GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Fakultät für
Physik

Master's Thesis

Measurement of the $t\bar{t}+Z$ Cross Section in the Four Lepton Channel Using a Deep Neural Network

Messung des $t\bar{t}+Z$ Wirkungsquerschnitts im vier-leptonischen Kanal mithilfe eines neuronalen Netzes

prepared by

Tomke Schröer

from Hamburg

at the II. Physikalischen Institut

Abstract

The production of top-quark pairs in association with a Z boson offers the opportunity to study the coupling of the Z boson to the top quark. This coupling is of interest because it is sensitive to the third component of the weak isospin and the hypercharge of the top quark. A measurement of the coupling probes the prediction of these parameters of the Standard Model of elementary particle physics (SM).

This thesis presents studies of the final states containing four electrically charged leptons. This channel is particularly pure but such events are rare. In order to increase the statistical precision, a deep neural network (DNN) was trained to perform the classification of events. The preselection is chosen to be loose to increase the acceptance of four-lepton events. Then, the separation of background and signal events is performed by the DNN. This way the statistical uncertainty of the cross section measurement is reduced by 7% to 35% depending on the used configuration for the measurement compared to a "cut and count" approach.

Within this thesis also the Z boson reconstruction is performed by a DNN instead of using a simple Z mass window based approach with the aim to increase the reconstruction efficiency. The reconstruction is required in order to study the coupling.

Zusammenfassung

Die Produktion von Top-Quark Paaren in Assoziation mit einem Z-Boson bietet die Möglichkeit die Kopplung des Z-Bosons an das Top-Quark näher zu untersuchen. Das ist insbesondere interessant, da die Kopplung von der dritten Komponente des schwachen Isospins sowie der Hyperladung des Top-Quarks abhängt und so die Vorhersage der Parameter des Standardmodells der Elementarteilchenphysik (SM) mithilfe der Messung der Kopplung überprüft werden kann.

Diese Arbeit präsentiert Studien des tetraleptonischen Kanal, also die Endzustände mit vier geladene Leptonen. Dieser Kanal ist sehr rein, doch die statistische Unsicherheit in diesem Kanal ist groß. Zur Erhöhung der statistischen Präzision wurde die Klassifizierung durch ein tiefes neuronales Netz (DNN) vorgenommen, wobei die Akzeptanz von Ereignissen durch eine schwache Vorselektion zunächst erhöht wird, um die tatsächliche Klassifizierung von dem DNN durchführen zu lassen. Auf diese Weise kann die statistische Unsicherheit der Messung des Wirkungsquerschnitts um 7% bis 35% in Abhängigkeit der verwendeten Messkonfiguration im Vergleich zu einem klassischen Ansatz, welcher orthogonale Schnitte auf relevante Observablen verwendet, gesenkt werden.

In dieser Arbeit wird auch die Rekonstruktion des Z-Bosons mithilfe eines DNN durchgeführt, anstatt einen einfachen Ansatzes basierend auf der Masse des Z-Bosons zu verwenden, mit dem Ziel die Rekonstruktionseffizienz zu erhöhen. Die Rekonstruktion wird zur Untersuchung der Kopplung benötigt.

Contents

Contents

1. Introduction

The field of particle physics focuses on the description of the smallest pieces of matter: the elementary particles. The Standard Model of elementary particle physics (SM) constitutes the best description of the elementary particles and their interactions so far. Even though there are still observations that cannot be described by the SM such as dark matter, it is able to describe large parts of our universe on the smallest scales with high accuracy. With the help of measurements, parameters of the SM can be determined with increasing precision. This way predictions of the SM are constantly probed, phenomena beyond the SM can be discovered, and the SM extended.

This thesis focuses on the process of top-quark pair production in association with a Z boson ($t\bar{t}Z$) with a final state including four electrically charged leptons. Instead of using a "cut and count" approach, i.e., defining regions by applying cuts on relevant observables, a Deep Neural Network (DNN) is used to classify events into signal and background.

The $t\bar{t}Z$ process is especially interesting, because it offers the chance to measure the coupling between the Z boson and the top quark. This coupling is sensitive to the third component of the weak isospin I_3 and the hypercharge Y of the top quark, and is therefore useful to probe the SM's prediction of these quantities.

To perform such a measurement it is necessary to reconstruct the $t - Z$ system. Therefore it has to be determined which electrically charged leptons originate from the Z boson and which do not. To perform the assignment of the leptons to the Z boson, a second DNN is used instead of using an approach based on the invariant masses of the lepton pairs. A DNN is capable of taking more kinematic features into account and can use correlations between them. This way it does not follow a strict set of rules for the assignment like the simple approach does but it is more flexible to adjust its reconstruction based on specific event kinematics.

In the last step, the results from the classification are used to perform a fit to measure the signal strength of $t\bar{t}Z$ production, $\mu_{t\bar{t}Z}$, and the normalisation of ZZ background process, N_{ZZ}, based on the classifier's output. These fits are compared to those performed on the regions defined by a "cut and count" approach. The influence of the statistical

and systematic uncertainties are discussed.

In Section 1.1 the SM is introduced briefly. Afterwards the particle which plays the most important role in the context of this master's thesis, the top quark, is introduced in Section 1.2. The process of top-quark pair production is described in Section 1.3 and the same process in association with a Z boson in Section 1.4. The LHC and the ATLAS detector are described in Chapter 2. A general overview of DNNs is given in Section 3.1, followed by the event selection in Section 3.2 and a description of the tetralepton channel of $t\bar{t}Z$ in Section 3.3. The actual analysis of the presented events starts with the classification of signal and background events in Section 3.4 followed by the Z boson reconstruction in Section 3.5. The final part of the analysis is the fit of the inclusive cross section in Section 3.6. A summary and outlook is given in Chapter 4.

1.1. The Standard Model

The SM [1–11] includes 12 fermions, the gauge bosons and the Higgs boson. The fermions are spin-$^1/_2$ particles and can be further split into leptons and quarks. In addition, leptons and quarks can be arranged in three generations which comprise two particles each. To every fermion exists an antifermion which has the same mass but opposite charge. Here, charge refers to the quantum numbers associated with the generators of a symmetry group.

A summary of all particles is given in Figure 1.1. The generations of leptons each include an electrically neutral neutrino, which is massless according to the SM, and a charged lepton. The first generation consists of the electron e and the electron neutrino ν_e, the second one of the muon μ and the muon neutrino ν_μ, and the third one of the tau-lepton τ and the tau-neutrino ν_τ.

The fermions are described by spinors. The spinors can be decomposed into a left- and a right-handed component referring to the eigenstates of the chirality operators given by

$$P_L = \frac{1}{2}(1 - \gamma^5), \quad P_R = \frac{1}{2}(1 + \gamma^5). \tag{1.1}$$

The γ^5 matrix is defined as $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ with the Dirac matrices γ^μ, where $\mu = 0, \ldots, 3$. In the relativistic case, the helicity of a particle, i.e., the projection of the spin on the axis of the direction of translation, equals its chirality.

The left-handed leptons of each generation differ regarding their quantum number of the third component of the weak isospin. In the following, when using "weak isospin", the third component is implied. While neutrinos carry a weak isospin of $I_3 = +^1/_2$, the
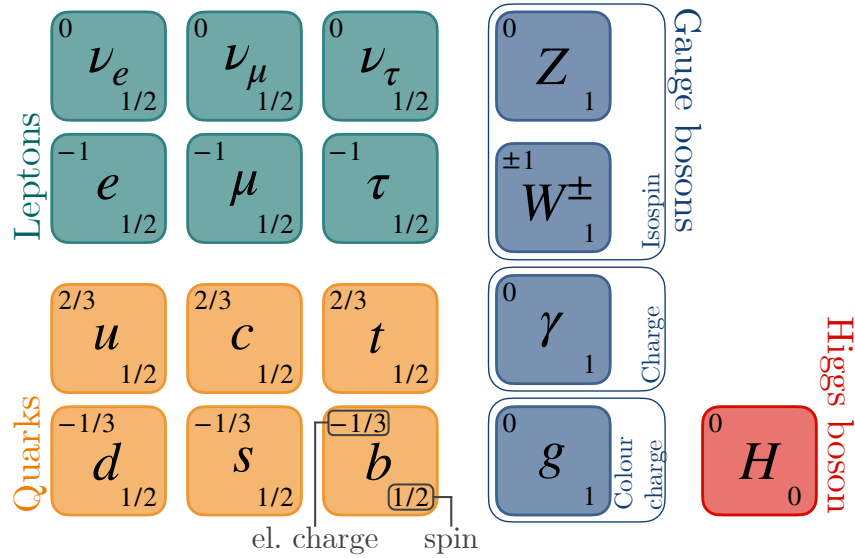
Figure 1.1.: The elementary particles of the Standard Model. The upper left part shows the leptons ordered by generations. The quarks are shown in the lower left and are ordered the same way. The bosons are split into the gauge bosons in blue with spin 1 and the Higgs boson in red with spin 0. The electric charge and spin are listed for all elementary particles. In addition, the quantum numbers the gauge bosons couple to are given.

electrically charged leptons with a charge of -1, given in fractions of the elementary charge e, have $I_3 = -1/2$. The same concept can be applied to the left-handed quarks. There, one distinguishes up-type quarks with an electric charge of $+2/3\,e$ and a weak isospin of $I_3 = +1/2$ and down-type quarks with an electric charge of $-1/3\,e$ and a weak isospin of $I_3 = -1/2$. The first generation comprises the up quark u and the down quark d, the second generation the charm quark c and the strange quark s, and the third one the top quark t and the bottom quark b. Right-handed fermions carry a weak isospin of $I_3 = 0$. This motivates the arrangement into doublets for the left-handed fermions and singlets

for the right-handed ones:

$$
\begin{aligned}
Q_L &= \begin{pmatrix} u_L \\ d_L \end{pmatrix} && \begin{pmatrix} c_L \\ s_L \end{pmatrix} && \begin{pmatrix} t_L \\ b_L \end{pmatrix} \\
u_R &= u_R && c_R && t_R \\
d_R &= d_R && s_R && b_R \\
L_L &= \begin{pmatrix} \nu_{e,L} \\ e_L \end{pmatrix} && \begin{pmatrix} \nu_{\mu,L} \\ \mu_L \end{pmatrix} && \begin{pmatrix} \nu_{\tau,L} \\ \tau_L \end{pmatrix} \\
\nu_R &= \nu_{e,R} && \nu_{\mu,R} && \nu_{\tau,R} \\
\ell_R &= e_R && \mu_R && \tau_R
\end{aligned}
\tag{1.2}
$$

Since the antiparticles always carry the opposite charge, right-handed antiparticles are arranged in doublets and left-handed antiparticles in singlets.

In addition to the particles, there are three elementary interactions included in the SM; the electromagnetic, the weak and the strong interaction. Gravity is not described by the SM.

The interactions are mediated by the gauge bosons. Gluons are the mediators of the strong interaction and photons of the electromagnetic interaction. The weak interaction is split into charged weak interactions, mediated by the electrically charged W^+ and W^- bosons, and neutral weak interactions, mediated by the electrically neutral Z boson. These interactions of the particles and gauge bosons within the framework of the SM are described by a quantum field theory with an underlying gauge symmetry of the form:

$$
SU(3)_C \times SU(2)_L \times U(1)_Y.
\tag{1.3}
$$

The $SU(3)_C$ describes the strong interaction. The gluons couple to the colour-charge C. Since gluons and quarks are the only particles carrying colour charge, gluons only interact with quarks or other gluons.

Leptons can only interact via the electroweak interactions which are described by the $SU(2)_L \times U(1)_Y$ gauge symmetry in the context of electroweak unification [1–3]. The L indicates the coupling to the left-handed fermions with the weak isospin I_3 as the corresponding quantum number. The weak isospin doublets are affected by the local gauge transformation of the $SU(2)_L$ while the singlets stay untouched. As a consequence, only the weak isospin doublets can participate in interactions fully described by the gauge symmetry of the $SU(2)_L$.

The quantum number associated with the $U(1)_Y$ is the hypercharge Y. It is defined

via the electric charge Q and the weak isospin of the fermion:

$$Q = I_3 + \frac{Y}{2}. \tag{1.4}$$

Interactions described by the $U(1)_Y$ also allow the interaction with weak isospin singlets. The strong interaction and the electroweak interactions are described in more detail in the following sections.

Even though the Higgs boson does not mediate any of the fundamental forces, it carries an important role in the SM. Massive particles acquire their mass through the interaction with the field of the Higgs boson. This procedure is described by the Brout-Englert-Higgs mechanism [12–15]. Direct mass terms of the electroweak gauge bosons would violate the gauge invariance of the underlying gauge group. Massive particles that preserve the invariance are therefore only possible through spontaneous symmetry breaking of the gauge symmetry of $SU(2)_L$.

1.1.1. The Strong Interaction

The strong interaction is described by the $SU(3)_C$ gauge group which couples to the colour charge C. There are three colour charges, red (r), blue (b), and green (g), and in addition three anti-colour charges ($\bar{r}$, $\bar{b}$, $\bar{g}$).

The term "colour charge" should only be understood as a label to describe the three orthogonal states of the $SU(3)$. The generators of the $SU(3)$ are the Gell-Mann matrices λ^α, where $\alpha = 1, \ldots, 8$. The corresponding current of the strong interaction is given by

$$j_q^\mu = \bar{u} c_j^\dagger \left[-\frac{1}{2} i g_S \lambda^\alpha \gamma^\mu \right] c_i u, \tag{1.5}$$

where u and $\bar{u}$ denote the quark wavefunctions and $c_{i/j}$ the colour wavefunction of the quarks. The eight physical gluons G_μ arising from the theory of quantum chromodynamics (QCD) [16, 17] are combinations of these orthogonal colour wavefunctions r, b and g:

$$G_\mu = r\bar{g}, \; g\bar{r}, \; r\bar{b}, \; b\bar{r}, \; g\bar{b}, \; b\bar{g}, \; \frac{1}{\sqrt{2}}(r\bar{r} - g\bar{g}), \; \frac{1}{\sqrt{6}}(r\bar{r} + g\bar{g} - b\bar{b}). \tag{1.6}$$

The strong interaction has an important aspect which makes it stand out compared to the others. While it is mostly experienced that forces like gravity decrease with increasing distance in daily life, the opposite is the case for the strong interaction. This behaviour is known as *confinement*: quarks cannot be observed as free particles but only as bound states in colour neutral hadrons. A hadron is colour neutral if it carries either three quarks

with a different (anti-)colour charge or two quarks where one quark carries a colour charge and the other one the corresponding anti-colour charge.

When two quarks are separated from each other, virtual gluons are exchanged between them and finally a new quark-antiquark pair is created. This process is repeated multiple times and multiple hadrons are created. This is known as *hadronisation.*

The strong coupling also depends on the energy range since it has a running coupling constant. In the low energy range, the coupling is stronger and then decreases with increasing energy. The coupling constant is approximately given by

$$\alpha(q^2) = \frac{\alpha(\mu^2)}{1 - \alpha(\mu^2)\frac{1}{3\pi}\ln\left(\frac{q^2}{\mu^2}\right)} \tag{1.7}$$

with the energy q^2 and a reference energy scale μ^2. This means in the high energy range the quarks are less bound through the strong interaction. This concept is called *asymptotic freedom.*

1.1.2. The Electroweak Interaction

The electroweak interaction was first described by Glashow [1], Weinberg [2] and Salam [3] with the GWS model. This model unifies the weak and electromagnetic interaction in the description as a $SU(2)_L \times U(1)_Y$ gauge group. The invariance under $SU(2)_L$ local gauge transformations is satisfied by the three gauge fields W_μ^k, $k = 1, \ldots, 3$. The physical fields of the charged $W^\pm$ bosons can be written as a linear combination of $W_\mu^{(1)}$ and $W_\mu^{(2)}$:

$$W_\mu^\pm = \frac{1}{\sqrt{2}}(W_\mu^{(1)} \mp W_\mu^{(2)}). \tag{1.8}$$

Since the $W^\pm$ bosons arise from the gauge invariance of the $SU(2)_L$, they only couple to left-handed fermions and right-handed antifermions. This means that the charged weak interaction mediated by the $W^\pm$ bosons is maximally parity violating.

While the lepton flavour is conserved when interacting via the $W^\pm$ bosons, this is not the case for quarks. The weak interaction mediated by the $W^\pm$ bosons is the only interaction changing the quark flavour. In addition it can change the generation of quarks, even though these processes are suppressed. This behaviour is described by the CKM matrix

which gives the rotation between the mass eigenstates and the weak eigenstates.

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{sb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix}. \tag{1.9}$$

Here, d', s', b' denote the weak eigenstates and d, s, b the mass eigenstates.

The coupling to the weak isospin spinor doublets given in Equation (1.2) leads to the following charged weak currents

$$j_-^\mu = \frac{g_W}{\sqrt{2}} \bar{L}_L \gamma^\mu \nu_L = \frac{g_W}{\sqrt{2}} \bar{L} \gamma^\mu \frac{1}{2}(1 - \gamma^5)\nu, \tag{1.10}$$

$$j_+^\mu = \frac{g_W}{\sqrt{2}} \bar{\nu}_L \gamma^\mu L_L = \frac{g_W}{\sqrt{2}} \bar{\nu} \gamma^\mu \frac{1}{2}(1 - \gamma^5)L, \tag{1.11}$$

$$j_-^\mu = \frac{g_W}{\sqrt{2}} V_{ij} \bar{u}_L^i \gamma^\mu d_L^j = \frac{g_W}{\sqrt{2}} V_{ij} \bar{u}^i \gamma^\mu \frac{1}{2}(1 - \gamma^5)d^j, \tag{1.12}$$

$$j_+^\mu = \frac{g_W}{\sqrt{2}} V_{ij}^* \bar{d}_L^j \gamma^\mu u_L^i = \frac{g_W}{\sqrt{2}} V_{ij}^* \bar{d}^j \gamma^\mu \frac{1}{2}(1 - \gamma^5)u^i, \tag{1.13}$$

where V_{ij} denotes the corresponding CKM matrix element and g_W the coupling strength. The gauge symmetry of the $U(1)_Y$ gives rise to the field B_μ which couples to the hypercharge Y of the particles. The mixing between $W_\mu^{(3)}$ and B_μ defines the physical fields A_μ and Z_μ which mediate the neutral electroweak interaction. A_μ refers to the photon γ and Z_μ to the neutral Z boson.

$$A_\mu = +B_\mu \cos\theta_W + W_\mu^{(3)} \sin\theta_W, \tag{1.14}$$

$$Z_\mu = -B_\mu \sin\theta_W + W_\mu^{(3)} \cos\theta_W \tag{1.15}$$

with the weak mixing angle θ_W. The photon couples to the electric charges of the particles and this way does not couple to neutrinos and itself since neutrinos and photons are electrically neutral. The corresponding current is given by

$$j_{\text{em.}}^\mu = g_e \bar{\Psi} \gamma^\mu \Psi. \tag{1.16}$$

with the Dirac spinor Ψ. The neutral weak current of the Z boson is given by

$$j_Z^\mu = \frac{1}{2} g_Z \bar{\Psi} \gamma^\mu (c_V - c_A \gamma^5)\Psi. \tag{1.17}$$

It couples to charged and neutral fermions. The current shows that the coupling has a $V - A$ structure which means that the neutral weak interaction has a vector and an

axial-vector component. Within the SM, the coupling parameters c_V and c_A depend on the charge and the third component of the weak isospin of the corresponding particle:

$$c_V = I_3 - 2Q \sin^2 \theta_W, \tag{1.18}$$

$$c_A = I_3. \tag{1.19}$$

One of the greatest achievements of Glashow, Weinberg and Salam was the unification of the weak and the electromagnetic interaction. The unification is described by the mixing angle θ_W, which describes the connection between the coupling constants of the electromagnetic and the weak interactions. The electromagnetic coupling constant g_e is given by $g_e = Qe$, where Q is the charge of the particle in units of the elementary charge. The connection of the coupling constants via the mixing angle θ_W is then given by:

$$e = g_W \sin \theta_W = g_Z \sin \theta_W \cos \theta_W. \tag{1.20}$$

In addition, the mixing angle connects the masses of the Z boson and the $W^\pm$ bosons via

$$m_W = m_Z \cos \theta_W. \tag{1.21}$$

1.2. The Top Quark

Even though the idea of a third generation of quarks was already brought up in 1973 by Kobayashi and Maskawa [18], it took more than twenty years to complete the third generation of quarks with the discovery of the top quark in 1995 by the CDF and DØ Collaborations at the TEVATRON [19, 20]. A first experimental indication for a third generation of quarks was the discovery of the τ lepton in 1975 [21] which extended the lepton sector in the SM by a third generation. This supported the idea of the existence of an additional quark generation. It took two more years until the first quark of the third generation was discovered: the bottom quark in 1977 [22]. The b-quark is the weak isospin partner of the top quark in the SM.

The top quark has some properties which make it a unique elementary particle. With a measured mass of $m_t = 173.34 \pm 0.27(\text{stat}) \pm 0.71(\text{syst})\,\text{GeV}$ [23] it is the heaviest elementary particle in the SM. Due to its high mass, the top-quark decay has a large phase space and this way decays before it hadronises. This means that there are no hadrons including top quarks but also that its properties can be studied directly from the decay products. All other quarks can only be observed as bound states in the form of hadrons which further decay into bundles of hadrons in the detector.
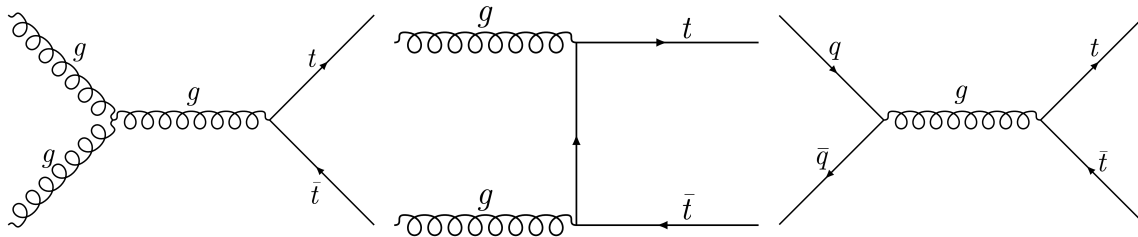
Figure 1.2.: There are two production modes of a top-quark pair. The left Feynman diagram and the one in the middle show gluon-gluon fusion, the right diagram illustrates quark-antiquark annihilation. At the LHC, where only protons and no anti-protons are brought to collision, gluon-gluon fusion is the dominant production mode.

Top quarks can be produced as single top quarks or as top-quark pairs. The studies of this thesis focus on the production of top-quark pairs in association with a Z boson. Hence, only pair production is described in detail.

1.3. Top-Quark Pair Production

A top-quark pair consists of a top quark and an antitop quark. There are two different production modes: gluon-gluon fusion and quark-antiquark annihilation which are shown in Figure 1.2. In proton-proton collisions at high energies the gluon-gluon fusion is the dominant production mode. This is because antiquarks are only available as sea quarks. The number of sea quarks carrying a small fraction of the total momentum is much smaller than the number of sea gluons. At high energies even a small fraction of the total energy is sufficient to produce a top-quark pair. Therefore, more gluons than quark-antiquark pairs that are capable of producing a top-quark pair are available. The LHC operates at energies of the order of magnitude of 1 TeV and proton-proton collisions are studied. Hence, about 90% of the top-quark pairs are produced via gluon-gluon fusion. The dominant decay process of a top quark is the decay into a W boson and a b-quark. Decays into other quarks are possible, but are highly suppressed by the off-diagonal entries of the CKM matrix. The top quark is the only quark that has a larger mass than the W boson. This means, while all other quarks can only decay into a virtual W boson, the W boson from the top quark is on-shell so that the process is not suppressed. The W boson either decays into two quarks or into a charged lepton and the corresponding neutrino. The decay into hadrons is called the hadronic decay and into leptons the leptonic decay.

Since two W bosons emerge from the decay of a top-quark pair, this leads to three different combinations. When both W bosons decay into quarks, this is called the all-hadronic decay, when both decay into leptons, the dilepton decay. A mixture of both

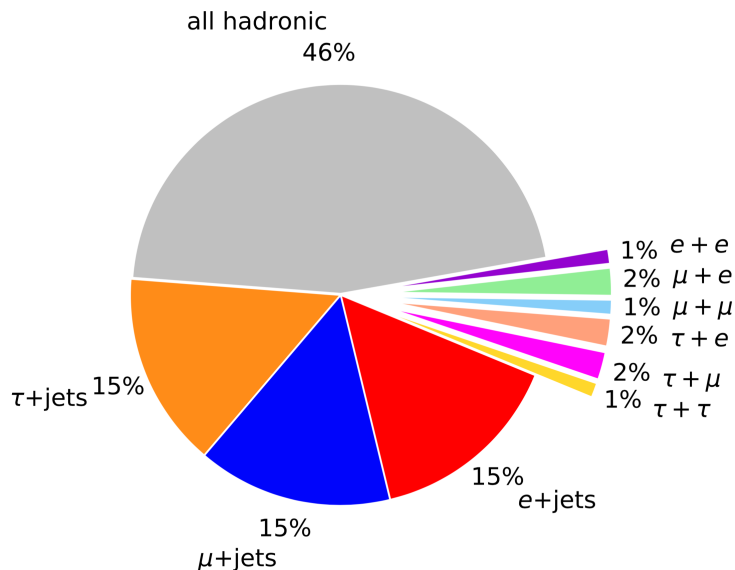***Figure 1.3.:*** The branching ratios of the decays of the W bosons in the decay of a top-quark pair.

decay modes is referred to as ℓ+jets. The branching ratios of the W boson decays are shown in Figure 1.3.

1.4. Top-Quark Pair Production in Association with a Z boson

During the production and decay of a top-quark pair, all quarks, leptons and W bosons can radiate a Z boson. Of most interest for this work are those events, where the Z boson couples directly to the top quark, because the coupling between the top quark and the Z boson is sensitive to the third component of the weak isospin I_3 and the hypercharge Y of the top quark. This can be seen in the following Lagrange density which can be obtained from the current given in Equation (1.17) with the vector and axial-vector coupling parameters c_V and c_A given in Equations (1.18) and (1.19):

$$\mathcal{L} \propto \bar{u}(p_t)[\gamma^{\mu}(c_V^t - c_A^t \gamma^5)]v(p_{\bar{t}})Z_{\mu} \tag{1.22}$$

with the spinors u and v of the top and antitop quark, respectively, and the field Z_{μ} describing the physical Z boson. The analysis of top-quark pair production in association with a Z boson ($t\bar{t}Z$) offers the opportunity to probe the SM prediction of this coupling. Contributions beyond the SM may change the strength and structure of the coupling
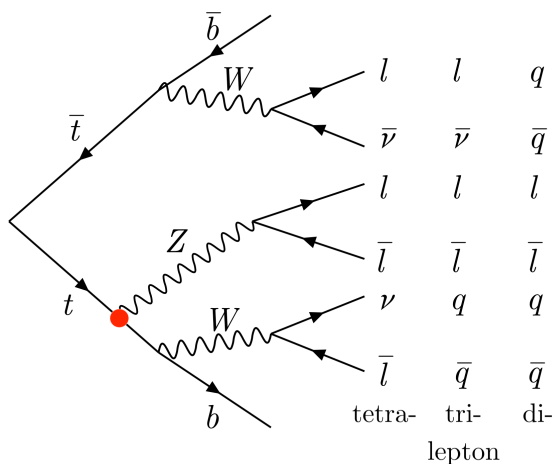
Figure 1.4.: A top-quark pair decays in most cases into a W boson and a b-quark. The additional Z boson decays into two leptons. Depending on the decay of the W bosons, the tetra-, tri- and dilepton channel is distinguished. The coupling of the top quark and the Z boson is marked with a red dot.

which would lead to deviations between the measurement and the SM prediction.

From an experimental point of view, usually only final state topologies, where the Z boson decays into two electrically charged leptons are considered. Decays into quarks or neutrinos are not taken into account. This leads to the three different topologies depending on the W boson decays which are shown in Figure 1.4. Depending on the number of electrically charged leptons, the decay channels are labelled as the dilepton, trilepton and tetralepton channels.

This thesis deals with the tetralepton channel. Since four-lepton final states are in general rare, the tetralepton channel benefits from a high purity. However, as a consequence, statistics in this channel are also low as shown in Figure 1.3. Only in 9% of the cases both W bosons decay leptonically. The trilepton and dilepton case carry similar branching ratios with 45% and 46%, respectively.

2. Experimental Setup

The experimental setup of the Large Hadron Collider (LHC), the ATLAS detector with its different components and the definition of the reconstructed objects is given in the following sections.

2.1. The Large Hadron Collider

The LHC is the most powerful hadron collider in the world and is nowadays capable of accelerating protons up to a beam energy of 6.5 TeV. It is located at the European Organisation for Nuclear Research (CERN, French: Conseil Européen pour la Recherche Nucléaire) in Geneva, Switzerland but also crosses the French border. The circular collider has a circumference of 26.7 km and is located approximately 100 m under the surface. This location provides a natural shielding from particles not originating from the collisions like cosmic radiation.

Before the collider ring was used for proton-proton collisions, the Large Electron-Positron Collider (LEP) was placed in the tunnel of the LHC and used for the collision of electrons and positrons.

Until now data at the LHC has been taken in two runs: Run 1 took place between 2010 and 2012 with protons at a centre-of-mass energy of 7 and 8 TeV. For Run 2 (2015-2018) the centre-of-mass energy was increased to 13 TeV. At four collision points detectors observe particles originating from high energetic collisions. The four experiments are called ATLAS [24], CMS [25], ALICE [26] and LHCb [27]. An overview of the CERN complex including the LHC and its preaccelerators is given in Figure 2.1. Before the accelerated hadrons are entering the LHC, they pass some smaller accelerators. This way they already enter with a beam energy of 450 GeV to be then accelerated even further until a centre-of-mass energy of 13 TeV is reached. After that, packages of particles, known as bunches, are brought to collision.

While LHCb and ALICE are specialised in the studies of *b*-quark physics and heavy-ion physics, respectively, the ATLAS and CMS experiments are multi-purpose experiments and can be used for a wide range of measurements. The largest detector at the LHC is the
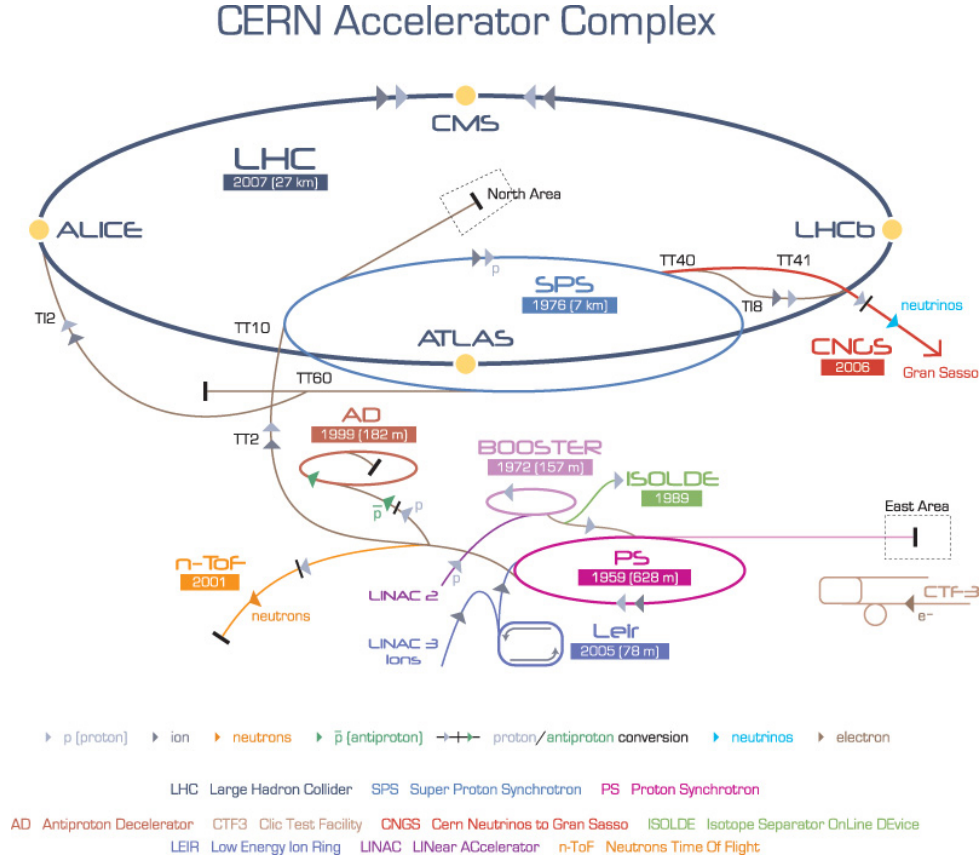
CERN Accelerator Complex

Figure 2.1.: A schematic overview of the CERN complex including the LHC and its preaccelerators. © CERN

ATLAS detector which is discussed in the following section.

2.2. The ATLAS Detector

The ATLAS detector has a diameter of 25 m, a length of 45 m and a weight of about 7000 tons. A schematic illustration of it can be seen in Figure 2.2.

It is arranged in a cylindric way around the beam pipe and can this way cover nearly the whole solid angle. This design motivates the position definition via angular variables. One of them is the pseudo rapidity η defined as

$$\eta = -\ln\left[\tan\frac{\theta}{2}\right] \tag{2.1}$$

with the angle θ between the particle track and the beam axis. For particles close to the beam line, the pseudorapidity η diverges and is zero for particle tracks orthogonal to the
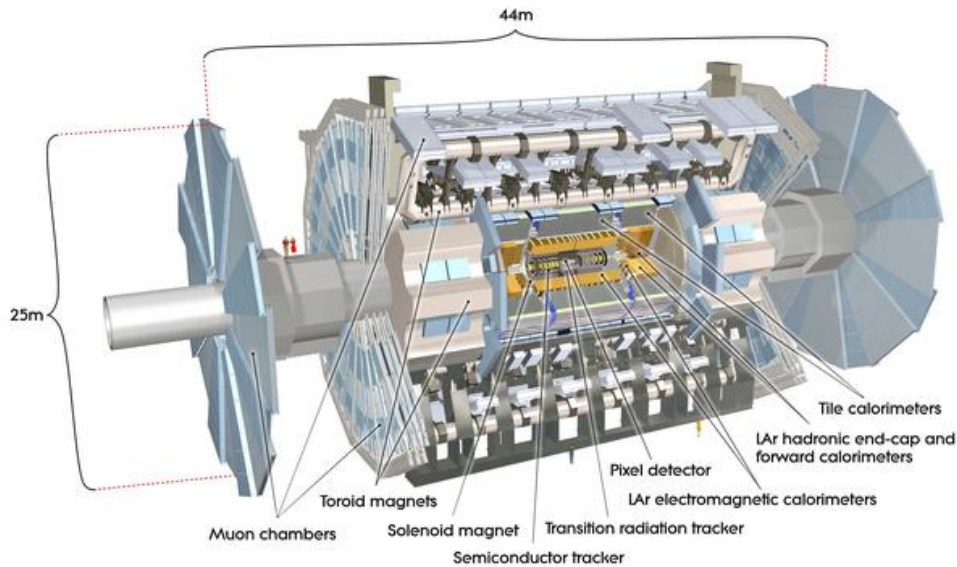
Figure 2.2.: A schematic illustration of the ATLAS detector showing its different layers used for the particle detection. To get an impression of the size of the detector, two people are drawn for comparison. © CERN

beam line. The azimuth angle ϕ is defined as the angle between the particle track and the axis pointing from the detector to the centre of the LHC tunnel. The ATLAS detector is symmetric in ϕ.

The particles produced in the collisions have to pass different layers of the detector which are optimised to measure individual properties of the particles sequentially. First, the particles pass the Inner Detector (ID) which consists of multiple layers and measures the direction, the momentum and the charge of the passing particles. The pixel detector is its first component. With its 92 million electronic readout channels, the pixels, it is capable of measuring the tracks of the passing particles that produce single hits. The second part is the semiconductor tracker which is a silicon microstrip tracker consisting of multiple modules that are arranged in four concentric barrels. The readout strips every 80 μm are used to provide a position measurement. The third component is the transition radiation tracker. It is based on the principle that radiation is emitted at the boundary of two media with a different refraction index. By measuring the deposited energy in the form of radiated photons, the tracker is, for example, used to differentiate between different particle types.

After that the particles enter the calorimeter system. It is designed to stop most particles so that their entire energy is deposited in the calorimeter. Two types of calorimeters are distinguished: the electromagnetic calorimeter, which measures the energy of electrons

and photons when interacting with matter, and the hadronic calorimeter, which does the same with hadrons when they interact with atomic nuclei.

In the electromagnetic calorimeter, high energetic electrically charged leptons and photons create electromagnetic showers. Through interaction with atomic nuclei, the photon converts into an electron-positron pair. This process is called pair production. These high energetic electrons and positrons radiate new photons via bremsstrahlung and the process is repeated. This way the particles create a cascade until the energy of the particles becomes too low to continue pair production and photon radiation.

The electromagnetic calorimeter used in ATLAS is a sampling calorimeter. This means the calorimeter consists of active layers that are used to measure the deposited energy and absorbing layers where the shower is created. As an active medium liquid argon is used, the absorbing layers are made from metal like tungsten, copper or lead.

The hadrons usually pass the electromagnetic calorimeter and enter the hadronic calorimeter. There, the hadrons can interact electromagnetically, but the dominant interaction is the strong interaction. This way a cascade of hadrons is created. It ends when the energy of the created hadrons is too low to create new hadrons via the strong interaction.

The hadronic calorimeter is designed as a sampling calorimeter, too, and comprises absorbing layers of steel and active plastic scintillating tiles. While electromagnetic calorimeters can also be designed as homogenous calorimeters, this is not of practical use for the hadronic calorimeters. In a homogenous calorimeter one material is used as the active and the absorbing material at the same time. Therefore, a compromise between both properties has to be made. The reason why only electromagnetic calorimeters are realised as homogenous calorimeters is the large difference between the radiation length X_0 for the electromagnetically interacting particles in the electromagnetic calorimeter and the nuclear interaction length λ for the strong interactions in the hadronic calorimeters. Due to the larger nuclear interaction length λ hadronic calorimeters need to be much larger than electromagnetic calorimeters and therefore homogenous hadronic calorimeters would simply need too much space.

Neutrinos cannot be stopped in the calorimeter system since they do not interact with matter. Also muons pass the calorimeters even though they interact with matter, because they are minimum ionising particles. They normally still leave a trace in the calorimeter.

The Magnet System bends the tracks of the particles and makes it possible to measure the charge and the momentum of the particles, which can be determined via the deflection radius. The system consists of the Central Solenoid Magnet and the Toroid Magnet, comprising the Barrel Toroid and the End-cap Toroid. The solenoid magnet has a thickness of only 4.5 cm to make sure that the interaction of the magnet with the studied particle is

kept as low as possible while providing the momentum measurement. The toroid magnet is especially used to determine the muons momentum.

The outermost part is the Muon Spectrometer. While all other particles besides muons and neutrinos are stopped in the calorimeter systems, muons are detected by the muon spectrometer since they are the only particles interacting with matter reaching it. The spectrometer consists of different components. The components leave some space in-between to provide a track measurement of the muons. They also measure the deflection radius of the muons in the magnetic field to measure the muon's charge and momentum.

2.3. Object Definitions

When talking about any particle in the context of an analysis, one normally refers to the reconstructed four-momentum of the particle based on the trace it leaves in the detector. For the reconstruction, there are criteria for each particle that have to be passed. The definition of the reconstructed particles used in the studies of this thesis is given in this section.

Electrons: electrons are identified via energy clusters in the calorimeter matched with one or multiple tracks in the ID. They need to have a transverse momentum of $p_T > 7\,\text{GeV}$, and $|\eta| < 2.47$ is required. If the pseudorapidity of the electron candidate lies in a range of $1.37 < |\eta| < 1.52$, it is discarded because this is the transition region between the endcap and the barrel of the electromagnetic calorimeter where the efficiencies of the measurement are decreased. The track of an electron candidate needs to have $|z_0 \sin\theta| < 0.5\,\text{mm}$ and $|d_0|/\sigma(d_0) < 5$. Here, z_0 describes the longitudinal impact parameter with respect to the reconstructed primary vertex, and d_0 is the transverse impact parameter with respect to the beam axis. The uncertainty of d_0 is denoted as $\sigma(d_0)$. These criteria make sure that the candidate originates from the primary vertex and this way suppresses contributions from electrons that are not part of the primary collision or candidates that fake electrons. For electrons the "MediumLH" identification criteria defined in Ref. [28] are applied.

Muons: muons are identified due to their tracks in the muon spectrometer and the calorimeter. These traces are matched with tracks in the ID that lie in a pseudorapidity range of $|\eta| < 2.5$. The requirement for the transverse momentum of $p_T > 7\,\text{GeV}$ is identical to the one used for electrons. Tracks associated with muons need to fulfil $|z_0 \sin\theta| < 0.5\,\text{mm}$ and $|d_0|/\sigma(d_0) < 3$ to ensure that the muon originates from the collision of the primary vertex. Muons have to pass the "Medium" identification requirements defined in Ref. [29].

Both, muons and electrons, need to fulfil isolation criteria to suppress the contribution

of fake leptons such as leptons originating from hadron decays. To do so, a cone is defined around the lepton candidates with a radius of $\Delta R = \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2}$. For electrons, the radius is determined as the minimum of $\Delta R = 10/p_{\mathrm{T}}$ and $\Delta R = 0.2$. Within this cone the transverse energy and the transverse momentum excluding the electron itself must be less than 6% of the transverse energy and transverse momentum of the electron candidate, respectively [28]. For muons, the cone radius is defined as the minimum of $\Delta R = 10/p_{\mathrm{T}}$ and $\Delta R = 0.3$ for a muon with $p_{\mathrm{T}} < 50\,\mathrm{GeV}$; for a transverse momentum larger than $50\,\mathrm{GeV}$, the cone radius is fixed to $\Delta R = 0.2$. The transverse momentum within this cone without the muon candidate plus $0.4\,E_T^{\mathrm{neflow}}$ must be smaller than 4.5% of the muon candidate's transverse momentum, where E_T^{neflow} is defined as the neutral particle-flow objects in a cone with $\Delta R = 0.2$ [29].

Jets: jets consist of bunches of hadrons that are created in the process of hadronisation. Since they comprise multiple particles, multiple tracks in the detector are included. The jet reconstruction is done by matching preselected tracks with one topological cluster and, based on a calculation of the expected energy, more clusters can be added to recover the full shower energy. The clustering is done with the anti-k_t algorithm [30] with a cone radius of $\Delta R = 0.4$. The jet candidates have to fulfil $p_{\mathrm{T}} > 25\,\mathrm{GeV}$ and $|\eta| < 2.5$. In addition, the EMPFlow jet vertex tagger (JVT) [31] is applied to jets with a $p_{\mathrm{T}} < 60\,\mathrm{GeV}$ to reduce the effects from multiple collisions within the same bunch crossing known as pile-up.

Jets including b-hadrons (b-jets): b-jets are identified using the machine-learning based DL1r high-level tagger [32]. There are different working points (WP) calibrated, which are defined as the efficiency to tag true b-jets. Calibrated WPs exist at 60%, 70%, 77% and 85% operating efficiency as measured in $t\bar{t}$ Monte-Carlo events.

3. Classification and Reconstruction of $t\bar{t}Z$ Events Using Deep Learning

This analysis comprises two parts: the classification of $t\bar{t}Z$ events and the Z boson reconstruction. The classification and reconstruction are done by two separate Deep Neural Networks (DNNs). Afterwards the results from the classification DNN is used for the measurement of the inclusive cross section of $t\bar{t}Z$.

The focus of this analysis is the tetralepton channel of $t\bar{t}Z$. Since this channel is particularly pure but has a limited number of signal events, the main aim of the classification is to keep the purity while increasing the statistical precision of the measurement by increasing the signal acceptance. Hence, only a loose preselection of events presented in Section 3.2 is applied. This way more events pass the selection and are used for the training of the DNN.

Afterwards, the signal region can be defined by using the output of the classification DNN. DNNs have the advantage that they do not simply apply orthogonal cuts on relevant observables but are able to take correlations between observables into account. With the help of the classification DNN as many signal events as possible, which may otherwise be lost when using a "cut and count" approach, should be kept. At the same time the background contribution is kept small such that the channel maintains its purity.

The Z boson reconstruction is performed by a DNN which is trained to assign the leptons originating from the Z boson to it. The aim is to outperform previously used strategies which simply used the invariant masses of all leptons for the assignment. This approach is described in more detail in Section 3.5. The Z boson reconstruction DNN may be able to increase the number of correctly assigned Z boson lepton pairs, because it does not only look at each invariant mass separately but can base its prediction on the event as a whole. This means it can either take other kinematics besides the invariant masses into account or it can correlate the invariant masses with each other and this way find certain schemes, e.g., regarding the distribution of invariant masses, which are useful
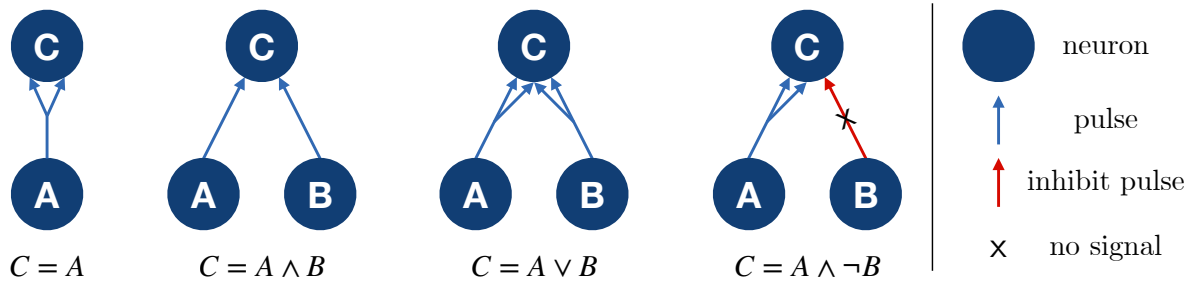
19

Figure 3.1.: Logical computations performed with neurons. The neuron C is active as soon as it receives at least two pulses simultaneously. The neurons A and B can either send none, one, two or an inhibit pulse. The inhibit pulse lowers the signal to be able to create the "not" computation.

for the assignment.

3.1. Deep Neural Networks

Neural networks are a powerful tool which, especially over the last years, gained more and more interest in science and our daily life. The inspiration was given by biological neurons. McCulloch and Pitts presented a simplified computational model of the functionality of biological neurons already in 1943 [33]. This simplified model, known as *artificial neuron*, has one or more binary inputs and one binary output. This way it is able to perform simple logical computations like "and", "or" and "not" as shown in Figure 3.1.

The simplest artificial neural network, meaning a network that can learn from its experience, is the perceptron invented in 1957 by Frank Rosenblatt [34]. It consists of one layer of linear threshold units (LTU). A single LTU is shown in Figure 3.2. These units take several weighted inputs, sum them up and output a step function. This way it is capable of performing simple binary classification. This idea was further enhanced by using multiple layers with modified units, which are explained in more detail in the following.

The first layer of a neural network is the input layer and the last one the output layer. All layers in between are called hidden layers. Neural networks with one input layer, at least two hidden layers and one output layer are called deep neural networks. The input layer is used to present the network all available information that it should base its prediction on. The output layer provides the network's prediction. In case of a classification DNN, the network calculates a probability for every case it should distinguish. The different cases are referred to as *classes*. In this thesis both, the classification and Z boson reconstruction, is based on separate probabilities for each class. For a binary classifica-
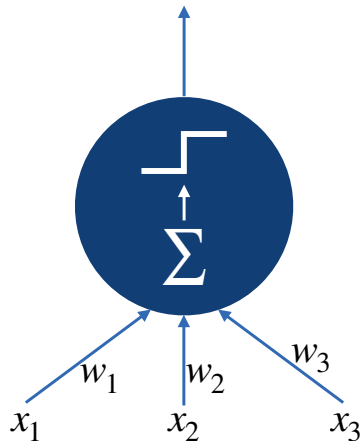
Figure 3.2.: A linear threshold unit sums the inputs x_i weighted with w_i. The output is given by a step function, i.e., is zero if the threshold of the step function is not passed and 1 if the weighted input is larger than the threshold. A perceptron consists of a layer of linear threshold units.

tion, as performed for the classification of signal and background events, one probability is calculated. Then, a probability close to one refers to class A and a probability close to zero to class B. For the Z boson reconstruction, multiple classes are used such that each class i has its own probability p_i, and the class k with $p_k = \max(\{p_i | i = 0, \ldots, N_{\text{classes}}\})$ with the number of classes N_{classes} is chosen.

The layers consist of *nodes*. A node is an instance where all inputs are summed up. All layers comprise multiple nodes except the output layer which can comprise only one node in the case of a binary classification. In general the number of nodes in the output layer depends on the number of classes the DNN should distinguish. The nodes of layer i are connected to nodes of layer $i - 1$ and $i + 1$. If they are connected to every node of the neighbouring layers, the network is called fully connected.

In addition, each layer has a bias node which is used as a linear shift by adding an input independent *bias term* between 0 and 1. It helps to keep a layer active when the input values are close to 0 so that the output of that layer would be close to 0 as well. The *bias node* is connected to every node of the following layer.

The connections between the nodes of neigbouring layers carry *weights*. The weights and the bias term are modified during the training of a DNN such that the networks prediction is improved.

For the training, labelled instances are used so that the true class of each instance is known. To be able to evaluate the modifications made during the training a metric to describe the network's performance is needed. This metric is the *loss function* $J(\vec{\theta})$ which

compares the DNN output with the true label. In the case of a binary classification as performed for the classification of signal and background events, the binary cross entropy $J(\vec{\theta})_{\text{bin}}$ is used given by

$$J(\vec{\theta})_{\text{bin}} = \frac{1}{m} \sum_{i=0}^{m} \left[y_i \log\left(\hat{y}_i\right) + (1 - y_i) \log\left(1 - \hat{y}_i\right) \right] \tag{3.1}$$

with the DNN prediction $\hat{y}_i$ and the label y_i of instance i in the training sample of size m and the vector of weights including the bias term $\vec{\theta}$. For a multi-class classification as used for the Z boson reconstruction the categorical cross entropy $J(\vec{\theta})_{\text{multi}}$ is used:

$$J(\vec{\theta})_{\text{multi}} = \frac{1}{m} \sum_{i=0}^{m} \sum_{j=0}^{N_{\text{classes}}} y_{i,j} \log\left(\hat{y}_{i,j}\right) \tag{3.2}$$

with the DNN prediction $\hat{y}_{i,j}$ and the label $y_{i,j}$ of instance i of class j and the number of classes N_{classes}.

The input of one node is given by the output of the nodes of the previous layer that it is connected to, but with the connection weight applied. The output of a node is calculated by the *activation function*. The Rectified Linear Unit function $ReLU(z)$ given in Equation (3.3) is used as the activation function for the hidden layers in this thesis. For the output layer the sigmoid function $\sigma_{\text{sig}}(z)$ given in Equation (3.4) is used for binary classification, for multi-class classification the softmax function $\sigma_{\text{soft}}(\vec{z})_i$ given in Equation (3.5) is used.

$$ReLU(z) = \max\left(0, z\right), \tag{3.3}$$

$$\sigma_{\text{sig}}(z) = \frac{1}{1 + \exp\left(-z\right)}, \tag{3.4}$$

$$\sigma_{\text{soft}}(\vec{z})_i = \frac{\exp\left(z_i\right)}{\sum_{j=0}^{N_{\text{inputs}}} \exp\left(z_j\right)} \tag{3.5}$$

with the total sum z of weighted inputs, the vector of weighted inputs $\vec{z}$ and the number of inputs N_{inputs}. z is used if the layer comprises only one node, e.g., for a binary classifier or when the outputs of each node in one layer are independent from each other as for the $ReLU$ activation function. The vector $\vec{z}$ is used if the outputs of the nodes are not independent as for the softmax function where the outputs add up to 1. The output of node i is then described by $\sigma_{\text{soft}}(\vec{z})_i$

The sigmoid activation function has the advantage that it does not diverge for large inputs but gives a value between 0 and 1, and this way can be easily used for binary classification. The softmax function is the counterpart for the multi-class classification.
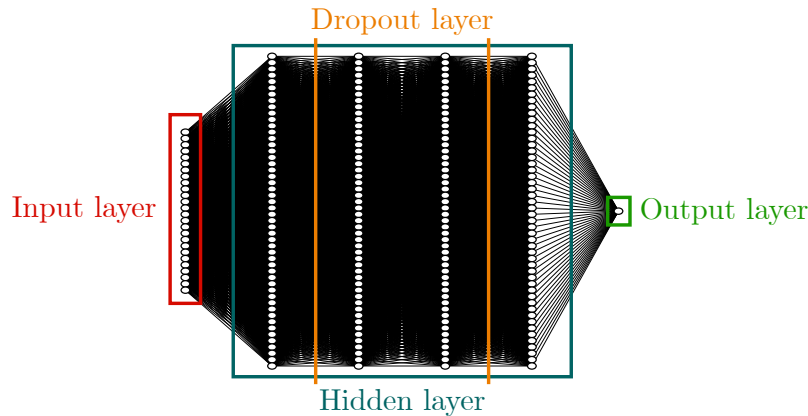
Figure 3.3.: Structure of a DNN with an input layer, multiple hidden layers and one output layer. During training, dropout layers can be used. The network is fully connected, i.e. each node is connected to every node of the two neighbouring layers.

Similar to the sigmoid function it provides a value between 0 and 1 but in addition it makes sure that the sum of all outputs is 1. Therefore the output of a layer using the softmax function can be interpreted as a probability distribution for the different classes.

During training, additional layers, the dropout layers, can be used to regularise the network to avoid that certain node connections gain a much higher impact than others [35]. In every training step every node in a dropout layer has a probability p to be dropped out, i.e., it does not receive any input and therefore does not produce any output. The connections to these nodes remain unchanged in this training step. In the next training step it has the same chance to be dropped out again. This procedure is only applied during the training, for the evaluation and application of the network it is no longer used.

Another regularisation technique to avoid vanishing and exploding gradients is batch normalisation [36]. The regularisation is done by zero centering and normalising and then scaling and shifting each input before or after the activation function is applied. To perform the zero centering and the normalisation the mean value and its standard deviation are calculated. To obtain the rescaled and normalised input vector, the mean is subtracted and the difference is divided by the standard deviation. Afterwards the input vector is scaled by an output scale parameter vector and shifted by the output shift parameter vector which contains the offset of every input of the layer. The scale and shift parameters are additional training parameters and therefore increase the complexity of the network which results in longer training times per epoch.

A schematic overview of an example structure of a DNN is given in Figure 3.3. The used framework for the architecture of the DNN is TENSORFLOW [37]. The optimiser used to

update the connection weights is the NADAM optimiser [38] which is the ADAM optimiser [39] but with incorporated Nesterov momentum [40] using the principle of stochastic gradient decent (SGD).

To get an unbiased sample of training instances, the instances are weighted in a way that every class is represented equally. To achieve this, the weight factor w_j^i of instance i of class j is modified as

$$w_j^i \longrightarrow \frac{w_j^i}{\sum_i^{N_j} w_j^i} \tag{3.6}$$

with the total number of events N_j in class j. This way it can be avoided that the DNN prefers one class due to the larger number of training instances and tends to classify new instances as this preferred class. The input variables x_i are scaled to a value between 0 and 1 like

$$x_i \longrightarrow \frac{x_i - \min(x)}{\max(x) - \min(x)}. \tag{3.7}$$

To validate the DNN, k-fold cross validation is used. This means that all events are split into k statistically independent subsets (folds) of equal size. Afterwards, $k-1$ folds are used for the training. The remaining one is used as a testing fold. From these $k-1$ folds, 20% of the events are set aside. They are only used as a validation subset. Both, the validation subset and the testing fold, are not used for the training of the DNN. The validation subset is used for an unbiased validation while the model is still in the process of training. The testing fold is used for the evaluation of the DNN after the training is finished. In total this means that the training is performed on a fraction of $(k-1)/k(100\% - 20\%)$ of the total number of samples. Every fold is once used as a testing fold such that the training procedure is repeated k times. This way k separate DNNs are trained. For the evaluation of the DNNs the events from each fold are classified by the DNN where the corresponding fold was the testing fold and therefore not used for the training. This way the evaluation is always performed on events the DNNs have not yet seen. For the actual application the classification is done by taking all DNNs into account and providing one complete prediction.

To construct a DNN that makes reliable predictions for different sets of training instances, overtraining should be avoided. Overtraining means that the DNN is trained in a way, that it provides precise predictions for the training instances, but is not able to generalise well so that the predictions for unseen data are significantly worse. There are different mechanisms to prevent overtraining. The general principle known as "early

stopping" is to stop the training before the DNN becomes sensitive to statistical fluctuations in the training set. In this case, the training stops if the improvement is less than a threshold Δ_{min} regarding the loss function evaluated for the validation subset for $n_{\text{epochs}}^{\text{earlyst.}}$ epochs. If this condition is not fulfilled after the maximal number of epochs $n_{\text{epochs}}^{\text{max}}$, the training stops, too.

3.2. Event Selection

The selected events for this analysis have to pass the following criteria. All events need to have a primary vertex and be part of the *good run list* which, for example, ensures that all detector parts were fully operational when the event was recorded. Exactly 4 electrically charged leptons are required, one or two of which must be matched with the single lepton or dilepton trigger, respectively. At least one opposite sign same flavour (OSSF) lepton pair needs to be part of the event. The leading and sub-leading leptons, i.e., the two leptons with the two highest transverse momenta, need to have p_{T} values of at least 17 GeV and 10 GeV, respectively. The remaining two leptons automatically have p_{T} values of at least 7 GeV since this is the minimal p_{T} required in the object definition of electrons and muons.

The number of jets is restricted to be between 2 and 4. More jets only occur with a small rate in $t\bar{t}Z$ final states. The upper bound is also motivated by the potentially worse modelling in regions with higher jet multiplicities. At least one b-jet passing the 85% WP of the DL1r high-level tagger is required. Also, the leptons assigned to the Z boson need to have an invariant mass $m_{\ell,\ell}^Z$ in a window around the Z mass of $m_Z \pm 20$ GeV with $m_Z = (91.1876 \pm 0.0021)$ GeV [41]. $m_{\ell,\ell}^Z$ is defined as the invariant mass of two leptons closest to the Z boson mass in the event. In addition the corresponding leptons need to have an opposite electric charge and the same flavour.

The event selection is chosen to be looser than the selection in the latest analysis performed by ATLAS [42], because the classification of events into signal and background events is performed by the classification DNN. The main differences are the wider window of ± 20 GeV instead of ± 10 GeV around m_Z in which the invariant mass $m_{\ell,\ell}^Z$ has to be, the looser lepton p_{T} cuts and the dropped $E_{\text{T}}^{\text{miss}}$ cuts, which were used to suppress the ZZ background introduced in the next section. Another difference is the cut on the number of jets. At least two jets were required but no upper restriction was used. Therefore this cut is looser than in the presented selection used for the classification DNN.

For the DNN the number of unweighted Monte-Carlo (MC) events is important, because this determines the size of the training set. In Table 3.1 the signal and background

Table 3.1.: Number of unweighted MC events for the $t\bar{t}Z$ and background samples for the classification DNN and the $t\bar{t}Z$ samples for the Z boson reconstruction DNN.

Sample	Number of unweighted MC events
$t\bar{t}Z$ (class. DNN)	63569
background (class. DNN)	197023
$t\bar{t}Z$ (reco. DNN)	64195

events with the presented selection applied are listed. These are the training samples for the classification DNN. The $t\bar{t}Z$ events with the selection used in the previous ATLAS analysis applied are shown, too, because these are the training events for the Z boson reconstruction DNN. The shown numbers include all events, i.e., also those events which are used as a validation subset or as the testing fold. The number of events used for the actual training procedure is therefore $(k-1)/k(100\% - 80\%)$ of the numbers given in Table 3.1 with the number of folds k as described in detail in the previous section.

Table 3.1 shows, that the number of unweighted $t\bar{t}Z$ events is smaller for the presented selection than for the tight selection used in the latest ATLAS analysis. This is caused by the tighter cut on the number of jets. Nevertheless, Figure 3.8 shows that the weighted number of events is still increased by applying the presented selection. Since this selection is used to open up the acceptance for the weighted events, it is still referred to as "loose selection" even if the acceptance is not increased for the unweighted MC events.

3.3. The Tetralepton Channel

In the tetralepton channel, both W bosons from the top quarks decay into electrically charged leptons and the corresponding neutrinos. The Z boson decays into an electrically charged lepton pair of the same flavour with opposite electric charge. If the Z boson decays into τ leptons, only the final state electrons and muons will be reconstructed. Then the leptons originating from the τ lepton decays do not need to have the same flavour. If the τ lepton decays into hadrons, the event is not part of the tetralepton channel, because only final state leptons are taken into account. Since the fraction of events where the Z boson decays into two τ leptons passing the preselection is of the order of magnitude of a few per mille, these events are not treated differently in both DNNs.

The main background processes are the production of a top quark and a W boson in association with a Z boson (tWZ) and diboson production with two Z bosons (ZZ)
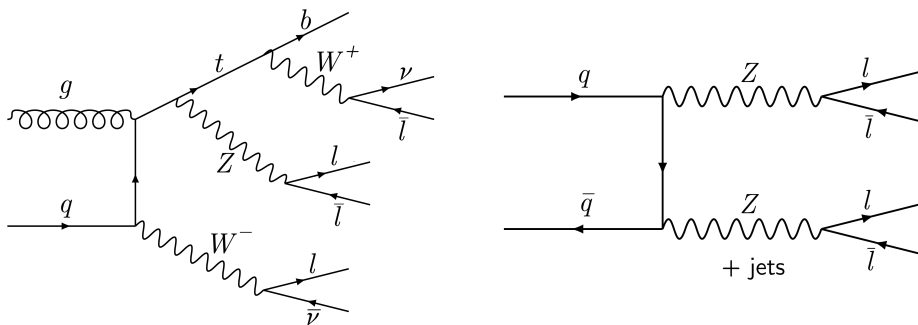
Figure 3.4.: The main background processes to $t\bar{t}Z$. On the left the production of a top quark and a W boson in association with a Z boson (tWZ) and on the right diboson production with two Z bosons (ZZ) and additional jets.

including additional jets. Here, one distinguishes light jets, c-jets and b-jets. While c-jets originate from c quarks, light jets originate from any quark besides c and b quarks or gluons. In this analysis especially b-jets are needed so that ZZ and $t\bar{t}Z$ have similar final states. Both background processes are shown in Figure 3.4. Besides ZZ and tWZ, top-quark pairs in association with a Higgs boson ($t\bar{t}H$) and events with fake leptons are part of the background. Fake electrons are, for example, additional photon emissions that shower in the electromagnetic calorimeter and are mistaken for electrons. Muons can be faked by charged hadrons whose lifetime is long enough to reach the muon spectrometer. Fake events are, for example, $t\bar{t}Z$ events of the trilepton channel with an additional fake lepton, so that these events pass the selection of the tetralepton channel. Other background processes with lower impact like diboson production with two W bosons in association with a Z boson are grouped together as "Others".

In the tetralepton channel different regions are distinguished. The differentiation is made based on the flavour of the leptons of the W bosons. If both leptons have the same flavour, the region is called the same flavour (SF) region. If they have a different flavour, the event is part of the different flavour (DF) region. This differentiation is made due to the different background compositions: ZZ is the main background in the SF region while it is negligible in the DF region, because in most cases ZZ decays into two OSSF lepton pairs. The only cases where the ZZ final state has exactly one OSSF lepton pair is where one Z boson decays into two τ leptons which decay into different flavour leptons or where one lepton flavour is incorrectly reconstructed.

Within the SF and DF regions, events can be discriminated further based on the electron and muon multiplicities. In the SF region, events with 4 electrons ($eeee$), 4 muons ($\mu\mu\mu\mu$) or two of each flavour ($ee\mu\mu$) can be distinguished. In the DF region, three electrons and one muon ($eee\mu$) or one electron and three muons ($e\mu\mu\mu$) are possible. A motivation

to subdivide into these regions is that electrons and muons have different resolutions regarding the reconstruction of their kinematic properties. For example the p_T of a muon in the high p_T range can be measured more precisely than for electrons [28, 43].

3.4. Classification of Signal and Background Events

The aim of the first part of this analysis is the classification of events into signal and background events. To do so, a DNN is defined and input observables are determined. Afterwards the network is trained and its performance evaluated. For the evaluation different performance metrics are used.

3.4.1. Classifier Architecture

Before the training the DNN architecture has to be defined. The network is chosen to have 4 hidden layers with 50 nodes per layer and a 4-fold cross validation is performed. This compact DNN structure was chosen to avoid overtraining from the beginning. Such a DNN has a stable performance which means that it provides reliable predictions for different sets of data samples. A hyperparameter optimisation to find the best DNN structure was not performed, since this is beyond the scope of this thesis. To regularise the network, dropout layers after the first and third hidden layer are used during the training process.

The second part of the DNN architecture is the choice of input observables. Based on these observables, the network should become sensitive to the differences between signal and background events.

To be able to reconstruct the kinematics of the event, multiple kinematic observables are used as input like the transverse momenta of the jets and leptons and the ΔR between the leptons and the b-jets.

Some observables used already require the reconstruction of the Z boson. At this point the two parts of the analysis overlap. During the development of the classification DNN, the leptons with an opposite electric charge, the same flavour and an invariant mass closest to the Z mass m_Z are assigned to the Z boson and are called "Z leptons" in the following. The remaining leptons are referred to as "non-Z leptons".

The observables used in the network using this Z boson reconstruction are the $\Delta\phi$ between the two Z leptons, the $\Delta\phi$ between the two non-Z leptons, the transverse momentum of the reconstructed Z boson as well as the transverse momentum and the invariant mass of the system of the two non-Z leptons.
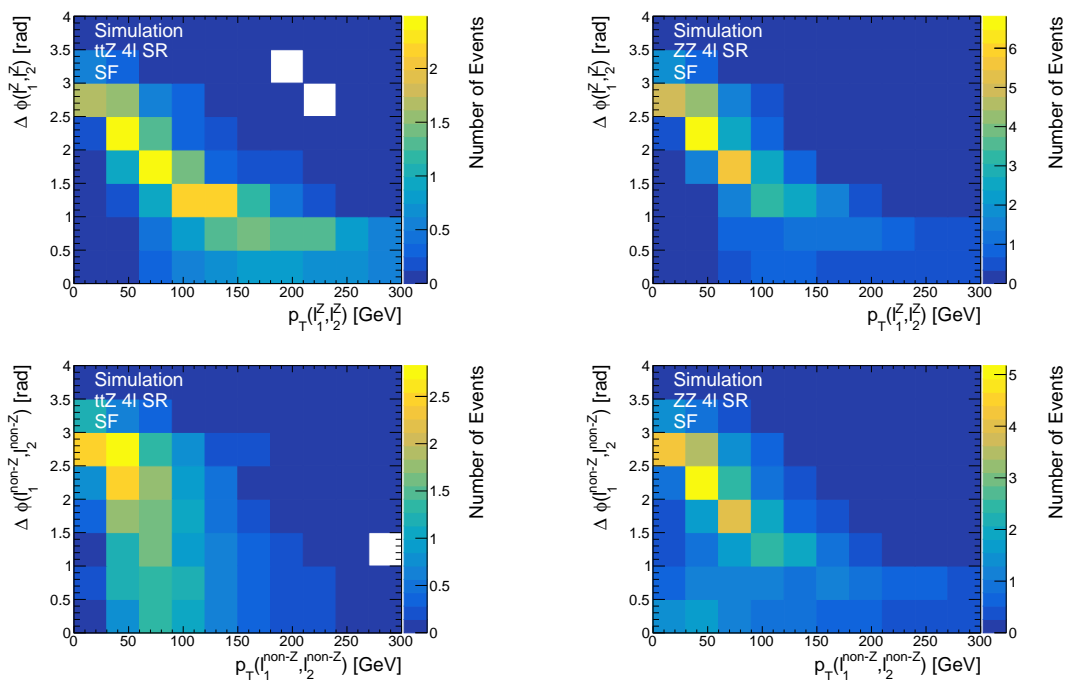
Figure 3.5.: The distributions of the number of events as a function of $\Delta\phi$ and p_T of the Z leptons in the upper plots and the non-Z leptons in the lower plots for ZZ and $t\bar{t}Z$. For ZZ the distributions look similar, because the leptons labelled as "non-Z leptons" actually originate from a Z boson, too. This way these observables can help to distinguish ZZ from the signal $t\bar{t}Z$.

These observables help to distinguish ZZ events from other events, because the major difference between ZZ and the other two main contributions, tWZ and the signal $t\bar{t}Z$, is that for ZZ the two leptons labelled as "non-Z leptons" originate from a Z boson, too. For tWZ they originate from the W boson decay and the top-quark decay, and for $t\bar{t}Z$ from the two top-quark decays. Thus, leptons from the ZZ decay labelled as Z leptons should have similar distributions of kinematic observables as those labelled as non-Z leptons. For tWZ and $t\bar{t}Z$ these distributions can look different. This is especially important for the SF region where ZZ is the dominant background.

To demonstrate the differences, the distributions of the number of events as a function of $\Delta\phi$ and the transverse momenta of the systems of the (non-)Z leptons are shown in Figure 3.5 for ZZ and $t\bar{t}Z$ events in the SF region. While the distributions of the Z and non-Z observables have a similar shape for ZZ, they differ from each other for $t\bar{t}Z$. This way the $\Delta\phi$ observables in combination with the transverse momenta help to distinguish ZZ from $t\bar{t}Z$.

Another difference between all three main processes is the number of b-jets. In $t\bar{t}Z$ two

29

b-jets are expected, one in tWZ and none in ZZ. In ZZ events at least one jet being mistagged as a *b*-jet or additional gluon emissions producing *b*-jets is necessary, since one *b*-jet passing the 85% WP is already required in the preselection of events. Hence, the pseudo-continuous *b*-tagging WP variable for the second jet is used as well. The ordering of the jets is done by the WP they pass, so that the pseudo-continuous *b*-tagging WP variable for the second jet returns the WP of the jet passing the second highest WP within this event. Since $t\bar{t}Z$ always includes two *b*-jets, it is expected that the second jet passes on average a higher working point than those of the background events, which do not necessarily include two *b*-jets.

Due to the W boson decays in $t\bar{t}Z$ and tWZ events, larger $E_{\mathrm{T}}^{\mathrm{miss}}$ is expected for in the final states of both processes compared to ZZ final states which motivates $E_{\mathrm{T}}^{\mathrm{miss}}$ as an additional input variable.

The last observable used is the jet multiplicity. Since different numbers of *b*-jets are expected for each process, also the number of jets in general are different. In combination with the assumption that in ZZ events jets are partly mistagged as *b*-jets, so that there is actually an additional jet from a gluon emission, this observable helps to separate signal and background events, too.

Different approaches are tested. They are based on the differences in SF regions and DF regions and the differences between the lepton flavour configurations. Depending on the exact region definitions, lepton flavour information is an additional input observable. In total three approaches are used to perform the classification.

- Training of one DNN including all events, no splitting into regions defined by the lepton flavour but using lepton flavour information as input observables,

- Training of two DNNs, one for the SF region and one for the DF region to take different background compositions into account and using lepton flavour information as input observables,

- Training of five DNNs, one for each lepton flavour composition to take differences in the reconstruction of the leptons into account, lepton flavour information not used as input observables.

The results are discussed in the following section.

3.4.2. Classifier Performance

The training of all DNNs for all approaches presented in the previous section was performed. The results of the evaluation of the DNNs for the SF region, the DF region and
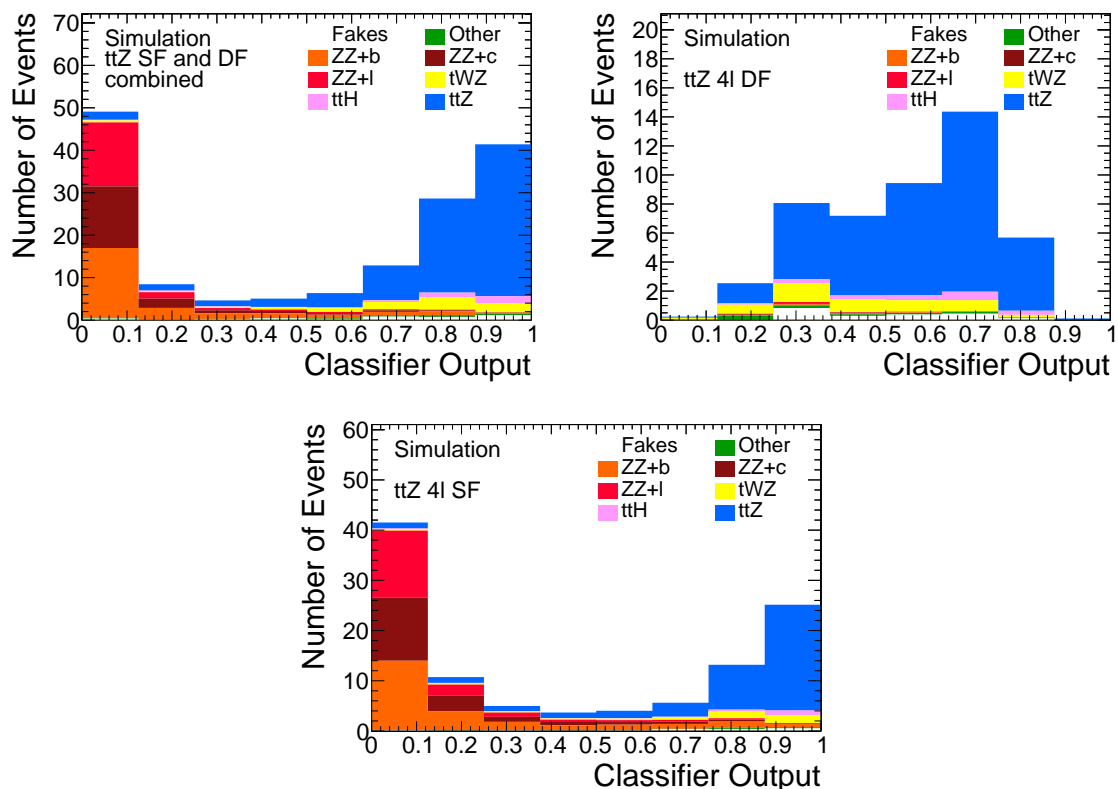
Figure 3.6.: The classifier outputs of the DNNs for the SF and DF regions and the DNN trained on the combination of both regions.

the combination of both, referred to as "combined region", are shown in Figure 3.6. These plots provide evidence that the ZZ contribution can be separated from the signal in the SF and combined region. In both regions, the signal $t\bar{t}Z$ peaks at a DNN output near 1 and decreases with decreasing DNN output. For the ZZ contribution the opposite is true. It peaks near 0 and decreases with increasing DNN output. For the DF region, the signal peak is at a DNN output around 0.7. The separation of signal and background events is not as clearly visible as for the other regions, but the DF region also has a smaller background contamination.

The distributions of the remaining DNNs for the regions defined via their electron and muon multiplicities can be found in Figure A.1. In general the distributions of the regions that are part of the SF and DF regions follow the shape of the distributions of the SF and DF regions, respectively.

To make a quantitative comparison of the performance of the DNNs, quantities that evaluate the DNN's performance are needed. One of those is the ratio of the total number of signal events and (the square root of) the total number of background events S/B ($S/\sqrt{B}$).
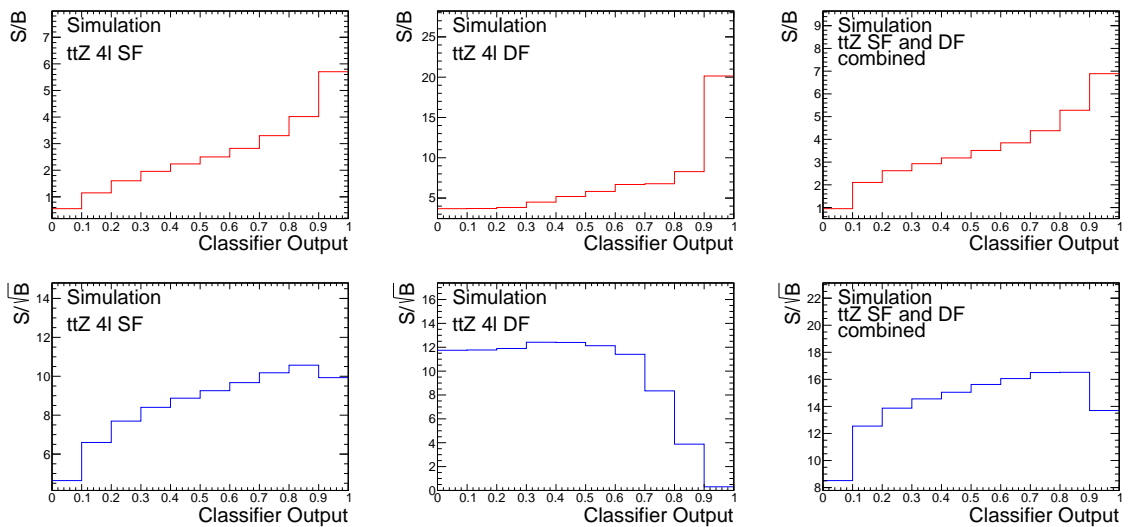
31

Figure 3.7.: The S/B and $S/\sqrt{B}$ ratios depending on the cut applied on the classifier output. All events that would pass the selection if a cut on the given classifier output was applied are used for the calculation of the ratios.

Figure 3.7 shows the ratios depending on a possible cut on the DNN output at a certain threshold DO_{cut}. That is, all events that have a higher DNN output than DO_{cut} are taken into account, and afterwards the ratios are calculated. Both metrics describing the same region need to be taken into account simultaneously. While S/B provides the degree of purity, $S/\sqrt{B}$ is an indication for the statistical significance in this region. When choosing a threshold for the region definition, the best compromise between both ratios must be found.

All S/B distributions increase with increasing DNN output while the $S/\sqrt{B}$ distributions first increase and, at a certain point, decrease again. This is expected, because signal events have on average a higher DNN output so that the purity increases with increasing DNN output. At some point the number of signal events is reduced by applying a cut on the DNN output while only few additional backgrounds events are cut away. This means the statistical significance is decreased which leads to a decreased $S/\sqrt{B}$.

The distributions for the regions defined via their electron and muon multiplicity can be found in Figure A.2. Their shapes are again similar when comparing them to the distribution of the SF and DF regions, but their performance is in general worse. For this reason the focus will lie on the approaches using one DNN for all events or two DNNs, one for the SF and one for the DF region.

To determine a first useful cut parameter DO_{cut} for the final region definition, the best compromise of both quantities can be chosen by eye. Limited by the binning, only a range

Table 3.2.: Comparison of the S/B and $S/\sqrt{B}$ ratios of the approaches using a DNN and using orthogonal cuts on relevant observables, referred to as "cut and count" approach.

Region	S/B (cut and count)	S/B (DNN) (+ improvement)	$S/\sqrt{B}$ (cut and count)	$S/\sqrt{B}$ (DNN) (+ improvement)
SF and DF combined	3.04	5.28 (+ 95%)	14.7	16.5 (+13%)
SF	2.77	5.70 (+106%)	9.7	9.9 (+ 3%)
DF	3.34	6.68 (+100%)	11.1	11.4 (+ 2%)

of the size of one bin can be chosen as a possible cut threshold.

For the combined region, the best compromise between both quantities is given by a cut between 0.8 and 0.9. For the SF region a cut between 0.9 and 1 and for the DF region between 0.6 and 0.7 is applied.

To evaluate the performance improvement in comparison to the "cut and count" approach used in the previous ATLAS analysis of the tetralepton channel, Table 3.2 compares the S/B and $S/\sqrt{B}$ of the different approaches. The same cuts as given in Ref. [42] were applied and the ratios calculated. Table 3.2 shows that the selected cut values DO_{cut} define regions with improved S/B for all DNNs in comparison to a "cut and count" approach. The same holds for $S/\sqrt{B}$, but compared to the combined region with an improvement of 13%, the improvement in the SF region and DF region is much smaller with 3% and 2%, respectively. In general, Figure 3.7 shows that independent of the applied cut the $S/\sqrt{B}$ of the SF and DF region is never significantly better than the "cut and count" approach. These results depend on the chosen binning for the S/B and $S/\sqrt{B}$ distributions. Therefore they can only give a general overview of the networks performance, the concrete values differ for different binning configurations.

For the DF region this result can be partly explained by the small fraction of signal events which is gained when applying the looser selection. Figure 3.8 shows the increase in signal and background events for both regions.

In the DF region the number of signal events is only increased by 0.8% when using the presented selection in comparison to the tighter selection for the "cut and count" approach. Hence, almost any cut on the DNN output of the DF region DNN results in a region with fewer signal events than the region resulting from a "cut and count" approach. In addition, the background contribution in the DF region does not accumulate in the lower DNN output region but is a flat distribution. Hence, the statistical significance cannot be improved by reducing the background contribution significantly either.
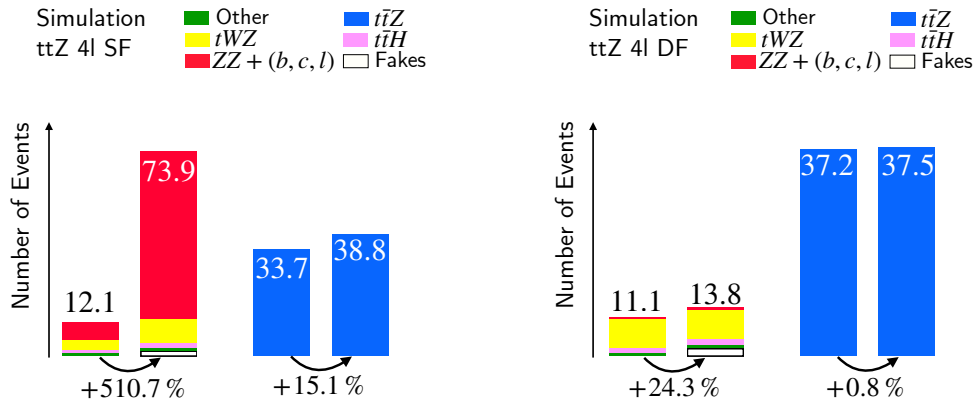
Figure 3.8.: Number and percentages of gained signal and background events when applying a looser selection in comparison to the previously performed analysis in the tetralepton channel of $t\bar{t}Z$ [42].

For the SF region the low increase can be explained by the larger number of background events. In the DF region the number of background events is only increased by 24.3%. In the SF region an increase of background events of 510.7% in comparison to an increase of signal events of 15.1% is observed. Therefore, many more background events have to be separated from the signal events to increase the statistical significance.

To evaluate the general performance of the DNN without choosing a specific cut parameter, the separation power $\mathcal{S}$ defined as

$$\mathcal{S} = \frac{1}{2} \sum_{i}^{N} \frac{(s_i - b_i)^2}{s_i + b_i} \tag{3.8}$$

with the total number of bins N, the background contribution b_i and the signal contribution s_i in bin i is used. The prefactor $^1\!/_2$ makes sure that, if all signal events lie completely in one bin and all background events lie completely in another one, the separation equals 1, or 100%.

Besides their performance in terms of the separation, it is also important that the DNNs are able to generalise. A suitable parameter to evaluate this is the "area under curve" (AUC) for the training and testing folds. This is the integral of the ROC curves which are defined as the true positive rate against the false positive rate. If the AUCs of the training samples and the testing samples are similar, the DNN does not overtrain. Fluctuations of the AUC indicate instability of the training and are therefore a hint for undertraining. Undertraining can occur due to a lack of training instances, if the training is stopped too early or when the model is too complex. The fluctuation is given by the standard deviation σ of the four folds.

Table 3.3.: Evaluation of the performance of the DNNs im terms of their separation power $\mathcal{S}$ and the AUC of the testing and training folds.

Region	separation power $\mathcal{S}$	AUC (training)	AUC (testing)
SF and DF combined	38.4%	0.760 ± 0.007	0.754 ± 0.022
SF	36.9%	0.820 ± 0.009	0.811 ± 0.018
DF	8.3%	0.544 ± 0.011	0.537 ± 0.009

The separation powers and means of the AUCs are summarised in Table 3.3. The same is done in Table A.1 for the remaining regions.

Both metrics lead to the same conclusion: the DNN for the DF region has the weakest performance of all networks. Nevertheless, Table 3.3 shows that the network generalises well. Both AUC values are close to each other and also the standard deviations are small.

In total the evaluation shows that the DNN approach is useful to increase the purity and the statistical significance in the combined region. The definition of a SF region and a DF region is useful, too, because regarding the $^S/_B$ ratio the improvement is in the same magnitude for all regions with around 100% improvement. There, the SF region shows most improvement. Also from a machine learning perspective the DNN for the SF region performs best in terms of the AUC and similar to the DNN for the combined region in terms of the separation power. The DNN for the DF region performs significantly worse.

The SF and DF regions suffer from a lower number of training events and a different signal and background composition compared to the combined region. While the DF region still maintains partly its degree of purity, it still cannot increase the number of signal events which leads to difficulties decreasing the statistical uncertainty. The SF region gains more signal events but also a large fraction of background events that needs to be separated to decrease the statistical uncertainty. The combined region benefits from the advantages of both regions: the number of signal events can be increased due to the SF region while the increased background contribution has a smaller impact due to the additional signal events from the DF region.

For the DNNs of the regions defined by the lepton multiplicities no further improvement is observed.

3.5. *Z* Boson Reconstruction

The second part of the analysis is the reconstruction of the Z boson. The classical approach determines the Z leptons by assigning the two leptons with identical lepton flavour and an opposite electric charge whose invariant mass is closest to the Z boson mass m_Z to the Z boson. This approach was used in the previous ATLAS analysis [42].

The advantage of using a DNN instead of the simple approach above is that a DNN can take more kinematic features and correlations between them into account. In addition, it may be useful to distinguish between the different lepton flavour configurations introduced in Section 3.4 for the following reason: the leptons originating from the Z boson need to have the same flavour. For the $ee\mu\mu$ region this means that there are only 2 possibilities to assign the leptons to the Z boson while there are 6 possibilities to choose from for the regions with 4 same-flavour leptons. This makes the reconstruction more challenging. Due to the different efficiencies in the lepton reconstruction for electrons and muons, see Ref. [28, 29], the electron's electric charge is more often misreconstructed in the high p_T range. This can decrease the efficiency of the assignment since the electric charge is used for the Z boson reconstruction. This has not been taken into account so far. By providing the lepton flavour information and electric charge as input observables the DNN may be able to learn these differences.

Before starting the training of a Z boson reconstruction DNN, the previously used approach was studied and evaluated to be able to compare both approaches. Since this approach predominantly focusses on the invariant masses of the leptons, it will be referred to as "invariant-mass approach".

The evaluation of the reconstruction performance is based on the knowledge of the true Z leptons. These leptons are determined by matching the reconstructed leptons with the leptons using truth information.

3.5.1. Matching of Reconstructed and Truth Leptons

For each signal event the information of the reconstructed kinematics and the truth information are available. This means for every true Z lepton, observables like the electric charge, the lepton flavour and its four-momentum are known. Only events where two leptons are labelled as true Z leptons are part of the dataset. Therefore, processes like the Drell-Yan process where a photon instead of a Z boson is radiated are not included.

With the help of a matching algorithm the reconstructed leptons which describe the two true Z leptons can be found. The matching is done by evaluating the geometrical distance between the truth leptons and the reconstructed leptons. To be matched a ΔR
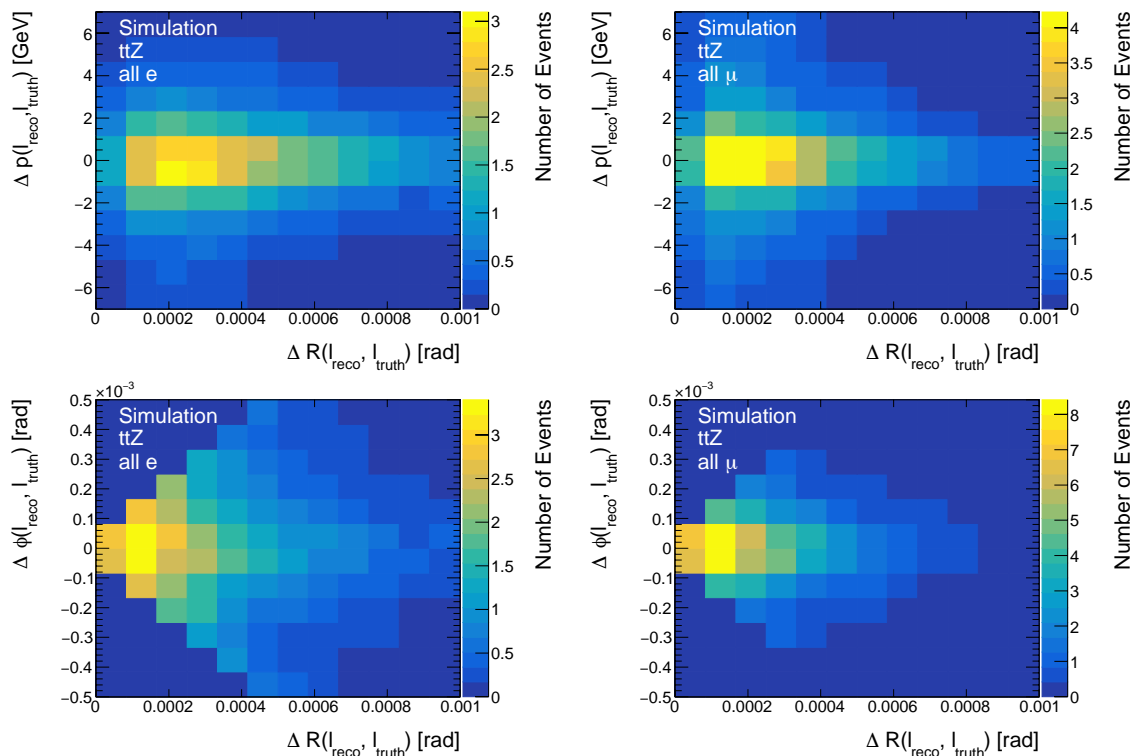
Figure 3.9.: The distributions of the number of events as a function of ΔR and Δp or $\Delta \phi$ between the matched leptons. On the left the distributions for all electrons in all events is shown and on the right the same for all muons.

smaller than 0.1 between two candidates is required. In addition, the candidates need to have the same electric charge and lepton flavour.

In the first step, all possible combinations fulfilling these criteria are determined. If a reconstructed or truth lepton is matched with more than one truth or reconstructed lepton, respectively, the combination with the smallest ΔR is chosen.

To evaluate this approach, distributions of other observables describing the lepton kinematics are useful to verify if the leptons not only match in ΔR, but also in, for example, the momenta p and the azimuthal angles ϕ. The number of events as a function of the differences in these observables and the ΔR between the reconstructed lepton and the truth lepton it was matched with are shown in Figure 3.9.

The distributions show that the momenta p and the azimuthal angle ϕ of the matched leptons differ by at most $7\,\mathrm{GeV}$ and $0.5 \times 10^{-3}\,\mathrm{rad}$, respectively. Therefore, it can be assumed that the matching algorithm is reliable. This was also checked for the pseudo-rapidity η leading to the same conclusion. The corresponding distribution is shown in Figure A.3

Table 3.4.: Efficiencies of the Z boson reconstruction when using the invariant-mass approach.

	$eeee$	$eee\mu$	$ee\mu\mu$	$e\mu\mu\mu$	$\mu\mu\mu\mu$
Z boson correctly reconstructed	86.5%	93.8%	92.8%	93.8%	83.3%
Z boson incorrectly reconstructed	13.5%	6.2%	7.2%	6.2%	16.7%

3.5.2. Evaluation of the Invariant-Mass Approach

To evaluate the invariant-mass approach, the same selection as in Ref. [42] is applied and the matching of reconstructed and truth leptons is performed. Afterwards the invariant-mass approach is used to determine the reconstructed Z leptons. If this approach identifies those reconstructed leptons as Z leptons which were matched with the true Z boson leptons, the event is labelled as correctly reconstructed. If one or both leptons are not correctly identified, it is labelled as incorrectly reconstructed. Cases where one or both of the truth leptons originating from the Z boson cannot be matched with reconstructed leptons are labelled as incorrectly reconstructed, too. In those cases it is generally possible that the leptons originating from the Z boson were chosen by the invariant-mass approach but this cannot be verified.

The reason for non-matched leptons can be a photon radiation which changes the momentum of the reconstructed lepton. The truth information does not include radiations. Therefore the four-momenta of the leptons differ. Another possibility is that the Z boson decays into tau leptons. The truth leptons are then labelled as tau leptons while the reconstructed leptons from the tau decays are labelled as electrons or muons. Due to the flavour constraint they cannot be matched.

The efficiency is used as an evaluation metric, i.e., the ratio of the number of events where the Z boson was reconstructed correctly to the total number of events. The evaluation was performed in each region defined by the lepton flavour multiplicity separately. The results are given in Table 3.4.

The efficiencies show how many additional events can potentially be reconstructed when using a different approach. Z bosons are mostly misreconstructed in the regions with four same-flavour leptons. This matches the expectation due to the larger number of possible combinations. The remaining three regions have comparable efficiencies, even though the reconstruction efficiency in the $ee\mu\mu$ channel is slightly smaller compared to the $eee\mu$ and $e\mu\mu\mu$ channel. The recoverable percentage of events are here in a range between $6.2\,\%$

Table 3.5.: The fractions of events where the *Z* boson was misreconstructed when using the invariant-mass approach split into different categories.

	eeee	*eeeμ*	*eeμμ*	*eμμμ*	*μμμμ*
two matches, one incorrect	6.7%	4.4%	–	4.6%	8.6%
two matches, both incorrect	2.6%	–	4.6%	–	4.4%
one match, correct	0.6%	0.5%	0.3%	0.4%	0.5%
one match, incorrect	0.4%	0.1%	0.3%	0.1%	0.6%
no match	3.2%	1.2%	2.0%	1.1%	2.8%

and 8.2 %.

The different reasons for choosing the wrong leptons for the reconstruction are shown in Table 3.5. Here, different cases were distinguished: the first case is that the two truth leptons originating from the *Z* boson were matched with two reconstructed leptons but one or both of them are not the ones chosen by the invariant-mass approach. The second one is that only one true *Z* lepton could be matched. In this case one distinguishes the matched reconstructed leptons, which are identical with one of the leptons chosen by the invariant-mass approach, and those which are not. The last case is that no match at all could be found.

Table 3.5 shows that in most cases the reason for an incorrectly reconstructed *Z* boson is that both true *Z* leptons were matched, but the wrong leptons were chosen by the invariant-mass approach. These events can be recovered when using a DNN for the reconstruction. In a significant fraction of events no match at all was found. These events should be excluded from consideration because it cannot be determined if the *Z* boson reconstruction was successful or not.

3.5.3. *Z* Boson Reconstruction DNN Architecture

While the classes for the classification are defined by the underlying event topology, there are different possibilities for the *Z* boson reconstruction. Two ways of defining the classes are presented.

One option is to use every combination of lepton pairs as classes. The leptons are

labelled via their transverse momenta p_T, with lepton 1 having the largest and lepton 4 the smallest transverse momentum. The classes can be defined by the lepton combination based on this labelling. This leads to 6 classes, namely *lepton 1 and lepton 2 originate from Z boson, lepton 1 and lepton 3 originate from Z boson* and so on.

The disadvantage of this approach is that it does not come with a physical motivation. The events of one class do not necessarily have more in common with events of their own class than with events from another class. Also, a completely different labelling of events can occur when using another observable than the transverse momentum p_T to define the lepton labels.

For this reason a second, physics motivated approach is tested. It is based on the invariant-mass approach. The first class is *both correct* with events where the invariant-mass approach chooses the correct leptons as Z leptons. The second class is *one correct* where only one of the leptons is correctly chosen and the third that no lepton is correct, referred to as *both incorrect.*

For both approaches, only events where both true Z boson leptons are matched can be assigned to a class, because otherwise the correct class cannot be determined. The events where an assignment is not possible are labelled as *no class.*

The approach using labels defined via the lepton numbering based on their transverse momenta is referred to as "lepton-numbering approach" or "lepton-numbering labels". The second approach making use of the invariant-mass approach is referred to as "approach using invariant-mass labels".

After the definition of labels the DNN itself and its input observables have to be defined. The Z boson reconstruction is only done for $t\bar{t}Z$ events making the available size of training samples four times smaller than for the classification. The exact numbers are given in Table 3.1. Hence, the reconstruction DNN is chosen to have only 3 layers with 30 nodes each. Two folds are used for the k-fold cross validation. Similarly to the classification DNN, dropout layers with an 0.3 dropout probability are used. A hyperparameter optimisation was not performed. The output layers have 3 or 6 nodes depending on the number of classes.

Each output node can be seen as a binary classifier determining whether the presented event is part of the class or not. This means for every class i exists a classifier i that ideally outputs 1 if the event is part of the class i and 0 if it is not.

The input observables differ depending on the used labels. For the lepton-numbering approach the invariant masses of lepton pairs, the electric charges and lepton flavour information are used as input observables. The invariant masses are calculated for leptons of the same lepton flavour and an opposite electric charge. If these criteria are not fulfilled,

it is set to zero. These observables are suitable inputs because the invariant masses of the leptons originating from the Z boson are expected to be close to the Z boson mass m_Z.

Lepton flavour information and electric charges of the leptons are important because based on these certain combinations can be excluded with high certainty since Z bosons decay into lepton pairs with same flavour and opposite charge. But as discussed previously, the electric charges, especially from electrons, and the lepton flavours can be misreconstructed. While the invariant-mass approach automatically excludes these events, the DNN is not forced to choose a combination where the electric charge and lepton flavour requirements are fulfilled. In addition to misreconstructed lepton flavours, events where the Z boson decays into two tau leptons can cause final states with Z leptons with a different lepton flavour. In these cases the tau leptons decay into a neutrino and a W boson with the W boson decaying further into an electron or muon and a neutrino. Nevertheless, the fraction of these events is in the order of a magnitude of per mille and therefore not treated separately.

For the approach using invariant-mass labels the transverse momenta p_{T}, the lepton flavour information, the electric charges, the pseudorapidity η and the azimuthal angles ϕ of all leptons are used as inputs. The invariant masses of the Z leptons and the non-Z leptons are used as well. These variables are defined by the invariant-mass approach.

The motivation behind a DNN using these labels is to determine if events of the classes *one correct* or *both incorrect* have kinematics that differ from those events where both leptons are correctly identified. To reconstruct the direction of the particles the pseudorapidity η and the azimuthal angles ϕ are used. In addition the invariant masses of the Z and non-Z leptons help to understand if the prediction of the invariant-mass approach fails because these masses are for example both close to each other or both far away from the Z boson pole mass m_Z. The electric charges and lepton flavour information are used for the same reason as for the previously presented DNN.

This means the main goal is to determine what the events look like where the invariant-mass approach fails, compared to those where it succeeds, and use these differences to correct the reconstruction algorithm. For events of the class *both incorrect* the two remaining leptons are chosen as Z leptons.

For events of the class *one correct*, an ambiguity remains for regions with four leptons with the same lepton flavour. Then, two lepton pairs fulfil the requirement of an opposite electric charge. In these cases, the lepton pair with an invariant mass closest to the Z boson pole mass m_Z can be chosen.

For three leptons with the same lepton flavour the correct lepton pair is well-defined because of the charge constraint. For two leptons of each flavour no *one correct* cases are
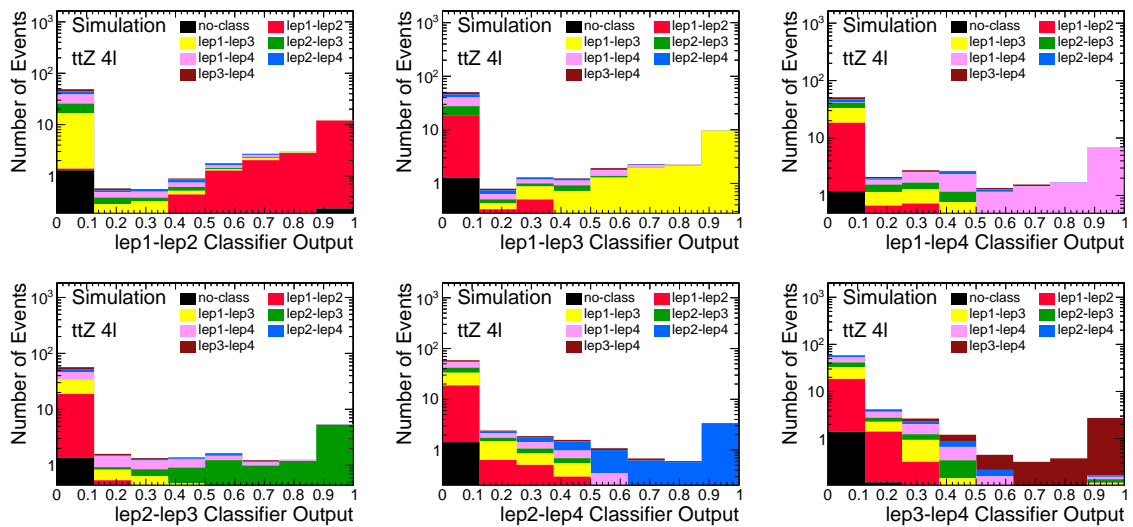
Figure 3.10.: Distributions of events for each classifier with lepton-numbering labels. Every node from the output layer belongs to one class and outputs the probability that the presented input belongs to the corresponding class.

possible due to the lepton flavour constraint.

3.5.4. Z Boson Reconstruction DNN Performance

For the Z boson reconstruction DNN, only events passing the selection used in the latest ATLAS analysis given in Ref. [42] are used as training samples. This way it is possible to make a direct comparison between both approaches. Before comparing with the invariant-mass approach it has to be determined which labelling is more useful.

First, the distributions of the output of each output node from both DNNs are evaluated. Figure 3.10 shows the distribution of events for each classifier of the DNN using lepton-numbering labels. The labels based on the invariant-mass approach lead to three output nodes which results in the three distributions given in Figure 3.11.

The comparison between all distributions indicates that the lepton-numbering approach leads to a better separation. Figure 3.10 shows that the events of the class i accumulate near 1 for the classifier i and only a small tail towards lower classifier outputs is visible. All other events are mostly close to 0. For the second set of classifiers in Figure 3.11 this trend is not observed. When comparing all three distributions only for the *both correct* classifier a trend towards higher output values for this class is observed. For the remaining classifiers a distinct separation is not visible and only few events reach a classifier output higher than 0.6.

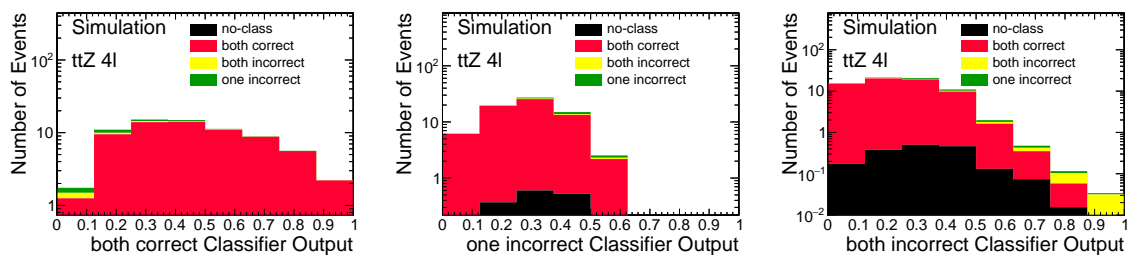To confirm the statement that the lepton-numbering approach performs better, the

Figure 3.11.: Distributions of events for each classifier with labels defined by the invariant-mass approach. Every node from the output layer belongs to one class and outputs the probability that the presented input belongs to the corresponding class.

Table 3.6.: Efficiencies of all approaches calculated by excluding those events which are not part of any class.

	$eeee$	$eee\mu$	$ee\mu\mu$	$e\mu\mu\mu$	$\mu\mu\mu\mu$
DNN (lepton-numbering labels)	84.9%	92.4%	95.0%	92.0%	84.0%
DNN (invariant-mass approach labels)	56.6%	71.2%	74.0%	59.6%	37.5%
Invariant-mass approach	90.3%	95.5%	95.3%	95.3%	86.6%

efficiencies of both networks are calculated. They are given in Table 3.6. In addition, the efficiencies of the invariant-mass approach are shown. They are calculated excluding the *no class* events. For these events, the true class is not known and therefore a prediction encoded as correct is not possible which means an efficiency of 100% could not be reached. For this reason the efficiencies of the invariant-mass approach differ from those given in Table 3.4, because there the *no class* events were taken into account to study the fraction of events where the matching fails.

The DNN with the approach using non-physics motivated labels reaches higher efficiencies than the labels with a physical motivation which matches the first impression based on Figures 3.10 and 3.11.

A reason for this may be the uneven distribution of events. Even though in both approaches the events are not distributed equally over all classes this effect has a higher impact for the DNN using invariant-mass labels. The distribution of events over the classes is shown in Figure 3.12.
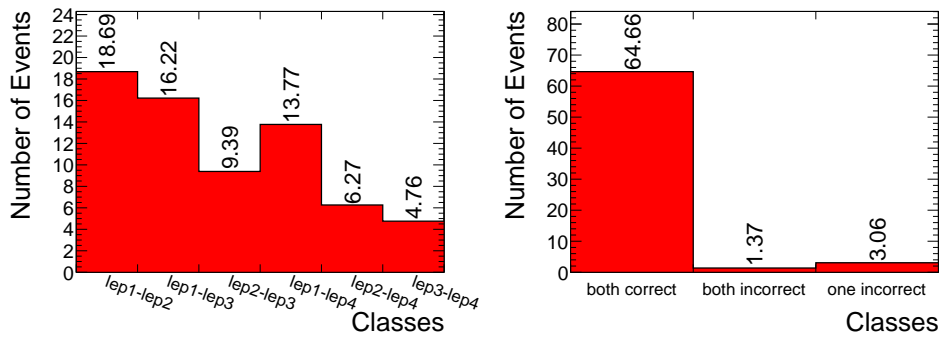
Figure 3.12.: Yields for each class of the DNN using lepton-numbering labels on the left and for the DNN using invariant-mass labels on the right.

To get an unbiased training set, the events are reweighted as given in Equation (3.6) but a significantly smaller number of training events still leads to less sensitivity on events of the corresponding class since the network gets trained on fewer training samples. The least populated class for the lepton-numbering approach is the class where lepton 3 and lepton 4 are the two Z boson leptons with a fraction of 6.9%. For the approach using labels based on the invariant-mass approach the class *both incorrect* only contributes 2.0% of events and the class *one class* only 4.4%. This means the least populated class of the lepton-numbering approach constributes more events than the two lowest populated classes of the approach using only three classes in total. Therefore, the DNN may not get enough training samples to become sensitive to the differences between the classes for the approach using invariant-mass labels.

To study whether the DNNs are able to generalise the AUC values are listed in Table 3.7. The upper block shows the performance of the DNN using lepton-numbering labels. The AUCs of the approach using invariant-mass labels are given in the part underneath. The AUC values for the approach using invariant-mass labels are significantly smaller than those from the lepton-numbering approach. This is in agreement with the observation of a weaker performance based on the efficiencies. In addition the uncertainties are larger and especially for the class *both incorrect* the AUC values for the training and testing dataset have a large discrepancy. This leads to the assumption that the DNN is not able to generalise and make a prediction as reliable for unseen events.

At the same time Table 3.7, shows that the AUCs of the training and testing samples for the approach using lepton-numbering labels are close to each other with differences between 0 and 0.004. All AUC values are maximally 1.3% away from 100% and therefore nearly reach the highest value possible. Also the uncertainties are much smaller compared to the first approach. This suggests a stable performance and that the classifiers are able

Table 3.7.: AUC values for the training and testing folds for all classifiers of both DNN approaches.

Class	lep1-lep2	lep1-lep3	lep2-lep3
AUC (Train)	0.992 ± 0.001	0.991 ± 0.002	0.992 ± 0.000
AUC (Test)	0.991 ± 0.002	0.990 ± 0.003	0.991 ± 0.002
Class	lep1-lep4	lep2-lep4	lep3-lep4
AUC (Train)	0.988 ± 0.003	0.991 ± 0.001	0.987 ± 0.001
AUC (Test)	0.988 ± 0.005	0.989 ± 0.001	0.983 ± 0.005
Class	both correct	one correct	both incorrect
AUC (Train)	0.760 ± 0.017	0.749 ± 0.014	0.853 ± 0.060
AUC (Test)	0.739 ± 0.015	0.705 ± 0.022	0.778 ± 0.060

to generalise. Due to the lower efficiencies and the less stable performance the approach using invariant-mass labels is discarded.

After choosing the approach for the DNN, its performance is compared to the invariant-mass approach which is not based on machine learning. The efficiencies of both approaches are compared.

Table 3.6 shows that the DNN is not able to outperform the invariant-mass approach. The largest discrepancy is seen in the *eeee* region with a difference of 5.4%. In the *eeµµ* region the discrepancy is only 0.3% but the DNN does not reach a larger efficiency than the invariant-mass approach in any region.

The reasons for the weaker performance are studied. The confusion matrix helps to understand in which cases the DNN in comparison to the invariant-mass approach fails to make the correct prediction. The matrix shows what fraction of events with a certain label is classified as what class. This means a perfect DNN would yield an identity matrix. The matrices for both approaches are shown in Figure 3.13.

For the invariant-mass approach the efficiency decreases only slightly when going from the lower left to the upper right corner. The DNN starts with a larger efficiency in the lower left corner, but the decrease is steeper such that the efficiency in the upper right corner is around 12.7% smaller than for the invariant-mass approach. The labels are ordered in a way that when going to the right and upwards the labels include more leptons with lower transverse momentum. This means that most confusion occurs when leptons with a low transverse momentum compared to the remaining leptons in the event originate from the *Z* boson. If those leptons with the largest transverse momentum of the event, i.e., lepton 1 and lepton 2, originate from the *Z* boson the prediction of the DNN
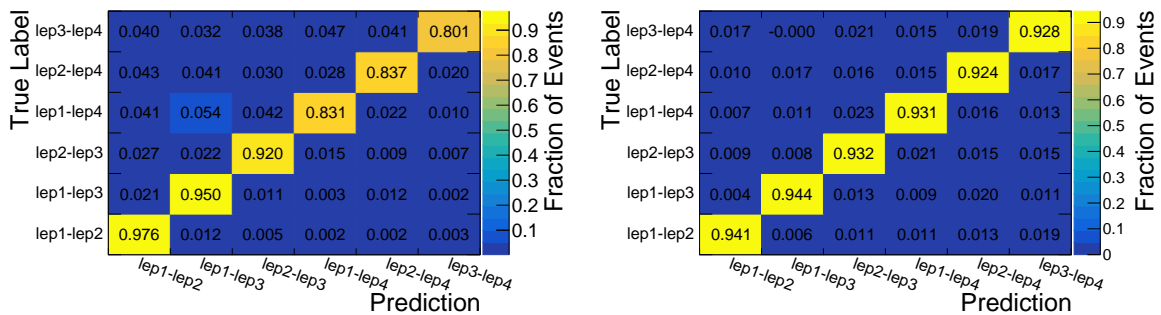
Figure 3.13.: The confusion matrices for the DNN approach on the left and the invariant-mass approach on the right.

is reliable and even better than the prediction of the invariant-mass approach.

To understand these events more precisely, it is useful to see in what p_T range the events are more likely misclassified. The events, where both approaches provide the same and correct prediction and those in which the invariant-mass approach is correct while the DNN is not, are most interesting. Since the DNN bases the prediction on the same information as the invariant-mass approach, it seems unintuitive that the DNN fails in those cases where the invariant-mass approach succeeds. The distributions of the leading lepton and the sub-leading lepton are shown in Figure 3.14. The distributions of the remaining leptons are shown in Figure A.4. The distribution of the p_T of lepton i only shows events where lepton i is one of the true Z leptons, e.g., the distribution of the transverse momentum of lepton 1 only includes events of the classes *lep1-lep2*, *lep1-lep3* and *lep1-lep4*. This way it can be seen if the wrong reconstruction occurs at specific Z boson kinematics.

Figure 3.14 shows that the distribution of events where the DNN struggles compared to the invariant-mass approach is shifted towards the lower p_T region. The distributions of events where both approaches make the same and correct prediction have a smaller maximum and, at least for the leading lepton, the maximum is shifted to the right. The remaining events are further distributed along the tail towards higher transverse momenta. Therefore, the DNN fails mostly in cases where the leptons carry a low p_T.

When the leading or the sub-leading leptons have a low transverse momentum, this also means, that all other leptons have an even lower p_T than these two and thus the leptons are close to each other in transverse momenta. This is the case for the shifted events in Figure 3.14 which means that the DNN also struggles with those events.

Another observable that helps to understand for which events the DNN does an incorrect prediction is the invariant mass of the (non-)Z leptons. The distribution of the invariant
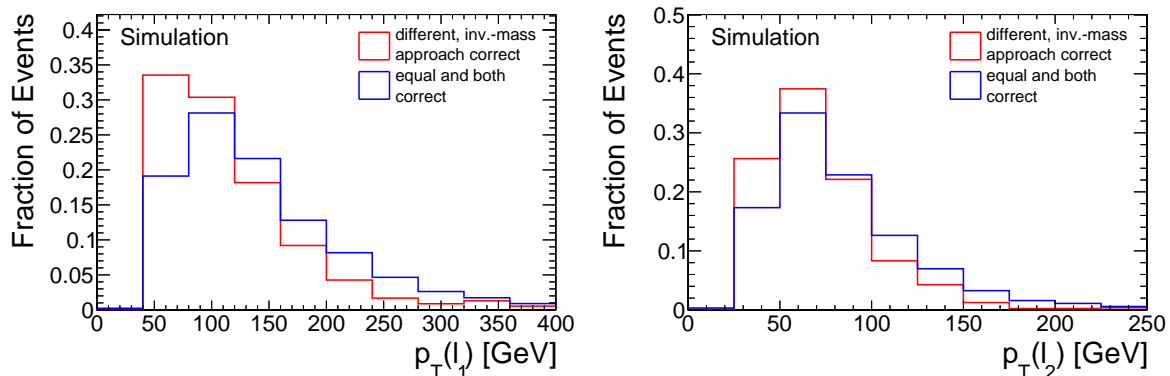
Figure 3.14.: The distribution of the transverse momenta p_T of the leading lepton on the left and the subleading lepton on the right. The events are split into those where the Z boson reconstruction DNN and the invariant-mass approach provide the same and correct prediction and those where only the invariant-mass approach makes the correct prediction.

mass of the Z leptons $m_{\ell,\ell}^Z$ as a function the invariant masses of the non-Z leptons $m_{\ell,\ell}^{\text{non-}Z}$ are shown in Figure 3.15. Only the distributions where the invariant-mass approach provides the correct prediction are shown. This makes sure that the leptons used for the calculation of the invariant masses are the true (non-)Z leptons.

As expected the number of events peaks around the Z boson mass for the invariant mass of the two Z leptons and decreases towards higher and lower masses steeply. The main differences between the distributions for the events, where both approaches provide the correct prediction and those where only the invariant-mass approach is correct, is along the axis of the invariant mass of the two non-Z leptons. For events, where the DNN is not able to determine the correct Z leptons, the number of events peaks around the Z boson mass for the invariant mass of the non-Z leptons. This means that events where both lepton pairs have an invariant mass close to the Z boson mass m_Z are those where the DNN has difficulties. This also matches the expectation. For the events where both approaches provide the correct prediction the maximum is below the Z boson mass m_Z. Nevertheless a tail towards higher masses is visible and therefore also events around the Z boson mass can be categorised correctly by both approaches.

These studies help to understand in which cases the DNN makes an incorrect prediction but it cannot explain why the prediction fails and therefore does not provide a solution to improve the network's performance. A possibility why the DNN has a lower performance is an insufficient size of the training data set. To check whether the efficiencies can be increased by using a larger training data set, the training of 10 networks is performed. The DNNs only differ in the number of training samples, starting with a DNN using 10%
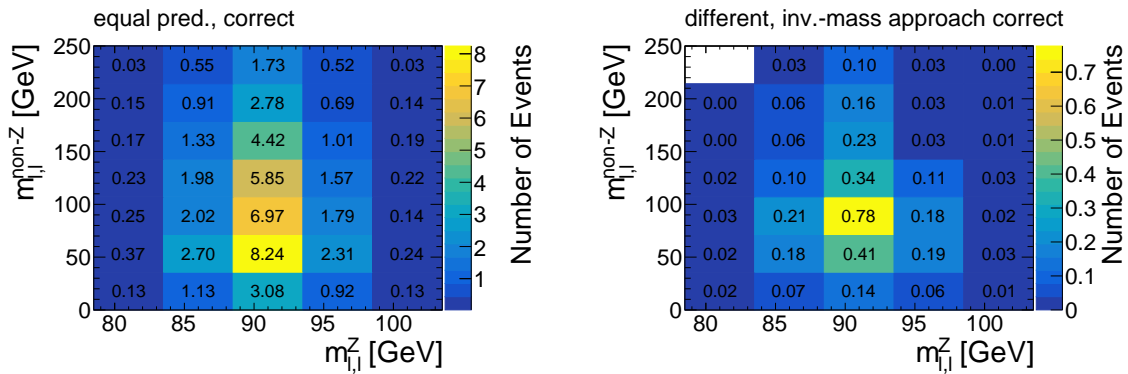
Figure 3.15.: The distribution of events depending on the invariant mass of the Z leptons and the invariant mass of the non-Z leptons.
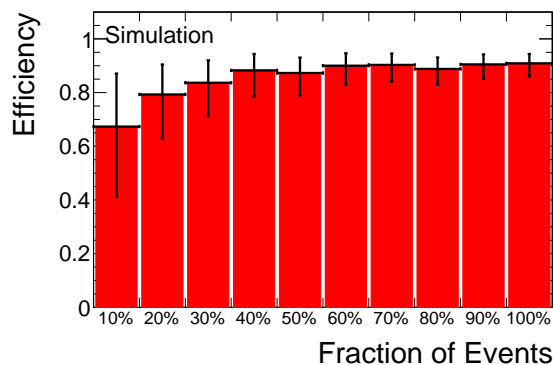

Figure 3.16.: Efficiency depending on the fraction of events used for the training of the DNNs.

of the available samples and increase this fraction in steps of 10% until 100% is reached. The resulting efficiencies are shown in Figure 3.16.

The distribution shows that the efficiency first increases but reaches a plateau at 60%. Therefore the reason for the lower efficiency is not the small training size. This also gets supported by the AUC values and their uncertainties which implies a stable performance and therefore no influence due to statistical fluctuations is expected. Other reasons for the worse performance are not found or studied.

The final conclusion is, that the studied DNN for the Z boson reconstruction does not provide an improvement. Nevertheless, the lepton-numbering approach is able to reach efficiencies only 0.3% up to 5.7% smaller than the invariant-mass approach. The approach using invariant-mass labels has much smaller efficiencies. For example, in the $\mu\mu\mu\mu$ region it only reaches a 37.5% efficiency. Also, the largest efficiency of 74% in the $ee\mu\mu$ region

is not in the range of the invariant-mass approach and the DNN using lepton-numbering labels. Another physics motivated approach where the events are more evenly distributed among the classes might be able to reach higher efficiencies.

It should be kept in mind that the invariant-mass approach already has a high efficiency. Figure 3.8 and Table 3.5 show that, for the SF region, 3.9 out of 33.7 and, for the DF region, 2.3 out of 37.2 event can maximally be gained. Here, also events with the label *no class*, i.e., events where the matching of truth and reconstructed leptons failed, are included even though in these cases a validation of the prediction is not possible. When excluding the events, where not all true Z leptons can be matched, only 2.7 out of 32.5 for the SF region and 1.7 out of 36.6 for the DF region can be recovered by using a different approach.

3.6. Measurement of the Inclusive Cross Section

The results from Section 3.4 can be used to measure the $t\bar{t}Z$ inclusive cross section by performing a profile likelihood fit and extracting the signal strength of $t\bar{t}Z$, $\mu_{t\bar{t}Z}$. The signal strength is defined by

$$\mu_{t\bar{t}Z} = \frac{\sigma_{\mathrm{meas}}}{\sigma_{\mathrm{SM}}} \tag{3.9}$$

with the measured cross section σ_{meas} and the SM prediction of the cross section σ_{SM}.

The expected uncertainty is studied with an Asimov fit: a pseudo dataset is generated based on the nominal predictions for signal and background from MC simulations. Since this implies that the simulated prediction equals the data, the fitted nuisance parameters (NPs) remain at their initial value of zero. The real collider data is not used.

Besides fits of the signal strength $\mu_{t\bar{t}Z}$, simultaneous fits of the signal strength $\mu_{t\bar{t}Z}$ and the normalisation of ZZ, N_{ZZ}, are performed.

The measurement is affected by several systematic uncertainties which are used to determine the total uncertainty for the fit. They are separated in different categories which are described in the following sections.

3.6.1. Instrumental Uncertainties

The instrumental uncertainties are related to the detector and take into account that the different detector components do not have an efficiency of 100% but carry a systematic uncertainty and uncertainties regarding the calibration.

For the integrated luminosity, an uncertainty of 1.7% is derived using the methodology given in Ref. [44]. This uncertainty is applied to all simulated events besides the fakes samples. The same is true for the uncertainty of the MC pileup which arises from the reweighting of the MC pileup distribution in order to match with data. The uncertainty is determined by varying the correction factors used for the reweighting within their uncertainties.

Jet uncertainties arise from the jet energy scale (JES), the jet energy resolution (JER) and JVT requirements. The JES including its uncertainties is determined in Ref. [45] by using the information from test-beam data and combining it with collision data and simulation resulting in 21 NPs. For the JER, 8 NPs are calculated. The uncertainty of the JES and the JVT are a function of the jet p_{T} and increase at low p_{T} values.

Several uncertainties affect the jet tagging. They are estimated for the correction factors applied to the MC simulation to reproduce the tagging algorithm performance in data. For b-tagging these uncertainties are propagated by 45 nuisance parameters. For the uncertainties of the mis-tagging of c- and ℓ-jets, i.e., jets not originating from b-quarks, 20 NPs are used for each of them. For b-jets the calibration and associated uncertainties are derived using dileptonic $t\bar{t}$ events [46]. The calibration of c-tagged jets is performed using W boson decays in $t\bar{t}$ events [47]. For light-flavour jets, dijets events [48] are used for the calibration. The uncertainties for b- and c-jets depend on the jet's transverse momentum p_{T} including bin-to-bin correlations. For the light jets the uncertainties depend on the jet's p_{T} and the jet's pseudorapidity η.

Lepton uncertainties are composed of the trigger, identification, reconstruction and isolation uncertainties. In addition, the energy scale and momentum are a source of uncertainty [29, 49]. The uncertainties for the scale and resolution of the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ are estimated with the help of $Z \to \mu^{+}\mu^{-}$ events [50].

3.6.2. Theoretical Uncertainties

Several theoretical uncertainties for every process are considered. For tWZ and $t\bar{t}Z$ the uncertainties of the factorisation μ_F and the renormalisation μ_R of the nominal MG5_AMC@NLO and PYTHIA 8 samples are taken into account. The evaluation of the uncertainties is performed by varying the scales once simultaneously and once individually by a factor of 2.0 and 0.5 compared to their default value in the event generator. The PDF uncertainties for tWZ and $t\bar{t}Z$ are calculated using the PDF4LHC15 prescription [51]. Furthermore, an uncertainty is applied for tWZ to take different treatments of interferences between tWZ and $t\bar{t}Z$ by the diagram removal schemes DR1 and DR2 into account [52].

Scale and PDF uncertainties are applied to the $t\bar{t}H$ cross section. A conservative uncertainty of 50% is applied to the cross section of the processes grouped as "Others" following the approach given in Ref. [53]. For the single measurement of $\mu_{t\bar{t}Z}$, conservative uncertainties on the normalisation of the cross section of $ZZ + \text{jets}$ are used as well. They are distinguished by the jet flavour. For $ZZ + b$ and $ZZ + c$ a 30% and 20% uncertainty are applied, respectively. For $ZZ + \ell$, an uncertainty of 50% is applied. The uncertainties for $ZZ + \text{jets}$ were evaluated in the latest measurement using the comparison of data and MC simulation of $Z + b/c$ events [54]. In the simultaneous fit of $\mu_{t\bar{t}Z}$ and N_{ZZ}, the uncertainties for $ZZ + \text{jets}$ are not applied.

3.6.3. Fit Results

Different configurations are used for the fitting procedure to be able to make a direct comparison between the DNN approach and the "cut and count" approach. First, Asimov fits are performed on a distribution using events in one bin with the selection from the previous ATLAS analysis applied, referred to as tight selection. Afterwards fits to the four-bin distribution of the leading lepton transverse momentum are performed using the same selection. This is compared to a fit using the events with the loose preselection presented in Section 3.2 and fitting the four-bin distribution from the SF and DF classifiers presented in Section 3.4. This way, a direct comparison between the "cut and count" approach and the DNN approach is possible. Configurations using more than four bins are not used because of the statistical limitation in the tetralepton channel of $t\bar{t}Z$.

The fits are performed in the SF and the DF regions. Further control regions are not used. The one-bin distribution of events passing the tight selection, the distribution of the transverse momentum of the leading lepton and the distributions of the classifier outputs with the loose selection are shown for every region in Figure 3.17. The distributions include all systematics mentioned in Sections 3.6.1 and 3.6.2, i.e., also the uncertainties on the $ZZ + \text{jets}$ cross section. Figure A.5 shows the same distributions without the cross section uncertainties, used for the simultaneous fit of $\mu_{t\bar{t}Z}$ and the normalisation N_{ZZ}.

First, the fit results using both, the signal strength $\mu_{t\bar{t}Z}$ and the normalisation of ZZ, as free floating parameters are presented. For the one-bin distribution with the tight selection, the following results are achieved:

$$\mu_{t\bar{t}Z}^{\text{tight, one bin}} = 1.00 \pm 0.22, \tag{3.10}$$

$$N_{ZZ}^{\text{tight, one bin}} = 1.00 \pm 1.61. \tag{3.11}$$

This measurement is used as a reference to evaluate the fit to the distribution of the
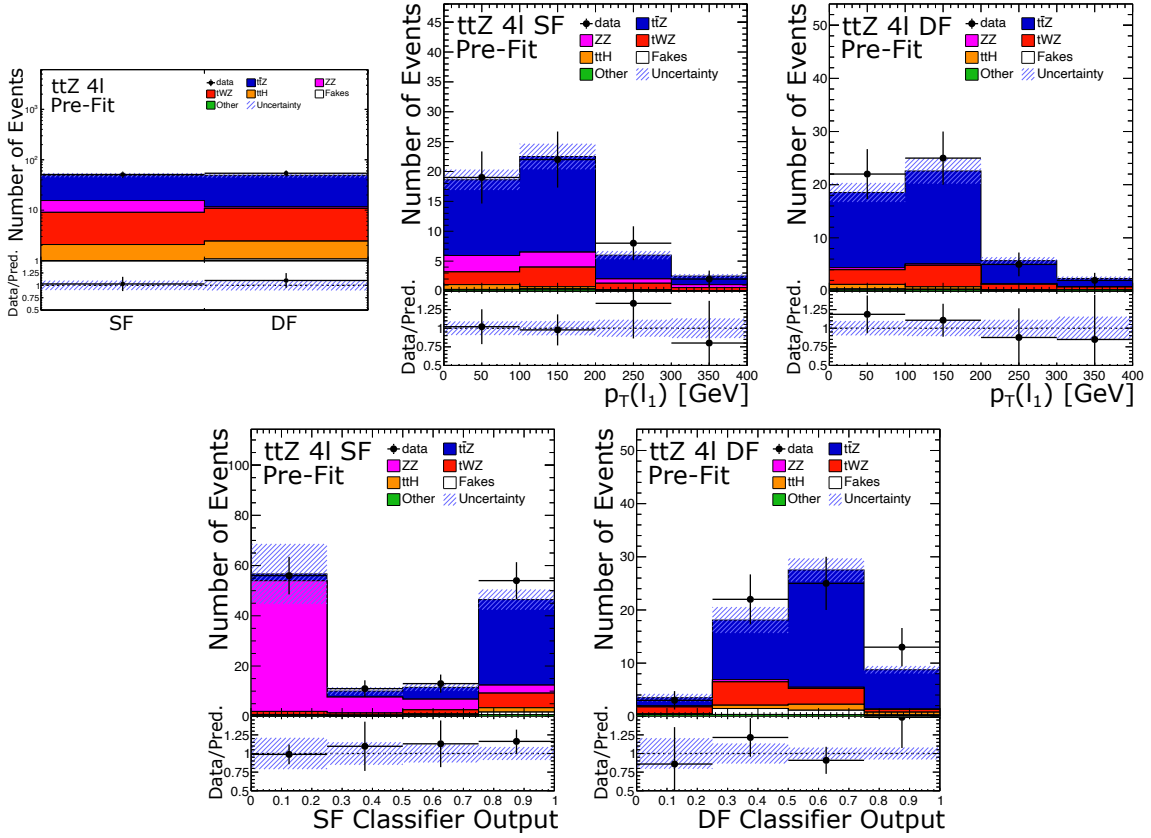
Figure 3.17.: The distributions for the different fitting configurations. The upper left plots shows the one-bin distributions and the two plots in the upper right the p_T distributions with the tight selection. The lower two plots show the distributions of the classifier outputs with the loose selection. All systematics mentioned in Sections 3.6.1 and 3.6.2 are included.

leading lepton p_T. The four-bin distribution leads to the following measurement:

$$\mu_{t\bar{t}Z}^{\text{tight, four bins}} = 1.00^{+0.23}_{-0.20}, \tag{3.12}$$

$$N_{ZZ}^{\text{tight, four bins}} = 1.00^{+1.62}_{-1.55}. \tag{3.13}$$

For the measurement of $\mu_{t\bar{t}Z}$, a larger upper uncertainty compared to the fit to the one-bin distribution is observed. Nevertheless, since the uncertainty shows an asymmetry, the total uncertainty is decreased by 2%. A constraint of the uncertainties for N_{ZZ} is not observed.

$$\mu_{t\bar{t}Z}^{\text{DNN, four bins}} = 1.00 \pm 0.15, \tag{3.14}$$

$$N_{ZZ}^{\text{DNN, four bins}} = 1.00 \pm 0.17. \tag{3.15}$$

A significant improvement is achieved for the signal strength of $t\bar{t}Z$. The uncertainties are in total reduced by 30%. The uncertainties for N_{ZZ} are reduced by 89% due to the larger number of ZZ events passing the loose selection. This means there are more events which are sensitive to the measured quantity. While the tight selection suppresses the ZZ contribution, the loose selection leads to a ZZ dominated SF region where most of the ZZ events accumulate in the first bin of the SF classifier output distribution. Therefore, the first bin, enriched in ZZ events, can be used as a control region for the measurement of N_{ZZ} and no definition of a separate region is needed.

In total the uncertainties of both fit parameters can be significantly reduced when using the observables defined by the classifier outputs as the fitted distributions.

For comparison, the fits using only the signal strength as a free floating parameter are performed. For the one-bin distribution with the tight selection the resulting uncertainties are given by

$$\mu_{t\bar{t}Z}^{\text{tight, one bin}} = 1.00 \pm 0.16. \tag{3.16}$$

By only using one free floating parameter the uncertainty can be decreased by 28%. For the fit using the four-bin distribution the uncertainties show again a shift towards larger values.

$$\mu_{t\bar{t}Z}^{\text{tight, four bins}} = 1.00^{+0.17}_{-0.15}. \tag{3.17}$$

In total the width of the uncertainty band remains unchanged and therefore no improvement is observed compared to the fit using one bin. The four-bin distribution provides additional shape information which carries an uncertainty. The shape information can lead to a more precise measurement but at the same time the uncertainties decrease the precision. The one-bin distribution does not have any shape information and therefore also their uncertainties do not apply. If these effects balance each other this can cause the unchanged width of the uncertainty band. Nevertheless, the uncertainties are reduced compared to those given in Equation (3.12).

The best performance is achieved by the fit using the distributions of the classifier outputs with an uncertainty decrease of 6.3% compared to the fit on the four-bin distribution:

$$\mu_{t\bar{t}Z}^{\text{DNN, four bins}} = 1.00 \pm 0.15. \tag{3.18}$$

The fit using only the signal strength of $t\bar{t}Z$ as a free floating parameter leads to no further improvement compared to the results using two free floating parameters given

Table 3.8.: Systematic and statistic uncertainties of the fits for every fit configuration. The upper part shows the uncertainties for the fits where $\mu_{t\bar{t}Z}$ and N_{ZZ} are used as free floating parameters and the lower part the fits where only $\mu_{t\bar{t}Z}$ is free floating.

	parameter	stat. unc.	syst. unc.
$\mu_{t\bar{t}Z}$ and N_{ZZ} free float.	$\mu_{t\bar{t}Z}^{\text{tight, one bin}}$	± 0.21	± 0.07
	$\mu_{t\bar{t}Z}^{\text{tight, four bins}}$	± 0.20	$^{+0.10}_{-0.03}$
	$\mu_{t\bar{t}Z}^{\text{DNN, four bins}}$	± 0.13	± 0.07
	$N_{ZZ}^{\text{tight, one bin}}$	± 1.60	± 0.17
	$N_{ZZ}^{\text{tight, four bins}}$	± 1.54	$^{+0.51}_{-0.19}$
	$N_{ZZ}^{\text{DNN, four bins}}$	± 0.14	± 0.10
$\mu_{t\bar{t}Z}$ free float.	$\mu_{t\bar{t}Z}^{\text{tight, one bin}}$	± 0.14	± 0.08
	$\mu_{t\bar{t}Z}^{\text{tight, four bins}}$	± 0.14	$^{+0.10}_{-0.07}$
	$\mu_{t\bar{t}Z}^{\text{DNN, four bins}}$	± 0.13	± 0.08

in Equation (3.14). In conclusion, the DNN performs best for all configurations but it does not make a difference if both parameters are fitted simultaneously or not. The DNN classifiers with the loose selection are able to constrain the uncertainties of the *ZZ* normalisation.

Due to the low number of signal events, it is useful to compare statistical and systematic uncertainties. They are given for every configuration in Table 3.8.

Table 3.8 shows that the statistical uncertainty dominates. In the simultaneous fits of $\mu_{t\bar{t}Z}$ and N_{ZZ}, the largest statistical uncertainties are for the one-bin distributions. In the single-fit configuration, the statistical uncertainties of the one- and four-bin distributions are the same. The smallest statistical uncertainties are in both configurations for the fits using the classifier outputs. In general it is plausible that the fits using the classifier outputs carry the smallest statistical uncertainty since the selection applied is looser in order to increase the acceptance of events and therefore the statistics.

The impact of the different systematics on the simultaneous measurement of the signal strength $\mu_{t\bar{t}Z}$ and N_{ZZ} and on the single measurement of $\mu_{t\bar{t}Z}$ are summarised in Table 3.9 for the fits using the distributions of the classifier outputs. They are grouped into categories. The impacts on the fits using the one-bin and four-bin distribution of the leading

Table 3.9.: Impact of the systematics on the fits using the four-bin distribution of the classifier outputs and the loose selection applied.

Uncertainty	impact [%] on measurement of		
	N_{ZZ}	$\mu_{t\bar{t}Z}$ (simultaneous)	$\mu_{t\bar{t}Z}$ (single)
Tagging	4.7	1.4	1.3
Jets	7.4	2.0	1.9
Lepton	5.0	5.6	5.6
Luminosity	1.7	2.1	2.1
$E_{\mathrm{T}}^{\mathrm{miss}}$	0.5	0.3	0.3
$t\bar{t}Z$ PDF	0.1	2.7	2.7
$t\bar{t}Z$ scale choice	0.1	0.4	0.4
Pileup	0.7	1.0	1.0
Background modelling	0.2	1.4	1.4
Trigger	0.1	0.1	0.1
ZZ cross section (single fit only)	–	–	1.4

lepton p_{T} are shown in Tables A.2 and A.3, respectively.

Table 3.9 shows that the lepton systematics have a large impact with 5.0% to 5.6% on all fits which make use of the distribution of the classifier outputs. For the measurement of $\mu_{t\bar{t}Z}$ the uncertainty of the PDF of $t\bar{t}Z$ has the second largest impact with 2.7% in both fitting configurations.

Not all of the NPs introduced in Sections 3.6.1 and 3.6.2 are used for the fitting procedure. In case their impact is negligible they are dropped to reduce complexity of the fit. This procedure is called pruning. If the NP's impact on the shape is smaller than 0.01% it is only considered for the normalisation and vice versa. In case the impact on both, the shape and the normalisation, is below 0.01% the NP is dropped completely. The pruning of the NPs used for the simultaneous fit using the classifier output distributions is given in Figure A.7. For the fit on $\mu_{t\bar{t}Z}$ only the pruning is given in Figure A.8.

With the help of the fits it is possible to constrain those NPs which were not dropped in the pruning procedure. The uncertainties for every parameter are given in Figure 3.18 for the single fit on the signal strength of $t\bar{t}Z$, $\mu_{t\bar{t}Z}$. This way Figure 3.18 also includes the uncertainties on the cross section of ZZ + jets. The NPs for the simultaneous fit of $\mu_{t\bar{t}Z}$ and N_{ZZ} are given in Figure A.6. The 1σ band is marked in yellow and the 2σ band in green. The parameters have an initial value of 0 represented by the black dot. This value stays unchanged by the design of an Asimov fit. The uncertainty is given by the black bar.
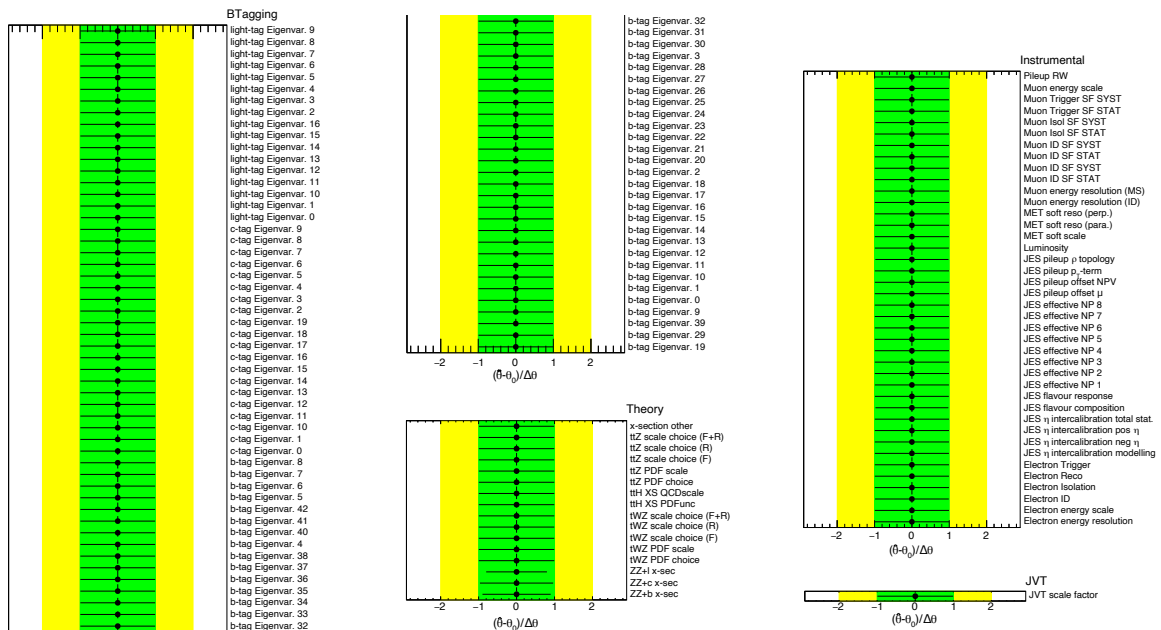
Figure 3.18.: The constraints on the NP used in the single fit of $\mu_{t\bar{t}Z}$. The yellow area shows the 1σ environment and the green area the 2σ environment. The bars show the uncertainty of the NP.

Most of the NPs are not constrained by the fit. Their error bars fill the whole 1σ environment. Only the cross section uncertainties of ZZ + jets are constrained. Most reduction is observed for $ZZ + \ell$ with 20.0% followed by $ZZ + b$ with 10.2% and $ZZ + c$ with 3.6%.

One reason that the other NPs are not constrained is the small available statistics. Even though it is increased by the loose selection used for the DNN, statistics are still the dominant uncertainty as shown in Table 3.8. For the ZZ contribution, a smaller systematic limitation is observed in the SF region which influences the capability of constraining the related systematics. In the SF region the number of ZZ events is 60.8 which helps to reduce the uncertainties of the NPs for the cross section of ZZ + jets through the fit. In addition, the uncertainties were chosen in a conservative manner with 50% (30%/20%) for $ZZ + \ell$ ($ZZ + b/+c$) such that a constraint can be expected.

To summarise the fit of the distributions from the classification classifier outputs have the smallest uncertainty. Here, no difference between the simultaneous fit of $\mu_{t\bar{t}Z}$ and N_{ZZ} and the single fit of $\mu_{t\bar{t}Z}$ was observed. Only within this fit a constraint on the uncertainty of N_{ZZ} due to the larger ZZ background contribution was achieved. Due to limited signal statistics, no NPs is constrained except for the ones corresponding to the uncertainties on the cross section of ZZ + jets.

4. Conclusion and Outlook

Within this thesis, the tetralepton channel of the production of a top-quark pair in association with a Z boson was studied. This channel suffers from limited statistics which motivates the use of a loose event selection and a DNN to separate signal events ($t\bar{t}Z$) from background events instead of applying cuts on relevant observables. Three approaches were studied, namely, using one DNN for all events, splitting into the SF and the DF region and training one DNN for each region or training 5 DNNs, one for each lepton flavour configuration. An improvement in terms of purity is observed compared to the "cut and count" approach for the DNN of the combined region with 95% and the two DNNs for the SF and DF region with 106% and 100% improvement, respectively. A direct comparison with the DNNs for the regions defined by the lepton flavour multiplicities is not performed because they did not show further improvement compared to the other options.

In terms of statistical significance, the DNN for the combined region showed the largest improvement with 12%. The evaluation of the separation power led to best results for the SF and the combined region with a separation power of 37% and 38%, respectively. The DNN of the DF has a significantly lower separation power with 8%. The AUC values provided best results for the DNN of the SF region followed by the DNN for the combined region. The DNN for the DF region showed the weakest performance with AUC values of 0.54. In total, the evaluation showed that the DNN for the DF region has overall the weakest performance, the DNNs for the SF region and combined region have comparable performances depending on the evaluated metric.

Nevertheless, both approaches are useful since they both lead to an improved classification of signal and background events. One reason for the weaker performance of the DNN for the DF region is the smaller number of gained signal events when applying the looser selection. By finding a selection which leads to a larger gain, the network's performance may be improved. In addition, the distribution of the DF classifier output for tWZ, which is the main background in the DF region, is mostly flat. The performance of the DNN can be improved by finding observables which are useful to distinguish $t\bar{t}Z$ and tWZ events and thus to increase the separation.

4. Conclusion and Outlook

The second part of the thesis dealt with the reconstruction of the Z boson in $t\bar{t}Z$ events. The approach using a DNN for the reconstruction was compared with the classical invariant-mass approach used in the latest ATLAS analysis [42]. Two approaches using two different sets of class labels were studied. The approach using lepton-numbering labels outperforms the approach using invariant-mass labels with 21.0% to 47.4% larger efficiencies depending on the lepton flavour region. Nevertheless the lepton-numbering approach is still not able to reach larger efficiencies than the invariant-mass approach. The discrepancies are between 0.3% and 5.4%. The events, which the DNN tends to misclassify while the prediction of the invariant-mass approach is correct, include preferentially Z leptons in the low p_T region or have the invariant mass of the two non-Z leptons close to the Z boson mass.

A DNN which is more sensitive to these cases might reach larger efficiencies. Also a more physics motivated approach with an even distribution of events over the classes can lead to an improvement. For both DNN approaches, the classification and the reconstruction DNNs, an improvement could be possible by performing a hyperparameter optimisation to find the best DNN architecture. Nevertheless it should be kept in mind that the Z boson reconstruction using the invariant-mass approach already reaches high efficiencies. The possible gain when excluding the no-class events is 2.7 out of 32.5 events for the SF region and 1.7 out of 36.6 events for the DF region.

In the final part of the thesis, the classification DNNs are used to perform Asimov fits of the signal strength of $t\bar{t}Z$, $\mu_{t\bar{t}Z}$, and the normalisation of ZZ, N_{ZZ}. Simultaneous fits of both parameters and only of the signal strength $\mu_{t\bar{t}Z}$ using the classifier outputs are compared to fits performed on the distribution of the p_T of the leading lepton with the tight selection applied. The results show that the fit performed on the distributions of the classifier outputs compared to the fit on the p_T distribution can decrease the uncertainties for $\mu_{t\bar{t}Z}$ from 21.5% (16.1%) to 15.0% for the simultaneous (single) fit. While the fit on the DNN distributions gave the same results for both fitting configurations, the single fit on the p_T distribution leads to improved results for $\mu_{t\bar{t}Z}$.

In addition, the statistical uncertainties can be reduced. The statistical uncertainty of $\mu_{t\bar{t}Z}$ is decreased from 20% to 13% for the simultaneous and from 14% to 13% for the single fit, respectively. Nevertheless the statistical uncertainties are the dominant uncertainties in all fits. This also influences the possible constraint of systematics. Only the cross section of ZZ for the single fits of $\mu_{t\bar{t}Z}$ can be constrained.

The SF region with the loose selection applied is dominated by the ZZ contribution. Due to the separation of $t\bar{t}Z$ and ZZ events by the DNN, the distribution of the classifier outputs has separate bins dominated by one of the processes. Therefore a definition of

an additional control region for the measurement of N_{ZZ} is not needed when using the classifier output distributions for the fitting.

Some systematics have not been taken into account yet. This includes the uncertainties on the PDF and the scale choice of ZZ, the uncertainties on the $t\bar{t}Z$ showering and other $t\bar{t}Z$ related systematics. An increased number of NPs would lead to a more sophisticated treatment of the JER and JES systematics. In a similar manner more detailed NP schemes for leptons and jets were not investigated.

Bibliography

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

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

[3] A. Salam, *Weak and Electromagnetic Interactions*, Conf. Proc. C **680519**, 367 (1968)

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

[5] H. Georgi, S. L. Glashow, *Unified weak and electromagnetic interactions without neutral currents*, Phys. Rev. Lett. **28**, 1494 (1972)

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

[7] H. D. Politzer, *Asymptotic freedom: An approach to strong interactions*, Phys. Rep. **14(4)**, 129 (1974)

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

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

[10] G. 't Hooft, M. Veltman, *Regularization and renormalization of gauge fields*, Nucl. Phys. B **44(1)**, 189 (1972)

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

[12] P. W. Anderson, *Plasmons, Gauge Invariance, and Mass*, Phys. Rev. **130** (1963)

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

Bibliography

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

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

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

[17] D. J. Gross, F. Wilczek, *Asymptotically free gauge theories. 1*, Phys. Rev. D **8**, 3633 (1973)

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

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

[20] S. Abachi, et al. (DØ Collaboration), *Observation of the Top Quark*, Phys. Rev. Lett. **74**, 2632 (1995)

[21] M. L. Perl, et al., *Evidence for Anomalous Lepton Production in e^+-e^- Annihilation*, Phys. Rev. Lett. **35**, 1489 (1975)

[22] S. W. Herb, et al., *Observation of a Dimuon Resonance at 9.5 GeV in 400 GeV Proton Nucleus Collisions*, Phys. Rev. Lett. **39**, 252 (1977)

[23] ATLAS, CDF, CMS and DØ Collaborations, *First combination of Tevatron and LHC measurements of the top-quark mass*, arXiv:1403.4427 (2014)

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

[25] CMS Collaboration, *The CMS experiment at the CERN LHC. The Compact Muon Solenoid experiment*, JINST **3**, S08004 (2008)

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

[27] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3**, S08005 (2008)

[28] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79(8)**, 639 (2019)

[29] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **81(7)**, 578 (2021)

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

[31] ATLAS Collaboration, *Tagging and suppression of pileup jets*, ATLAS-CONF-2014-018 (2014)

[32] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, ATL-PHYS-PUB-2017-013 (2017)

[33] W. S. McCulloch, W. Pitts, *A logical calculus of the ideas immanent in nervous activity*, Bull. Math. Biophys. **5**, 115 (1943)

[34] F. Rosenblatt, *The perceptron, a perceiving and recognizing automaton (Project Para)*, Technical Report 85-460-1, Cornell Aeronautical Laboratory (1957)

[35] G. Hinton, et al., *Dropout: A Simple Way to Prevent Neural Networks from Overfitting*, J. Mach. Learn. Res. **15**, 1929 (2014)

[36] S. Ioffe, C. Szegedy, *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*, in *Proceedings of the 32nd International Conference on Machine Learning*, ICML 2015 (Lille, France, July 6–11, 2015), ed. by F. Bach and D. Blei

[37] M. Abadi, et al., *TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems* (2015), software available from tensorflow.org

[38] T. Dozat, *Incorporating nesterov momentum into adam*, in *Proceedings of the 4th International Conference on Learning Representations*, ICLR 2016 (San Juan, Puerto Rico, May 2–4, 2016)

[39] D. Kingma, J. Ba, *Adam: A Method for Stochastic Optimization*, in *Proceedings of the 3rd International Conference on Learning Representations*, ICLR 2015 (San Diego, CA, USA, May 7–9, 2015), ed. by Y. Bengio and Y. LeCun

Bibliography

[40] Y. Nesterov, *A method for solving the convex programming problem with convergence rate $O(1/k^2)$*, Dokl. Akad. Nauk SSSR **269**, 543 (1983)

[41] P.A. Zyla et al. (Particle Data Group), *Review of Particle Physics*, Prog. Theor. Exp. Phys. **2020**, 083C01 (2020)

[42] ATLAS Collaboration, *Measurements of the inclusive and differential production cross sections of a top-quark-antiquark pair in association with a Z boson at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Eur. Phys. J. C **81**, 737 (2021)

[43] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV*, Eur. Phys. J. C **76(5)**, 292 (2016)

[44] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **76(12)**, 653 (2016)

[45] ATLAS Collaboration, *Jet energy scale measurements and their systematic uncertainties in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **96(7)**, 072002 (2017)

[46] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with $t\bar{t}$ events in pp collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **79(11)**, 970 (2019)

[47] ATLAS Collaboration, *Measurement of b-tagging Efficiency of c-jets in $t\bar{t}$ Events Using a Likelihood Approach with the ATLAS Detector*, ATLAS-CONF-2018-001 (2018)

[48] ATLAS Collaboration, *Calibration of light-flavour jet b-tagging rates on ATLAS proton-proton collision data at $\sqrt{s}$ =13 TeV*, ATLAS-CONF-2018-006 (2018)

[49] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data*, JINST **14(12)**, P12006 (2019)

[50] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s} = 13$ TeV*, Eur. Phys. J. C **78(11)**, 903 (2018)

[51] J. Butterworth, et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43**, 023001 (2016)

[52] F. Demartin, et al., *tWH associated production at the LHC*, Eur. Phys. J. C **77(1)**, 34 (2017)

[53] ATLAS Collaboration, *Measurement of the $t\bar{t}Z$ and $t\bar{t}W$ cross sections in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Rev. D **99(7)**, 072009 (2019)

[54] ATLAS Collaboration, *Measurements of the production cross-section for a Z boson in association with b-jets in proton-proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **07**, 044 (2020)

A. Supplementary Material

Table A.1.: Evaluation of the performance of the DNNs in terms of their separation power $\mathcal{S}$ and the AUC of the testing and training folds for the regions defined by the lepton flavour multiplicity.

Region	separation power $\mathcal{S}$	AUC (training)	AUC (testing)
$eeee$	35.9%	0.772 ± 0.016	0.742 ± 0.078
$eee\mu$	10.3%	0.535 ± 0.011	0.522 ± 0.053
$ee\mu\mu$	38.4%	0.832 ± 0.013	0.834 ± 0.042
$e\mu\mu\mu$	8.45%	0.584 ± 0.016	0.546 ± 0.029
$\mu\mu\mu\mu$	39.0%	0.831 ± 0.015	0.815 ± 0.054

Figure A.1.: The classifier outputs of the DNNs for the regions defined via the electron and muon multiplicity.
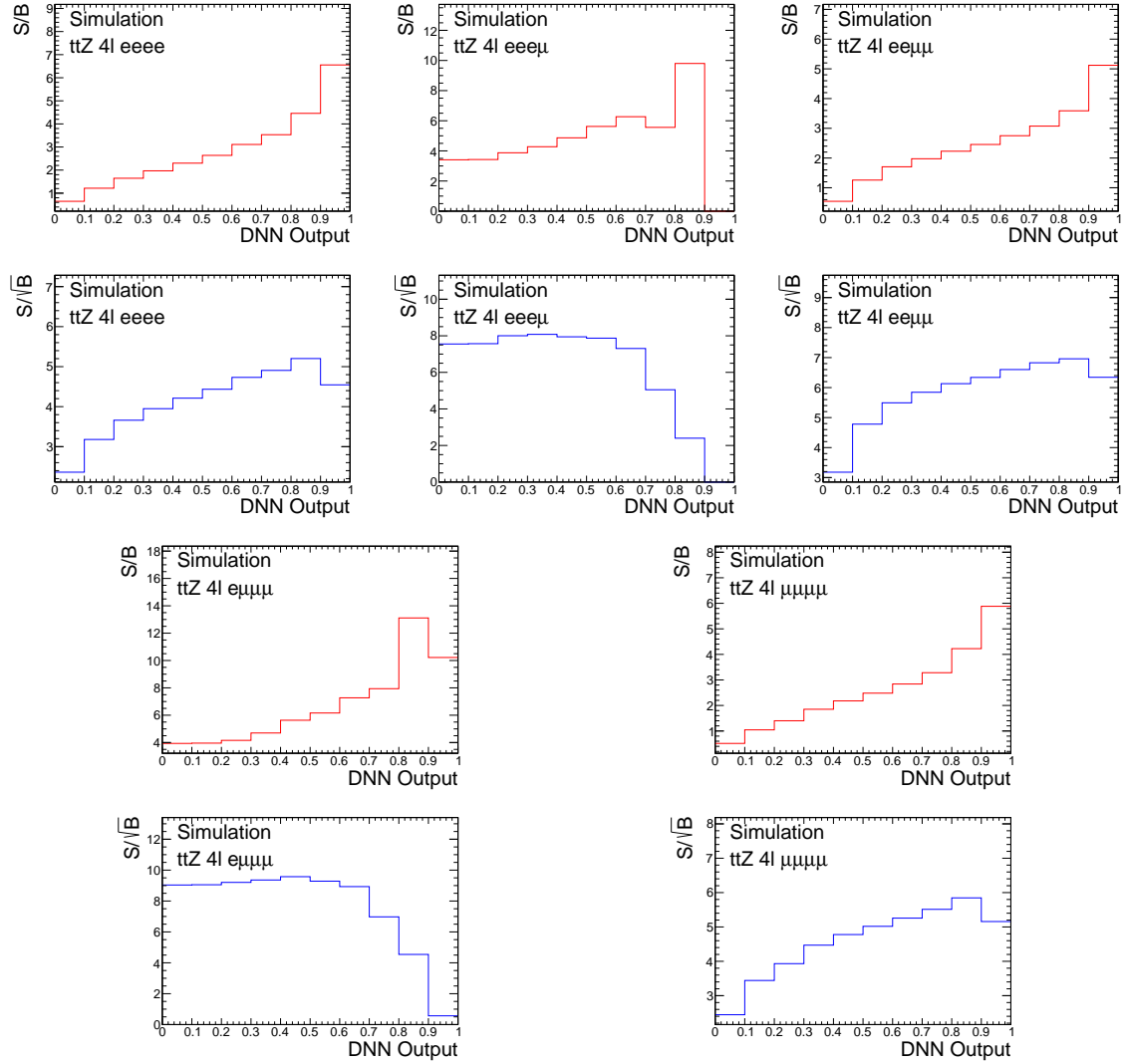
Figure A.2.: The S/B and $S/\sqrt{B}$ ratios depending on the cut applied on the classifier output. All events that would pass the selection if a cut on the given classifier output was applied are used for the calculation of the ratios.

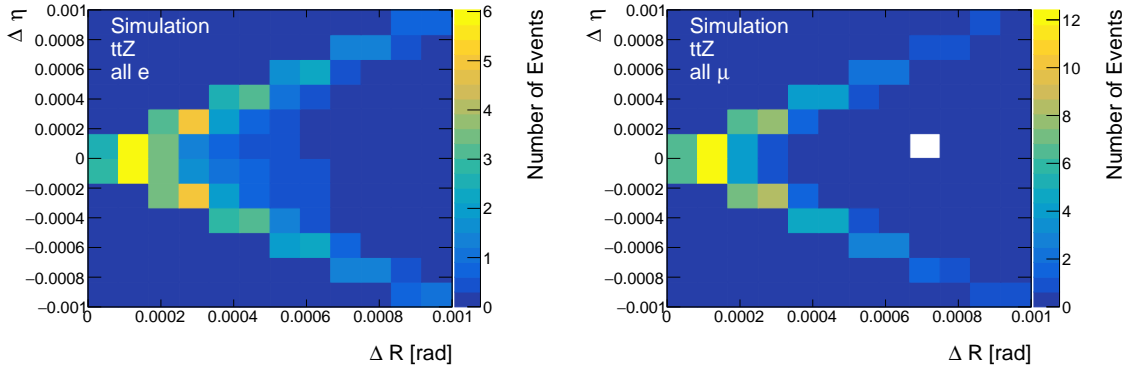***Figure A.3.:*** The distributions of the number of events as a function of ΔR and $\Delta\eta$ between the matched leptons. On the left the distributions for all electrons in all events is shown and on the left the same for all muons.
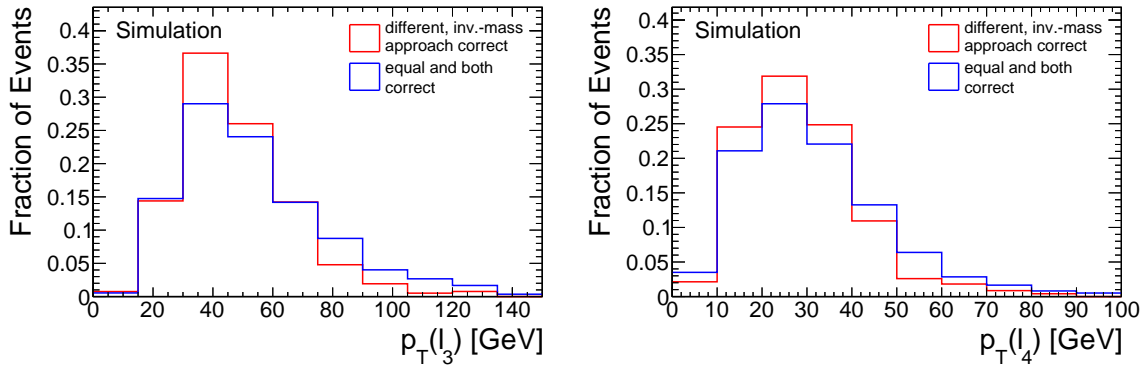

Figure A.4.: The distribution of the transverse momenta p_T of the third on the left and the fourth lepton on the right. The events are split into those where the Z boson reconstruction DNN and the invariant-mass approach provide the same and correct prediction and those where only the invariant-mass approach makes the correct prediction.

Table A.2.: Impact of the systematics on the fits using the one-bin distributions and the tight selection applied.

Uncertainty	impact [%] on measurement of		
	N_{ZZ}	$\mu_{t\bar{t}Z}$ (simultaneous)	$\mu_{t\bar{t}Z}$ (single)
Tagging	5.4	1.2	1.7
Jets	7.6	1.8	2.5
Lepton	4.9	5.6	6.0
Luminosity	1.4	2.2	2.3
E_T^{miss}	4.6	0.1	0.4
$t\bar{t}Z$ PDF	1.4	2.9	2.8
$t\bar{t}Z$ scale choice	2.5	0.3	0.3
Pileup	2.1	1.3	1.1
Background modelling	1.9	1.4	1.4
Trigger	–	0.1	0.1
ZZ cross section (single fit only)	–	–	1.8

Table A.3.: Impact of the systematics on the fits using the four-bin distributions of the leading lepton p_T and the tight selection applied.

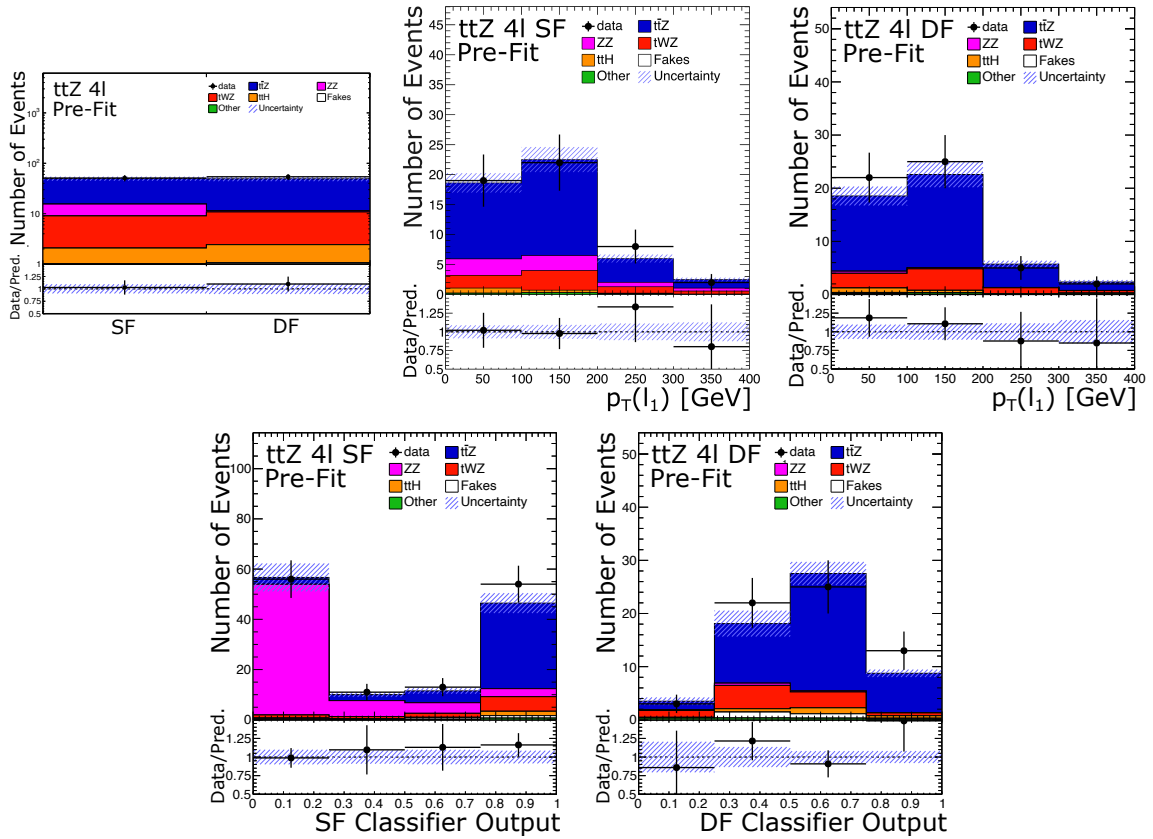| Uncertainty	impact [%] on measurement of		
	N_{ZZ}	$\mu_{t\bar{t}Z}$ (simultaneous)	$\mu_{t\bar{t}Z}$ (single)
Tagging	5.3	1.2	1.7
Jets	7.6	1.8	2.5
Lepton	5.9	5.4	6.0
Luminosity	1.5	2.2	2.3
E_T^{miss}	4.7	0.1	0.3
$t\bar{t}Z$ PDF	1.0	2.8	2.7
$t\bar{t}Z$ scale choice	2.4	0.3	0.3
Pileup	2.0	1.3	1.1
Background modelling	1.6	1.3	1.3
Trigger	0.1	0.1	0.1
ZZ cross section (single fit only)	–	–	1.8

Figure A.5.: The distributions for the different fitting configurations. The upper left plots shows the one-bin distributions and the two plots in the upper right the p_T distributions with the tight selection applied. The lower two plots show the distributions of the classifier outputs with the loose selection. All systematics besides the uncertainties on the cross sections of ZZ+jets are included.
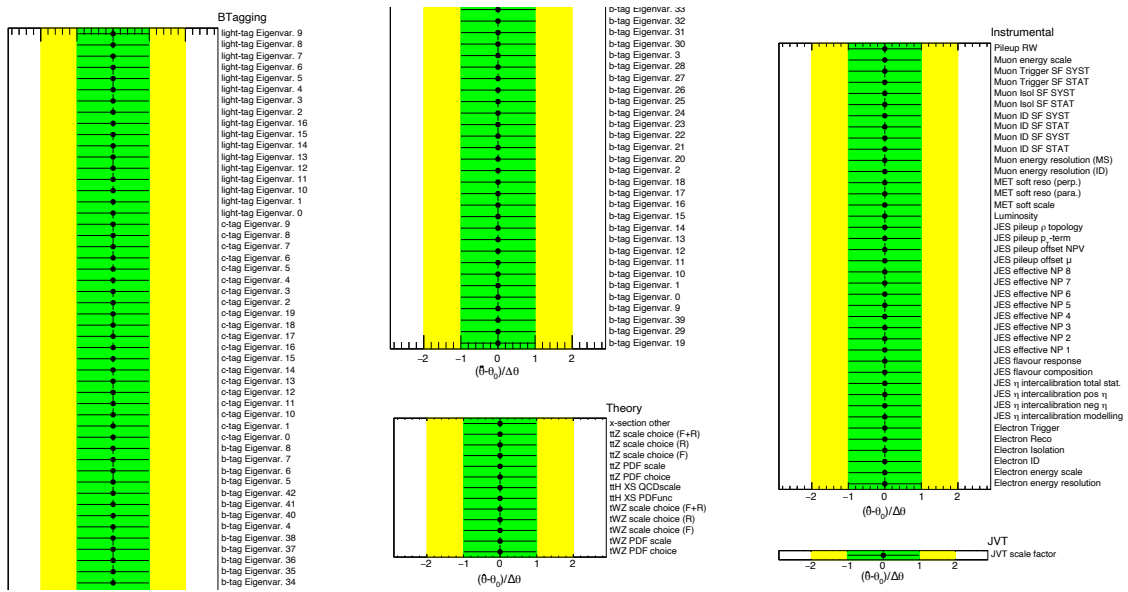
Figure A.6.: The constraints on the NP used in the simultaneous fit of $\mu_{t\bar{t}Z}$. The yellow area shows the 1σ environment and the green area the 2σ environment. The bars show the uncertainty of the NP.
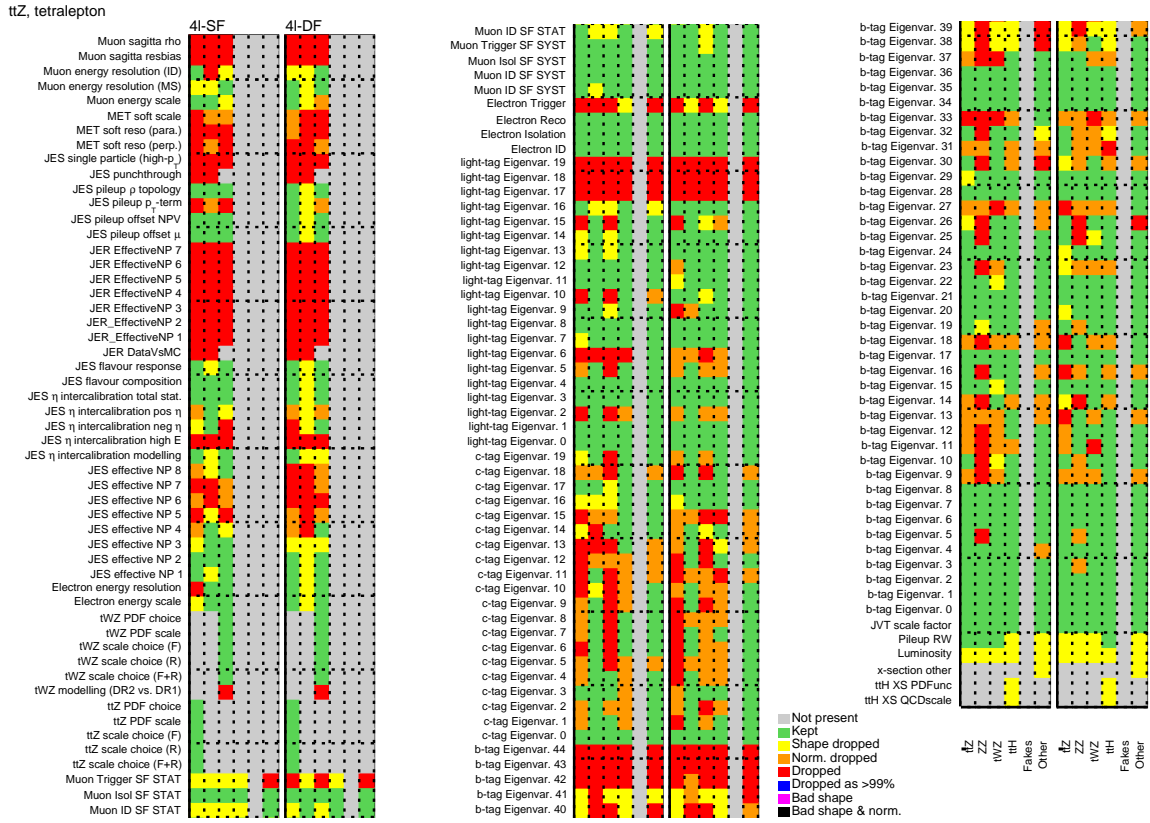
Figure A.7.: The pruning status of every uncertainty used for the simultaneous fit of $\mu_{t\bar{t}Z}$ and N_{ZZ}.
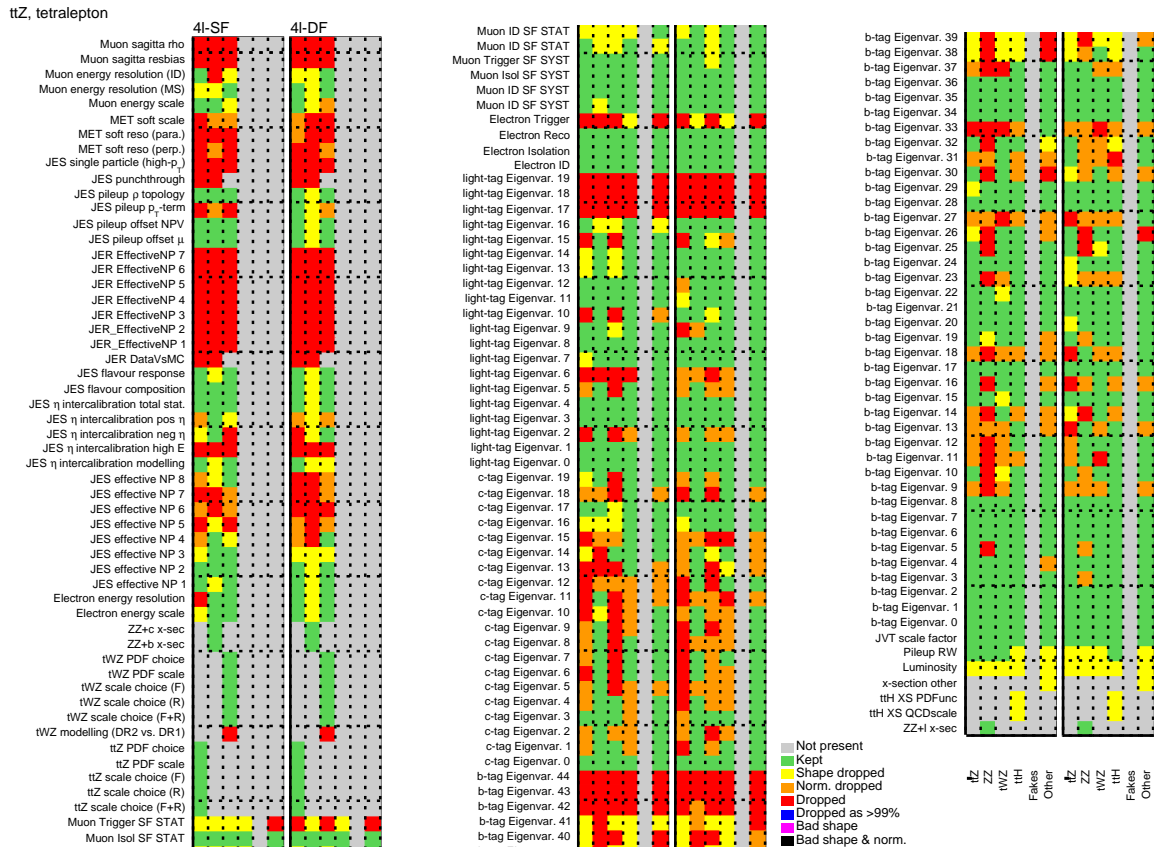
Figure A.8.: The pruning status of every uncertainty used for the single fit of $\mu_{t\bar{t}Z}$.

Acknowledgements

First of all I would like to thank Prof. Dr. Arnulf Quadt for offering the opportunity of working in his working group and the supervision of my master's thesis. It was a pleasure being part of the group. I learned a lot and improved my skills over several discussions during the meetings and facing new challenges every week. In addition, I am thankful for the support regarding my next steps in academia. I also would like to thank Dr. Lisa Shabalina for our weekly discussions and your support and feedback during the whole year. Your input helped me a lot to keep focus on the important things and to understand the scientific workflows. Another important person in the context of this master's thesis was Steffen Korn. Special thanks go to you. You were always there for every kind of questions or problems but also motivated me to solve problems by myself which helped me a lot to evolve over the last year. I would like to thank Dr. Knut Zoch and Dr. Jelena Jovicevic for their input and for supporting the supervision of my thesis. Especially I would like to thank those three people who proofread every chapter of my thesis. I would like to thank Prof. Dr. Stan Lai for agreeing to be the second referee.

All members of the working group who were not only there for help regarding every kind of problems but also always up for social activities and spending time together will make this time a memory that I will look back on happily. It was a pleasure meeting you. Last but not least I would like to thank my family and friends who supported me not only over the last year but during my whole time in Göttingen.

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

Darüberhinaus erkläre ich, dass diese Abschlussarbeit nicht, auch nicht auszugsweise, im Rahmen einer nichtbestandenen Prüfung an dieser oder einer anderen Hochschule eingereicht wurde.

Göttingen, den 25. April 2022

(Tomke Schröer)