,\text{MET}}$ [GeV]</td>
<td>$&lt; 130$</td>
<td>–</td>
</tr>
<tr>
<td>$am_{T_2}$ [GeV]</td>
<td>$&lt; 180$</td>
<td>–</td>
</tr>
</tbody>
</table>

**Table 9:** The definitions of the control regions for the $tW0L$ channel. The CR$_{tW0L}$ ($t\bar{t}$) reaches a 83% purity in $t\bar{t}$ events while the CR$_{tW0L}$ ($Z + \text{jets}$) reaches a 85% purity in $Z + \text{jets}$ events.

the $tW0L$ channel and an individual normalisation parameter is applied for this background. In order to select a high number of $t\bar{t}$ events where both top quarks decay leptonically, but one is mismeasured by the detector, the selection on the $am_{T2}$ variable is inverted. Furthermore, in order to stay statistically independent while
increasing the $t\bar{t}$ purity, the selection on the $p_T$ of the $b$-jet with the second highest $p_T$ is inverted as well leading to a CR with a 95% $t\bar{t}$ purity with around 1000 events. The second CR is designed for the $W$ + jets process where the region is defined with one lepton in the final state. The statistical independence to the $tW_{1L}$ signal regions as well as a high purity is achieved by using low $m_T^{lep,MET}$ values and reversing the requirement on the $m_{W}^{had}$ variable. Similarly to the $Z +$ jets CR of the $tW_{0L}$ channel, no tagged $W$-boson is required in order to achieve a stable fit. This results in a CR with around 1000 events and a 61% $W$ + jets purity. The summary of these regions can be found in Table 10 while the $E_T^{miss}$ distributions are shown in Figure 48.

Finally, two CRs are added for the single top process as well as the $t\bar{t}Z$ process. A CR for the $t\bar{t}Z$ process is chosen because it is a dominant background in the $SR_{lep,top}$ part of the $tW_{1L}$ channel. This CR is using the definition from Ref. [156] with the only exception being that in this analysis a single lepton trigger is used instead of a di-lepton trigger. This CR has three leptons in the final state. Two of these leptons are assumed to originate from the $Z$-boson and either the only lepton pair with same flavour, but opposite charge is used or the one closest to the $Z$ mass is chosen to result from the $Z$-boson decay. Like in the CR for the $Z +$ jets process, the two leptons originating from the $Z$-boson are treated as part of the missing transverse momentum. With this strategy, most SM processes except the $t\bar{t}Z$ process

Figure 47: The $E_T^{miss}$ and $E_T^{miss}$ distributions for the (a) $t\bar{t}$ and (b) $Z +$ jets CRs of the $tW_{0L}$ analysis channel.
### Table 10: The definitions of the control regions for the $tW_{1L}$ channel. The CR$_{tW_{1L}}$ ($t\bar{t}$) reaches a 95% purity in $t\bar{t}$ events while the CR$_{tW_{1L}}$ ($W$+jets) reaches a 61% purity in $W$+jets events.

<table>
<thead>
<tr>
<th>Variable</th>
<th>CR$<em>{tW</em>{1L}}$ ($t\bar{t}$)</th>
<th>CR$<em>{tW</em>{1L}}$ ($W$+jets)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trigger</td>
<td>$E_T^{miss}$</td>
<td>$E_T^{miss}$</td>
</tr>
<tr>
<td>$E_T^{miss}$ [GeV]</td>
<td>$\geq 250$</td>
<td>$\geq 250$</td>
</tr>
<tr>
<td>$S_{E_T^{miss}}$</td>
<td>-</td>
<td>$\geq 15$</td>
</tr>
<tr>
<td>$\min[\Delta\phi(\text{jet}_{1-4}, E_T^{miss})]$</td>
<td>$\geq 0.5$</td>
<td>$\geq 0.5$</td>
</tr>
<tr>
<td>Number of baseline leptons</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Number of signal leptons</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$p_{T1}^l$ [GeV]</td>
<td>$\geq 30$</td>
<td>$\geq 30$</td>
</tr>
<tr>
<td>Number of signal jets</td>
<td>$\geq 3$</td>
<td>$\geq 3$</td>
</tr>
<tr>
<td>$p_{T1}^l$ [GeV]</td>
<td>$\geq 30$</td>
<td>$\geq 30$</td>
</tr>
<tr>
<td>$p_{T2}^l$ [GeV]</td>
<td>$\geq 30$</td>
<td>$\geq 30$</td>
</tr>
<tr>
<td>$p_{T3}^l$ [GeV]</td>
<td>$\geq 30$</td>
<td>$\geq 30$</td>
</tr>
<tr>
<td>Number of $b$-tagged jets</td>
<td>$\geq 2$</td>
<td>$\geq 1$</td>
</tr>
<tr>
<td>$p_{T1}^{b1}$ [GeV]</td>
<td>$\geq 50$</td>
<td>$\geq 50$</td>
</tr>
<tr>
<td>$p_{T2}^{b1}$ [GeV]</td>
<td>$\geq 50$</td>
<td>$\leq 50$</td>
</tr>
<tr>
<td>$N_J^{J/R=1.0}$</td>
<td>-</td>
<td>$= 0$</td>
</tr>
<tr>
<td>$m_T$ [GeV]</td>
<td>$\geq 130$</td>
<td>$\in [40, 100]$</td>
</tr>
<tr>
<td>$am_{T2}$ [GeV]</td>
<td>$&lt; 180$</td>
<td>$\geq 180$</td>
</tr>
<tr>
<td>$m_{had}^W$ [GeV]</td>
<td>-</td>
<td>$&lt; 60$</td>
</tr>
</tbody>
</table>

**Figure 48:** The $E_T^{miss}$ distributions for the (a) $t\bar{t}$ and (b) $W$+jets CRs of the $tW_{1L}$ analysis channel.
can be removed from the selection resulting in a 69% purity with around 209 total events. While the single top background is not a dominant background in either of the signal regions that are analysed, the uncertainties in its modelling can become large enough to influence the results of the overall fit. Therefore, in order to reduce the impact of these uncertainties, a dedicated CR is defined as well. In order to achieve a high enough purity, this CR is designed in a two lepton channel like the CR for the $Z+jets$ process. However, in this case, the leptons are not treated as a part of the missing transverse momentum. For a high purity, the main variables that are used are the invariant mass between the two leptons $m_{ll}$ which is set to be outside the mass range for the $Z$-boson as well as $m_{T2}$ to reduce $W+jets$ events. Further variables like $m_{bl}^{\text{min}}$ and $m_{bl}^{t}$ are used as well. This results in a CR with around 637 total expected events and a 62% single top purity. The summary of both CR definitions can be found in Table 11 while the $E_{T}^{\text{miss}}$ and $E_{T,\text{ll}}^{\text{miss}}$ distributions are presented in Figure 49.

It can be observed in the single top CR that the number of data events is significantly lower than the number of expected events. This difference will affect the normalisation parameter, but is expected in a phase space with high missing transverse energy. The reason for this disagreement between data and MC events will be explained in Section 10.1 when discussing the normalisation parameter.
<table>
<thead>
<tr>
<th>Variable</th>
<th>CR (Single $t$)</th>
<th>CR ($t\bar{t}Z$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_T^{\text{miss}}$ [GeV]</td>
<td>$\geq 200$</td>
<td>–</td>
</tr>
<tr>
<td>$E_{T,\text{ll}}^{\text{miss}}$ [GeV]</td>
<td>–</td>
<td>$\geq 140$</td>
</tr>
<tr>
<td>$\text{min}[\Delta \phi(\text{jet}_{1-4}, E_T^{\text{miss}})]$</td>
<td>$\geq 0.5$</td>
<td>–</td>
</tr>
<tr>
<td>Number of baseline leptons</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Number of signal leptons</td>
<td>2 (OS)</td>
<td>3 (at least one SF-OS pair)</td>
</tr>
<tr>
<td>$p_T^{l_1}$ [GeV]</td>
<td>$\geq 25$</td>
<td>$\geq 30$</td>
</tr>
<tr>
<td>$p_T^{l_2}$ [GeV]</td>
<td>$\geq 20$</td>
<td>$\geq 20$</td>
</tr>
<tr>
<td>$p_T^{l_3}$ [GeV]</td>
<td>–</td>
<td>$\geq 20$</td>
</tr>
<tr>
<td>Number of signal jets</td>
<td>$\geq 1$</td>
<td>$\geq 3$</td>
</tr>
<tr>
<td>$p_T^{j_1}$ [GeV]</td>
<td>$\geq 50$</td>
<td>$\geq 30$</td>
</tr>
<tr>
<td>$p_T^{j_2}$ [GeV]</td>
<td>–</td>
<td>$\geq 30$</td>
</tr>
<tr>
<td>$p_T^{j_3}$ [GeV]</td>
<td>–</td>
<td>$\geq 30$</td>
</tr>
<tr>
<td>Number of $b$-tagged jets</td>
<td>$\geq 1$</td>
<td>$\geq 2$</td>
</tr>
<tr>
<td>$p_T^{b_1}$ [GeV]</td>
<td>$\geq 50$</td>
<td>$\geq 30$</td>
</tr>
<tr>
<td>$p_T^{b_2}$ [GeV]</td>
<td>–</td>
<td>$\geq 30$</td>
</tr>
<tr>
<td>$m_{ll}^{\text{miss}}$ [GeV]</td>
<td>$\geq 40$, $\notin [71, 111]$ if SF</td>
<td>$\in [71, 111]$ for at least one SF-OS pair</td>
</tr>
<tr>
<td>$m_{T2}$ [GeV]</td>
<td>$\leq 100$</td>
<td>–</td>
</tr>
<tr>
<td>$m_{T1}^{\text{min}}$ [GeV]</td>
<td>$&gt; 170$</td>
<td>–</td>
</tr>
<tr>
<td>$m_{T1}^{\text{max}}$ [GeV]</td>
<td>$&gt; 150$</td>
<td>–</td>
</tr>
</tbody>
</table>

**Table 11:** The definitions of the control regions for the $t\bar{t}Z$ and single top background. The CR (Single $t$) reaches a 62% purity in single top events while the CR ($t\bar{t}Z$) reaches a 69% purity in $t\bar{t}Z$ events.

### 9.2 Validation Regions

Before the normalisation parameters, that will be retrieved from the CRs, can be applied to different fits in the SRs, it first needs to be analysed if these normalisation factors are consistent with the predicted background in a signal region like scenario. Therefore, a validation region (VR) for every dominant SM process is defined as well which is kinematically closer to the signal regions than the CRs but where the signal contamination is still low. Therefore, the fit results from the background-only fit are applied to the VR definitions and the difference between data and MC prediction.
is analysed. If the difference is not significant, the normalisation parameters can be applied to the signal regions with a higher signal contamination. Otherwise, the background estimation strategy has to be changed. Like before, one VR for every one of the most dominant background processes is defined with the exception being the $t\bar{t}$ background and the $t\bar{t}Z$ background in the $tW_{1L}$ channel. In the first case, two $t\bar{t}$ VRs are defined, one for the $SR_{tW_{1L}}^{\text{had.top}}$ and one for the $SR_{tW_{1L}}^{\text{lep.top}}$ region. For the $t\bar{t}Z$ background, no VR is defined since it was shown in Ref. [156] that the background and data show a good agreement.

For the $tW_{0L}$ channel, one VR is defined for the $t\bar{t}$ process. In comparison to the CR, this region is defined in a zero lepton channel and is designed very close to the signal region, but a few changes were applied to the definitions in Table 5 in order to increase the $t\bar{t}$ purity and reduce the signal contamination. First, as it was observed in Figure 37, the $t\bar{t}$ process peaks at lower values of $S_{E^{\text{miss}}_T}$. Therefore, this variable selection is inverted. However, since this variable is shown to be mismodelled at low values of $S_{E^{\text{miss}}_T}$, it was decided to set the selection of this variable in between the values of 10 and 14. The second difference uses the observations from Figure 41 in order to increase the purity of the $t\bar{t}$ background. This figure showed that a well defined peak of $t\bar{t}$ events can be observed around the top mass in the $m_{W\text{Tagged},b_1}$ variable. Therefore, the VR removes the requirement for $\Delta R_{W\text{Tagged},b_1}$ and inverts the selection on $m_{W\text{Tagged},b_1}$ to remain statistically independent to the $tW_{0L}$ signal region. This definition creates a region with around 263 events, a $t\bar{t}$ purity of 78% and a maximum signal contamination of 6%.

For the other two dominant backgrounds in the signal regions of the $tW_{0L}$ channel, which are the $Z + \text{jets}$ and $W + \text{jets}$ processes, a combined VR is designed which aims at increasing the purity for both backgrounds at the same time. The reason for this is a combination of the challenging design for a zero lepton $W + \text{jets}$ VR and the fact that regions with a high amount of $W + \text{jets}$ and $Z + \text{jets}$ events in a zero lepton channel are kinematically very similar. The VR for these two backgrounds features two changes when compared to the signal region defined before. First, the
requirement for a tagged $W$-boson is removed. However, if one is present in the event, the $\Delta R_{W\text{Tagged,}b_{1}}$ requirement is inverted to reduce the signal contamination.

For the second difference to the signal region, two new variables are introduced. The first one is $\min[\Delta \phi(j_{\text{all}}, E_{T}^{\text{miss}})]$ which is replacing the additional requirement on $\min[\Delta \phi(j_{1-4}, E_{T}^{\text{miss}})]$ variable and which includes all signal jets for calculating the minimum angular distance between the jets and $E_{T}^{\text{miss}}$. The second variable is $\Delta R_{j_{1},j_{2}}$ which is an especially effective variable to increase the $Z + \text{jets}$ purity against the $\bar{t}t$ background, because the jets in this process are a product of the recoil against the $Z$-boson where the jets in the $\bar{t}t$ background are results from the top decay. This definition creates a region with 3748 events, a purity of combined $Z + \text{jets}$ and $W + \text{jets}$ processes of 67% and a maximum signal contamination of 9%. However, the signals that lead to such a high contamination are individual ones while most of the signal contamination lies around 5% which is low enough for this VR definition to be used. A summary of these two VRs is presented in Table 12. The $E_{T}^{\text{miss}}$ distributions for these validation regions are shown in Figure 50.

Four additional VRs are defined. Three of those are designed for the $tW_{1L}$ channel. For the $\bar{t}t$ background, two VRs are used, one for the $\text{SR}_{tW_{1L}}^{\text{lep,} \bar{t}t}$ and one for the $\text{SR}_{tW_{1L}}^{\text{had,} \bar{t}t}$ region since $\bar{t}t$ is the main background in both channels. The main difference is the inversion of the $amT_{2}$ variable in both cases in comparison to the

Figure 50: The $E_{T}^{\text{miss}}$ distributions for the (a) $\bar{t}t$ and (b) $V + \text{jets}$ VRs of the $tW_{0L}$ analysis channel.
Table 12: The definitions of the validation regions for the tW_{0L} channel. The VR_{tW_{0L}} (tt) reaches a 78% purity in tt events while the VR_{tW_{0L}} (V + jets) reaches a 67% purity in V + jets events.
Figure 51: The $E_{\text{miss}}$ distributions for the (a) had.top $t\bar{t}$, (b) lep.top $t\bar{t}$, (c) $W + \text{jets}$ VRs of the $tW_{1L}$ analysis channel as well as (d) the single top VR.

this region. The difference to the SR_{had,top} region is achieved by requiring a $m_{T,\text{lep,MET}}$ value between 40 and 100 GeV increasing the purity of $W + \text{jets}$ in this one lepton channel leading to a region with around 1000 events, a 26% $W + \text{jets}$ purity and a maximum signal contamination of 0.6%. The final validation region is designed for the single top background. It was decided to design this region in a one lepton channel only and not design two VRs for the $tW_{0L}$ and $tW_{1L}$ individually since the CR is defined with two leptons in the final state and the main goal of this CR is to reduce the effect of the systematic uncertainties of the single top modelling. In order to decrease most of the backgrounds and reduce the signal contamination, two changes to the signal regions are applied. First, this VR requires a second $b$-jet in addition as well as a lower $m_{T,\text{lep,MET}}$ value compared to the signal regions. However,
Table 13: The definitions of the validation regions for the tW1L channel as well as the VR for the single top process. The \( \text{VR}^{\text{lep.top}}_{\text{tW1L}} \) (tt) and \( \text{VR}^{\text{had.top}}_{\text{tW1L}} \) (tt) reach a 88% and 89% purity in \( \bar{t}t \) events respectively, the \( \text{VR}_{W+\text{jets}} \) reaches a 26% purity in \( W+\text{jets} \) events and the VR (Single t) reaches a 41% purity in single top events.

<table>
<thead>
<tr>
<th>Variable</th>
<th>( \text{VR}^{\text{lep.top}}_{\text{tW1L}} ) (tt)</th>
<th>( \text{VR}^{\text{had.top}}_{\text{tW1L}} ) (tt)</th>
<th>( \text{VR}_{W+\text{jets}} ) (W+jets)</th>
<th>VR (Single t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( E_T^{\text{miss}} ) [GeV]</td>
<td>( \geq 250 )</td>
<td>( \geq 250 )</td>
<td>( \geq 250 )</td>
<td>( \geq 250 )</td>
</tr>
<tr>
<td>( S_E^{\text{miss}} )</td>
<td>( \geq 15 )</td>
<td>–</td>
<td>( \geq 15 )</td>
<td>( \geq 15 )</td>
</tr>
<tr>
<td>( \min[\Delta \phi(\text{jet}_{1-4},E_T^{\text{miss}})] )</td>
<td>( \geq 0.5 )</td>
<td>( \geq 0.5 )</td>
<td>( \geq 0.5 )</td>
<td>( \geq 0.5 )</td>
</tr>
<tr>
<td>Number of baseline leptons</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Number of signal leptons</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>( p_T^1 ) [GeV]</td>
<td>( \geq 30 )</td>
<td>( \geq 30 )</td>
<td>( \geq 30 )</td>
<td>( \geq 30 )</td>
</tr>
<tr>
<td>Number of signal jets</td>
<td>( \geq 2 )</td>
<td>( \geq 3 )</td>
<td>( \geq 3 )</td>
<td>( \geq 3 )</td>
</tr>
<tr>
<td>( p_T^1 ) [GeV]</td>
<td>( \geq 50 )</td>
<td>( \geq 50 )</td>
<td>( \geq 50 )</td>
<td>( \geq 50 )</td>
</tr>
<tr>
<td>( p_T^2 ) [GeV]</td>
<td>( \geq 30 )</td>
<td>( \geq 30 )</td>
<td>( \geq 30 )</td>
<td>( \geq 30 )</td>
</tr>
<tr>
<td>Number of b-tagged jets</td>
<td>( \geq 1 )</td>
<td>( \geq 1 )</td>
<td>( \geq 1 )</td>
<td>( \geq 2 )</td>
</tr>
<tr>
<td>( p_T^b ) [GeV]</td>
<td>( \geq 50 )</td>
<td>( \geq 50 )</td>
<td>( \geq 50 )</td>
<td>( \geq 50 )</td>
</tr>
<tr>
<td>( p_T^{b2} ) [GeV]</td>
<td>( \leq 50 )</td>
<td>( \leq 50 )</td>
<td>( \leq 50 )</td>
<td>( \leq 50 )</td>
</tr>
<tr>
<td>( m^{\text{lep.min.}}_T ) [GeV]</td>
<td>( \geq 130 )</td>
<td>( \geq 200 )</td>
<td>( \in [40,100] )</td>
<td>( &lt; 100 )</td>
</tr>
<tr>
<td>( m_{\text{jet}_1,\text{jet}_1} ) [GeV]</td>
<td>( \geq 200 )</td>
<td>( &lt; 200 )</td>
<td>( &lt; 200 )</td>
<td>–</td>
</tr>
<tr>
<td>( am_{T2} ) [GeV]</td>
<td>( &lt; 180 )</td>
<td>( &lt; 180 )</td>
<td>( \geq 180 )</td>
<td>( \geq 180 )</td>
</tr>
<tr>
<td>( m_{\text{had}}^{W} ) [GeV]</td>
<td>–</td>
<td>–</td>
<td>( \geq 60 )</td>
<td>–</td>
</tr>
</tbody>
</table>

a region definition like this also increases the \( \bar{t}t \) background. Therefore, in order to reduce the number of \( \bar{t}t \) events, the \( am_{T2} \) selection is still applied to remove \( \bar{t}t \) events with two leptons where one of the leptons is mismeasured by the detector. This results in a region with 717 events, a 41% single top purity and a maximum signal contamination of 1.3%. All of these VRs are summarised in Table 13. The \( E_T^{\text{miss}} \) distributions are shown in Figure 51 for the tW1L channel VRs and the single top VR.
9.3 Systematic uncertainties

As described in Chapter 3 the systematic uncertainties influence the likelihood function as nuisance parameters and are included together with the control and signal regions into the “HistFitter” setup. In Section 3.2 it was explained that the “HistFitter” framework determines the impact of these systematic uncertainties by varying the values of each uncertainty in an interval around a central value with an upper limit on the variation of $1\sigma$. This is done by including generated MC events of three categories. The first “nominal” category describes the central value and for each systematic uncertainty an “up” and “down” variation is implemented additionally which represent the $+1\sigma$ and $-1\sigma$ variation of the systematic uncertainty. In this section, an overview over the included systematic uncertainties is presented as well as a short description of how these are measured. Section 9.3.1 thereby describes the evaluation of the experimental uncertainties that result from the measurement of particles in the detector while Section 9.3.2 describe theoretical uncertainties that result from the modelling of the MC samples for signals and SM backgrounds.

9.3.1 Experimental uncertainties

The most dominant experimental uncertainties in this analysis are the uncertainties related to jets since these are the main physics objects that are used in the $tW_0L$ channel and important objects in the $tW_{1L}$ channel. Uncertainties on jet measurements are provided for each definition of jets independently. This means that uncertainties for jets with a radius parameter of $R = 0.4$ and $R = 1.0$ are included and treated independently. For jets, two main categories make up the dominant part of the uncertainties which are uncertainties on the jet energy scale calibration (JES) and jet energy resolution (JER) that were presented in detail in Section 7.2.3. The uncertainties of the JES calibration are used as nuisance parameters in this
analysis. In total, there are 80 JES systematic uncertainties which result mostly from the in situ calibration and assumptions made in the MC simulations. All of these uncertainties are derived in bins of $p_T$ with $\eta = 0$ and bins of $\eta$ with $p_T = 80$ GeV which can be applied to the jets used in the analysis [134]. The systematic uncertainty on the JER is determined in a comparison between MC simulation and data. A smearing factor is applied to the jets in the MC simulation which is a random number from a Gaussian distribution with the width

$$\sigma = \sqrt{(\sigma_{\text{data}} + \Delta\sigma_{\text{data}})^2 - \sigma_{\text{data}}^2}$$

(101)

where $\sigma_{\text{data}}$ is the jet energy resolution measured in data events and $\Delta\sigma_{\text{data}}$ being its uncertainty resulting from the method described in Section 7.2.3. Comparing the difference between the jet energy resolution with and without smearing factor is determined as the systematic uncertainty on the JER [157, 158].

For the jets with a radius parameter of $R = 1.0$, very similar methods are used for the JES and JER systematic uncertainties. In addition to that, since JMS and JMR corrections are applied to large-$R$ jets as well, uncertainties on these corrections are included. For the JMS, the uncertainty on the correction described in Section 7.2.3 as well as other inaccuracies of the $R_{\text{trk}}$ method define the systematic uncertainties of this corrections [141]. For the JMR, the systematic uncertainties are determined similarly to the JER. A randomly generated value from a Gaussian distribution is applied on the mass resolution defined in Equation (88). The difference in JMR this results in is taken as the systematic uncertainty [142].

Another uncertainty related to jets is the uncertainty on the $W$-tagging. The uncertainties in this case are determined when validating the taggers in MC simulations and data. They can be determined for the efficiency of the tagger itself which is defined as

$$\epsilon_{\text{MC}} = \frac{N_{\text{tagged truth } W}}{N_{\text{total truth } W}}$$

(102)
for MC simulations and as

\[
\epsilon_{\text{data}} = \frac{N_{\text{tagged}}^{\text{data}} - N_{\text{tagged}}^{\text{truth non-}W}}{N_{\text{data}}^{\text{total}} - N_{\text{truth non-}W}^{\text{total}}} \quad (103)
\]

in data. \(N_{\text{truth W}}^{\text{total}}\) in this case is the number of all events that are \(W\)-boson matched before \(W\)-tagging while \(N_{\text{tagged}}^{\text{truth W}}\) describes the number of \(W\)-tagged events. Similar for data, \(N_{\text{data}}^{\text{total}}\) and \(N_{\text{data}}^{\text{tagged}}\) describe the number of all data events and the number of tagged data events. However, in the data case, the number of these events is reduced by \(N_{\text{truth non-}W}^{\text{total}}\) and \(N_{\text{tagged}}^{\text{truth non-}W}\) in order to account for the background events without a matched \(W\)-boson before and after applying the tagging algorithm.

Aside from the efficiency, uncertainties can also be determined on scale factors that are applied in order to account for differences between the efficiencies of the tagger in data and MC [159]. This is done in a very similar way for the uncertainties on scale factors in the \(b\)-tagging algorithm [160]. Other systematic uncertainties that are also applied in the “HistFitter” setup, but which are not as dominant can be summarised as uncertainties on the pile-up, the luminosity and the missing transverse energy. Furthermore, since the \(tW_{1L}\) channel includes lepton objects as well as some of the control regions, lepton efficiency, identification, calibration, isolation and trigger uncertainties are also considered.

9.3.2 Theoretical uncertainties

The estimation of the theoretical uncertainties is very similar between the different SM background and signal samples. However, depending on the generators that were used to generate the MC events, a few differences in the uncertainty calculation are defined. For \(V + \text{jets}\) SM backgrounds, i.e. \(W + \text{jets}\) and \(Z + \text{jets}\), the events were generated in the SHERPA MC generator. The uncertainties on the modelling for these MC events can be estimated by varying four different parameters and comparing the
results to the generated events without any variation of these parameters. These parameters include the matrix element matching scale (ckkw), the resummation scale (qsf) and the renormalisation ($\mu_R$) and factorisation ($\mu_F$) scale. The ckkw scale is used for the calculation of the overlap between the jets from the matrix element calculation and parton showering. Its nominal value is set to 20 GeV while the two variations change this parameter to 30 GeV and 15 GeV. The qsf describes the scale that is used for the resummation of soft gluon emission. The nominal value for this resummation scale $\mu_{qsf}$ is varied when estimating the uncertainties to double and half of its value. The last two scales for events generated by SHERPA describe the scale for the strong coupling constant which is used in the calculation of the underlying hard process ($\mu_R$) as well as the scale used for the calculation of the parton density functions ($\mu_F$). Similar to the qsf uncertainties, the nominal value for these scales is varied to double and half its value when estimating the uncertainties. However, these uncertainties are estimated together in a 7-point variation scheme. For this process, in addition to either one of these scales being varied individually, the uncertainties are also estimated by varying both of these parameters to double or half their value at the same time resulting in a total of seven different estimations where the first one represents the nominal generation where neither of the two parameters is varied [161, 162]. For all of these six variations of parameters, an envelope strategy is used in this analysis to determine the uncertainty. In this case, the highest difference in either of these variations to the nominal result is chosen and included as both the “up” and “down” variation.

For the $t\bar{t}$ and single top background, the 7-point variation scheme is used as well to estimate the uncertainties on the renormalisation and factorisation scales. However, uncertainties due to the modelling of the hard scatter process and the parton shower hadronisation are estimated by comparing the samples used in this analysis listed in Table 3 to alternatively generated samples. These alternative samples are a combination of samples generated with MADGRAPH5_aMC@NLO and POWHEG interfaced to PYTHIA8 and HERWIG7 [163] respectively. The uncertainty is derived
as

$$\sigma_{\text{sys}} = \frac{N_{\text{nominal}} - N_{\text{alternative}}}{N_{\text{nominal}}}$$

(104)

by comparing the predicted events numbers $N$ in nominal and alternative samples. However, due to the fact that the detector response in the alternative samples was simulated with the ATLFAST-II simulation, regions where $W$-tagging is utilised are compared on “TRUTH” level as the modelling of the $W$-tagging in the ATLFAST-II simulation is not reliable. Furthermore, in the case of the single top background, another uncertainty to the ones mentioned is measured in this comparison to alternative samples. When generating single top samples with a $W$-boson in the final state, the generated events overlap with events coming from $t\bar{t}$ processes. This overlap has to be removed in the single top generation to avoid double counting so that the single top samples can be later combined with $t\bar{t}$ background samples [164]. There are two methods for removing the overlap. The one used in the nominal single top samples is the diagram removal method (DR). In this method, all Feynman-diagrams in the NLO amplitudes with a $W$-boson in the final state and that are resonant to $t\bar{t}$ and single top production, are removed from the calculation. The second method that is used in the alternative samples, is the diagram subtraction (DS). This method modifies the cross-section of the NLO single top processes with a $W$-boson in the final state. For this, a subtraction that cancels the $t\bar{t}$ contribution is implemented into the calculation [165]. By comparing nominal and alternative samples in the described way, the resulting uncertainty includes the effect the different overlap removal methods have on the modelling. Finally, for $t\bar{t}$ and single top background samples, uncertainties on the initial (ISR) and final state radiation (FSR) are included. This is done by varying different internal weights in PYTHIA8 when generating $t\bar{t}$ and single top samples and comparing the generated samples with the nominal values [166].

For $t\bar{t}V$ processes, uncertainties are determined in a similar way. Parton shower and radiation uncertainties are determined by varying internal weights in PYTHIA8 and by using HERWIG7 instead of PYTHIA8 respectively. Additional uncertainties are
determined by comparing the nominal generated samples to samples generated with
the \textsc{Sherpa} generator. The uncertainties for the $t\bar{t}V$ background are all determined
at "TRUTH" level.

For the signal process, systematic modelling uncertainties are determined as well.
The renormalisation and factorisation uncertainties are evaluated with a 3-point
variation where one point represents the nominal values and the other two points
represent the simultaneous variation of $\mu_R$ and $\mu_F$ by a factor of 2 and 0.5 respectively.
Furthermore, radiation uncertainties (ISR and FSR) are estimated in the same way
as the background by varying internal weights in \textsc{Pythia8}. The uncertainty on every
of those variations $\sigma_{\text{var}}$ is evaluated through

$$
\sigma_{\text{var}} = \frac{N_{\text{nominal}} - N_{\text{variation}}}{N_{\text{nominal}}} \quad (105)
$$

by using the event numbers $N$ for the nominal and varied parameters similar to the
background uncertainty estimation. However, since the signal process uncertainties
are only included in the model-dependent fit, they are combined for every signal
process to reduce the number of uncertainties. This is done through the error
propagation method

$$
\sigma_{\text{tot}} = \sqrt{\sum_{i=0}^{N} (\sigma_{\text{var},i})^2} \quad (106)
$$

where the upper and lower uncertainties are then represented by half the value of
$\sigma_{\text{tot}}$. 

10 Results

The signal regions defined in Chapter 8 and the control and validation regions as well as the systematic uncertainties defined in Chapter 9 are included in the statistical tool “HistFitter” [72] which performs three fits that provide the results of the analysis. The first fit is the background-only fit which provides normalisation factors from the control regions which are later used as set parameters. Furthermore, the background-only fit also uses all the fitted parameters to provide a first result in the signal regions. By comparing the fitted SM background predictions to the measured data, it can be analysed if the background-only hypothesis can be excluded. If there is no significant excess observed, two additional fits are performed. These two fits provide upper limits for different signal processes. In the model-independent fit described in Section 10.2, the upper limits for a generic signal are calculated while the model-dependent fit presented in Section 10.3 provides upper limits for the specific signal processes of the underlying 2HDM+a model.

10.1 Background-only fit

The defined CRs, VRs and SRs together with the normalisation factors of the backgrounds are included in “HistFitter” with the systematic uncertainties as nuisance parameters. The normalisation factors as well as nuisance parameters are fitted at the same time in the CRs and SRs while the results are applied to the VRs to validate the background normalisation. In this section, the results of this fit are
First, the effect of the systematic uncertainties is shown. In Figure 52, the impact of the different groups of systematic uncertainties on the overall uncertainty are presented. It can be observed that the overall uncertainty is dominated by detector and modelling uncertainties. The detector and object reconstruction related uncertainties are mainly dominated in the tW_{0L} channel by W-tagging and JER uncertainties. For the modelling uncertainties, the tW_{0L} channel is dominated by the modelling of Z+jets and single top processes. This is expected since Z+jets is one of the main backgrounds in the tW_{0L} signal regions and as mentioned before, single top modelling uncertainties are expected to be high even though they were reduced in this analysis through the inclusion of the single top CR. In the case of the tW_{1L} channel, the detector uncertainties are dominated by JER and JES uncertainties. The W-tagging uncertainties are also a dominant part of the SR_{lep.top} region, but not the SR_{had.top} region since the SR_{had.top} region does not utilise W-tagging. Similarly, the modelling uncertainties are expected to change due to different dominant backgrounds in this region. Instead of Z+jets, \ttbar and W+jets modelling uncertainties are dominant in the tW_{1L} channel. Like in the tW_{0L} channel, the single top modelling uncertainty also contributes significantly to this uncertainty group.
<table>
<thead>
<tr>
<th>Normalisation parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_{tW0L}$</td>
<td>$0.96 \pm 0.12$</td>
</tr>
<tr>
<td>$\mu_{tW1L}$</td>
<td>$0.92 \pm 0.06$</td>
</tr>
<tr>
<td>$\mu_{W+ jets}$</td>
<td>$0.98 \pm 0.07$</td>
</tr>
<tr>
<td>$\mu_{Z+ jets}$</td>
<td>$1.08 \pm 0.10$</td>
</tr>
<tr>
<td>$\mu_{Single t}$</td>
<td>$0.43 \pm 0.13$</td>
</tr>
<tr>
<td>$\mu_{t\bar{t}Z}$</td>
<td>$1.18 \pm 0.17$</td>
</tr>
</tbody>
</table>

Table 14: Normalisation parameters for the main backgrounds in the analysis for the tW$_{0L}$ and tW$_{1L}$ channels.

Figure 53: The $E_T^{\text{miss}}$ and $E_T^{\text{miss}}$ distributions for the $t\bar{t}$ (upper plots) and $Z + \text{jets}$ CRs (lower plots) of the tW$_{0L}$ analysis channel. On the left side, the distributions are shown before the normalisation parameters of the fit are applied. On the right side, it shows the distributions with applied normalisation parameters and fitted uncertainties.
Figure 54: The $E_{\text{miss}}$ distributions for the $t\bar{t}$ (upper plots) and $W+$ jets CRs (lower plots) of the $tW_{1L}$ analysis channel. On the left side, the distributions are shown before the normalisation parameters of the fit are applied. On the right side, it shows the distributions with applied normalisation parameters and fitted uncertainties.

The second result consists of the normalisation parameters. In this analysis, the background-only fit results in six parameters, one for every CR while these parameters are fitted in all CRs at the same time. The resulting normalisation parameters are shown in Table 14 and the effect of this normalisation on the distributions of $E_T^{\text{miss}}$ and $E_{T,\text{ll}}^{\text{miss}}$ in all CRs and VRs are shown in Figure 53 - Figure 58 where a comparison is done between the distributions before and after applying the normalisation factors and the fitted uncertainties. It can be observed that most of the normalisation parameters are at a value which is in a 1$\sigma$, at maximum 2$\sigma$, interval around 1 with the exception being the single top normalisation parameter. However, a low
normalisation parameter for the single top MC predictions can be expected in this phase space. The reason for this is the diagram removal method described in Section 9.3.2 to remove overlaps with the $t\bar{t}$ background. In the phase space of this analysis, the predictions of both methods, the DR and the DS, are not very good resulting in such a high difference from 1 for the normalisation factor. Therefore, the normalisation factor for the single top background has to be further validated through the single top VR.

The figures show how the normalisation parameters impact the difference between data and MC predictions. Especially in the case of the single top background,
Figure 56: The $E_{\text{T}}^\text{miss}$ distributions for the $t\bar{t}$ (upper plots) and $V + jets$ VRs (lower plots) of the $tW_{0L}$ analysis channel. On the left side, the distributions are shown before the normalisation parameters of the fit are applied. On the right side, it shows the distributions with applied normalisation parameters and fitted uncertainties.

this effect can be observed the strongest due to the normalisation factor having the highest difference from 1. Furthermore, it can also be observed in Figure 56 - Figure 58 which show the distributions for the VRs that the normalisation factors can get validated when comparing data against the MC predictions in these regions while also accounting for uncertainties. This can be observed in a similar way in Figure 59 which shows all CRs and VRs in a histogram with one bin respectively. The first six bins represent the CRs which show no difference between data and MC predictions. This is expected as the number of CRs is the same as the number of normalisation parameters. But the VRs also show no significant deviation between
Figure 57: The $E_T^{\text{miss}}$ distributions for the SR$^{\text{had.top}}_T \bar{t}t$ (upper plots), the SR$^{\text{lep.top}}_T \bar{t}t$ (middle plots) and $W+\text{jets}$ VRs (lower plots) of the $tW_{1L}$ analysis channel. On the left side, the distributions are shown before the normalisation parameters of the fit are applied. On the right side, it shows the distributions with applied normalisation parameters and fitted uncertainties.
Figure 58: The $E_{\text{T}}^{\text{miss}}$ distributions for the single top VR. On the left side, the distributions are shown before the normalisation parameters of the fit are applied. On the right side, it shows the distributions with applied normalisation parameters and fitted uncertainties.

Figure 59: A summary of the comparison between data and MC predictions for all CRs and VRs in a one bin histogram after applying the fit results from the background-only fit. The lower panel shows the significant deviation between data and MC predictions in terms of $\sigma$.

data and MC predictions. The highest difference can be observed in the VR for the $tW_{\text{qL}} \bar{t}t$ background which is at around 1$\sigma$. However, this difference is still low enough to validate the fitted normalisation parameters shown in Table 14.

Finally, the background-only fit also provides a first analysis of the signal regions.
Figure 60: A summary of the comparison between data and MC predictions for all signal region bins after applying the fit results from the background-only fit. The lower panel shows the significant deviation between data and MC predictions in terms of $\sigma$.

By comparing data and MC predictions in the signal region bins after applying all of the fit results of the background-only fit, it can be analysed if a signal is observed. This result is shown in Figure 60. The figure shows no significant excess between data and MC predictions in any of the signal region bins which means that there is no new physics measured in the data. Therefore, results are then interpreted as limits on the cross section of a signal in two ways. The model-independent fit of Section 10.2 provides upper limits on the cross section and event numbers of a generic signal while the model-dependent fit of Section 10.3 provides upper limits on the cross section of the 2HDM+a signals. However, there is one additional detail that can be observed in this result. A significant deficit ($\sim 2.5\sigma$) between data and MC predictions in the fourth bin of the tW$_{0L}$ channel can be observed. However, considering that the third and fifth bin do not show a similar behaviour, this is considered as a statistical fluctuation. A total overview over the number of fitted background events in comparison to the number of measured data events for all signal regions can be found in Table 15 and Table 16.
First, a model-independent result is provided. This describes the potential upper limits on the number of signal events without such a generic signal being excluded through the results of the analysis. Table 17 shows the results of this fit. The
interpretation of these results will be discussed in this section.

Since there is no distribution in $E_T^{\text{miss}}$ for a generic signal, the signal channels are split into inclusive region as it can be observed in the first column. This means that the second line takes the sum of all observed and expected events in the $tW_{0L}$ channel bins into the calculation. The third line takes the sum of all events from the second to last bin of the $tW_{0L}$ channel bins and so on. The sum of these observed and expected events are shown in the second and third column. The columns four to six describe different versions of 95% confidence level (CL) upper limits of a potential generic signal. The fourth line describes the upper limit of the visible cross section in fb for such a generic signal taking the observed and expected events into account. The fifth and sixth column describe two different ways of upper limits on the number of signal events from this generic signal sample. The fifth column takes both, the expected events as well as observed events into account. The sixth column calculates

<table>
<thead>
<tr>
<th>Signal channel</th>
<th>Obs.</th>
<th>SM exp.</th>
<th>$\langle \epsilon \sigma \rangle_{\text{obs}}^{95}$ [fb]</th>
<th>$S_{\text{obs}}^{95}$</th>
<th>$S_{\text{exp}}^{95}$</th>
<th>$CL_B$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$SR_{tW0L}(E_T^{\text{miss}} \geq 250 \text{ GeV})$</td>
<td>133</td>
<td>147 ± 15</td>
<td>0.21</td>
<td>29</td>
<td>36_{-10}^{+14}</td>
<td>0.24</td>
</tr>
<tr>
<td>$SR_{tW0L}(E_T^{\text{miss}} \geq 330 \text{ GeV})$</td>
<td>66</td>
<td>82.8 ± 8.9</td>
<td>0.11</td>
<td>15.5</td>
<td>23.7_{-6.6}^{+9.4}</td>
<td>0.09</td>
</tr>
<tr>
<td>$SR_{tW0L}(E_T^{\text{miss}} \geq 400 \text{ GeV})$</td>
<td>33</td>
<td>42.0 ± 5.6</td>
<td>0.08</td>
<td>11.7</td>
<td>16.4_{-4.8}^{+6.8}</td>
<td>0.15</td>
</tr>
<tr>
<td>$SR_{tW0L}(E_T^{\text{miss}} \geq 500 \text{ GeV})$</td>
<td>8</td>
<td>16.6 ± 2.3</td>
<td>0.04</td>
<td>5.4</td>
<td>9.7_{-2.8}^{+4.3}</td>
<td>0.03</td>
</tr>
<tr>
<td>$SR_{tW0L}(E_T^{\text{miss}} \geq 600 \text{ GeV})$</td>
<td>6</td>
<td>7.0 ± 1.7</td>
<td>0.05</td>
<td>6.5</td>
<td>7.4_{-2.1}^{+3.3}</td>
<td>0.38</td>
</tr>
<tr>
<td>$SR_{tW1L}(E_T^{\text{miss}} \geq 250 \text{ GeV})$</td>
<td>239</td>
<td>237 ± 25</td>
<td>0.42</td>
<td>58</td>
<td>57_{-15}^{+21}</td>
<td>0.53</td>
</tr>
<tr>
<td>$SR_{tW1L}(E_T^{\text{miss}} \geq 300 \text{ GeV})$</td>
<td>130</td>
<td>121 ± 17</td>
<td>0.33</td>
<td>46.4</td>
<td>40_{-11}^{+15}</td>
<td>0.67</td>
</tr>
<tr>
<td>$SR_{tW1L}(E_T^{\text{miss}} \geq 350 \text{ GeV})$</td>
<td>69</td>
<td>65.8 ± 8.8</td>
<td>0.19</td>
<td>26.3</td>
<td>24.1_{-6.0}^{+9.5}</td>
<td>0.60</td>
</tr>
<tr>
<td>$SR_{tW1L}(E_T^{\text{miss}} \geq 400 \text{ GeV})$</td>
<td>40</td>
<td>36.7 ± 9.0</td>
<td>0.17</td>
<td>23.7</td>
<td>21.7_{-5.8}^{+7.8}</td>
<td>0.62</td>
</tr>
<tr>
<td>$SR_{tW1L}(E_T^{\text{miss}} \geq 450 \text{ GeV})$</td>
<td>25</td>
<td>20.2 ± 8.4</td>
<td>0.16</td>
<td>22.0</td>
<td>19.1_{-5.0}^{+6.4}</td>
<td>0.69</td>
</tr>
<tr>
<td>$SR_{tW1L}(E_T^{\text{miss}} \geq 500 \text{ GeV})$</td>
<td>9</td>
<td>6.4 ± 2.3</td>
<td>0.07</td>
<td>10.2</td>
<td>8.6_{-2.3}^{+3.8}</td>
<td>0.74</td>
</tr>
</tbody>
</table>

Table 17: Results of the model-independent fit. The first column describes the signal channel, the second and third column the observed and expected MC events after the background-only fit. The fourth and fifth columns present the 95% CL upper limits on visible cross section and the observed number of events while the sixth column shows the 95% CL upper limit on the number of signal events depending on the expected background events in a ±1σ deviation. The last column describes the confidence level for the background-only hypothesis $CL_B$. 
this upper limit by only using the expected number of background events by assuming a $\pm 1\sigma$ interval around this expected number of events for the observed events. The last column describes the CL for the background-only hypothesis. The lower this value, the higher the possibility that the background-only hypothesis can be rejected. This can be best observed in the fifth line which uses the two last bins of the $tW_{0L}$ channel. The fourth bin in particular was the bin that showed a significant deficit between data and MC predictions in Figure 60. Therefore, the background-only hypothesis is less likely when analysing this bin of the analysis. However, since the difference is a deficit and the potential signal is always treated as an addition to the SM events, this alone can not be used as an indication of new physics.

### 10.3 Model-dependent results

The model-dependent results analyse the potential exclusion limits for signals in the 2HDM+a model described in Section 2.3.2. Since the cross sections and even the distributions of the signals are dependent on different parameters, the exclusion limits are analysed in two different parameter planes. First, the signals in this analysis are always analysed with a $\sin\theta$ value of $1/\sqrt{2}$. In the case that the signal was generated with $\sin\theta = 1/(2\sqrt{2})$, the cross section is recalculated to represent $\sin\theta = 1/\sqrt{2}$. Furthermore, the masses of the two charged Higgs bosons, the CP-odd Higgs boson and the heavy CP-even and neutral Higgs boson are set to be equal $m_{H^\pm} = m_A = m_H$. The first plane is a $m_a - m_H$ parameter plane with the mass of the mediator $a$ on the x-axis and the mass of the heavy CP-even Higgs boson on the y-axis. The mass of the mediator $a$ can be found in the interval of 100 GeV to 450 GeV while the mass of the heavy Higgs boson is set between 400 GeV and 2000 GeV. In this case the $\tan\beta$ value is set to 1. The second parameter plane is a $m_H - \tan\beta$ plane with the mass of the heavy CP-even Higgs boson on the x-axis and the $\tan\beta$ value on the y-axis. The mass of the heavy Higgs boson is again set between 400 GeV and 2000 GeV while the value of $\tan\beta$ is varied between 0.3 and
The expected and observed exclusion contours for (a)+(b) the parameter plane \(m_a - m_H\) (c)+(d) the parameter plane \(m_H - \tan\beta\) with \(m_a = 250\) GeV and (e)+(f) the parameter plane \(m_H - \tan\beta\) with \(m_a = 150\) GeV. The figures (a),(c) and (e) show the exclusion contours with only the DM\(_t\) samples while (b),(d) and (f) show the exclusion contours for the combined DM\(_t\) and DM\(t\bar{t}\) samples. The solid lines show the observed exclusion contour and the dashed lines the expected exclusion contours. The yellow, green and red lines describe the exclusion contours of the individual tW\(_{2L}\), tW\(_{0L}\) and tW\(_{1L}\) channels while the blue contour line shows the combined exclusion contour of all three channels. The yellow band describes the 1\(\sigma\) uncertainty on the expected exclusion contour of all three channels combined.

Figure 61:
This parameter plane has two sets, one where the mass of the mediator is set to 250 GeV and one where its mass is set to 150 GeV. In addition to that, every plane is accompanied by a second plane with the same parameters. However, in this case, not only signals with a single top quark in the final state are taken into account. While this analysis was designed to maximise the sensitivity to DM final states in association with one single top quark and a W-boson, signals with DM and two top quarks in the final state are also partially sensitive due to the similarity in the signature. Therefore, a combination of both signatures might improve the exclusion limits. However, as these signals are generated in a different framework, they have to be reweighted in order to emulate a final state in the 2HDM+a model as it was described in Chapter 6.

The exclusion contours of the individual $tW_{0L}$, $tW_{1L}$ and $tW_{2L}$ channels as well as the exclusion contour for the combined analysis of all three channels is shown in Figure 61. In this figure, both versions are shown, the version where just the signals with a single top quark in the final state are included as well as the version where the DM+$tW$ final state is combined with the DM+$t\bar{t}$ final state. It can be observed in these figures that the combined exclusion is enhanced for high masses of $H$ as well as low values of $\tan \beta$ when including the DM+$t\bar{t}$ final state. Furthermore, it can be observed that the $tW_{0L}$ and $tW_{1L}$ channel are the dominant channels at high masses of $H$ and $a$ and their combination even allows an exclusion at high $\tan \beta$ values. The reason why an exclusion at $\tan \beta = 30$ is possible results from the increase in cross section for the signal processes at higher $\tan \beta$ values which was shown in Figure 11. The $tW_{2L}$ channel in comparison mainly contributes to the combined exclusion at low values for the mass of $H$. With all three channels combined, values up to $m_H = 1500$ GeV, as well as $m_a = 400$ GeV and $\tan \beta = 30$ can be excluded which is an improvement to the exclusion of the previous analysis shown in Figure 62 that did not utilise the $W$-tagging and the $tW_{0L}$ channel [29].

Another observation is that the $tW_{0L}$ channel is the only channel where the observed exclusion contour is enhanced in comparison to the expected exclusion contour. The
reason for this is the fact that the $tW_0L$ channel showed a deficit in data in one of the signal bins which results in a higher observed exclusion when combining all of the signal regions since the expected exclusion limit assumes data events similar to the expected background events.

In comparison with results in similar analyses of the CMS experiment, the results of this thesis also provide a better exclusion. Figure 63 shows the results of a DM search in association with a single top quark and a top quark pair from Ref. [167]. However, in difference to this analysis, CMS used a simplified model with four free parameters which are the mass of the mediator $a$ ($m_a$), the mass of the DM particles which is set to 1 GeV ($m_\chi$) and the coupling of the mediator to the SM quarks as well as the DM particles $g_q = g_\chi = 1$. The second difference is that the CMS analysis only uses 35.9 fb$^{-1}$ of data compared to the 139.0 fb$^{-1}$ used in this analysis. The CMS results show that the DM$t$ analysis channel alone can not exclude any mediator in the mass window analysed. However, when combining the results from the DM$t$ and DM$t\bar{t}$ analysis channel, mediator masses up to 300 GeV can be excluded. In comparison, the analysis presented in this thesis allowed exclusions up to $m_a = 400$ GeV.

Figure 62: Previous results of the analysis for DM in association with a top quark and a $W$-boson with the combination of the channels with one lepton and two leptons in the final state. (a) shows the exclusion limits for the 2HDM+$a$ signals in the $m_a$-$m_H$ plane with $\tan \beta = 1$ while (b) shows the exclusion limits in the $m_H$-$\tan \beta$ plane with $m_a = 250$ GeV [29].
Overall, data are found to be in agreement with the expected SM background in the SRs. Upper limits at 95% confidence level (CL) are computed on the ratio between the measured and theoretical cross sections for a scalar or pseudoscalar mediator and either the $t/\bar{t}$ or $t/\bar{t}$ production process. The expected limit for the $t/\bar{t}$ signal alone is depicted by the blue dash-dotted line, while the expected limit for the $t/\bar{t}$ signal model when compared to the $t/\bar{t}$ signal model alone is depicted by the blue dash-dotted line, while its relative to the theory predictions, shown for the scalar (left) and pseudoscalar (right) models.

Figure 6: The expected and observed 95% CL limits on the DM production cross-section, relative to theory predictions for DM$t$ (blue), DM$\bar{t}$ (red) and combined (black) processes. The results are based on a simplified model with four free parameters: the mass of the mediator ($m_\chi$), the mass of the DM ($m_\chi = 1$ GeV) and the couplings of the mediator to the SM and DM $g_q = g_\chi = 1$. [167].
11 Summary and Outlook

This thesis presented an analysis for the search of DM with data from the ATLAS detector at the LHC where the DM in the final state was produced in association with one top quark and one $W$-boson. More precisely, the focus was on implementing an analysis channel with zero leptons in the final state and showing that it is comparable to the already analysed channels with one and two leptons in the final state. First, the analysis strategy was presented with the $W$-tagging being the main discriminator for this analysis channel exploiting the fact that the $W$-boson not originating from the top quark can be expected to have a high $p_T$. This resulted either from the high mass of the intermediate Higgs boson from the 2HDM in comparison to the mass of the mediator $a$ or the back to back production of the $W$-boson against a mediator with a higher mass. The inclusion of the $W$-tagging algorithm allowed for a very good discrimination between signal and SM background events and through splitting the signal region into five regions depending on the $E_T^{\text{miss}}$ value, this allowed high sensitivity significances for most analysed signal processes. Together with the other two analysis channels, a number of control regions was defined to constrain the SM backgrounds and which entered the background-only, the model-independent and model-dependent fit. The result of the background-only fit showed no significant excess in either of the signal regions but one significant deficit in one of the $tW_{0L}$ signal region bins. However, through comparison to the signal regions with similar $E_T^{\text{miss}}$ values, it was determined that this deficit is most likely a statistical fluctuation. Since the background-only hypothesis could not be
rejected, an exclusion fit for a model-independent signal and the signals presenting the 2HDM+a model was done. The results from this fit showed a higher exclusion of signal processes compared to the previous analysis [29], especially in the regions with high masses of $H^\pm$ and $a$. Furthermore, the combined results from the $tW_{0L}$, $tW_{1L}$ and $tW_{2L}$ channels allowed an exclusion of points with $\tan\beta = 30$ for the first time in this model as it was theorised in Ref. [67]. Finally, the results also showed the impact of the new channel with zero leptons in the final state in comparison to the other two analysis channels. Especially at high masses of $H^\pm$, the new channel was comparable to the channel with one lepton in the final state which was previously the dominant channel of the analysis.

In the future, the analysis can be further improved in two ways. The first is by including more data from the ATLAS detector from future measurements. Since the number of events for both SM backgrounds as well as signal processes are directly proportional to the luminosity, the sensitivity of the signal process increases even when keeping the same signal region definitions. Exemplary, this can be shown with the event number values of Table 6 from Section 8.4. In the current analysis, the total number of SM background events is 155.1 while the number of signal events for a signal process with $m_a = 250$ GeV, $m_{H^\pm} = 800$ GeV and $\tan\beta = 1$ is 77.9. This results in an expected signal sensitivity of 2.04 assuming a 20% uncertainty on the background events. However, if the luminosity was doubled, then this would result in two times the amount of background and signal events which therefore leads to an expected signal sensitivity of 2.11. While the improvement seems low, the significance is very dependent on the actual uncertainty on the background events which is often lower than 20% as Figure 52 showed for most signal region bins.

The second possible improvement of the analysis is by improving the signal region definiton. One of the main SM backgrounds included $\tau$ leptons that were decay products of either the top quark or a $W$-boson. These types of leptons were not analysed in more detail in this analysis. Hadronically decaying $\tau$ leptons have unique properties that differentiate them from other hadronised jets and it can therefore be
determined if a jet originated from a $\tau$ lepton. This algorithm is described in detail in Ref. [168] and it could be used to further discriminate background from signal events by using specific $\tau$ lepton properties in either the background or signal. Analysing the properties of $\tau$ leptons between the signal and SM background would improve the signal sensitivity in two ways. First, the difference in properties between $\tau$ leptons in the signal and in $t\bar{t}$ and $W+\text{jets}$ background events could be used to reject more of these backgrounds since these were the second and third most dominant backgrounds in the final signal region definition. Furthermore, the existence of $\tau$ leptons in the final state could also allow for the rejection of the most dominant background in the $tW_{0L}$ channel, $Z+\text{jets}$, as this background mainly consists of $Z \to \nu\nu$ events which does not include a $\tau$ lepton.
Bibliography


# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Elementary Particles of the SM</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>Examples for possible interactions between fermions with gauge bosons</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>as the carrier of the charge</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Examples for possible interactions between gauge bosons</td>
<td>8</td>
</tr>
<tr>
<td>4</td>
<td>Higgs Potential in cases with $\mu^2 &lt; 0$</td>
<td>13</td>
</tr>
<tr>
<td>5</td>
<td>Comparison between theory and observations of galaxy rotation curves</td>
<td>18</td>
</tr>
<tr>
<td>6</td>
<td>Example of gravitational lensing around galaxy cluster “CL0024+17”</td>
<td>20</td>
</tr>
<tr>
<td>7</td>
<td>The temperature fluctuations in the cosmic microwave background</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>measured with WMAP</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Temperature power spectrum $D_{TT}^{TT}$ from the Planck 2018 data</td>
<td>22</td>
</tr>
<tr>
<td>9</td>
<td>Processes for the production of DM in the final state</td>
<td>27</td>
</tr>
<tr>
<td>10</td>
<td>Feynman diagrams for the production of DM in association with a</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>single top quark</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Cross section for the production of DM from $pp$ collisions in association</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>with a top quark at 14 TeV as a function of $\tan\beta$</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>A schematic overview of the analysis strategy in the “HistFitter” framework</td>
<td>36</td>
</tr>
<tr>
<td>13</td>
<td>The CERN accelerator complex</td>
<td>44</td>
</tr>
<tr>
<td>14</td>
<td>Overview of the ATLAS detector</td>
<td>46</td>
</tr>
<tr>
<td>15</td>
<td>Diagram of the interaction of different particles in the layers of the</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td>ATLAS detector</td>
<td></td>
</tr>
</tbody>
</table>

187
<table>
<thead>
<tr>
<th>Page</th>
<th>Description</th>
<th>Appendix</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>The inner detector of the ATLAS detector</td>
<td>49</td>
</tr>
<tr>
<td>17</td>
<td>The Insertable B-Layer of the ID</td>
<td>51</td>
</tr>
<tr>
<td>18</td>
<td>The Calorimeter System of the ATLAS detector</td>
<td>52</td>
</tr>
<tr>
<td>19</td>
<td>The muon system of the ATLAS detector</td>
<td>54</td>
</tr>
<tr>
<td>20</td>
<td>Overview of the trigger system at the ATLAS detector</td>
<td>56</td>
</tr>
<tr>
<td>21</td>
<td>A visualisation of the possible patterns at the LHC</td>
<td>61</td>
</tr>
<tr>
<td>22</td>
<td>Ratio of the run-integrated track counting luminosity calculated in the 2016 selection, the 2017 selection and the 2017 + Si hit selection</td>
<td>64</td>
</tr>
<tr>
<td>23</td>
<td>2D histogram showing the amount of events with specific values of $\langle \mu \rangle$ for two different selections in a LHC fill with low average interaction rate</td>
<td>66</td>
</tr>
<tr>
<td>24</td>
<td>Ratio between the $\langle \mu \rangle$ values of different track selections using two different correlation factors</td>
<td>67</td>
</tr>
<tr>
<td>25</td>
<td>1D histogram of the square root of the $\chi^2$ values where one event represents the $\sqrt{\chi^2}$ value for the ratio of $\langle \mu \rangle$ values between two selections for an individual BCID</td>
<td>69</td>
</tr>
<tr>
<td>26</td>
<td>Ratio of the run-averaged pile-up parameter $\langle \mu \rangle$ between the 2016 selection, 2017 + Si hit selection and the 2017 selection in dependence on the position along the bunch string</td>
<td>71</td>
</tr>
<tr>
<td>27</td>
<td>Ratio of the run-averaged pile-up parameter $\langle \mu \rangle$ between the 2016 selection, 2017 + Si hit selection and the 2017 selection in dependence on the bunch position inside the bunch train</td>
<td>72</td>
</tr>
<tr>
<td>28</td>
<td>Distribution of the fitted values for the ratio between the $\langle \mu \rangle$ values of the 2016 selection and 2017 selection in dependence of the average interaction rate of the LHC fill</td>
<td>74</td>
</tr>
<tr>
<td>29</td>
<td>Distribution of the fitted values for the ratio between the $\langle \mu \rangle$ values of the 2017 + Si hit selection and 2017 selection in dependence of the average interaction rate of the LHC fill</td>
<td>75</td>
</tr>
</tbody>
</table>
30 Visualisation of the time over threshold (ToT) in the ATLAS pixel detector ........................................... 76
31 An illustration of the production of a MC event ................. 80
32 Kinematic variables after applying the preselection requirements ... 109
33 The three decay channels for the SM $t\bar{t}$ background ........ 110
34 The two decay channels of the $\tau$ lepton ......................... 110
35 TRUTH level analysis for the SM $t\bar{t}$ background after applying the preselection requirements .................. 112
36 The $m_T(b_1, E_T^{\text{miss}})$ variable after applying the preselection requirements for all SM backgrounds as well as four signal processes ... 114
37 The $S_{E_T^{\text{miss}}}$ variable after applying the preselection requirements and $m_T(b_1, E_T^{\text{miss}}) \geq 180$ GeV for all SM backgrounds as well as four signal processes .............................. 115
38 The $\min[\Delta\phi(jet_{1-4}, E_T^{\text{miss}})]$ variable after applying the preselection requirements and $m_T(b_1, E_T^{\text{miss}}) \geq 180$ GeV and $S_{E_T^{\text{miss}}} \geq 14$ for all SM backgrounds as well as four signal processes .... 116
39 The number of tagged $W$-bosons after applying the preselection requirements and $m_T(b_1, E_T^{\text{miss}}) \geq 180$ GeV, $S_{E_T^{\text{miss}}} \geq 14$ and $\min[\Delta\phi(jet_{1-4}, E_T^{\text{miss}})] \geq 0.9$ for all SM backgrounds as well as four signal processes ....... 117
40 The angular distance between the tagged $W$-boson and the $b$-jet with the highest $p_T$ after applying the preselection requirements and $m_T(b_1, E_T^{\text{miss}}) \geq 180$ GeV, $S_{E_T^{\text{miss}}} \geq 14$, $\min[\Delta\phi(jet_{1-4}, E_T^{\text{miss}})] \geq 0.9$ and $N_{W-\text{tagged}}^{J^{R=1.0}} \geq 1$ for all SM backgrounds as well as four signal processes ................................. 119
41 The reconstructed mass of the tagged $W$-boson and the $b$-jet with the highest $p_T$ ............................................ 119
42 The distribution for the $E_T^{\text{miss}}$ variable after applying the preselection and all selection requirements ......................... 121

189
The distribution of the mass between the lepton and the $b$-jet with the highest $p_T$ and the distribution of the mass between the $b$-jet with the highest $p_T$ and the jet with the highest $p_T$ that is not a $b$-jet.

The $m_{\text{lep,MET}}^{T}$ variable after the preselection and $am_{T2}$ variable after preselection and $m_{\text{lep,MET}}^{T} \geq 120$ GeV.

The distribution for the number of tagged $W$-bosons in the $\text{SR}_{\text{lep}}^{\text{lep,top}}$ selection and the $E_T^{\text{miss}}$ variable in the $\text{SR}_{\text{had}}^{\text{lep,top}}$ selection.

Overview over the phase space for the signal, control and validation regions.

The $E_T^{\text{miss}}$ and $E_{T,\text{ll}}^{\text{miss}}$ distributions for the $t\bar{t}Z$ and $Z$ + jets CRs of the $tW_{0L}$ analysis channel.

The $E_T^{\text{miss}}$ distributions for the $t\bar{t}$ and $W$ + jets CRs of the $tW_{1L}$ analysis channel.

The $E_T^{\text{miss}}$ and $E_{T,\text{ll}}^{\text{miss}}$ distributions for the single top and $t\bar{t}Z$ CRs.

The $E_T^{\text{miss}}$ distributions for the $t\bar{t}$ and $V$ + jets VRs of the $tW_{0L}$ analysis channel.

The $E_T^{\text{miss}}$ distributions for the $\text{SR}_{\text{had}}^{\text{lep,top}} t\bar{t}$, $\text{SR}_{\text{had}}^{\text{lep,top}} t\bar{t}$, $W$ + jets VRs of the $tW_{1L}$ analysis channel as well as the single top VR.

Relative contributions of uncertainties from different sources in percent on the total background yield in each signal region of the two analysis channels.

The $E_T^{\text{miss}}$ and $E_{T,\text{ll}}^{\text{miss}}$ distributions for the $t\bar{t}$ and $Z$ + jets CRs of the $tW_{0L}$ analysis channel.

The $E_T^{\text{miss}}$ distributions for the $t\bar{t}$ and $W$ + jets CRs of the $tW_{1L}$ analysis channel.

The $E_T^{\text{miss}}$ and $E_{T,\text{ll}}^{\text{miss}}$ distributions for the single top and $t\bar{t}Z$ CRs.

The $E_T^{\text{miss}}$ distributions for the $t\bar{t}$ and $V$ + jets VRs of the $tW_{0L}$ analysis channel.
The $E_T^{\text{miss}}$ distributions for the $\text{SR}_{tW_{1L}}^{\text{had, top}} t\bar{t}$ the $\text{SR}_{tW_{1L}}^{\text{lep, top}} t\bar{t}$ and $W + \text{jets}$ VRs of the $tW_{1L}$ analysis channel 

58 The $E_T^{\text{miss}}$ distributions for the single top VR 

59 A summary of the comparison between data and MC predictions for all CRs and VRs in a one bin histogram after applying the fit results from the background-only fit 

60 A summary of the comparison between data and MC predictions for all signal region bins after applying the fit results from the background-only fit 

61 The expected and observed exclusion contours for the parameter plane $m_a - m_H$, the parameter plane $m_H - \tan\beta$ with $m_a = 250$ GeV and the parameter plane $m_H - \tan\beta$ with $m_a = 150$ GeV 

62 Previous results of the analysis for DM in association with a top quark and a $W$-boson with the combination of the channels with one lepton and two leptons in the final state 

63 The expected and observed 95% CL limits on the DM production cross-section relative to theory predictions for DM$t$, DM$t\bar{t}$ and combined processes
# List of Tables

1. Proportionality of the couplings to down-type quarks ($\zeta_d$), up-type quarks ($\zeta_u$) and leptons ($\zeta_l$) depending on the 2HDM type .......................... 26
2. Differences in the track definitions for the three selections .................. 63
3. Overview over the generators and precision used for the SM background samples ................................................................. 83
4. The preselection for the $tW_{0L}$ and $tW_{1L}$ channels of the analysis .... 108
5. The selection requirements for the $tW_{0L}$ channel .......................... 120
6. Expected yields before fitting the results for all SM backgrounds as well as one specific signal process in the $tW_{0L}$ channel bins .................. 121
7. The selection requirements for the $tW_{1L}$ channel split into the SR$_{tW_{1L}}^{lep,top}$ and SR$_{tW_{1L}}^{had,top}$ channel .................................................. 124
8. Yields for all SM backgrounds as well as one specific signal process in the SR$_{tW_{1L}}^{lep,top}$ channel and the SR$_{tW_{1L}}^{had,top}$ bins ........................................ 126
9. The definitions of the control regions for the $tW_{0L}$ channel .......... 130
10. The definitions of the control regions for the $tW_{1L}$ channel .......... 132
11. The definitions of the control regions for the $t\bar{t}Z$ and single top background ................................................................. 134
12. The definitions of the validation regions for the $tW_{0L}$ channel .... 137
13. The definitions of the validation regions for the $tW_{1L}$ channel as well as the VR for the single top process ........................................ 139
14 Normalisation parameters for the main backgrounds in the analysis
for the \( tW_{0L} \) and \( tW_{1L} \) channels . . . . . . . . . . . . . . . . . . . . . . . 149
15 Fitted event yields of the prediction for the different SM backgrounds
as well as the measured data in the signal regions of the \( tW_{0L} \) channel 156
16 Fitted event yields of the prediction for the different SM backgrounds
as well as the measured data in the signal regions of the \( tW_{1L} \) channel 156
17 Results of the model-independent fit . . . . . . . . . . . . . . . . . . . . . . . . . 157
Acknowledgements

Finally, the thesis is finished and I want to take the opportunity to thank everyone who helped and supported me over the last three and a half years. The first thanks obviously goes to Prof. Dr. Beate Heinemann and DESY and its employees in general who made it possible for me to write this thesis. I want to thank Ingrid Gregor for being a cool group leader of the ATLAS group at Hamburg DESY and who was always easy to talk to. Ingrid always made the meetings on Friday an interesting event even during the hard times of Covid. I also want to thank the luminosity group at ATLAS which helped me with my qualification task. I want to thank Valerie Lang who was the leader of the luminosity group but also Richard Hawkings for giving me opportunities to present my results at ICHEP 2021 and especially Witold Kozanecki who knew everything there was to know about luminosity measurement and who I could always ask if there was a problem to solve.

I want to further thank the other members of our analysis team which were working with me from Hamburg but also Zeuthen like Lars Rickard Strom who would always help out in our analysis when we needed him. I want to also thank the other PhD students like Marianna Liberatore, Martin Habedank, Judith Höfer and of course Ben Bruers who worked very closely with me on the analysis. Ben, it was very easy to work with you and it was always fun even if our code may have needed a bit of cleaning every once in a while. This easy and fun work was obviously only possible because we had very good supervisors as well. Alvaro Lopez Solis organised our analysis and while we sometimes might have needed a bit more communication, we
were able to come this far, because you kept it all together. Thank you for that. And this thanks of course also goes to our other two supervisors: Priscilla Pani who was the supervisor from Zeuthen and who was always there to ask the important questions to show what was still missing and of course Claudia Seitz who I want to thank very much for enduring me over the whole time. Thank you Claudia for helping me and more importantly for also pushing me to always keep up with the work in the analysis, the qualification task and everything else while also making sure to take vacations every once in a while. Thank you, because I could not imagine a better supervisor.