GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN IN PUBLICA COMMODA SEIT 1737

Fakultät für Physik

Master's Thesis

Joint Measurement of the $t\bar{t}Z$ and tZq Process with the ATLAS Detector at $\sqrt{s} = 13$ TeV

Gemeinsame Messung des $t\bar{t}Z$ und tZq Prozesses mit dem ATLAS Detektor bei $\sqrt{s} = 13$ TeV

prepared by

Sebastian Alexander Preuth

from Hannover

at the II. Institute of Physics

Thesis number:	II.Physik-UniGöMSc-2023/04
Thesis period:	1st October 2023 until 31st March 2024
First referee:	Prof. Dr. Arnulf Quadt
Second referee:	Prof. Dr. Ariane Frey

Abstract

Die Untersuchung der Kopplung von Top-Quarks an Z-Bosonen spielt eine entscheidende Rolle bei der Validierung des *Standardmodells der Elementarteilchenphysik* (SM) und der Identifizierung potenzieller Abweichungen, die auf Physik jenseits des SM hinweisen könnten. Zwei Prozesse, die eine solche Kopplung aufweisen, sind Top-Quark-Paarproduktion zusammen mit einem Z-Boson ($t\bar{t}Z$) und Produktion einzelner Top-Quarks zusammen mit einem Z-Boson. In dieser Arbeit werden beide Prozesse in einer gemeinsamen Messung im trileptonischen Kanal untersucht. Dazu werden Daten vom Run 2 des ATLAS-Detektors am *Large Hadron Collider* (LHC) verwendet, die bei einer Schwerpunktsenergie von $\sqrt{s} = 13$ TeV produziert wurden. Dieser Datensatz umfasst 140 fb^{-1}.
Es werden verschiedene Techniken genutzt, um Untergrundprozesse von nicht-prompten Leptonen zu minimieren. Anschließend wird ein *tiefes neuronales Netzwerk* (DNN) verwendet, um die Signale von $t\bar{t}Z$-, tZq- und Diboson-Ereignissen zu separieren. Es wird ein Maximum-Likelihood-Fit mit einem Asimov-Datensatz durchgeführt. Das Ergebnis sind die Signalstärken der $t\bar{t}Z$- und tZq-Wirkungsquerschnitte, deren statistische Unsicherheiten mit anderen Analysen vergleichbar sind.

Abstract

The investigation of the coupling of top quarks to Z bosons plays a crucial role in validating the *Standard Model of Elementary Particle Physics* (SM) and the identification of potential deviations that could point to physics beyond the SM. Two processes that exhibit such coupling are top quark pair production associated with a Z boson ($t\bar{t}Z$) and production of single top quarks associated with a Z boson. This work investigates both processes in a joint measurement in the trileptonic channel. Data from Run 2 of the ATLAS detector at the *Large Hadron Collider* (LHC), produced at a centre-of-mass energy of $\sqrt{s} = 13$ TeV, is used for this measurement. This data set comprises 140 fb^{-1}.
Various techniques are used to minimise background processes of non-prompt leptons. Afterwards, a *deep neural network* (DNN) is used to separate the signals of $t\bar{t}Z$, tZq and diboson events. A maximum likelihood fit is performed with an Asimov data set. The results are the signal strengths of the $t\bar{t}Z$ and tZq cross-sections, with statistical uncertainties comparable to other analyses.

Contents

1. Introduction

Physics is the science that studies the most fundamental phenomena of nature. To explain their properties and behaviour, physicists found mathematical concepts describing matter, energy, interactions of elementary particles, space and time. Many see Galileo Galilei as the father of physics. He was followed, for example, by Isaac Newton, who formulated the law of gravitation and Albert Einstein, the founder of the general and special theory of relativity [1, 2]. In the field of nuclear and atomic physics, there were renowned physicists such as Marie Curie, who researched radioactivity [3], Ernest Rutherford, who discovered in a scattering experiment that atoms have a nucleus [4], and many others. In the 20$^{\text{th}}$ century, the *Standard Model of Particle Physics* (SM) [5–16] was developed. It is a theory describing three of the four known fundamental forces and the elementary particle spectrum. Nevertheless, it is incomplete. Phenomena like Dark Matter [17, 18] or neutrino oscillations [19–23] are not described by the SM. In addition, there is no accurate theory describing all four fundamental forces. Gaining a more profound understanding of the necessary expansion of the SM to incorporate these phenomena is essential.

In this work, two processes are investigated jointly: the production of top quark pairs and single top quarks, both occurring in association with a Z boson. These processes are referred to as $t\bar{t}Z$ and tZq. They are essential for studying the electroweak sector and potential deviations from SM predictions.

The joint measurement is performed using the Run 2 dataset and an analysis model compatible with a future Run 3 measurement. The dataset corresponding to a centre-of-mass energy of $\sqrt{s} = 13\,\text{TeV}$ and an integrated luminosity of $140\,\text{fb}^{-1}$ is used to extract the $t\bar{t}Z$ and tZq production signal strengths. The analysis targets the trilepton final state. To improve the background estimation, which is dominated by diboson production, *deep neural networks* (DNNs) are used.

2. The Standard Model of Particle Physics

This chapter introduces the fundamentals of the *Standard Model of Particle Physics* (SM), including a description of the particle spectrum and the fundamental forces. Subsequently, the main properties of the top quark are presented, and the top quark production processes relevant to this work are described.

2.1. The Standard Model

The SM is a quantum field theory with the gauge symmetry $SU(3)_C \times SU(2)_L \times U(1)_Y$. The $U(1)$ is a unitary group over a 1-dimensional Hilbert space. The $SU(n)$ is the special unitary group of degree n. Important for the SM are the cases $n = 2$ and $n = 3$. The indices denote the colour charge C, the coupling to only left-handed fermions L and the hypercharge Y. The gauge groups correspond to the respective fundamental interactions described by the SM. This is the strong interaction for the $SU(3)_C$ and the electroweak interaction for the $SU(2)_L \times U(1)_Y$. The latter unifies the weak and the electromagnetic force [5–7]. The fourth and last known fundamental force is gravity, which the SM does not explain.

Particle Spectrum

The SM describes all known elementary particles. These include five spin-1 vector bosons (gluon, photon, Z boson and $W^{\pm}$ boson), the Higgs boson, six quarks (up, down, charm, strange, top and bottom) and six leptons. Each vector boson is a mediator of its corresponding fundamental force. The gluon mediates the strong force, the photon mediates the electromagnetic force, and the massive Z and $W^{\pm}$ bosons mediate the weak force. The Higgs boson results from the existence of the Higgs field. Figure 2.1 shows an overview of the elementary particles of the SM.
Quarks and leptons have a half-integer spin. Thus, they are fermions. In the case of the leptons, a distinction is made between three charged leptons: the electron (e^{-}), the muon

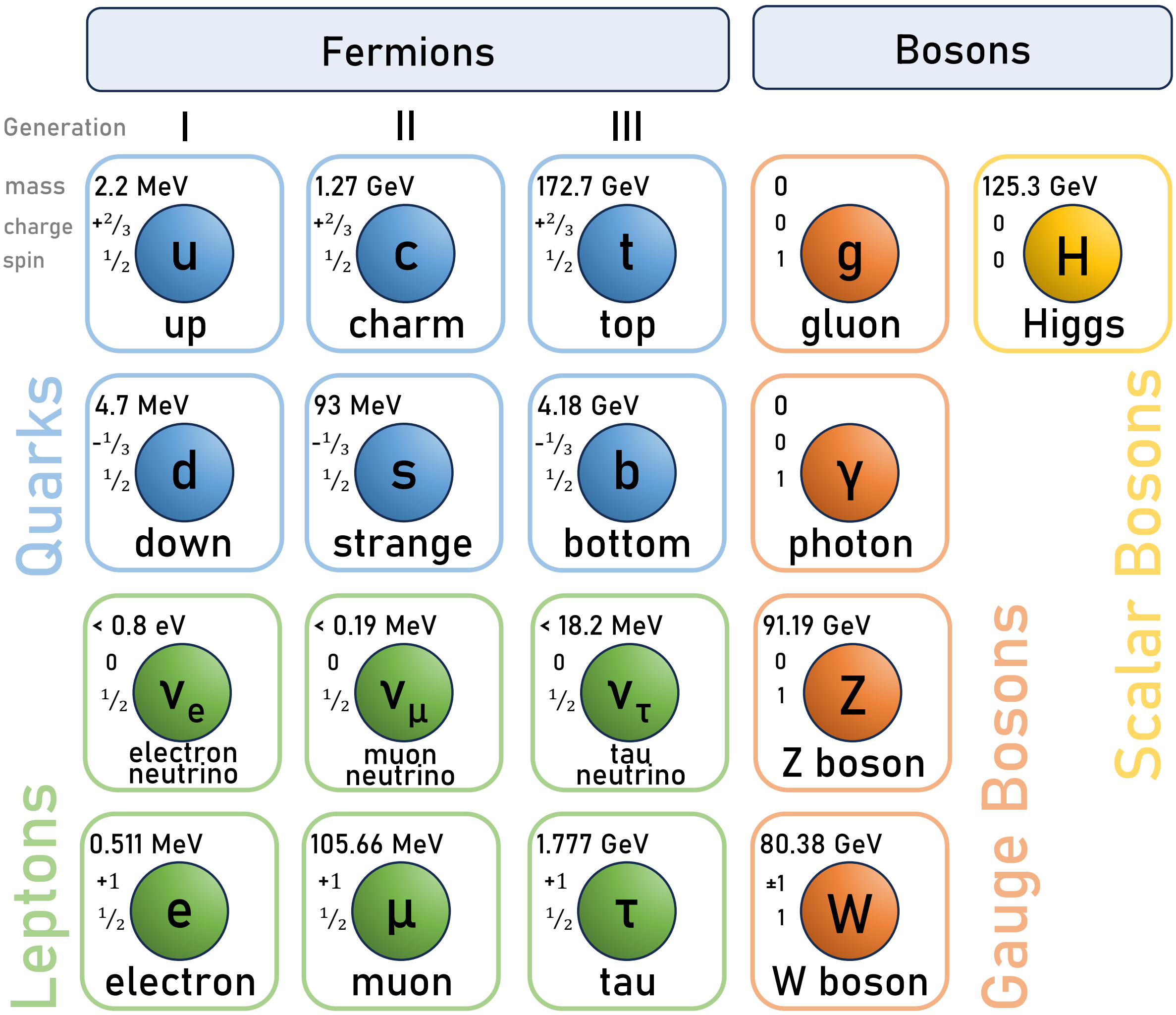


Figure 2.1.: Overview of the particle spectrum of the Standard Model, with the three fermion generations on the left side and the bosons on the right side. At the top left of each particle, its mass, electric charge, and spin are shown. The mass values are taken from Ref. [24].

(μ^-), and the tau (τ^-) and three electrically neutral neutrinos ν_e, ν_μ, and ν_τ. While neutrinos only interact via the weak force, the charged leptons additionally couple to the electromagnetic interaction. Quarks couple to all three fundamental forces described by the SM and occur with two different electrical charges $q_u = +^2\!/_3\,e$ and $q_d = -^1\!/_3\,e$, depending on whether they are an up- or a down-type quark. In addition, they have red, green, and blue colour charges.

Furthermore, there is the Higgs boson with spin-0. It interacts with every elementary particle that has a mass. These are the quarks, the charged leptons, the $W^\pm$ bosons, the Z boson, and the Higgs boson itself.

Local Gauge Invariance

The Dirac Lagrangian

$$\mathcal{L}_{\text{Dirac}} = \underbrace{i\overline{\psi}\gamma^{\mu}\partial_{\mu}\psi}_{\text{kin. term}} - \underbrace{m\overline{\psi}\psi}_{\text{mass term}} \tag{2.1}$$

is the Lagrangian density of a fermion field ψ and its Hermitian conjugate $\overline{\psi}$. The matrix γ^{μ} stands for the μ-th Dirac matrix with $\mu = 0, ..., 3$. It is composed of a kinetic and a mass term. The Lagrangian, as it is, is not invariant under a local $U(1)$ phase transformation. For this, an additional gauge term $\mathcal{L}_{\text{GT}} = -q\left(\overline{\psi}\gamma^{\mu}\psi\right)B_{\mu}$ must be added, which is called the interaction term. Here, q is a constant factor. B_{μ} is a vector field which transforms to $B_{\mu} \rightarrow B_{\mu} + \partial_{\mu}\lambda(x)$ and couples to the spinor fields. The Lagrangian density of the spin-1 gauge field is

$$\mathcal{L}_{\text{spin}-1} = \underbrace{-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}}_{\text{kin. term}} + \underbrace{\frac{1}{2}m_B^2 B^{\nu}B_{\nu}}_{\text{mass term}} \ . \tag{2.2}$$

The electromagnetic field strength tensor $F^{\mu\nu}$ is defined as $F^{\mu\nu} = \partial^{\mu}B^{\nu} - \partial^{\nu}B^{\mu}$, where m_B is the mass of the field B_{μ}. To remain invariant under local gauge transformation, the field must be massless. Physically, this is interpreted as a photon field. Combining all these terms, we obtain the Lagrangian density of *Quantum Electrodynamics* (QED)

$$\mathcal{L}_{\text{QED}} = i\overline{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\overline{\psi}\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \left(q\overline{\psi}\gamma^{\mu}\psi\right)B_{\mu} \,. \tag{2.3}$$

Similar calculations can be done for a local $SU(2)$ transformation. Here, the gauge field $\vec{W}_{\mu} = \left(W_{\mu}^1, W_{\mu}^2, W_{\mu}^3\right)^T$ is the vector of the three weak isospin fields W_{μ}^i. The Lagrangian

$$\mathcal{L} = i\overline{\psi}\gamma^{\mu}\partial_{\mu}\psi - \frac{1}{4}\vec{W}^{\mu\nu}\vec{W}_{\mu\nu} - \frac{q}{2}\left(\overline{\psi}\gamma^{\mu}\vec{\sigma}\psi\right)\vec{W}_{\mu} \tag{2.4}$$

includes the Pauli matrices $\vec{\sigma}$ and the tensor $W_i^{\mu\nu} = \partial^{\mu}W_i^{\nu} - \partial^{\nu}W_i^{\nu} - q\epsilon_{ijk}W_j^{\mu}W_k^{\nu}$.

The Electroweak Unification

The projection operators $P_{\text{L/R}} = (\mathbb{1} \mp \gamma^5)/2$ can project a spinor into its left-handed (LH) and right-handed (RH) components, where $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ is the chirality operator. For relativistic particles, the chirality equals the helicity, which is the projection of the particles' spin on its momentum. Wu and Goldhaber discovered that the charged currents of the weak force only couple to LH particles and RH antiparticles. Thus, it is maximally parity-violating [25, 26]. The neutral currents are also parity-violating, but couple to both

LH and RH particles. However, their coupling to LH particles is stronger.
The weak isospin T is a quantum number that is needed to describe the coupling between particles and the weak force. Whenever the weak isospin is mentioned in this thesis, it refers to its third component T_3. Since the charged currents of the weak interaction do not couple to RH particles, their weak isospin is $T_3 = 0$. Neutrinos and up-type quarks have an isospin of $T_3 = +1/2$. Leptons with negative electric charge and down-type quarks have $T_3 = -1/2$. For antiparticles, the weak isospin has the respective opposite sign. Thus, the lepton and the quark pairs form weak isospin doublets:

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

The down-type quarks are marked with a prime to indicate a weak eigenstate, not a mass eigenstate. The weak eigenstates are combinations of mass eigenstates of different quarks. This relation is given by the Cabibbo–Kobayashi–Maskawa (CKM) matrix [27, 28]

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \underbrace{\begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix}}_{\text{CKM matrix}} \begin{pmatrix} d \\ s \\ b \end{pmatrix} . \tag{2.5}$$

Due to this relation, quarks can couple to quarks from different generations, although these couplings are suppressed. They can only occur with the exchange of a $W^\pm$ boson and are proportional to $|V_{ij}|^2$.
The electroweak interaction unifies the weak and the electromagnetic interaction. Thus, it obeys an $SU(2)_L \times U(1)_Y$ symmetry. The theory was first introduced by Glashow, Weinberg and Salam [5–7]. While the $SU(2)_L$ gauge fields $\vec{W}_\mu$ couple to the weak isospin, the $U(1)_Y$ gauge field B_μ couples to the weak hypercharge $Y_W = 2(Q - T_3)$. This quantum number relates the weak isospin and the charge Q of a particle in terms of the elementary charge e. The mixing of the gauge fields leads to four physical fields. These are those of the $W^\pm$ bosons

$$W^\pm_\mu = \frac{1}{\sqrt{2}} \left(W^1_\mu \mp i W^2_\mu \right) , \tag{2.6}$$

those of the photon A_μ and the Z boson Z_μ:

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_\mathrm{W} & \sin\theta_\mathrm{W} \\ -\sin\theta_\mathrm{W} & \cos\theta_\mathrm{W} \end{pmatrix} \begin{pmatrix} B_\mu \\ W^3_\mu \end{pmatrix} . \tag{2.7}$$

Thus, the physical fields for the neutral bosons are obtained by rotating the $\left(B_\mu, W^3_\mu\right)$ plane by the Weinberg angle θ_W. This leads to the physical electromagnetic and weak neutral current

$$j^\mu_\mathrm{elm} = Qe\bar{\psi}\gamma^\mu\psi \tag{2.8}$$

$$j^\mu_Z = \frac{1}{2}g_Z\bar{\psi}\gamma^\mu(c_V - c_A\gamma^5)\psi\,, \tag{2.9}$$

where the coupling to the Z boson $g_Z = g_W/\cos\theta_\mathrm{W} = e/(\cos\theta_\mathrm{W}\sin\theta_\mathrm{W})$ is directly connected to the $W^\pm$ boson coupling g_W, the elementary charge and the Weinberg angle. The vector and axial-vector couplings

$$c_V = T_3 - 2Q\sin^2\theta_W \tag{2.10}$$

$$c_A = T_3 \tag{2.11}$$

provide information about the amount of parity violation. It is maximal for the coupling to neutrinos. The $V - A$ structure is also reflected in the corresponding part of the SM Lagrangian

$$\mathcal{L}_\mathrm{SM} \supset \bar{\psi}\left(\gamma^\mu(c_V - c_A\gamma^5)\right)\psi Z_\mu\,, \tag{2.12}$$

which describes the coupling of fermions to the Z boson.
The electroweak Lagrangian

$$\mathcal{L}_\mathrm{EW} = -\frac{1}{4}\vec{W}^{\mu\nu}\vec{W}_{\mu\nu} - \frac{1}{4}B^{\mu\nu}B_{\mu\nu} + i\sum_{\psi=\ell,q}\overline{\psi}\gamma^\mu D_\mu\psi + \mathcal{L}_\mathrm{h} + \mathcal{L}_\mathrm{y} \tag{2.13}$$

contains no mass terms for the gauge bosons. These would spoil the symmetry. Instead, they obtain their mass via coupling to the Higgs field, described by the Higgs mechanism and electroweak symmetry breaking. This is why $\mathcal{L}_\mathrm{EW}$ contains a Higgs term $\mathcal{L}_\mathrm{h}$ and a Yukawa term $\mathcal{L}_\mathrm{y}$. The Higgs mechanism, introduced by Brout and Englert [29], by Higgs [30] and by Guralnik, Hagen and Kibble [31], introduces a complex scalar field ϕ called the Higgs field. The gauge bosons receive their mass by coupling to this field. The Higgs Lagrangian contains the Higgs potential

$$V(\phi) = -\mu^2\phi^\dagger\phi + \lambda\left(\phi^\dagger\phi\right)^2 \tag{2.14}$$

with the parameters μ and λ. For $\lambda > 0$ and $\mu^2 < 0$, the potential has the shape of a Mexican hat leading to a degenerated ground state $|\langle 0|\phi|0\rangle| = \sqrt{\frac{\mu^2}{2\lambda}}$. From this point, the potential is no longer symmetric. Expanding the potential from its ground state leads to new fields and, in the end, to the existence of the Higgs boson. The Higgs boson was discovered in 2012 by the ATLAS and CMS collaborations [32, 33].

The Strong Interaction

The strong interaction is a fundamental force described by the theory of *Quantum Chromodynamics* (QCD), which is a $SU(3)_C$ gauge theory. The gauge group couples to the colour charge C which occurs in three distinct states: red (r), blue (b) and green (g), accompanied by their corresponding anti-colour charges ($\bar{\mathrm{r}}$, $\bar{\mathrm{b}}$, $\bar{\mathrm{g}}$). The $SU(3)_C$ has eight generators, which are the Gell-Mann matrices λ^a, where $a = 1, \ldots, 8$. Thus, there are eight gluon fields G^a_μ, corresponding to eight gluons. They are represented as combinations of the eight colour states:

$$G^a_\mu = rg,\, g\bar{r},\, r\bar{b},\, br,\, g\bar{b},\, bg,\, \frac{1}{\sqrt{2}}(r\bar{r} - g\bar{g}),\, \frac{1}{\sqrt{6}}(r\bar{r} + g\bar{g} - b\bar{b}). \tag{2.15}$$

The QCD Lagrangian

$$\mathcal{L}_{\mathrm{QCD}} = \sum_{\psi = q} \bar{\psi}\left(i\gamma^\mu D_\mu - m_\psi\right)\psi - \frac{1}{4} G^{\mu\nu a} G^a_{\mu\nu} \tag{2.16}$$

describes the dynamics of quarks and gluons, where $\sum_{\psi=q}$ stands for the sum over the quark fields and $D_\mu = \partial_\mu + \frac{ig_s}{2}\lambda_a G^\mu_a$ is the gauge covariant derivative. The gluon field strength tensor is given by

$$G^a_{\mu\nu} = \partial^\mu G^a_\nu - \partial_\nu G^a_\mu - g_s f_{abc} G^b_\mu G^c_\nu \tag{2.17}$$

with the structure constants f_{abc}. The strong coupling parameter is denoted with g_s.
In contrast to other forces like the electromagnetic force or gravity, the strong force increases with increasing distance. This phenomenon is known as confinement and implies that quarks remain permanently confined within colour-neutral hadrons. This could be a colour anti-colour pair or a triplet of three different (anti-)colours. Tetra- and pentaquarks are also possible [34–36]. When quarks are separated, they exchange virtual gluons until they form hadrons. This process is known as hadronisation and results in the formation of hadronic jets, which are high-energy showers of hadrons.
Nevertheless, the coupling strength of quarks decreases with increasing energy scale q^2. Thus, quarks interact weakly if q^2 is large. This behaviour is called asymptotic freedom and was first discovered by Gross, Wilczek and Politzer [12, 13]. The energy-dependent

strong coupling constant can be approximated by

$$\alpha(q^2) = \frac{\alpha(\mu^2)}{1 - \frac{\alpha(\mu^2)}{3\pi}\ln\left(\frac{q^2}{\mu^2}\right)}, \tag{2.18}$$

where μ^2 is a reference energy scale. The small coupling at large energy scales allows perturbative calculations.

2.2. The Top Quark

The top quark is the quark that was discovered last. This was achieved in 1995 by the CDF and DØ [37, 38] collaborations at the Tevatron. It is the heaviest known elementary particle. Its mass of 172.69 ± 0.30 GeV [24] is 42 times larger than the mass of the bottom quark ($m_b = 4.18^{+0.03}_{-0.02}$ GeV [24]), which is the next heaviest quark. This leads to a tiny lifetime of $\tau_t \approx 5 \times 10^{-25}$ s. Since the average time of hadronisation is $\tau_{\text{had}} \approx 3 \times 10^{-24}$ s, it decays before it can hadronise. In over 99% of cases, the decay products are a $W^{\pm}$ boson and a bottom quark. The $W^{\pm}$ boson from the top quark decay decays either leptonically, i.e. into a charged lepton and its respective neutrino, or hadronically, i.e. into an up-type and a down-type quark. The Feynman diagrams of a hadronic and a leptonic top quark decay are shown in Figure 2.2

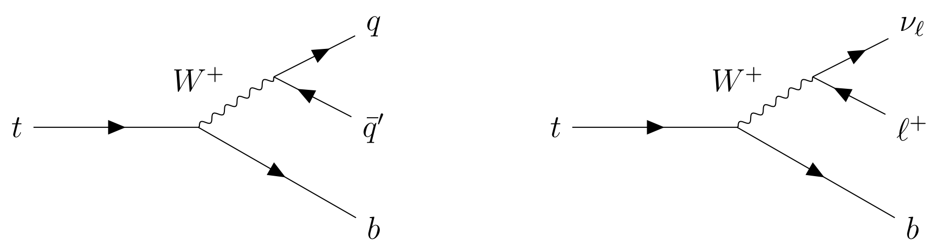


Figure 2.2.: Feynman diagrams of a leptonic (left) and hadronic top quark decay.

Top Quark Pair Production at the LHC

Electron-positron colliders have never reached enough centre-of-mass energy to produce top-antitop quark pairs ($t\bar{t}$). Due to the high centre-of-mass energy required to produce $t\bar{t}$ pairs, they can only be produced in hadron colliders like the *Large Hadron Collider* (LHC) with the current accelerator technology. Therefore, top quarks are mainly produced via the strong interaction. In a top quark pair production, a distinction between the alljets (46%), the lepton + jets (45%) and the dilepton channel (9%) is made. These

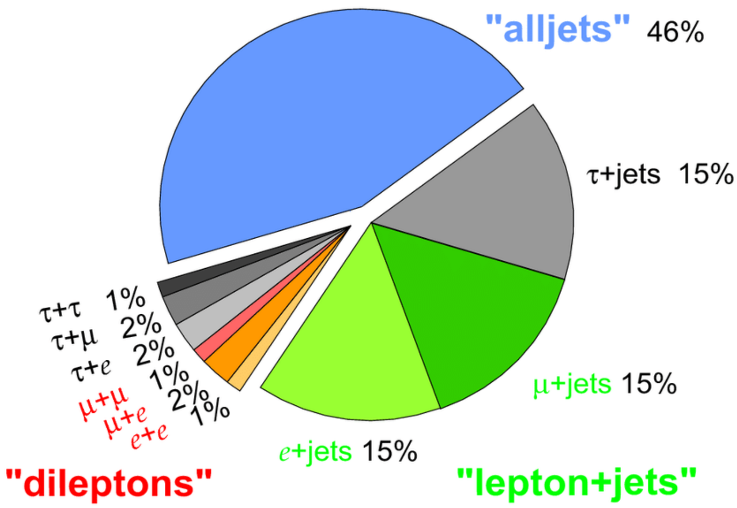


Figure 2.3.: Pie chart of the branching ratios of the $W^{\pm}$ boson decay channels resulting from a $t\bar{t}$ decay.

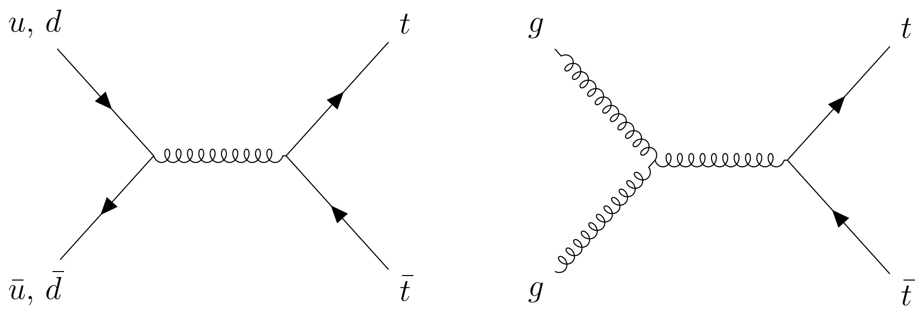


Figure 2.4.: Top quark pair production via a quark-antiquark pair annihilation (left) and gluon fusion (right). For gluon fusion, t-channel diagrams also contribute to the $t\bar{t}$ production.

channels correspond to the respective decay products of the $W^{\pm}$ boson from the top quark decay. Figure 2.3 shows a pie chart of the branching ratios of these channels. Finally, the top quark is reconstructed by measuring the four momenta of its decay products. Since neutrinos are not directly detected but still can be part of the top quark decay, they are reconstructed using missing transverse momentum p_{T}. This is explained in Chapter 3.3. In Run 2 of the LHC, top quarks are produced with proton-proton collisions at a centre-of-mass energy of $\sqrt{s} = 13$ TeV. Top quark pairs are mostly produced via gluon fusion. But also quark-antiquark annihilation can lead to a $t\bar{t}$ pair. Feynman diagrams of both channels are shown in Figure 2.4. Naively, it could be assumed that protons are composed of two up quarks and one down quark. But these are only the valence quarks. Protons also carry sea partons like other quarks, antiquarks and gluons. The probability density of finding a certain parton with a momentum fraction x in a proton is described by the

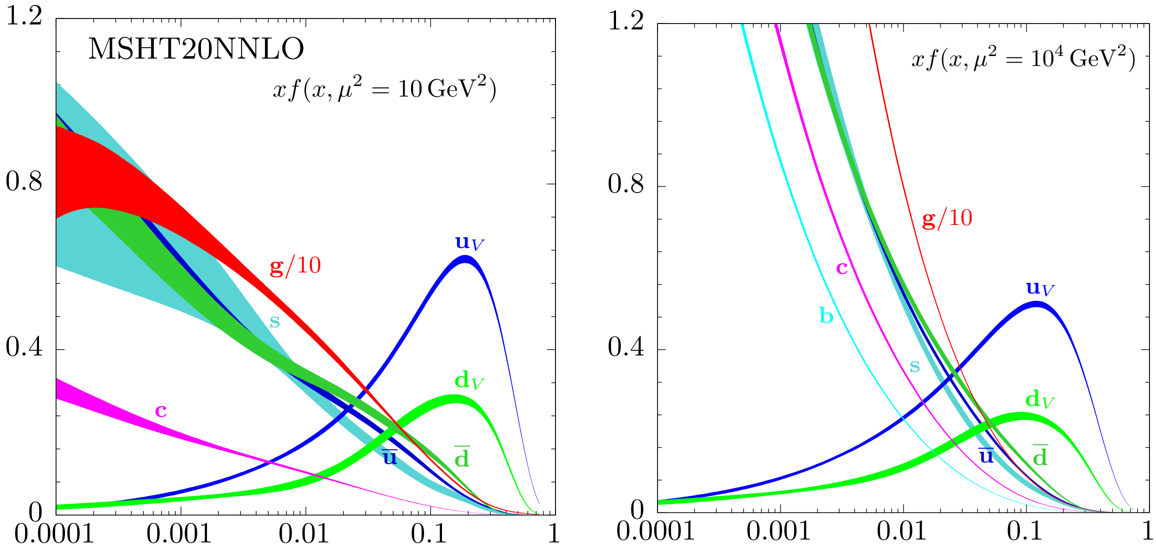


Figure 2.5.: Parton distribution function $f(x, \mu^2)$ multiplied by x of several partons in a proton as a function of the momentum fraction x evaluated at $\mu^2 = 10$ GeV^2 (left) and $\mu^2 = 10^4$ GeV^2. The valance quarks are denoted with V. The figures show the functions obtained in the MSHT20 global analysis at next-to-next-to-leading order (NNLO) [51].

parton distribution functions (PDFs) of a proton [39–51]. Figure 2.5 shows them for several partons in a proton at two different energy scales μ. At high energies like at the LHC, it is sufficient to have partons which do not carry the total momentum of the proton to produce $t\bar{t}$ pairs. In this region, an interaction between two gluons is more probable than an interaction between a quark and an antiquark. This is why gluon fusion plays the main role at the LHC with approximately 90 % of the $t\bar{t}$ production [24].

Top Quark Pair Production in Association with a Z Boson

At the LHC, the centre-of-mass energy is large enough to produce a top quark pair associated with a Z boson. This can be radiated by any quark, lepton, or $W^{\pm}$ boson involved in the $t\bar{t}$ production or decay. Of particular interest is the direct coupling between a top quark and a Z boson. The main production channels at the LHC of this process are shown in Figure 2.6. The direct tZ coupling is sensitive to the vector and axial components introduced in Equations 2.9, 2.10 and 2.11. These are directly connected to the weak isospin and the charge of the top quark. Investigating *top quark pair production in association with a Z boson* ($t\bar{t}Z$) presents an opportunity to test the Standard Models' prediction for this coupling. Any deviations between the measurement and the SM prediction could indicate contributions beyond the SM that alter the coupling strength and structure.

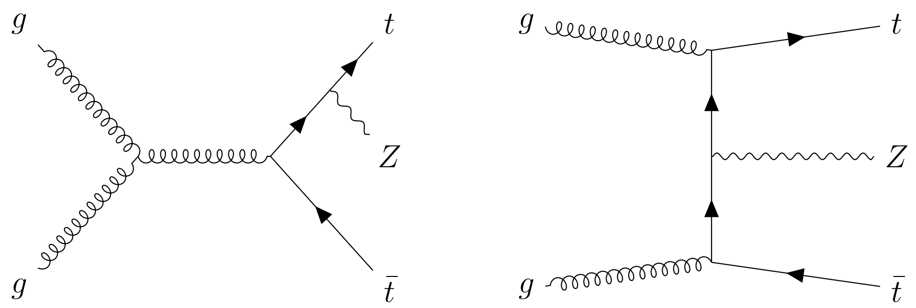


Figure 2.6.: Top quark pair production associated with a Z boson via gluon fusion. The left diagram is the dominant one and has a direct coupling of a top quark and a Z boson. The right diagram has a minor contribution, and the Z boson is produced via $W^{\pm}$ fusion.

Table 2.1.: Illustration of the decay chain in the trileptonic channel of the $t\bar{t}Z$ process. It shows the process shown in Figure 2.6 on the left

<table>
<tr><th>state</th><th colspan="8">particles</th></tr>
<tr><td>$t\bar{t}$ pair</td><td colspan="5">t</td><td colspan="3">$\bar{t}$</td></tr>
<tr><td>after tZ coupling</td><td colspan="3">t</td><td colspan="2">Z</td><td colspan="3">$\bar{t}$</td></tr>
<tr><td>primary decay products</td><td rowspan="2">b</td><td colspan="2">W^+</td><td rowspan="2">ℓ^+</td><td rowspan="2">ℓ^-</td><td colspan="2">W^-</td><td rowspan="2">$\bar{b}$</td></tr>
<tr><td>trilepton final state</td><td>ℓ'^+</td><td>$\nu_{\ell'}$</td><td>$\bar{q}_u$</td><td>q_d</td></tr>
</table>

This analysis does not investigate all final states of the $t\bar{t}Z$ process. This is due to the immense hadronic background in analysing proton-proton collisions. The analysed channel is the trilepton channel. Thus, the Z boson decays into two charged leptons $\ell^{\pm}$. Note that whenever this report refers to charged leptons, it means muons or electrons unless stated otherwise. Tau leptons have a very short lifetime, so the detector only detects their decay products. If they decay leptonically, there are two neutrinos. If they decay hadronically, there is one neutrino and two hadronic jets. Both would make a reconstruction of the $\tau^{\pm}$ extremely difficult. In the trileptonic channel, one $W^{\pm}$ decays leptonically, the other hadronically. Thus, the final state particles are two bottom quarks, three charged leptons, one neutrino and two quarks. For a better overview, an example of the decay chain of the trileptonic channel is given in Table 2.1.

Single Top Quark Production in Association with a Z Boson

If only one top quark is produced, it is called single top quark production. This process was already analysed by the DØ and CDF collaborations [52, 53], which were experiments

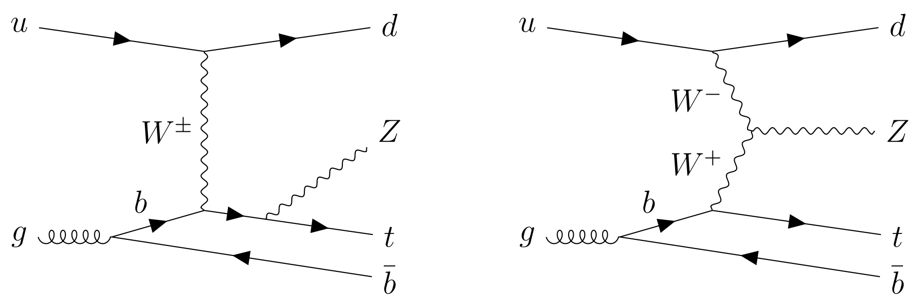


Figure 2.7.: Single top quark production associated with a Z boson via gluon fusion. The s-channel (left) and the t-channel (right) are shown.

at the proton-antiproton collider Tevatron. In addition to single top quark production, the ATLAS and CMS collaborations also measure *single top quark production in association with a Z boson* [54–57]. This process is referred to as tZq and happens when one of the particles radiates a Z boson. It is mainly produced in interactions of an up and a bottom quark. The latter can be a product of a gluon splitting. In addition to the top quark and the Z boson, another quark is produced in this process. Figure 2.7 shows two possible Feynman diagrams. As visible from the diagram, this is an electroweak process. Therefore, it is rarer than the $t\bar{t}Z$ production.

Since the probability of finding a gluon in a parton is very high for a small momentum fraction, the gluons in such interactions are typically soft. Moreover, the dominant tZq production is a t-channel, which is a scattering process. Therefore, the light quark in the final state usually occurs in the forward direction and is detected as a forward jet.

For the same reasons mentioned in Chapter 2.2, hadronic final states are not investigated in the framework of this analysis. Instead, the trilepton channel is analysed. The Z boson decays leptonically in the trilepton channel, similar to the top quark. Therefore, the final state includes a bottom quark, three charged leptons, a neutrino, and another quark. The other bottom quark resulting from gluon splitting typically has a high pseudorapidity $|\eta|$. As a result, it may not always be detected or b-tagged due to its close distance to the beamline.

3. Experimental Setup

This chapter introduces the experimental setup, including a description of the LHC, the ATLAS detector, and the object definitions.

3.1. The Large Hadron Collider

The *Large Hadron Collider* (LHC) [58] is the largest synchrotron on Earth. It is located at the *European Organization for Nuclear Research* (CERN). Figure 3.1 shows an overview of the accelerator complex, including several experiments, detectors and preaccelerators. Lying 100 m under the surface, the accelerator has a circumference of 26.7 km. To achieve a centre-of-mass energy of 13 TeV, the particles pass smaller accelerators like the Super Proton Synchrotron (SPS). Therefore, they can enter the LHC with an energy of 450 GeV. In the LHC, dipole magnets bend the beam, while quadrupole magnets focus it.
There are four interaction points where particles collide. At each of them, there is one of the LHCs' four main detectors: ATLAS, CMS, LHCb and ALICE [59–62]. While ATLAS, CMS and LHCb investigate proton-proton collisions, ALICE analyses quark-gluon plasma produced by lead-lead and lead-proton collisions. LHCb is a forward detector and studies the behaviour of B mesons and CP-violation. ATLAS and CMS are 4π detectors. This means they cover an area around the interaction point that is as large as technically possible. They are thus designed to observe all possible decay products. In Run 2, they analysed proton-proton collisions with $\sqrt{s} = 13$ TeV.

3.2. The ATLAS Detector

ATLAS stands for **A** **T**oroidal **L**HC **A**pparatu**S**. With a height of 25 m, a length of 44 m and a weight of approximately 7000 tonnes, it is the largest detector at the LHC. A schematic overview is shown in Figure 3.2.
To reconstruct the properties of particles, ATLAS measures their momentum, energy, and charge. The particles' mass can be calculated from the momentum and energy. The transverse momentum p_{T}, which is the projection of the momentum of a particle on the

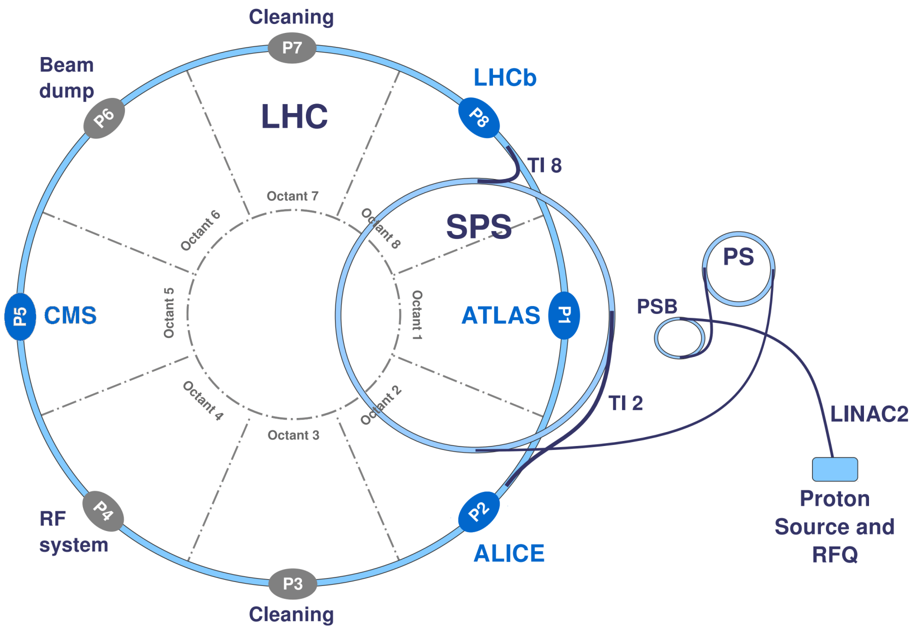


Figure 3.1.: Schematic overview of the LHC and the preaccelerators, located at the CERN accelerator complex. The preaccelerators are the Radio Frequency Quadrupole (RFQ), the linear accelerator (LINAC2), the Proton Synchrotron Booster (PSB), the Proton Synchrotron (PS) and the Super Proton Synchrotron (SPS). ©Jasmin Gruschke

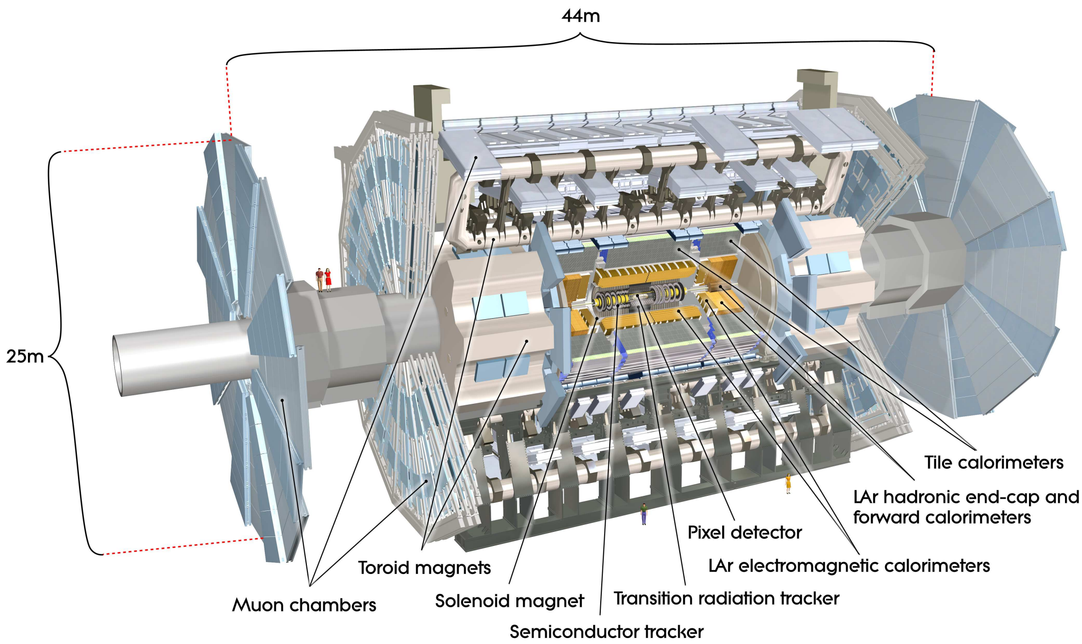


Figure 3.2.: Schematic overview of the ATLAS detector located at the LHC at CERN. The onion-like structure of the sub-detectors covers almost the entire solid angle around the interaction point. ©CERN

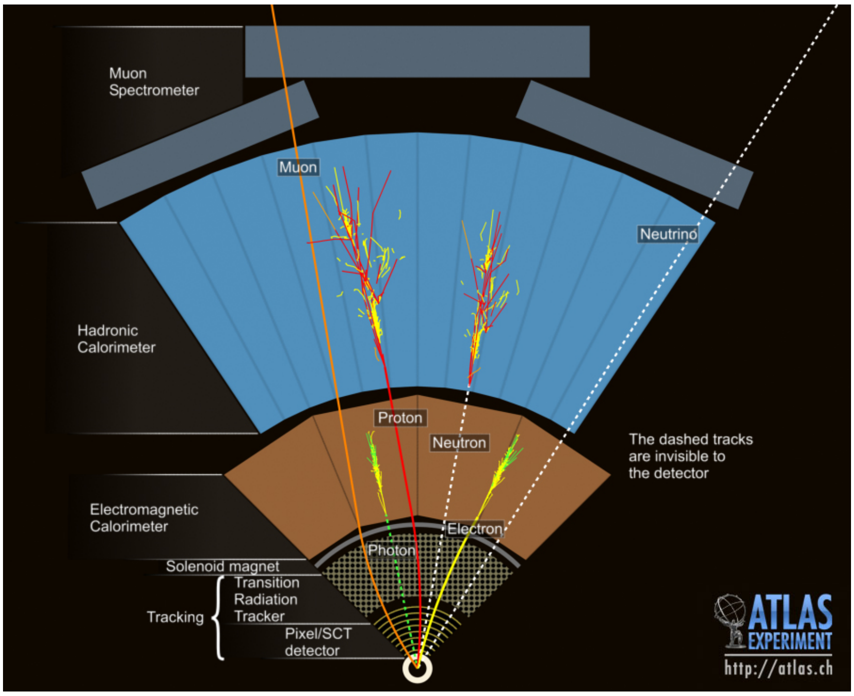


Figure 3.3.: Cross-section view of the layers in the ATLAS detector with example signatures of different particles. ©CERN

plane transverse to the beam axis, provides information about neutrinos. Since neutrinos cannot be detected directly by ATLAS, missing p_{T} indicates their existence.

The ATLAS detector consists of several layers, forming an onion-like structure. The interaction point is surrounded by the inner detector, followed by the electromagnetic and the hadronic calorimeter. The muon spectrometer is in the outer layer. Figure 3.3 shows a cross-section view of the detector layers. They are described in detail in the following sections.

The coordinate system of the ATLAS detector originates at the interaction point. The z-axis is defined by the beam-axis, with the x-y plane transverse to the beam axis. The azimuthal angle ϕ is measured circumferentially around the beam axis, and the polar angle θ is determined as the angle from the beam axis. To express the polar angle, the pseudorapidity $\eta = -\ln\tan(\theta/2)$ is commonly employed, as differences in pseudorapidity remain invariant under Lorentz transformations along the z-axis.

Inner Detector

The Inner Detector (ID), shown in Figure 3.4, tracks charged particles within the pseudorapidity range $|\eta| < 2.5$. A solenoid magnet surrounding the ID provides a 2 T axial magnetic field, which bends the trajectory of charged particles under the Lorentz force.

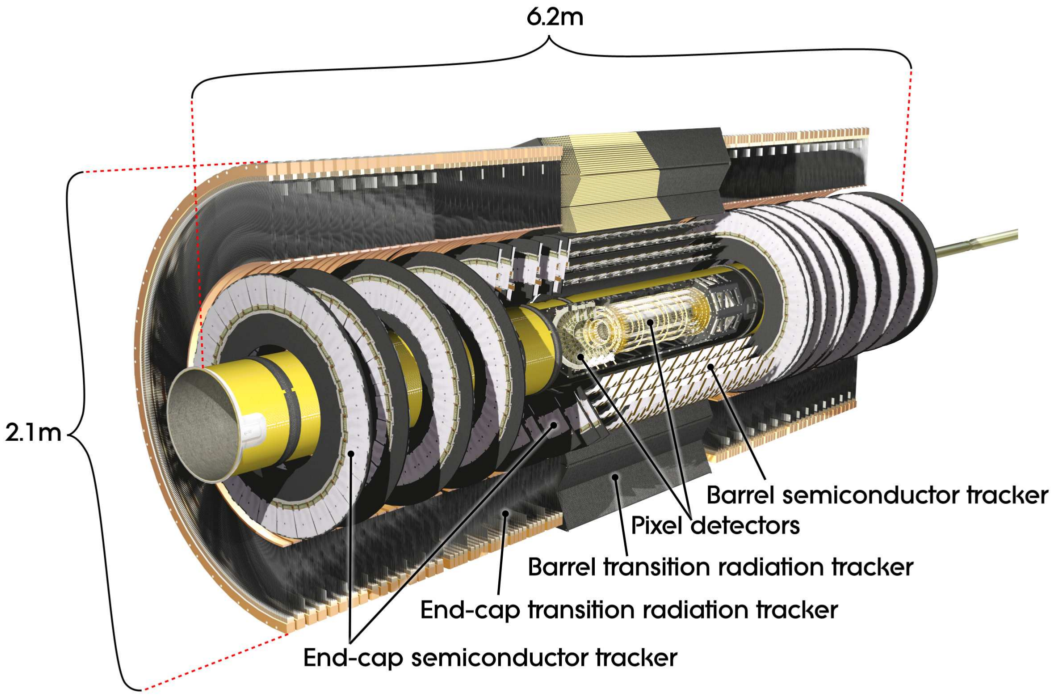


Figure 3.4.: Schematic overview of the Inner Detector (ID) of the ATLAS detectors. The interaction point is covered by Pixel detectors, the Semiconductor Tracker (SCT) and the Transition Radiation Tracker (TRT). SCT and TRT are composed of several barrel and end-cap components. ©CERN

The ID consists of three high-precision detectors, including the silicon pixel detector, which covers the vertex region. It is followed by the silicon microstrip tracker and the Transition Radiation Tracker (TRT). The latter enables radially extended track reconstruction up to $|\eta| = 2.0$ and provides electron identification information.

Calorimeter System

The calorimeter system, shown in Figure 3.5, is divided into two parts - the electromagnetic and the hadronic calorimeter. It covers the pseudorapidity range up to $|\eta| < 4.9$. The high-granularity lead/liquid-argon (LAr) calorimeters in the barrel and end-cap are used for electromagnetic calorimetry within the range of up to $|\eta| < 3.2$. There is also a thin LAr presampler that covers up to $|\eta| < 1.8$ for precise energy corrections. On the other hand, the hadronic calorimeter system uses a steel/scintillator-tile calorimeter within the range of up to $|\eta| < 1.7$, divided into three barrel structures. Furthermore, two copper/LAr hadronic end-cap calorimeters are used. The solid angle coverage is further expanded using forward copper/LAr and tungsten/LAr calorimeter modules optimised for electromagnetic and hadronic measurements, respectively.

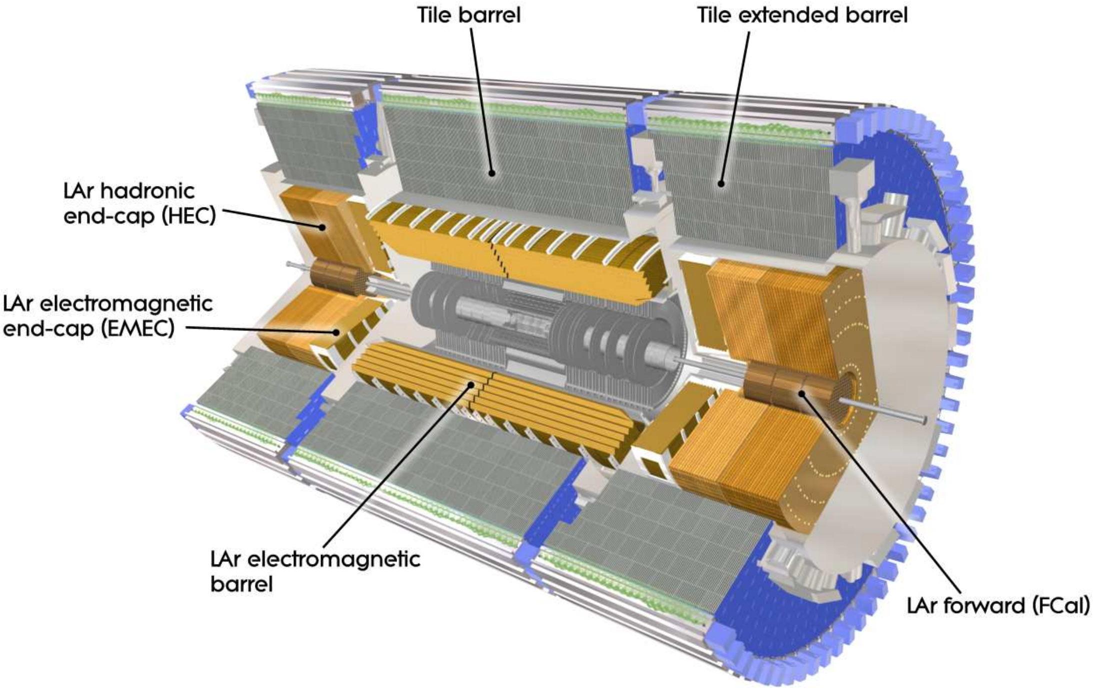


Figure 3.5.: Schematic overview of the calorimeter system of the ATLAS detector, including the LAr electromagnetic and hadronic calorimeter. Both parts are composed of several barrel and end-cap components. ©CERN

Muon Spectrometer

The muon spectrometer consists of two main components: the trigger and high-precision tracking chambers. The spectrometer is used to measure the deflection of muons in a magnetic field that is created by superconducting air-core toroids. The field integrals range from 2.0 to 6.0 Tm. The precision chambers cover the region $|\eta| < 2.7$. They include three layers of monitored drift tubes. Cathode-strip chambers are used in the forward region, where background levels are highest. Resistive-plate chambers are employed for the muon trigger system, which operates within the range $|\eta| < 2.4$ in the barrel region, while thin-gap chambers are used in the end-cap regions.

Data Selection

ATLAS uses a multi-level trigger system to choose interesting events for further analysis. The first-level trigger operates at a rate below 100 kHz and selects events from the 40 MHz bunch crossings. It is implemented in custom hardware. On the other hand, the high-level trigger is implemented in software and reduces the event rate to about 1 kHz. This allows events to be recorded and analysed in detail.

3.3. Object Reconstruction

Essentially, detectors like the ATLAS detector measure tracks and energy deposits. The energy of a particle shower can be calculated from the energy deposits. Tracks are formed by connecting hits in the ID, which record the positions of charged particles as they pass through it. The Lorentz force acting on charged particles causes a curvature of these tracks in the magnetic field of the detector. By accurately measuring the curvature of the tracks, the momentum and transverse momentum of the particles can be determined. Specific requirements must be applied to identify a particle. The selection criteria should minimise the background while maximising the signal-to-noise ratio to ensure accurate and reliable particle identification.

Electrons

Electron candidate reconstruction involves identifying energy deposit clusters in the electromagnetic calorimeter. These clusters are subsequently matched to tracks in the ID. Candidates must satisfy specific criteria, $p_{\mathrm{T}} > 7\,\mathrm{GeV}$, $|\eta| < 2.47$, and they must pass a "Medium" likelihood-based identification requirement [63, 64]. Electron candidates with clusters within the transition region between the barrel and the end-cap, defined as $1.37 < |\eta| < 1.52$, are excluded. The track associated with the electron must satisfy two conditions. First, the longitudinal impact parameter z_0 relative to the reconstructed primary vertex must satisfy the requirement $|z_0 \sin(\theta)| < 0.5$ mm. Second, the transverse impact parameter d_0 relative to the beam axis, must meet the criterion $|d_0|/\sigma(d_0) < 5$, where $\sigma(d_0)$ represents the uncertainty in d_0.

Muons

Muon candidates in the pseudorapidity range $|\eta| < 2.5$ are reconstructed from tracks in the muon spectrometer that are matched to tracks in the ID. These candidates must fulfil $p_{\mathrm{T}} > 7\,\mathrm{GeV}$ and the "Medium" identification requirements defined in Refs. [65, 66]. These identification requirements involve the significance of the charge-to-momentum ratio q/p and the number of hits in different ID and muon spectrometer subsystems. Additionally, the track associated with the muon candidates is required to satisfy $|z_0 \sin(\theta)| < 0.5\,\mathrm{mm}$ and $|d_0|/\sigma(d_0) < 3$.

Isolation Criteria

Isolation criteria are applied to both electrons and muons in the selection process. For electrons, the criterion involves two components. First, it requires that the sum of the

transverse momenta of tracks within a variable-size cone centred around the electron, excluding tracks originating from the electron itself, must not exceed 6% of the electron p_{T}. The cones' radius $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ is given as the minimum of $\Delta R = 10\,\mathrm{GeV}/p_{\mathrm{T}}$ and $\Delta R = 0.2$. Second, to exclude clusters originating from the electron itself, the sum of the transverse energy from topological clusters in the calorimeter within a cone of $\Delta R = 0.2$ around the electron should be less than 6% of the electrons' p_{T}. Topological clusters are groups of neighbouring calorimeter cells formed based on the spatial and geometric characteristics of energy deposits.
For muons, a similar criterion is applied. The radius of the track isolation cone is determined as the minimum value between $\Delta R = 10\,\mathrm{GeV}/p_{\mathrm{T}}$ and $\Delta R = 0.3$. Furthermore, the sum of the p_{T} of tracks within a variable-size cone around the muon, excluding the muons' own track, must not exceed 6% of the muon p_{T}.

Jets

Jets are created when quarks undergo hadronisation, which causes the formation of collimated sprays of hadrons. They are not single objects but clusters that can be reconstructed using the anti-k_t algorithm [67] with a radius parameter $R = 0.4$. The jet energy scale is calibrated based on 13 TeV data and simulation [68]. To be considered for this analysis, only jet candidates with $|\eta| < 4.5$ and $p_{\mathrm{T}} > 25$ GeV are included.
To account for the impact of jets arising from other proton-proton collisions that take place at the same time as the collision under study, an additional criterion is applied. This criterion involves using a likelihood-based jet-vertex-tagging method on jets that are characterised by having a pseudorapidity $|\eta|$ less than 2.5 and a transverse momentum p_{T} that is less than 120 GeV [69]. If the jets are located in the forward direction with $|\eta| > 2.5$, then they must meet the requirements of the "Medium" working point of the forward jet vertex tagger, as defined in Ref. [70].

b-Jets

Jets that originate from *b*-hadrons are known as *b*-jets. These hadrons have a relatively long lifetime due to the high CKM suppression of the decay of bottom quarks, which is evident from the corresponding CKM matrix element $V_{tb} \approx 1$. They also exhibit a characteristic topology. Identifying *b*-jets requires the use of the DL1d *b*-tagging algorithm [71]. This algorithm employs a DNN which considers various factors, such as the jets' transverse momentum and pseudorapidity, and outputs from the DIPS, JetFitter, and SV1 algorithms [72–74]. These algorithms provide information about the *c*-jet probability and the distance between the primary and secondary vertex. The DL1d *b*-tagging algorithm comes with four calibrated working points (WPs) at 60 %, 70 %, 77 % and 85 %. Each is defined based on the efficiency to tag true b-jets. In an analysis with 100 real b-jets, an 85 % WP means that the DL1d tagger tags 85 of them. It is, therefore, the loosest WP, while 60 % is the tightest WP.

Missing Transverse Momentum

Missing transverse momentum with magnitude $p_{\mathrm{T}}^{\mathrm{miss}}$ is calculated by taking the negative vector sum of the transverse momenta of all the physics objects that have been selected and calibrated. These objects include electrons, photons, muons, and jets. The calculation considers low-momentum tracks from the primary vertex that are not associated with any previously mentioned objects. These tracks are referred to as 'soft terms' [75]. In addition, the missing transverse momentum is used to reconstruct neutrinos, which cannot be directly detected.

4. Deep Neural Networks

Neural networks (NNs) are a tool for solving complex problems across various fields, including computer vision, natural language processing, and pattern recognition. They are inspired by the structure and functions of biological nervous systems like the human brain. That makes them capable of learning complicated patterns from large datasets. Warren McCulloch and Walter Pitts introduced the first simplified mathematical model of a neuron [76]. A McCulloch-Pitts-Neuron uses binary inputs and creates a binary output. Analogous to biological neural networks, inhibitory signals can be processed. The model can perform basic logical calculations. Another crucial step in the evolution of NNs is the *Perceptron* computing model, which Frank Rosenblatt developed in 1957 [77]. It can carry out basic binary classifications. The increasing computing power of computers over time and the development of training NNs using *backpropagation* [78] makes it possible to analyse huge amounts of data. Therefore, NNs have proven useful for data analyses in collider experiments. However, the basic structure of imitating a brain is still the same.
A basic modern NN consists of three types of layers: the input layer, the output layer, and the hidden layers, each with several nodes. The input layer is a vector $\vec{x}^{(0)}$, which in this analysis contains information about particle properties of an event in a *pp* collision. The output layer $\vec{y}$ indicates whether the NN has recognised the pattern of specific processes. A NN with multiple hidden layers is called a *Deep Neural Network* (DNN). Figure 4.1 shows a schematic overview of a DNN consisting of two hidden layers with six nodes each, five input parameters, and three output parameters.

4.1. Structure of a Deep Neural Network

Consider a DNN with L hidden layers and an input vector $\vec{x}^{(0)}$. In this notation, the upper index denotes the l^{th} hidden layer where $l = 0$ is the input layer. To calculate the entries of the first hidden layer $x_j^{(1)}$, the input layer entries $x_i^{(0)}$ are multiplied with weights $w_{ij}^{(0)}$. The weights connect each node of a layer l with each node of a layer $l+1$. In the scheme of Figure 4.1, they are represented by lines. A high value of $w_{ij}^{(l)}$ indicates a high influence of the parameter $x_i^{(0)}$ on the output. Furthermore, a bias $b_j^{(0)}$ is added, and an activation function $f^{(0)}$ is applied. Hence, the entries of the first layer vectors are calculated with

$$x_j^{(1)} = f^{(0)} \left(w_{ij}^{(0)} x_i^{(0)} + b_j^{(0)} \right) . \tag{4.1}$$

The bias is an offset to move the entire activation function. Activation functions introduce non-linearities into the network, allowing it to learn complex patterns in the data. They determine whether or at which intensity a neuron should be activated. The activation function which Perceptron uses is the binary step function $H(x_j^{(l)})$. Other commonly used activation functions are the sigmoid function $\sigma_{\text{sig}}(x_j^{(l)})$, the hyperbolic tangent $\tanh(x_j^{(l)})$, the *Rectified Linear Unit* (ReLU) function

$$\text{ReLU}(x_j^{(l)}) = \begin{cases} 0 & \text{for } x_j^{(l)} \leq 0 \\ x_j^{(l)} & \text{for } x_j^{(l)} > 0 \end{cases} \tag{4.2}$$

and the *softmax* function

$$\sigma_{\text{soft}}(x_j^{(l)}) = \frac{\exp\left(x_j^{(l)}\right)}{\sum_{k=1}^{n} \exp\left(x_k^{(l)}\right)} , \tag{4.3}$$

which is a generalisation of the sigmoid function. The activation functions not shown here are defined in Equations A.1, A.2 and A.3 in Appendix A. Applying the softmax function on the layer $\vec{x}^{(l)}$ with n nodes, its entries are transformed to a value between 0 and 1 with the property $\sum_{j=1}^{n} \sigma_{\text{soft}}(x_j^{(l)}) = 1$. Thus, the activated outputs sum up to one. The ReLU and sigmoid functions are shown in Figure 4.2.

The expression of Equation 4.1 can be generalised for two layers l and $l-1$ with m and n nodes using the matrix of weights

$$W^{(l)} = \begin{pmatrix} w_{11}^{(l)} & w_{12}^{(l)} & \cdots & w_{1n}^{(l)} \\ w_{21}^{(l)} & w_{22}^{(l)} & \cdots & w_{2n}^{(l)} \\ \vdots & \vdots & \ddots & \vdots \\ w_{m1}^{(l)} & w_{m2}^{(l)} & \cdots & w_{mn}^{(l)} \end{pmatrix} \tag{4.4}$$

and vector notation. Thus, the vector of layer l is calculated with

$$\vec{x}^{(l)} = f^{(l-1)} \left(W^{(l-1)} \cdot \vec{x}^{(l-1)} + \vec{b}^{(l-1)} \right) . \tag{4.5}$$

This calculation is iterated L times until the output vector $\vec{y} = f^{(L)} \left(W^{(L)} \cdot \vec{x}^{(L)} + \vec{b}^{(L)} \right)$ is obtained.

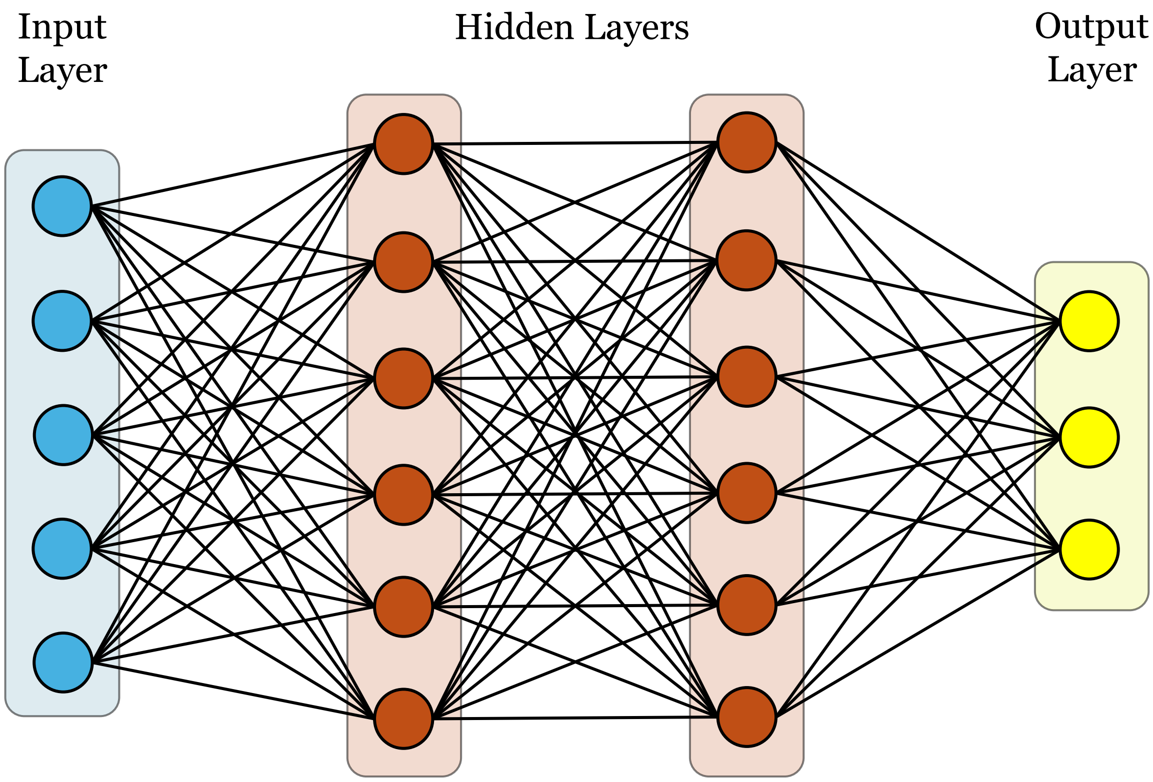


Figure 4.1.: Schematic overview of a deep neural network consisting of two hidden layers with six nodes each, five input parameters, and three output parameters. The connections between each node represent the weights.

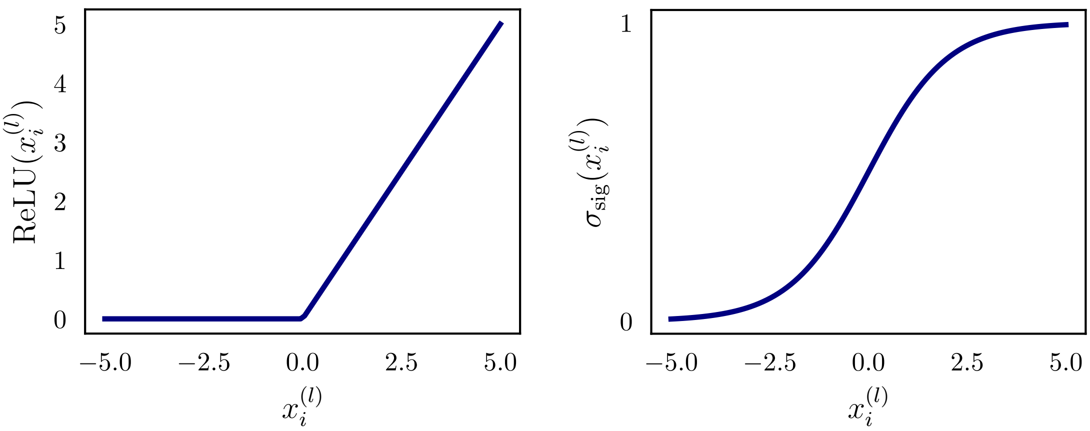


Figure 4.2.: Rectified Linear Unit $\mathrm{ReLU}(x_j^{(l)})$ (left) and sigmoid $\sigma_{\mathrm{sig}}(x_j^{(l)})$ (right) activation functions for input values $x_j^{(l)}$ between -5 and $+5$. The ReLU function outputs 0 for negative inputs and retains positive inputs unchanged, resulting in a piecewise linear activation, while the sigmoid function smoothly transitions from 0 to 1 as the input varies.

4.2. Training and Regularisation Techniques

Before a DNN can recognise patterns, it must learn from data. This process is called *training*, based on the backpropagation algorithm [78]. The training of a DNN involves the dynamic adjustment of its parameters $w_{ij}^{(l)}$ and $b_j^{(l)}$ to minimise the difference between its predicted outputs y_k and the truth training data $\hat{y}_k$. This is done by the minimisation of the *loss function* $C(\vec{y}, \hat{\vec{y}})$. The loss function is a measure of the quality of an output. Examples of loss functions are the mean squared error $\mathrm{MSE}(\vec{y}, \hat{\vec{y}}) = \sum_i (\hat{y}_i - y_i)^2$ and the categorical cross-entropy

$$\mathrm{CE}(\vec{y}, \hat{\vec{y}}) = -\sum_{i=1}^{N} \hat{y}_i \cdot \ln y_i \ . \tag{4.6}$$

The gradient descent of $C(\vec{y}, \hat{\vec{y}})$ with respect to the weights and biases is calculated and used to adjust them. This step is reiterated for each layer from the output to the first hidden layer. The model parameters are adjusted using a stochastic gradient descent (SGD) algorithm. This process is repeated several times until the loss function minimises. Each repetition is called an epoch.

Methods to avoid Overfitting and Stabilisation Techniques

Problems which can occur in the training are *underfitting* and *overfitting*. A model is underfitted if it cannot give accurate predictions on the data it is trained on. In this case, the training was stopped too early, or the DNN model is not complex enough. To achieve more complexity, layers or nodes are added to the DNN. Overfitting occurs when the model accurately predicts or classifies data included in the training set, but is incapable of classifying data on which it was not trained. To prevent overfitting, a randomly chosen subset of the training data, the validation set, is separated and not used in the backpropagation algorithm. Instead, the loss of the validation set is calculated after each epoch. In contrast to the loss of the training set, which is monotonically decreasing with an increasing number of epochs, the validation loss has a minimum. This is illustrated in Figure 4.3. On the left side of the minimum, the model is underfitted, and on the right side, it is overfitted. In a model which classifies data accurately, the training stops close to the minimum. To achieve this, two parameters are set up: the patience P and Δ_{min}. The latter is the minimum change in the monitored quantity to be considered as an improvement. If the number of epochs since an improvement is higher than the patience, the model assumes that the minimum validation loss has been reached and the training is stopped.

A common regularisation technique to stabilise a DNN and avoid overfitting is the im-

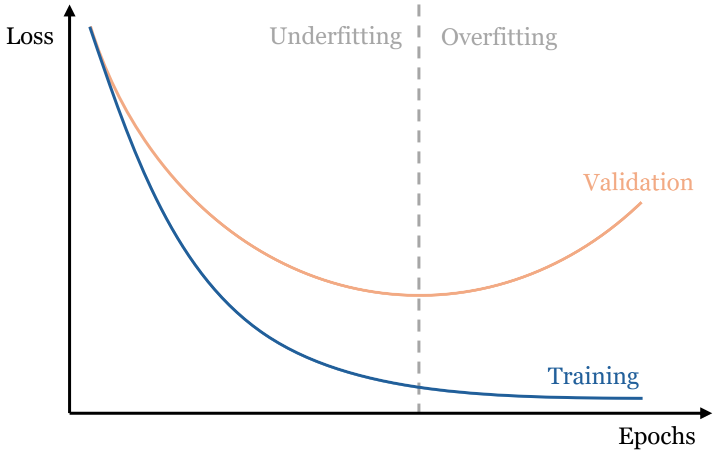


Figure 4.3.: Schematic illustration of the loss for the training and validation set depending on the epochs. Optimally, training is stopped when the loss of the validation set is minimal. The minimum is marked with a dotted line. If training is stopped significantly before it, this is referred to as underfitting. If it is stopped significantly after the minimum, it is overfitting.

plementation of *dropout layers* [79]. A dropout layer randomly ignores a subset of nodes in a given layer during training. This prevents the dropped-out nodes from participating in producing a prediction on the data. As a result, a new, slightly modified network architecture is built in each epoch, and the network learns to generate robust predictions without certain inputs. Furthermore, it prevents the permanent vanishing or disproportionate increase in the weights.

Another DNN regulation method is *batch normalisation* [80]. Batch normalisation layers normalise the activated output of a layer. The normalisation ensures that the mean output is close to 0 and the standard deviation is close to 1. This approach simplifies the learning of appropriate weights and biases. The batch normalised output is

$$x_i = \frac{a_i - \mu_{\mathrm{Batch},i}}{\sqrt{\sigma^2_{\mathrm{Batch},i} + \epsilon}} \cdot \gamma + \beta \tag{4.7}$$

where batch i has a mean $\mu_{\mathrm{Batch},i}$ and a standard deviation $\sigma^2_{\mathrm{Batch},i}$. The scaling factors γ and β are also trained parameters and $\epsilon = 10^{-3}$ is used for computational stability.

k-fold Cross-Validation

k-fold cross-validation is a technique to improve and evaluate machine learning models. The data, including the real and training data, is divided into k subsets, each of the same size. These so-called *folds* are separated into a training sample and a testing sample, which contains 25 % of the folds' data. The testing sample is never used for training and

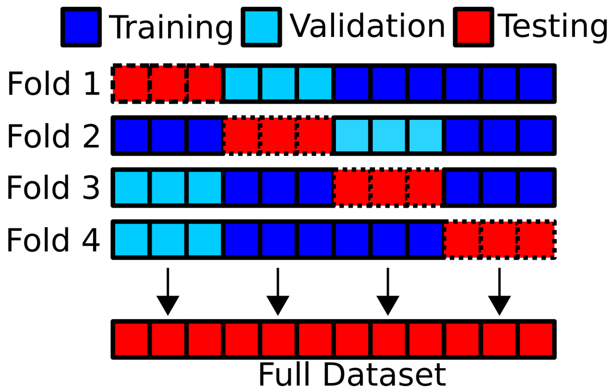


Figure 4.4.: Schematic illustration of k-fold cross-validation. The data is divided into k equal folds, with 25% allocated for testing. Each fold serves as training, validation and testing data.

serves for model evaluation. A certain fraction is used for validation from the training sample, and the rest is used for training. Now, there are k models on which the DNN is trained. For each model, the DNN trains on $k-1$ folds and tests the testing sample of the fold that is left over. Figure 4.4 shows a schematic overview of the k-folding procedure.

4.3. Model Evaluation

Model evaluation is crucial as it provides insights into the performance and reliability of DNNs. A possibility for evaluating a model is to compare the DNN scores for training and testing data. Discrepancies between training and testing set performances indicate issues such as overfitting or underfitting. Additionally, the ability to generalise can be assessed.

An evaluation of the general DNN performance gives the per-sample *separation power*

$$\mathcal{S} = \frac{1}{2} \sum_{i=1}^{N_{\text{bins}}} \frac{(S_i - B_i)^2}{S_i + B_i} \,. \tag{4.8}$$

S_i and B_i represent the signal and background events in bin i respectively, with N_{bins} being the total number of bins. It indicates whether a parameter is a useful discriminator between signal and background events, with higher values indicating stronger separation between the two. If all signal and background events lie in distinct bins, the separation power equals 1. In an accurately trained DNN, the DNN score clearly distinguishes between signal and background events, leading to a high separation power value.

An observable that indicates the efficiency of a classifier is the *receiver operating characteristic* (ROC) curve. A ROC curve shows the dependence of the true positive rate (TPR)

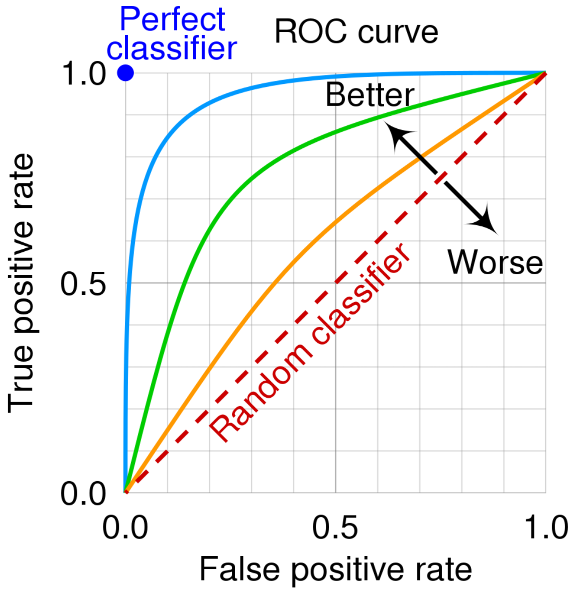


Figure 4.5.: Schematic illustration of a receiver operating characteristic (ROC) curve. The area under the ROC curve (AUC) summarises the classifiers' performance with random classification yielding AUC = 0.5 (red dotted line). With increasing classifier performance the AUC rises (orange, green and blue lines). The blue dot indicates a perfect classification achieving AUC = 1.

against the false positive rate (FPR). Its integral is the *area under the curve* (AUC). The AUC is approximated using

$$\mathrm{AUC} = \frac{1}{2} \sum_{i=1}^{N_{\mathrm{bins}}-1} \left(\mathrm{TPR}_i + \mathrm{TPR}_{i+1}\right) \left(\mathrm{TFR}_{i+1} - \mathrm{TFR}_i\right) \ . \tag{4.9}$$

A perfect classifier has AUC = 1. If the classifier does not recognise patterns and outputs random values, the AUC is 0.5. This scheme is illustrated in Figure 4.5. The AUC is also used to assess the impact of a certain parameter of a DNN output. Therefore, the input set of the parameter of interest is shuffled, and the AUC is calculated. This value is compared to $\mathrm{AUC}_{\mathrm{nom}}$, which is the AUC for an unshuffled set. The impact of the parameter on the classifier is expressed using the *nomalised permutation importance*

$$\Delta\mathrm{AUC} = \frac{\mathrm{AUC}_{\mathrm{nom}} - \mathrm{AUC}}{\mathrm{AUC}_{\mathrm{nom}}} \ . \tag{4.10}$$

5. Existing Analyses of the $t\bar{t}Z$ and tZq Processes and their Limitations

There are already measurements of the $t\bar{t}Z$ and tZq processes by ATLAS [81, 82]. Furthermore, there is a combined measurement of these processes by CMS [83]. All three measurements use data samples from LHC Run 2. Thus, they correspond to proton-proton collisions with a centre-of-mass energy of $\sqrt{s} = 13$ TeV and an integrated luminosity of 139 fb^{-1} for the ATLAS analyses and 138 fb^{-1} for the CMS analysis. In this chapter, the measurements and the limitations are presented and discussed.

5.1. $t\bar{t}Z$ Analysis by ATLAS

The analysis described in Ref. [81] measures the differential and inclusive cross-section of the $t\bar{t}Z$ process. The measurement targets the trilepton and the tetralepton final state.
A background event is an event which differs from the investigated process but has the same final state or shows identical signatures in the detector. In this analysis, a distinction between two different backgrounds is made. On one side are prompt lepton backgrounds. These leptons originate from the primary process. On the other side, there are backgrounds with non-prompt leptons, also referred to as “fake” leptons. Non-prompt leptons are particles that satisfy the identification criteria of leptons, described in Section 3.3, but do not originate directly from the primary collision. Instead, they arise from interactions of other particles produced in the collision. For example, non-prompt leptons could be true leptons from weak decays in jets, photons, pions or misidentified detector signatures. Relevant prompt lepton backgrounds are diboson production in association with jets ($WZ/ZZ +$ jets), single top production in association with a $W^{\pm}$ and a Z boson (tWZ), and the tZq process. The main non-prompt lepton background is top quark pair production with a non-prompt lepton ($t\bar{t}$ + fake).
So-called signal and control regions are defined to reduce the uncertainty of the measure-

ments. Signal regions (SR) are designed to select the signal and minimise background contamination. Control regions (CR) are used to enhance the selection of the main background sources and adjust systematic uncertainties. Both depend on kinematic variables like the number of charged leptons, their transverse momentum, missing transverse energy, flavour combinations of the non-Z leptons and the multiplicity of untagged jets N_{jets} and b-jets $N_{b\text{-jets}}$. Additionally, the regions depend on the number of opposite sign same flavour (OSSF) lepton pairs, opposite sign different flavour (OSDF) lepton pairs and the Z mass window $\Delta m_Z = \left| m_{\ell\ell}^Z - m_Z \right|$. The latter is the difference between the true Z boson mass m_Z and the reconstructed potential Z boson mass, where $m_{\ell\ell}^Z$ is the mass of an OSSF pair. An OSSF pair with the requirement $\Delta m_Z < 10$ GeV indicates the existence of an on-shell Z boson. If more than one OSSF pair fulfils this requirement, the one with the invariant mass closest to m_Z is considered to originate from the Z decay. To avoid contributions from resonant particles, the mass of every OSSF pair combination m_{OSSF} has to be higher than 10 GeV. There are also requirements on the leptons' transverse momenta $p_{\text{T}}(\ell_i)$ to reduce background from non-prompt leptons.

The SRs for the tetralepton channel are shown in Table 5.1. The names of the four regions indicate the number of leptons, whether the non-Z leptons have the same or different flavours and the b-jet multiplicity. The invariant mass of the non-Z leptons is denoted with $m_{\ell\ell}^{\text{non-}Z}$. Although the process requires two b-jets, events with one tagged jet at 85 % WP are also selected in the SRs. This is because b-jets are occasionally mistagged. Additionally, events with more than two untagged jets are allowed. This is due to QCD radiation, which frequently occurs in strong processes like the $t\bar{t}Z$ process. The distinction of OSSF and OSDF pairs for the non-Z lepton pairs is necessary because an OSDF is a veto for the $ZZ+$jets background but it is possible for the $t\bar{t}Z$ process.

Analysis and Results

For the inclusive $t\bar{t}Z$ cross-section measurement, the number of events in the trilepton and tetralepton SRs and in two particular CRs is fitted simultaneously, performing a profile-likelihood fit. Figure 5.1 shows the number of events in these regions and the deviation from the SM prediction. The result of the inclusive cross-section is $\sigma_{t\bar{t}Z} = 0.99 \pm 0.05\,(\text{stat.}) \pm 0.08\,(\text{syst.})$ pb. A more recent ATLAS analysis using the full dataset from Run 2 with a luminosity of $\mathcal{L} = 140$ fb^{-1} additionally includes the dilepton channel. The measured cross-section is $\sigma_{t\bar{t}Z} = 0.86 \pm 0.04\,(\text{stat.}) \pm 0.04\,(\text{syst.})$ pb [84]. Both measurements are in agreement with the SM prediction $\sigma_{t\bar{t}Z}^{\text{theo}} = 0.86^{+0.09}_{-0.09}$ fb [85].

Measurements performed at particle level capture the actual properties of particles, providing insights into the fundamental physics processes. In contrast, measurements at the

Table 5.1.: Common selections and signal regions of the tetraleptonic channel in the $t\bar{t}Z$ analysis described in Ref. [81]. The missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$ requirement is applied for the same flavour regions to reduce ZZ background. High $E_{\mathrm{T}}^{\mathrm{miss}}$ indicates a neutrino that is not produced in ZZ processes.

Variable	4ℓ-SF-$1b$	4ℓ-SF-$2b$	4ℓ-DF-$1b$	4ℓ-DF-$2b$
N_ℓ	$= 4$			
	≥ 1 OSSF lepton pair with $\lvert m_{\ell\ell}^{Z} - m_Z\rvert < 10$ GeV for all OSSF combinations: $m_{\mathrm{OSSF}} > 10$ GeV			
$p_{\mathrm{T}}(\ell_i)$	$> 27, 20, 10, 7$ GeV			
$\ell\ell^{\mathrm{non}-Z}$	e^+e^- or $\mu^+\mu^-$	e^+e^- or $\mu^+\mu^-$	$e^\pm\mu^\pm$	$e^\pm\mu^\pm$
$E_{\mathrm{T}}^{\mathrm{miss}}$	> 100 GeV, if $\lvert m_{\ell\ell}^{\mathrm{non}Z} - m_Z\rvert \leq 10$ GeV	> 50 GeV, if $\lvert m_{\ell\ell}^{\mathrm{non}Z} - m_Z\rvert < 10$ GeV	-	-
	> 50 GeV, if $\lvert m_{\ell\ell}^{\mathrm{non}Z} - m_Z\rvert > 10$ GeV	-	-	-
N_{jets}	≥ 2	≥ 2	≥ 2	≥ 2
$N_{b\text{-jets}}$	$= 1$	≥ 2	$= 1$	≥ 2

detector level include the effects of the experimental apparatus on the observed data. This is why examining new SM extensions using detector level measurements is nearly impossible. However, this is possible with measurements at particle level. For the differential cross-section measurement, a Bayesian unfolding procedure is performed [86]. This removes detector distortions and leads to a result at particle level. The differential cross-section regarding the variable X_i is calculated using

$$\frac{d\sigma_{\mathrm{tt}Z}}{dX_i} = \frac{1}{\mathcal{L} \cdot \mathcal{B} \cdot \Delta X_i \cdot \epsilon_{\mathrm{eff}}^{i}} \sum_j [\mathcal{M}^{-1}]_{ij} \cdot f_{\mathrm{acc}}^{j} \cdot \left(N_{\mathrm{obs}}^{j} - N_{\mathrm{bkg}}^{j}\right) \tag{5.1}$$

where $\mathcal{L}$ is the Luminosity, $\mathcal{B}$ is the branching ratio of the investigated $t\bar{t}Z$ channels and the $\epsilon_{\mathrm{eff}}^{i}$ are efficiency correction terms. The index i denotes the bin at particle level and j at detector level. The variables in the summation are the acceptance corrections f_{acc}^{j}, the number of observed particles N_{obs}^{j} and the number of background particles N_{bkg}^{j}. Additionally, there is the migration matrix $\mathcal{M}$. It quantifies the detector response. A matrix inversion is performed to remove the detector distortions, leading to a result at particle level. Figure 5.2 (a) shows the migration matrix of the transverse momentum of

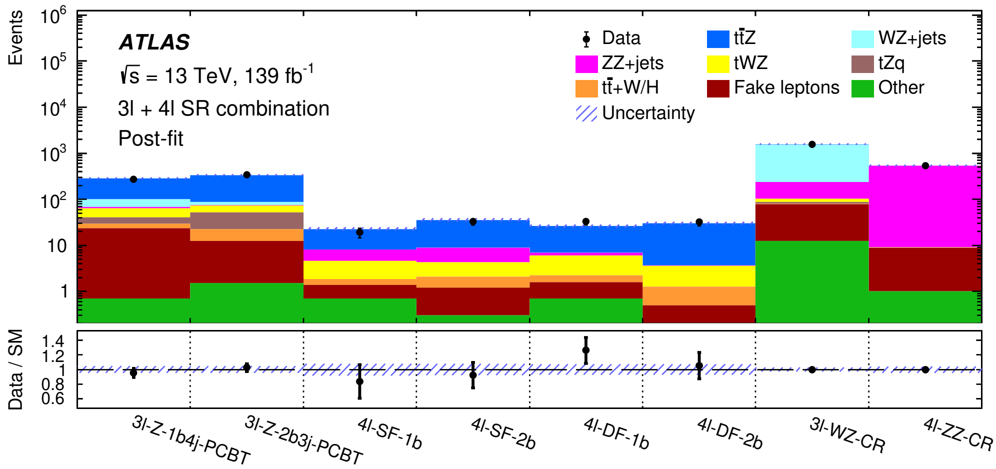
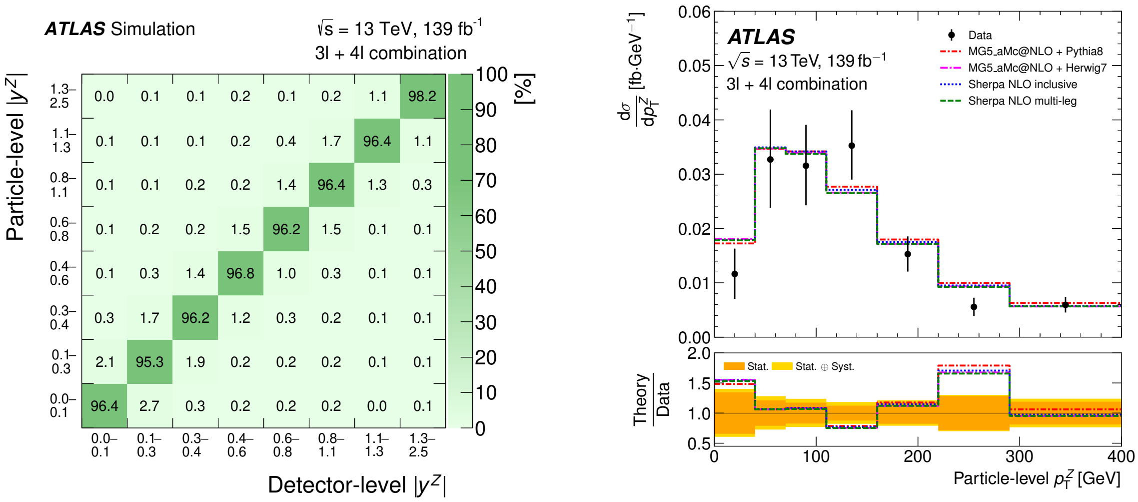


Figure 5.1.: Number of events for the signal regions and the WZ/ZZ + jets control regions after the combined fit [81].



(a) *(b)*

Figure 5.2.: **(a)** Migration matrix of the transverse momentum of the Z boson p_{T}^{Z} in the combination of the trilepton and the tetralepton regions for different p_{T}^{Z} bins. **(b)** Measured differential cross-section at particle level and several theoretical predictions [81].

the Z boson p_{T}^{Z} in the combination of the trilepton and the tetralepton regions for different p_{T}^{Z} bins. The entries are the fraction of events at particle level that are reconstructed at detector level. If the values of the diagonal terms of the migration matrix equal 100 %, it indicates perfect reconstruction of events without any migration or smearing between bins, suggesting ideal agreement between particle level and detector level measurements. Figure 5.2 (b) shows the measured differential cross-section at particle level and several theoretical predictions. The result is in agreement with the SM. The limiting factor for the differential measurement is statistical uncertainty.

5.2. tZq Analysis by ATLAS

The analysis described by Ref. [82] measures the total cross-section of the tZq process in the trileptonic channel. The relevant prompt lepton backgrounds are diboson and $t\bar{t}$ production associated with a Z, $W^{\pm}$ or a Higgs boson. Non-prompt lepton backgrounds are $Z+$jets, $t\bar{t}$ and $tW^{\pm}$ all with an additional fake lepton. The main backgrounds arise from diboson and $t\bar{t}Z$ production.

The SRs and CRs are shown in Table 5.2. They all have the same requirements for the number of leptons and the jet and lepton transverse momentum. The regions are denoted with the nomenclature njmb, where n refers to the jet multiplicity, i.e. the number of untagged jets and b-jets, and m is the b-jet multiplicity. Note that the b-jet referred to here is the b-jet resulting from the top quark decay, not from the gluon splitting (see Figure 2.6). To reconstruct the Z boson, all SRs require an OSSF lepton pair. Like in the $t\bar{t}Z$ analysis, the lepton pairs with the closest invariant mass to m_Z are selected as the OSSF pair from the Z boson. The other lepton is used to reconstruct the $W^{\pm}$ boson. Again a Z mass window of $\Delta m_Z < 10$ GeV is applied. The top quark is reconstructed, summing up the four-momenta of the b-jet and the $W^{\pm}$ boson. Further requirements on the SRs and CRs include the jet and b-jet multiplicity, the number of OSDF lepton pairs, and the pseudorapidity of untagged jets and b-jets. Moreover, events with two untagged jets are selected due to possible QCD radiation.

Analysis and Results

Neural Networks (NNs) are used to calculate the background estimation. The total cross-section is calculated, performing a simultaneous binned likelihood fit of the SRs and CRs. Monte Carlo (MC) distributions are used for the signal and background predictions. The measured total cross-section $\sigma_{tZq}^{\text{meas}} = 97 \pm 13\,(\text{stat.}) \pm 7\,(\text{sys.})\,\text{fb}$ is in agreement with the SM prediction $\sigma_{tZq}^{\text{theo}} = 102^{+5}_{-2}\,\text{fb}$ [82]. Figure 5.3 shows the number of events for the reconstructed top quark mass m_t and the deviations from the SM prediction in the SR 2j1b. The processes which contribute the most to the backgrounds are Z + jets, diboson processes with light flavour jets (VV+LF), diboson processes with heavy flavour jets (VV+HF) and $t\bar{t}Z + tWZ$. Heavy flavour jets are jets from bottom or charm quarks. Jets from lighter quarks are called light flavour jets. The figure shows the expected peak at the top quark mass m_t.

Table 5.2.: Common selections, signal, and control regions of the tZq analysis described in Ref. [82].

Exactly 3 leptons with $\lvert\eta\rvert < 2.5$ $p_{\mathrm{T}}(\ell_1) > 28$ GeV, $p_{\mathrm{T}}(\ell_2) > 20$ GeV, $p_{\mathrm{T}}(\ell_3) > 20$ GeV, $p_{\mathrm{T}}(\mathrm{jet}) > 35$ GeV			
SR 2j1b	CR diboson 2j0b	CR $t\bar{t}$ + fake 2j1b	CR $t\bar{t}Z$ 3j2b
≥ 1 OSSF pair $\lvert m_{\ell\ell}^{Z} - m_Z\rvert < 10$ GeV 2 jets, $\lvert\eta\rvert < 4.5$ 1 b-jet, $\lvert\eta\rvert < 2.5$	≥ 1 OSSF pair $\lvert m_{\ell\ell}^{Z} - m_Z\rvert < 10$ GeV 2 jets, $\lvert\eta\rvert < 4.5$ 0 b-jets	≥ 1 OSDF pair No OSSF pair 2 jets, $\lvert\eta\rvert < 4.5$ 1 b-jet, $\lvert\eta\rvert < 2.5$	≥ 1 OSSF pair $\lvert m_{\ell\ell}^{Z} - m_Z\rvert < 10$ GeV 3 jets, $\lvert\eta\rvert < 4.5$ 2 b-jets, $\lvert\eta\rvert < 2.5$
SR 3j1b	CR diboson 3j0b	CR $t\bar{t}$ + fake 3j1b	CR $t\bar{t}Z$ 4j2b
≥ 1 OSSF pair $\lvert m_{\ell\ell}^{Z} - m_Z\rvert < 10$ GeV 3 jets, $\lvert\eta\rvert < 4.5$ 1 b-jet, $\lvert\eta\rvert < 2.5$	≥ 1 OSSF pair $\lvert m_{\ell\ell}^{Z} - m_Z\rvert < 10$ GeV 3 jets, $\lvert\eta\rvert < 4.5$ 0 b-jets	≥ 1 OSDF pair No OSSF pair 3 jets, $\lvert\eta\rvert < 4.5$ 1 b-jet, $\lvert\eta\rvert < 2.5$	≥ 1 OSSF pair $\lvert m_{\ell\ell}^{Z} - m_Z\rvert < 10$ GeV 4 jets, $\lvert\eta\rvert < 4.5$ 2 b-jets, $\lvert\eta\rvert < 2.5$

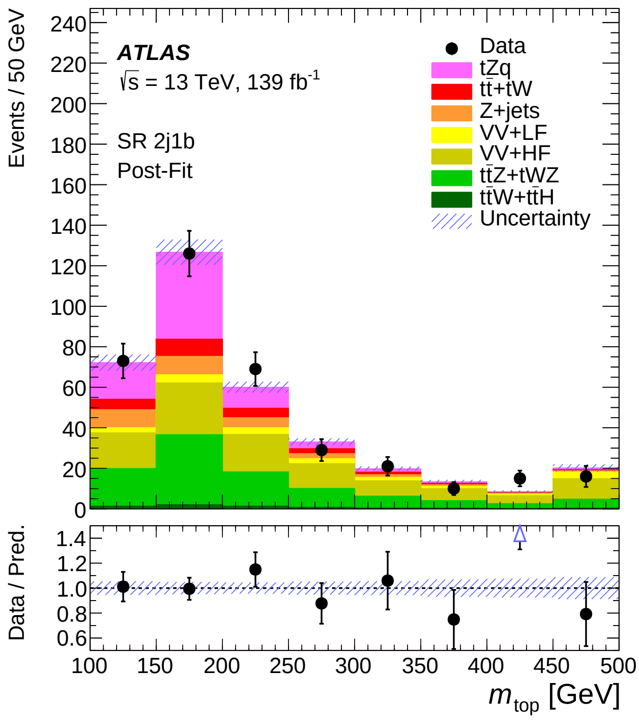


Figure 5.3.: Data and prediction comparison in the 2j1b region. The displayed variable is the reconstructed top quark mass m_t. The dominant backgrounds are diboson (VV) and $t\bar{t}Z$ + tWZ [82].

5.3. Combined Analysis of the tZq, the $t\bar{t}Z$ and the tWZ channel by CMS

In the analysis of Ref. [83], performed by the CMS Collaboration, the processes tZq, $t\bar{t}Z$ and, tWZ with trileptonic and tetraleptonic final states are measured jointly within the framework of an Effective Field Theory (EFT). An EFT is a theoretical framework that simplifies a complex theory by describing its low-energy behaviour through effective interactions, capturing the essential physics. The underlying theory is characterised by the energy scale Λ. The effective Lagrangian is

$$\mathcal{L}_{\text{eff}} = \mathcal{L}_{\text{SM}} + \sum_i \frac{c_i}{\Lambda^2} \mathcal{O}_i \,, \tag{5.2}$$

where c_i are the so-called Wilson coefficients and $\mathcal{O}_i$ are dimension six operators. If there are *Beyond Standard Model* (BSM) particles that are too massive to be produced on-shell at today's colliders, new phenomena could be visible at loop level. This is why the CMS analysis searches for deviations from the SM in dedicated observables. The combined analysis is of interest because the processes would be sensitive to the same EFT. Therefore, no SM assumptions are imposed on any of the three investigated processes.
Multivariate Analyses (MVA) techniques based on machine learning, like NNs, are used to enhance the sensitivity to BSM phenomena appearing from the EFT operators. A simultaneous fit to data in six event regions leads to results for the Wilson coefficients on detector level, which are $c_i = 0$ in the SM. Figure 5.4 shows the post-fit data-to-simulation comparisons depending on the NN output for the Wilson coefficient c_{tZ} in the $t\bar{t}Z$ (left) and tZq (right) SR. The lower panel shows that the EFT contribution rises with increasing NN output. In this analysis, the NN output is high if data does not conform to normal patterns. Thus, the distribution indicates that the observed data exhibits deviations from normal patterns when the contributions from EFT are high. This aligns with the expected outcome. The results are in agreement with the SM at a 95% confidence level.

5.4. Limitations

The presented analyses have advantages, disadvantages, and limitations. For the $t\bar{t}Z$ and the tZq process, background modelling is a limiting class of systematic uncertainties [81–84]. Additionally, the ATLAS tZq measurement is limited by statistical uncertainties [82]. Another disadvantage is the separate measurement of both processes. If there is an EFT contribution originating from an operator affecting a coupling that, in turn, impacts both

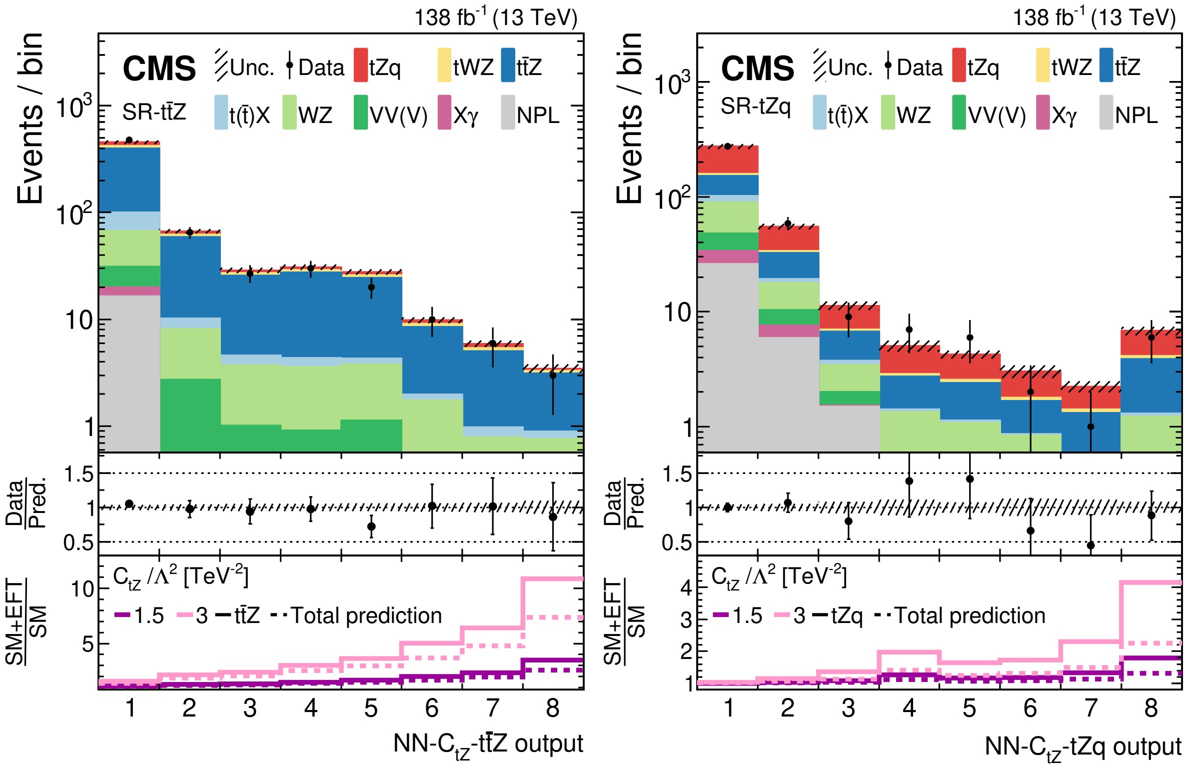


Figure 5.4.: Post-fit data-to-simulation comparisons depending on the NN output for the Wilson coefficient c_{tZ} in the $t\bar{t}Z$ (left) and tZq (right) SR. In both regions, the EFT contribution rises with increasing NN output [83].

the $t\bar{t}Z$ and tZq processes, access to the tZq component is restricted since it is removed from the data before unfolding. Therefore, the EFT effects might cancel each other out. In the analysis by the CMS collaboration [83], the processes are measured jointly, but the results are on detector level and model-dependent. This limits future reinterpretation.

6. Analysing the $t\bar{t}Z$ and tZq Process Jointly in the Trileptonic Channel

This chapter describes the analysis of the joint measurement of the $t\bar{t}Z$ and tZq processes in the trileptonic channel. The data was taken by ATLAS during the LHC Run 2 representing $\mathcal{L} = 140\ \text{fb}^{-1}$ [87]. Two SRs are set up for $t\bar{t}Z$ and tZq as well as a CR for the diboson background. For this, a DNN is used. The data is compared to the expectations from *Monte Carlo* (MC) simulations. Afterwards, a profile likelihood fit is performed.

6.1. Signal and Background Modelling using Monte Carlo Samples

Monte Carlo (MC) simulation is an established and widely used method in particle physics to simulate events from signal and background processes.
Although in true data it cannot be determined with certainty whether a process is a signal or a background process, it can be stated how many events of different processes are expected to be detected by the detector. The basic idea behind MC simulations is to generate numerous simulated events based on theoretical models and known probability distributions. These simulated events are generated by event generators, which simulate the collision processes and subsequent particle decays. By comparing the simulated events to experimental data, theoretical predictions can be validated or excluded.

Background Sources

The primary background sources in this analysis are processes with a similar final state like $t\bar{t}Z$ and tZq. These are in particular processes with three leptons. Events with two vector bosons, i.e. $W^{\pm}$ or Z bosons, are called diboson events (VV). If they come with additional jets, they are referred to as $VV+b$, $VV+c$ and $VV+l$, depending on whether the respective jet is a b-, c- or light jet. The analysis includes contributions from diboson processes with three or four charged leptons in the final state. Therefore, the notation

VV refers to $W^{\pm}Z$ or ZZ. Here, leptons also include $\tau^{\pm}$ leptons since they can decay leptonically into lighter leptons and the respective neutrinos. However, this process is suppressed due to requirements explained in Section 6.3. Diboson events are the largest background contribution in the joint $t\bar{t}Z$ and tZq analysis.

The second-largest contribution comes from events with non-prompt leptons, referred to as "fakes". These are events where a non-prompt lepton is reconstructed as a prompt lepton. Non-prompt leptons are particles not produced directly in the primary interaction. They instead originate from decays of other particles like heavy hadrons or result from misidentification. The non-prompt lepton processes included in the analysis are the production of $t\bar{t}$ pairs, of $t\bar{t}$ pairs associated with a photon ($t\bar{t}\gamma$), of two vector bosons, where one of them decays hadronically the other leptonically, and of single Z boson production where the Z boson decays leptonically. The decay into $\tau^{\pm}$ leptons is also considered here. In addition, each lepton simulated by an event generator contains information about whether it is a prompt or a non-prompt lepton. If there is a non-prompt lepton in a simulated event, this event is also classified as a fake event independent of the event class.

Other processes with trileptonic final states considered in the analysis are single top quark production in association with a $W^{\pm}$ and a Z boson (tWZ) and $t\bar{t}$ production associated with a $W^{\pm}$ boson ($t\bar{t}W$). The analysis does not include the simultaneous production of top quark pairs and a Higgs boson ($t\bar{t}H$) since no suitable MC samples are available. Minor contributions arise from producing three top quarks or three vector bosons. These processes are referred to as "other".

Monte Carlo Samples

The software framework `AnalysisTop Release 22` is used for the production of `ntuples`. These are data files storing the results of experimental data or MC event productions. They are created from data and MC samples, each having a different name, which always follows the same pattern. The name of an MC sample starts with the MC campaign, which is `MC20` here, followed by the centre-of-mass energy and the dataset identifier (DSID). The DSID is a number defining the simulated process. There is a cross-section σ and a k-factor for each DSID. The sample name also includes the event or matrix element generator, the PDF set and, if needed, the tune. The event generators relevant for this analysis are `Pythia8` [88, 89] and `Sherpa 2.2` [90]. The matrix element generator `MadGraph` [91, 92] is used for a few samples. After the generator, the decay chain is indicated in the sample name followed by the format `DAOD_PHYS`. The name ends with so-called tags. The `e-tag` defines the event generation, and the `s-tag` the full simulation. In this analysis, the tags `s3681` and `s3797` are used, indicating the detector simulation with `Geant4` [93–95].

No `a-tag` exists because only full simulation MC samples are used. The reconstruction tags `r13167`, `r13144`, and `r13145` stand for the MC campaigns `MC20a`, `MC20d`, and `MC20e`. The campaigns refer to the data-taking periods 2015 and 2016 (`MC20a`), 2017 (`MC20d`), and 2018 (`MC20e`). These are distinguished because there are different settings for the detector, trigger selection and pileup in the periods. The `p-tag` specifies the skim derivation for `PHYS` formats. Only samples with the tag `p5855` are used. All considered MC samples are listed in Tables B.1 and B.2 in Appendix B.

Reweighting

As MC events can be produced in any quantity for each sample, they are weighted so that the distribution of simulated events matches the expected distribution of real events. This weighting accounts for differences in cross-sections, detector efficiencies and other factors between the simulated and real data. The total weight of an MC event

$$w = w_{\mathrm{MC},i} \cdot w_{\text{x-sec}} \cdot w_{\mathrm{lep,SF}} \cdot w_{b\text{-tag,SF}} \cdot w_{\mathrm{pileup}} \cdot w_{\mathrm{year}} \cdot \mathcal{L} \tag{6.1}$$

is a product of several other weights and the luminosity $\mathcal{L} \approx 140\ \mathrm{fb}^{-1}$. Note that the weights in this chapter are unrelated to the DNN modelling weights. Since not every MC event is created equally, they come with a weight $w_{\mathrm{MC},i}$, which relates the event to the other simulated events of the event generator. The cross-section weight

$$w_{\text{x-sec}} = \frac{\sigma \cdot k}{\sum_{i=1}^{N} w_{\mathrm{MC},i}} \tag{6.2}$$

is applied to scale the simulated event to its predicted distribution, where σ is the theoretical assumption of the cross-section of the simulated process. The k-factor is needed to recover the best theory prediction. The adjustment is due to next-to-next-to-leading-order (NNLO) terms and electroweak corrections. The weights $w_{\mathrm{lep,SF}}$ and $w_{b\text{-tag,SF}}$ are scale factor (SF) weights and correct differences between simulated and real lepton detection and b-tagging. The pileup weight w_{pileup} aims to reproduce the distribution of pileup observed in real data. Pileup describes additional pp interactions occurring in the same bunch crossing as the primary interaction of interest. The last factor $w_{\mathrm{year}} \cdot \mathcal{L}$ includes the luminosity of the MC campaigns `MC20a`, `MC20d` and `MC20e`.

6.2. Particle Reconstruction

Particles like Z bosons and top quarks cannot be directly detected but can be reconstructed through their decay products. In both processes $t\bar{t}Z$ and tZq, the trileptonic channel requires a Z boson decaying into an opposite sign same flavour lepton (OSSF) pair. For each possible OSSF combination, the invariant mass $m_{\ell\ell}$ is calculated. The ℓ^+-ℓ^- combination with the invariant mass $m_{\ell\ell}^Z$ closest to $m_Z = 91.19$ GeV [24] is chosen to originate from an on-shell Z boson decay. The difference between m_Z and the reconstructed Z mass $m_{\ell\ell}^Z$ is denoted with $\Delta m_Z = \left|m_{\ell\ell}^Z - m_Z\right|$. The third lepton $\ell_{\text{non-}Z}$ is then assumed to be a product of a leptonic top quark decay.
Since the top quark decay chain is more complicated, its reconstruction is more difficult. A distinction is made for hadronically and leptonically decaying top quarks. For the latter, a reconstruction of the neutrino three-momentum $\vec{p_\nu}$ is necessary. Here, the charged lepton is the one which is not part of the OSSF pair from the Z decay. Since there is only one neutrino in the trileptonic channels of the $t\bar{t}Z$ and tZq process, it is assumed that the transverse momentum of the neutrino $p_{\nu,\mathrm{T}} = \sqrt{p_{\nu,x}^2 + p_{\nu,y}^2}$ equals the missing transverse energy $E_{\mathrm{T}}^{\text{miss}}$. Furthermore, the z-component of the neutrino momentum $p_{\nu,z}$ is estimated assuming an on-shell $W^\pm$ boson with mass $m_W = 80.377$ GeV [24]. This leads to the quadratic equation

$$\underbrace{\left(E_\ell^2 - p_{\ell,z}^2\right)}_{a} \cdot p_{\nu,z}^2 - \underbrace{2kp_{\ell,z}}_{b} \cdot p_{\nu,z} + \underbrace{E_\ell^2 p_{\nu,\mathrm{T}}^2 - k^2}_{c}\,, \tag{6.3}$$

which depends on the energy E_ℓ and the z-component of the momentum $p_{l,z}$ of the lepton from the top quark decay. The value k is defined as $k = \frac{m_W^2}{2} + p_{\nu,x}p_{\ell,x} + p_{\nu,y}p_{\ell,y}$. For $E_\ell^2 \neq p_{\ell,z}^2$, the solutions are $p_{\nu,z} = \frac{1}{2a} \cdot \left(-b \pm \sqrt{b^2 - 4ac}\right)$. Physically, the equation has two real solutions for $p_{\nu,z}$. However, due to previous assumptions, there could also be imaginary solutions. These assumptions are, e.g. an on-shell $W^\pm$ boson, that the missing transverse energy comes only from a neutrino, or that the charged leptons are correctly assigned to the Z boson. In the case that $b^2 < 4ac$, $p_{\nu,\mathrm{T}}$ is rearranged such that $b^2 = 4ac$ and Equation 6.3 has a real solution. The smaller solution for $|p_{\nu,z}|$ is selected since neutrinos from top quarks are mostly transverse and not longitudinal to the beamline. Summing up, the four-momentum of the neutrino and the corresponding lepton leads to the four-momentum of the $W^\pm$ boson. The $W^\pm$ boson is paired with a b-jet. Note that whenever b-jets or tagged jets are mentioned in the analysis, this refers to a tagged jet at 85 % WP. If there is only one tagged jet, this jet is directly assigned to the leptonically decaying $W^\pm$ boson. If there is more than one tagged jet, the two with the highest WP

are determined. In cases where two jets have the same DL1d WP, the jet with the higher p_{T} value is chosen as the leading b-jet. The other is called the sub-leading b-jet. Out of these two jets, the one with the closest angular distance $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}$ to the lepton $\ell_{\text{non-}Z}$ is paired with the leptonically decaying $W^{\pm}$ boson.
To reconstruct the hadronically decaying $W^{\pm}$ boson, the four-momenta of all jets are summed up except for the tagged jet, which is assigned to the other $W^{\pm}$ boson. If there are more than four jets, the two jets which lead to the best $W^{\pm}$ boson mass are selected, excluding the leading and sub-leading b-jet. In the end, the b-jet which is left is assigned to the hadronically decaying $W^{\pm}$ boson.
The mass of the top quark m_t is calculated from its four-momentum, which is the sum of the four-momenta of the $W^{\pm}$ boson and the b-jet. A distinction is made between the mass of the leptonically decaying top quark m_t^{lep} and the hadronically decaying top quark m_t^{had}.

6.3. Event Selection

The goal is to get a pure SR of $t\bar{t}Z$ and tZq events and to reduce backgrounds that dominantly result from diboson events and events with non-prompt leptons, referred to as fake leptons. Therefore, requirements on various parameters are applied to increase the ratios S/B, $S/\sqrt{B}$ and S/F, where S is the number of signal events, i.e. $t\bar{t}Z$ and tZq events, B is the number of background events, and F is the number of events including non-prompt leptons.

Preselection

The preselection is applied to every event, independent of the SRs and CRs. Since the analysis targets the trileptonic channel, only events with $N_\ell = 3$ charged leptons are selected. Furthermore, there are requirements on the jet multiplicity $N_{\text{jets}} \geq 2$ and b-jet multiplicity at 70 % WP $N_{b\text{-jet},70\%} \geq 1$. The 70 % WP is chosen to reduce diboson background. There are no upper bounds on the jet multiplicities to include events with QCD radiation and mistagged jets.

Strategies to Select Events including a Z Boson Reduce Background from Non-Prompt Leptons

To select processes including a Z boson, only events with at least one OSSF pair are considered. To reduce background from non-prompt leptons, the requirement $\Delta m_Z < 10$ GeV is applied on the Z boson mass window Δm_Z. For every OSSF pair combination, the

invariant mass of the two leptons must be larger than 10 GeV. This avoids contributions from resonant processes.
In addition, to further minimise the background from non-prompt leptons, requirements to the transverse momenta of the leptons $p_\mathrm{T}(\ell_i)$ are applied, as non-prompt leptons typically occur at low p_T. The indices $i \in \{1,2,3\}$ represent the leading, the sub-leading and the lepton with the lowest transverse momentum. They are estimated in the procedure, in which the efficiencies

$$Z(X) = \frac{S}{X}\left(1 - \frac{L}{S_0}\right)\frac{X_0}{S_0} \tag{6.4}$$

of different $p_\mathrm{T}(\ell_i)$ requirements are calculated for different variables X. These are the number of background events B, its square root $\sqrt{B}$ and the number of events with non-prompt leptons F. Here, S_0, B_0 and F_0 refer to the number of signal, background, and fake events with the loosest $p_\mathrm{T}(\ell_i)$ requirement $p_\mathrm{T}(\ell_1) > 27$ GeV and $p_\mathrm{T}(\ell_2) = p_\mathrm{T}(\ell_3) > 7$ GeV where the trigger defines $p_\mathrm{T}(\ell_1) > 27$ GeV. The number of lost signal events for each combination is $L = S - S_0$. In Equation 6.4, the expression of the efficiency is chosen to maximise $^S/_X$. It is normalised by the respective ratio of the loosest requirements so that $Z(X) > 1$ indicates an improvement in efficiency.
The efficiencies are calculated for every possible combination of $p_\mathrm{T}(\ell_1) > 30$, 27 GeV and $p_\mathrm{T}(\ell_{2,\,3}) > 25$, 20, 15, 10, 7 GeV with $p_\mathrm{T}(\ell_1) > p_\mathrm{T}(\ell_2) > p_\mathrm{T}(\ell_3)$. A summary of them, their total yields, the ratios $^S/_X$ and the efficiencies can be found in Tables C.1 and C.2 in Appendix C. The goal is to decrease the number of background events. The focus lies on the reduction of background from non-prompt leptons, MVA are performed to reduce diboson background (see Chapter 6.4). The cut combination $p_\mathrm{T}(\ell_i) > 30$, 20, 15 GeV is added to the event selection because it has the maximal values of $Z(B) = 1.56$ and $^S/_{\sqrt{B}} = 23.21$. Additionally, the efficiencies $Z(\sqrt{B}) = 1.11$ and $Z(F) = 4.31$ are relatively high compared to the efficiencies of other combinations. Combinations with larger $Z(F)$ are not selected due to a high loss of signal events. Other promising combinations like $p_\mathrm{T}(\ell_i) > 27$, 15, 15 GeV and $p_\mathrm{T}(\ell_i) > 30$, 20, 10 GeV are also not included in the event selection due to their small $Z(F)$. By applying the lepton p_T requirements, $^S/_B$ increases by 79 %, $^S/_{\sqrt{B}}$ by 25 %, and $^S/_F$ by 389 % loosing 12 % of the signal events. Table 6.1 summarises the preselection and event selection, including a description of the reasons for selecting these requirements. In the following, they are referred to as *common selections*. The fraction of the simulated processes after applying the selection is shown in a pie chart in Figure 6.1.

Table 6.1.: Requirements of the event and preselection of the joint $t\bar{t}Z$ and tZq analysis and their intention. In the following, they are referred to as common selections. The constraints limit the number of leptons N_ℓ, jets N_{jets}, b-jets $N_{b\text{-jets}}$, OSSF pairs N_{OSSF}, the Z boson mass window Δm_Z, the minimal invariant mass of an OSSF pair $m_{\ell\ell}^{\text{min}}$ and the transverse momenta of the leptons $p_{\text{T}}(\ell_i)$.

selection / exclusion	requirement
selects trilepton final state	$N_\ell = 3$
selects jets with sufficient p_{T}	$p_{\text{T}}(j_i) \geq 25$ GeV
selects minimal (b-)jet multiplicity for tZq	$N_{\text{jets}} \geq 2$, $N_{b\text{-jet},70\%} \geq 1$
selects events with a Z boson	$N_{\text{OSSF}} \geq 1$ $\Delta m_Z < 10$ GeV
excludes non-resonant processes	$m_{\ell\ell}^{\text{min}} > 10$ GeV
excludes high amount of non-prompt leptons	$p_{\text{T}}(\ell_i) > 30, 20, 15$ GeV

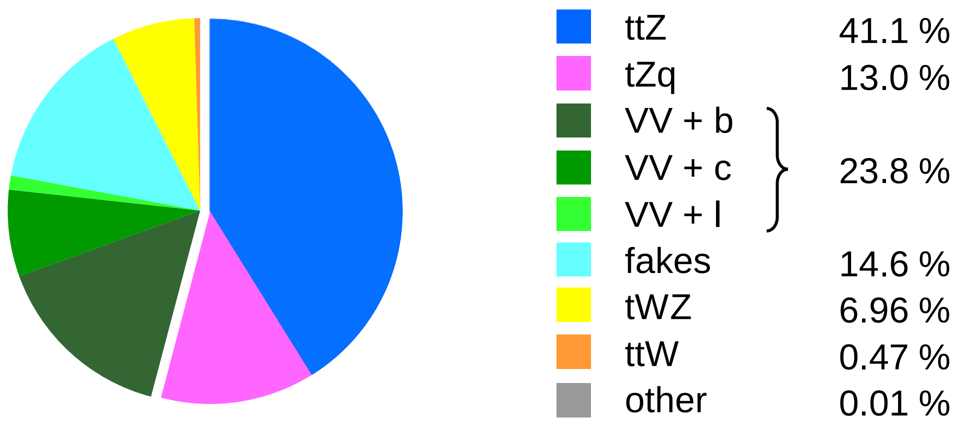


Figure 6.1.: Pie chart of the total number of simulated events after applying the common selections. The signal is separated from the background. The diboson processes (VV) and processes including non-prompt leptons (fakes) have the highest background contributions. Processes labelled "other" are so rare that they are almost invisible in this illustration.

6.4. Classification of $t\bar{t}Z$, tZq, and Diboson Events using a Deep Neural Network

This analysis aims to distinguish between signal events and background events from diboson processes. The latter is the dominant background in this analysis. A *Deep Neural Network* (DNN) is used to define SRs and CRs. This involves selecting appropriate input parameters. The next step is to train the DNN, evaluate its performance and define SRs and CRs using the DNN classification.

Architecture of the Deep Neural Network

In this approach, a DNN uses the MC generated data described in Chapter 6.1 to classify $t\bar{t}Z$, tZq and diboson events. The same weights mentioned in Equation 6.1 are applied, and only events passing the common selections listed in Table 6.1 are considered for the training. The output layer of the DNN consists of three nodes representing whether the model recognises a $t\bar{t}Z$, tZq or diboson pattern. The output parameters are the $t\bar{t}Z$ score $O_{t\bar{t}Z}$, the tZq score O_{tZq}, and the diboson score O_{VV}. The DNN is composed of three hidden layers, each with 30 nodes. All activation functions of the hidden layers are the ReLU function which is defined in Equation 4.2. The activation function of the output layer is the softmax function (see Equation 4.3) ensuring that scores are values between 0 and 1 and the sum equals 1.
The loss function is the categorical cross-entropy $\mathrm{CE}(\vec{y}, \hat{\vec{y}})$ defined in Equation 4.6. The loss is minimised, and the weights and biases are updated using the Adam optimiser with Nesterov momentum called Nadam optimiser [96–98]. The optimiser uses the principle of SGD.
For the training process, k-fold cross-validation is applied. The dataset is separated into $k = 4$ folds. In each fold, $25\,\%$ of the data is used for testing. Another $25\,\%$ is separated from the remaining subset for validation. The DNN is therefore trained on $75\,\% \cdot 75\,\% = 56.25\,\%$ of the MC simulation. The validation is performed with a patience of $P = 100$ epochs and a minimal difference in the validation loss of $\Delta_{\mathrm{min}} = 10^{-4}$. To stabilise the classifier, batch normalisation is applied after the first hidden layer. Additionally, the nodes of the same layer have a $20\,\%$ dropout probability.
Most of the diboson background is due to processes including a $W^{\pm}$ and a Z boson, as it is the only diboson process with a trileptonic channel. Thus, the input parameters for the DNN are selected to discriminate $t\bar{t}Z$, tZq and WZ + jets processes. In total, there are 20 input parameters, for example, the number of jets N_{jets}, and b-jets $N_{b\text{-jets}}$. The (b-)jet multiplicities are mainly selected to distinguish between $t\bar{t}Z$ and tZq since more jets are

Table 6.2.: Input parameters of the DNN which classifies $t\bar{t}Z$, tZq and diboson processes and their rank in the SRs and the diboson CR. The variable $j_{b,1}$ ($j_{b,2}$) denotes the leading (sub-leading) b-jet. An untagged jet is denoted as $j_{l,c}$.

Variable	Definition	Rank		
		SR-$t\bar{t}Z$	SR-tZq	CR-VV
N_{jets}	Number of jets	2	3	6
$N_{b\text{-jets}}$	Number of tagged jets	5	14	11
$j_{b,1}$ tag. WP	WP of the leading b-jet	8	10	5
$j_{b,2}$ tag. WP	WP of the sub-leading b-jet	6	13	8
$p_{\text{T}}(j_1)$	p_{T} of the leading jet	4	2	3
$p_{\text{T}}(j_2)$	p_{T} of the sub-leading jet	11	6	2
$p_{\text{T}}(j_3)$	p_{T} of the jet with the 3$^{\text{rd}}$ highest p_{T}	1	8	4
$\|\eta(j_1)\|$	$\|\eta\|$ of the leading jet	10	7	13
$\|\eta(j_2)\|$	$\|\eta\|$ of the sub-leading jet	18	16	20
$\|\eta(j_3)\|$	$\|\eta\|$ of the jet with the 3$^{\text{rd}}$ highest p_{T}	14	15	16
$\|\eta^{\max}(j_{l,c})\|$	$\|\eta\|$ of the untagged jet with the highest $\|\eta\|$	12	9	12
m_t^{had}	top mass from the hadronic top decay	3	5	17
m_t^{lep}	top mass from the leptonic top decay	9	1	1
$H_{\text{T,jet}}$	scalar sum of the p_{T} of all jets	16	11	10
$H_{\text{T,lep}}$	scalar sum of the p_{T} of all leptons	7	4	9
$p_{\text{T}}(\ell_1)$	p_{T} of the leading lepton	19	17	14
$p_{\text{T}}(\ell_2)$	p_{T} of the sub-leading lepton	15	19	18
$p_{\text{T}}(\ell_3)$	p_{T} of the lepton with the lowest p_{T}	17	20	19
$p_{\text{T}}(\ell_{\text{non-}Z})$	p_{T} of the lepton from the top decay	13	12	7
$E_{\text{T}}^{\text{miss}}$	missing transverse energy	20	18	15

expected for $t\bar{t}Z$ than for tZq. Other input parameters are the reconstructed top quark masses m_t^{lep} and m_t^{had}, indicating a leptonically or a hadronically decaying top quark. A list of all input parameters and their definition can be found in Table 6.2.

Classifier Performance and Evaluation

In Figure 6.2, the yields of the DNN outputs in their fit regions are shown. The respective ratio of $t\bar{t}Z$, tZq and diboson events rises with increasing DNN scores. Figure 6.3 shows the relation between the DNN scores in a two-dimensional distribution. Since the softmax activation function (Equation 4.3) is applied to the outputs, the sum of scores

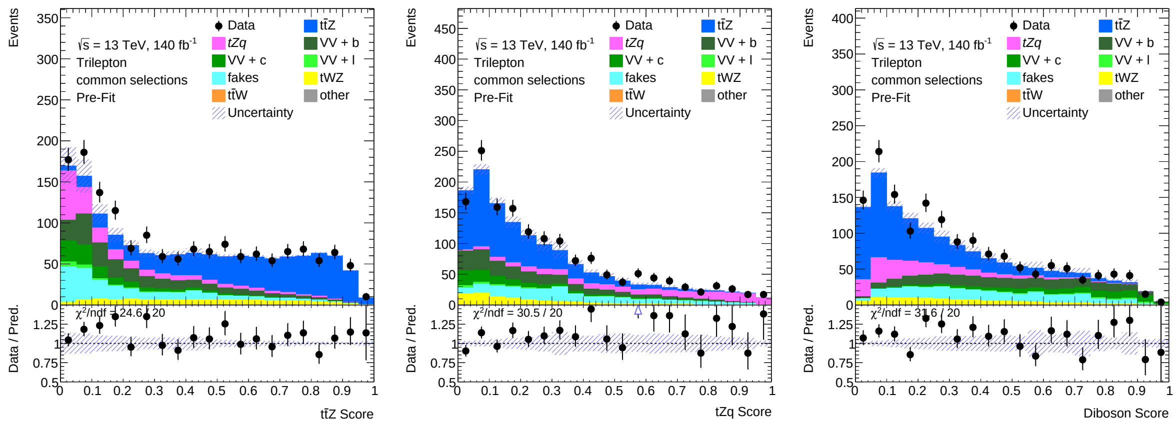


Figure 6.2.: Comparison of the $t\bar{t}Z$ (left), tZq (centre) and diboson score (right) in the common selections. With increasing DNN score, the respective MC yields rise. The lower panels show the data-MC agreement. Arrows indicate data out of the domains' range. Statistical and systematic uncertainties, introduced in Chapter 6.5, are included in the uncertainty band.

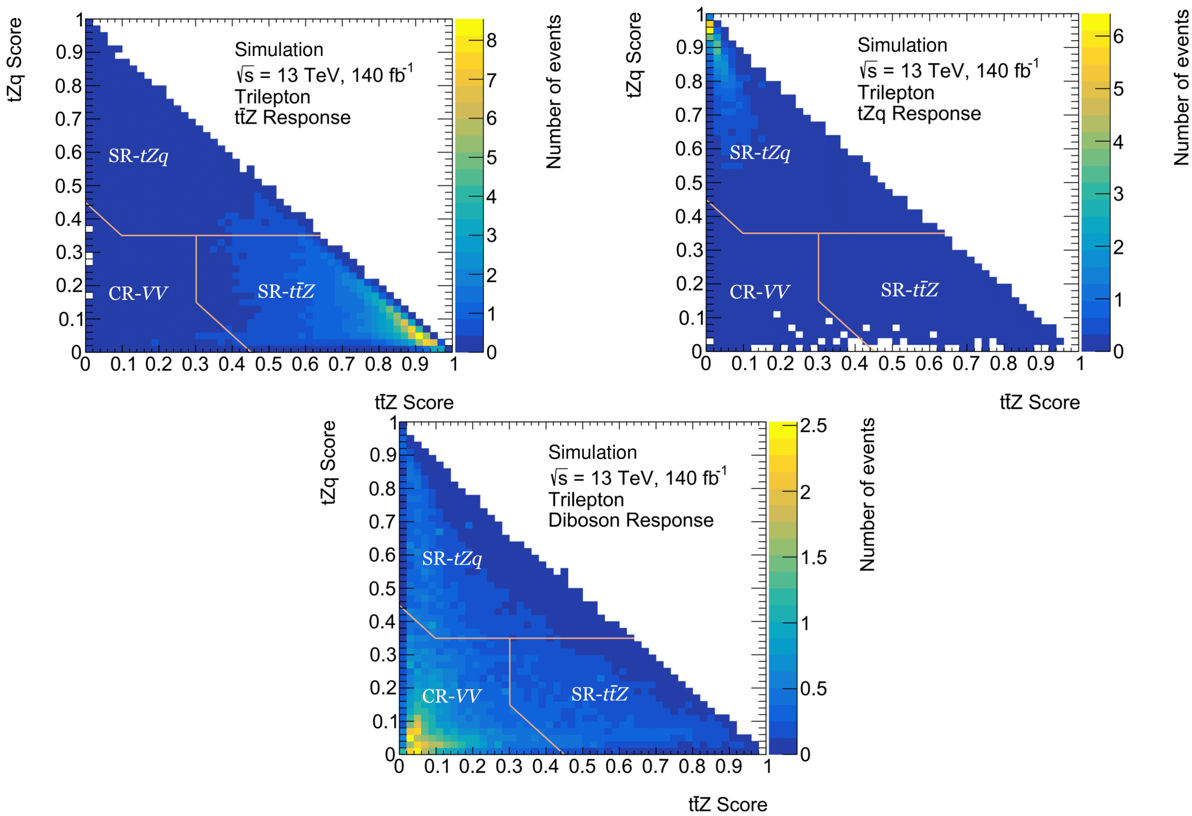


Figure 6.3.: Number of simulated $t\bar{t}Z$ (top left), tZq (top right) and diboson (bottom) events for different combinations of the DNN scores illustrated in two-dimensional distributions. A diagonal line indicates a constant value of the diboson score. The orange lines indicate the regions SR-$t\bar{t}Z$, SR-tZq and CR-VV.

always equals 1. Therefore, a diagonal line from the top left to the bottom right indicates a constant value of the score not plotted on the x- or y-axis, which is, in Figure 6.3, the diboson score. If the number of events in the lower-left corner is high, there are many diboson events. The two-dimensional distributions show the expected output values of an effectively trained classifier since the number of events in the respective corners is high.

The loss function of fold 1 is shown in Figure 6.4. In contrast to the scheme in Figure 4.3, the validation loss is higher than the initial training loss. This is due to the dropout layers. With more epochs, the training loss exceeds the validation loss and converges. The training is stopped before the validation loss decreases. Thus, the graph does not show an indication of overfitting.

Figure 6.5 shows training and testing comparisons of the diboson score. The distributions on the left-hand side display the fraction of events of the diboson score for signal and background events in fold 1. The signal is maximal for small $O_{t\bar{t}Z}$, showing an effective separation of $t\bar{t}Z$ and tZq events from the background. The distribution of O_{VV} shows a plateau for medium values. For small O_{VV}, they correspond mostly to background from tWZ and non-prompt processes. The larger the diboson score, the higher the fraction of diboson events. The comparison of training and testing data does not show a significant deviation.

The distributions on the right-hand side of Figure 6.5 display the ROC curves of the diboson score for all folds. Additionally, the AUC for training and testing is illustrated. As expected, the AUC is slightly higher for training than for testing. The AUC averaged over the four folds for training and testing is shown in Table 6.3. The testing AUC is approximately the same for the $t\bar{t}Z$ score with $\mathrm{AUC}_{t\bar{t}Z} = 0.866$ and the tZq score with $\mathrm{AUC}_{tZq} = 0.862$. The value for the diboson score is slightly lower at $\mathrm{AUC}_{VV} = 0.845$. Table 6.3 additionally displays the separation power $\mathcal{S}$ of each DNN score defined in Equation 4.8. For the $t\bar{t}Z$ and tZq scores, they are 34.9 % and differ only in the second decimal place, while for the diboson score, the value is significantly smaller at 25 %. The evaluation of the ROC curves, the AUC, and the separation power demonstrate precise discrimination power of $t\bar{t}Z$, tZq, and diboson events by the DNN model.

Based on the normalised permutation importance ΔAUC, defined in Equation 4.10, it is determined which DNN input parameters contribute most to the classification. This is visualised in bar charts in Figure 6.6. The parameter with the highest ΔAUC for the tZq and the diboson classification is the leptonic top quark mass m_t^{lep}. For the $t\bar{t}Z$ classification, it is the transverse momentum of the sub-sub-leading jet $p_{\mathrm{T}}(j_3)$. It is followed by the number of jets N_{jets} and the hadronic top quark mass m_t^{had}. The transverse momenta of the three jets with the highest transverse momentum $p_{\mathrm{T}}(j_i)$ have a large impact on all

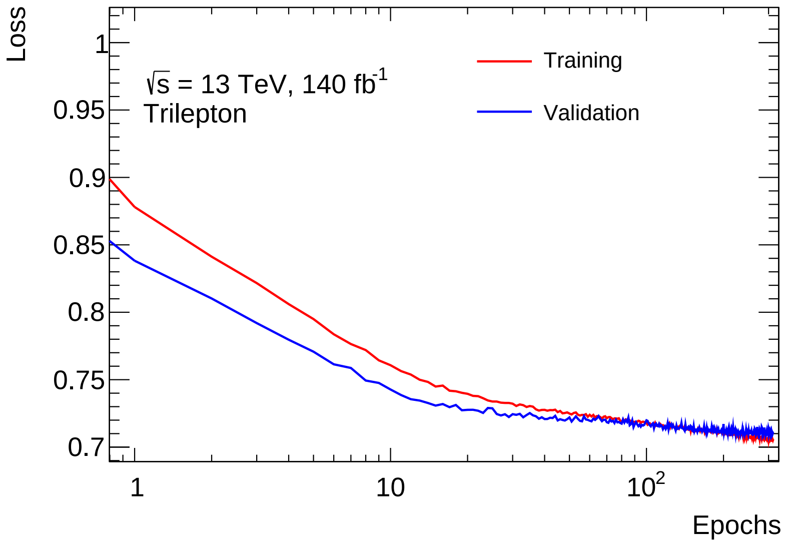


Figure 6.4.: Loss functions for fold 1. The shapes are comparable with the shapes of Figure 4.3. At the beginning, the training loss (red) is higher than the validation loss (blue) due to dropout features. At the end of the training, the validation loss is higher than the training loss and reaches its minimum.

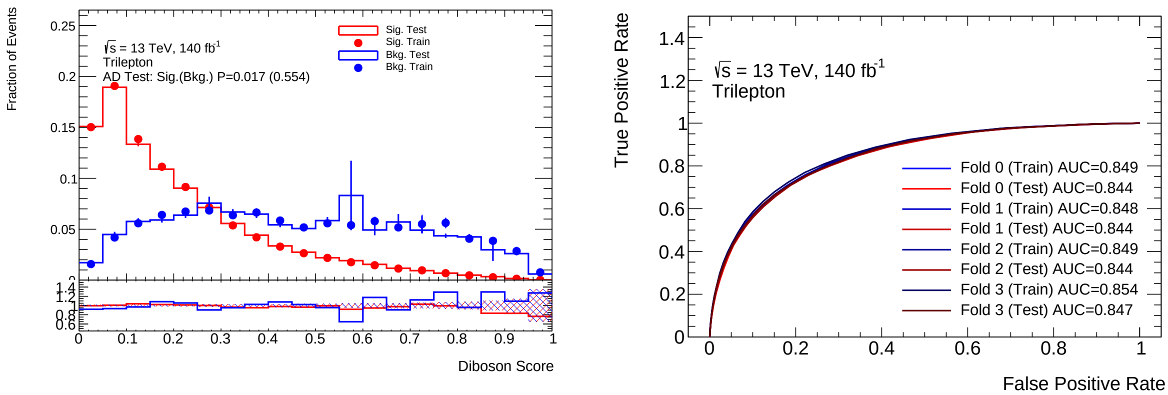


Figure 6.5.: Fraction of Signal (blue) and background (red) events, depending on O_{VV} for fold 1 (left) and receiver operating characteristic (ROC) curves (right) for the diboson output values. Both distributions compare data and testing. The lower band of the left distribution shows the training over testing ratio. The shaded error bands are the statistical uncertainty. The comparison of training and testing data does not show a significant deviation. The ROC curves form a distinct arc, comparable with the arcs of Figure 4.5.

Table 6.3.: Separation power $\mathcal{S}$, testing AUC and training AUC for the $t\bar{t}Z$, tZq and diboson score. For each AUC, the average value of the four folds is taken. The higher the parameters, the more accurate the classifier.

DNN Score	Separation Power	AUC (training)	AUC (testing)
$t\bar{t}Z$	34.9 %	0.869	0.866
tZq	34.9 %	0.874	0.862
Diboson	25.8 %	0.857	0.845

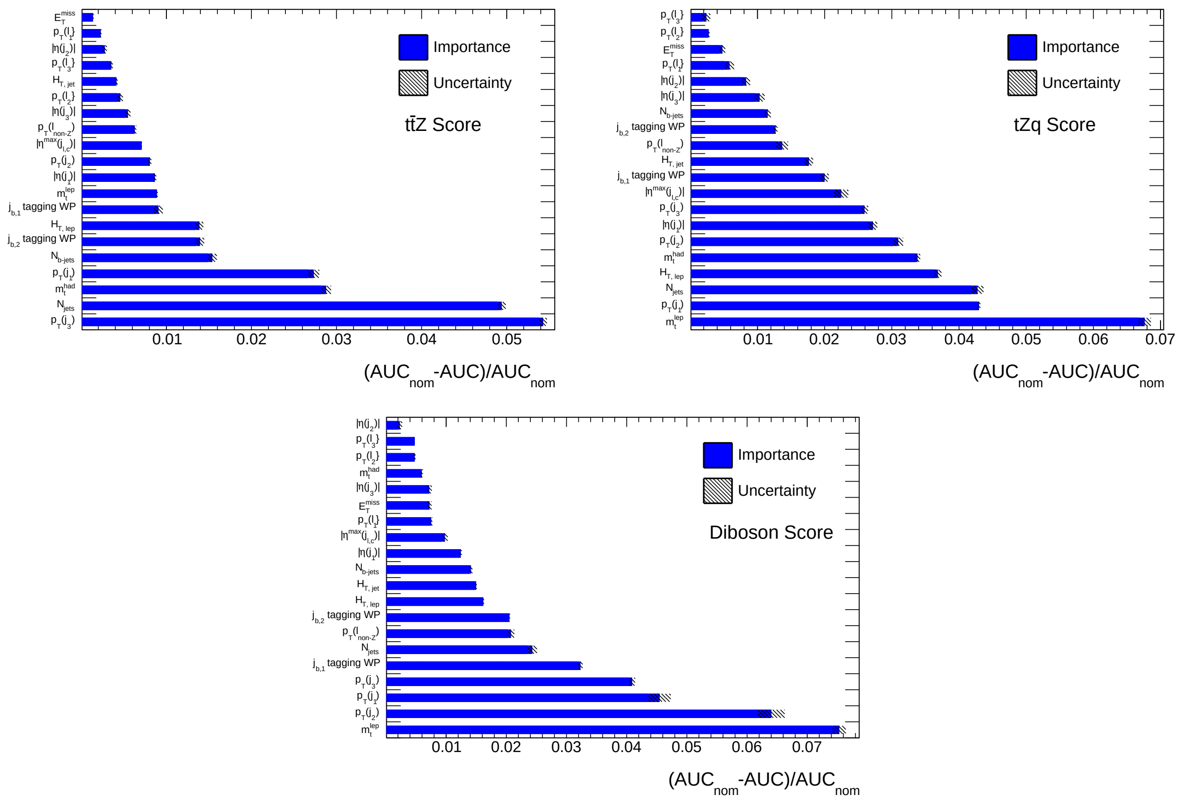


Figure 6.6.: Normalised permutation importance of the $t\bar{t}Z$ (top left), tZq (top right) and diboson score (bottom) for every input variable considered in the DNN. The parameter $(\mathrm{AUC}_{\mathrm{nom}} - \mathrm{AUC})/\mathrm{AUC}_{\mathrm{nom}}$ is a measure of how much the respective input parameter has contributed to the classification.

classifications. Minor contributions come from the lepton transverse momenta $p_{\mathrm{T}}(\ell_i)$, the pseudorapidities $\eta(j_2)$ and $\eta(j_3)$, and the missing transverse energy $E_{\mathrm{T}}^{\mathrm{miss}}$. The pseudorapidities $\eta(j_1)$ and $\eta^{\max}(j_{l,c})$ are mostly relevant for the tZq classification. All parameters, their definitions, and ranks for each classifier are listed in Table 6.2.
For additional DNN evaluation distributions and illustrations, refer to Figures D.1 and D.2 in Appendix D.

Definition of Signal and Control Regions

Orthogonal SRs and CRs are defined in using the DNN scores $O_{t\bar{t}Z}$, O_{tZq}, and O_{VV}. They are named SR-$t\bar{t}Z$, SR-tZq and CR-VV. A method similar to the reduction of fake events, described in Chapter 6.3, is used. A scheme of the selection process is illustrated in Figure 6.7. At first, a threshold of the diboson score is determined. Every event below this threshold remains available for the SRs. Events with a diboson score above are included in CR-VV. An analogous procedure is performed with the remaining SR candidates. Events not passing a requirement on the tZq score are candidates for SR-$t\bar{t}Z$, while the others are added to SR-tZq. From the potential SR-$t\bar{t}Z$ events, only those with high enough $t\bar{t}Z$ score are included in SR-$t\bar{t}Z$. The rest is added to CR-VV.
The requirements are determined by calculating several parameters of interest for different DNN scores in steps of 0.05. These are S/B, $S/\sqrt{B}$, $Z(B)$ and $Z(\sqrt{B})$ where the efficiency Z is defined in Equation 6.4. All events with a diboson score ≥ 0.55 are excluded from the SRs and included in the CR-VV because it maximises $Z(\sqrt{B})$ losing only 7.5 % of the signal events. The threshold for the tZq SR and $t\bar{t}Z$ SR candidates are chosen by evaluating the distributions of the same parameters as before using the tZq score as an upper limit for SR-$t\bar{t}Z$ and as a lower limit for SR-tZq. In this procedure, either tZq or $t\bar{t}Z$ are treated as signal events and the respective remaining events as background. The threshold is set at $O_{tZq} = 0.35$ because the efficiencies for $t\bar{t}Z$ and tZq are roughly the same at this point. The values for the tZq efficiencies for this condition are slightly higher than those for the $t\bar{t}Z$ process, since tZq events occur much less frequently. In addition, the $t\bar{t}Z$ score is used to improve the $t\bar{t}Z$ SR. For this region, only events with $O_{t\bar{t}Z} \geq 0.30$ are considered. Events not passing any of the mentioned conditions are added to CR-VV. Since the regions are orthogonal, they can be visualised as a triangle with different DNN scores on the x- and y-axis. This is illustrated in the distributions of Figure 6.3.

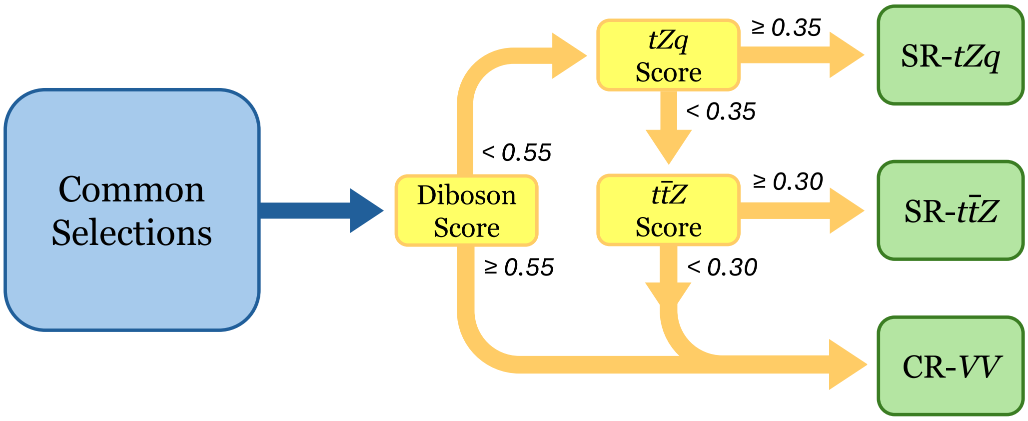


Figure 6.7.: Schematic overview of the SR and CR definitions. After the common selections, only events with $O_{VV} < 0.55$ are selected for the SRs. If $O_{tZq} \geq 0.35$ the event is added to SR-tZq. If $O_{tZq} < 0.35$ and $O_{t\bar{t}Z} \geq 0.35$ it is added to SR-$t\bar{t}Z$. The diboson CR includes all events which do not fulfil the mentioned requirements.

6.5. Uncertainties

Understanding statistical and systematic uncertainties is crucial since they define the accuracy of a measurement. Systematic uncertainties are categorised into instrumental and theoretical uncertainties. The former relates to the experimental setup, while the latter arises from the MC simulations. This chapter describes the applied systematic uncertainties in the analysis.

Intrumental Uncertainties

Instrumental uncertainties arise from detector effects, reconstruction algorithms, and calibration procedures. This is because the detection of objects and object classification efficiency differs in different parts of the detector.
Lepton reconstruction and identification uncertainties primarily affect the efficiency of correctly identifying leptons in the detector. These uncertainties arise from energy deposition, track reconstruction, and isolation requirements. Scale factors (SFs) are applied to quantify these uncertainties derived from simulation studies. SFs associated with electron reconstruction and identification are applied to all simulated events. For muons, the SFs are not applied to fake events. These are the systematic and statistical SFs for identifying low and high transverse momenta, isolation, and track-to-vertex association. In addition, lepton trigger efficiency uncertainties are accounted for by SFs related to electron trigger efficiency and muon trigger efficiency. These SFs address discrepancies in the efficiency of the lepton trigger and are only applied on prompt lepton events.

The uncertainty associated with the luminosity measurement is $\pm 0.83\,\%$ [87]. It is applied to all simulated events. There are additional uncertainties related to jets and b-tagging, but they are not available and hence are not considered in this analysis.

Theoretical Uncertainties

Theoretical uncertainties arise from the predictions of theoretical models used to simulate the processes. These uncertainties can originate from several sources, such as uncertainties in cross-sections, PDFs, or higher-order perturbative calculations. These are considered by applying normalisation uncertainties to the background samples.
A systematic uncertainty of $\pm 50\,\%$ is applied for the fake lepton background. The value relies on studies described in Ref. [81]. The tWZ cross-section uncertainty is set to $15\,\%$ based on an analysis performed in Ref. [84]. The contributions of $t\bar{t}W$ and other processes to the simulated events are less than $1\,\%$. Therefore, a normalisation of $\pm 50\,\%$ is assigned.
In addition to the systematic uncertainties, MC statistical uncertainties are applied for each sample. They arise from a finite number of MC events.

6.6. Performing a Profile Likelihood Fit

The parameters of interest (POI) for the fits are the signal strengths of the SRs $\mu_{t\bar{t}Z}$ and μ_{tZq}, and the normalisation factor of the CR $\mathcal{N}_{VV}$. The signal strength $\mu = {}^{\sigma_{\text{meas}}}/_{\sigma_{\text{SM}}}$ is defined as the ratio of the measured cross-section σ_{meas} and the cross-section expected from the SM, σ_{SM}. The fitting procedure is a binned maximum profile-likelihood fit performed by the framework `TRExFitter`. The maximised likelihood is

$$\begin{aligned} L = & \prod_{i \in \text{bin}} P\left(n_i | \mu_{t\bar{t}Z} \cdot S_i^{t\bar{t}Z}(\vec{\theta}) + \mu_{tZq} \cdot S_i^{tZq}(\vec{\theta}) + \mathcal{N}_{VV} \cdot B_i^{VV}(\vec{\theta}) + B_i^{\text{non-}VV}(\vec{\theta})\right) \\ & \cdot \prod_{j \in \text{syst}} G\left(\theta_{0,j} | \theta_j, \Delta\theta_j\right) \end{aligned} \tag{6.5}$$

where P and G are Poisson and Gauß distributions. The variables n_i represent the data yields, while S_i^r and B_i^r represent the expected number of signal events and background events in bin i and region r. The notation $B_i^{\text{non-}VV}$ refers to background events from non-diboson processes in bin i. The vector $\vec{\theta}$ contains *nuisance parameters* (NPs) that can affect the signal and background events. The Gaussian distribution has a mean $\theta_{0,j} = 0$ and a standard deviation of $\Delta\theta_j = 1$.
This analysis does not use the real data from the ATLAS experiment for the fit. Instead, the MC predictions are used to create a pseudo data set. The NPs are expected to

fulfil $\left(\hat{\theta}-\theta_0\right)=0$ with the best fit value $\hat{\theta}$. Thus, the expected signal strength equals 1. This fitting procedure is called an *Asimov fit*. The uncertainties of the estimated signal strength are the percentage values of the true cross-section uncertainties. The Asimov data is fitted to $O_{t\bar{t}Z}$ in SR-$t\bar{t}Z$, O_{tZq} in SR-tZq, and O_{VV} in CR-VV. The results are the parameters $\mu_{t\bar{t}Z}$, μ_{tZq} and $\mathcal{N}_{VV}$. To estimate the impact of an NP on a POI, the expression $\left(\hat{\theta}-\theta_0\right)/\Delta\theta$ is calculated.

Fit Results

The simultaneous profile-likelihood fit of the SRs and CR is performed, using the Asimov pseudo-dataset and maximising the likelihood L from Equation 6.5. Figure 6.8 compares data and post-fit MC simulations in the SRs and CR for the respective fitting parameters. None of them shows a significant deviation between data and MC simulation. The results of the Asimov fit are:

$$\mu_{t\bar{t}Z}=1.00^{+0.07}_{-0.07} = 1.00^{+0.07}_{-0.06}\,(\text{stat.})\,^{+0.03}_{-0.03}\,(\text{syst.}) \tag{6.6}$$

$$\mu_{tZq}=1.00^{+0.17}_{-0.16} = 1.00^{+0.14}_{-0.14}\,(\text{stat.})\,^{+0.10}_{-0.09}\,(\text{syst.}) \tag{6.7}$$

$$\mathcal{N}_{VV}=1.00^{+0.17}_{-0.16} = 1.00^{+0.11}_{-0.10}\,(\text{stat.})\,^{+0.14}_{-0.12}\,(\text{syst.})\,. \tag{6.8}$$

The relative total uncertainty of $\mu_{t\bar{t}Z}$ is $\pm 7\,\%$. The largest contributions are statistical uncertainties ($+7\,\%$, $-6\,\%$). The impacts of each systematic uncertainty on $\mu_{t\bar{t}Z}$ and μ_{tZq} are ranked in Figure 6.9. The ranking of $\mathcal{N}_{VV}$ is shown in Figure E.2. The highest impact on the $t\bar{t}Z$ signal strengths' uncertainty is the simulation of tWZ and fake events. The tZq signal strength and $\mathcal{N}_{VV}$ have both a total uncertainty of $+17\,\%$ and $-16\,\%$. Thus, the $t\bar{t}Z$ measurement is more precise than the tZq measurement. This is expected from other analyses [81–84]. For the tZq measurement, statistical uncertainties dominate with $14\,\%$. The biggest impact on the statistical uncertainties is the simulation of fake events. Note that systematic uncertainties associated with jets and b-tagging are not applied in this analysis. Thus, the actual systematic uncertainty is probably higher. A comparison of these with other measurements is, therefore, not appropriate. However, a comparison of the statistical uncertainties is sensible.

The inclusive $t\bar{t}Z$ cross-section measurements, described in Refs. [81] and [84], have both relative statistical uncertainties of $6\,\%$ in the trileptonic channel. This is slightly more accurate than this measurement. In the tZq analysis from Ref. [82] the relative statistical uncertainty is $13.4\,\%$. This is, again, slightly more precise than this analysis. The systematic uncertainty of the tZq analysis of $7.2\,\%$ is significantly lower than that of this measurement, although not all systematic uncertainties are applied.

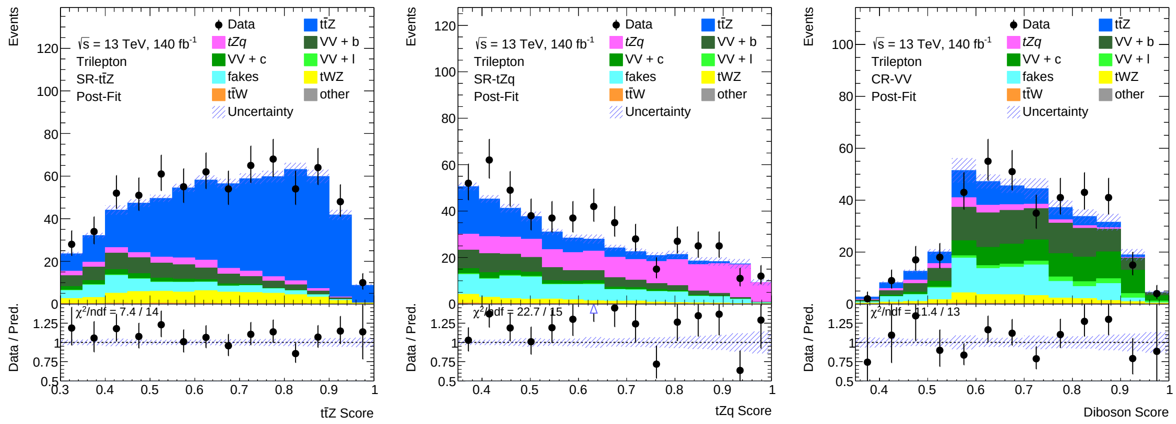


Figure 6.8.: Distribution of the $t\bar{t}Z$ (left), tZq (centre), diboson score (right) in the regions SR-$t\bar{t}Z$, SR-tZq and CR-VV. The lower panels show the data-MC agreement. Arrows indicate data outside the panels' range. Statistical and systematic uncertainties are included in the uncertainty band.

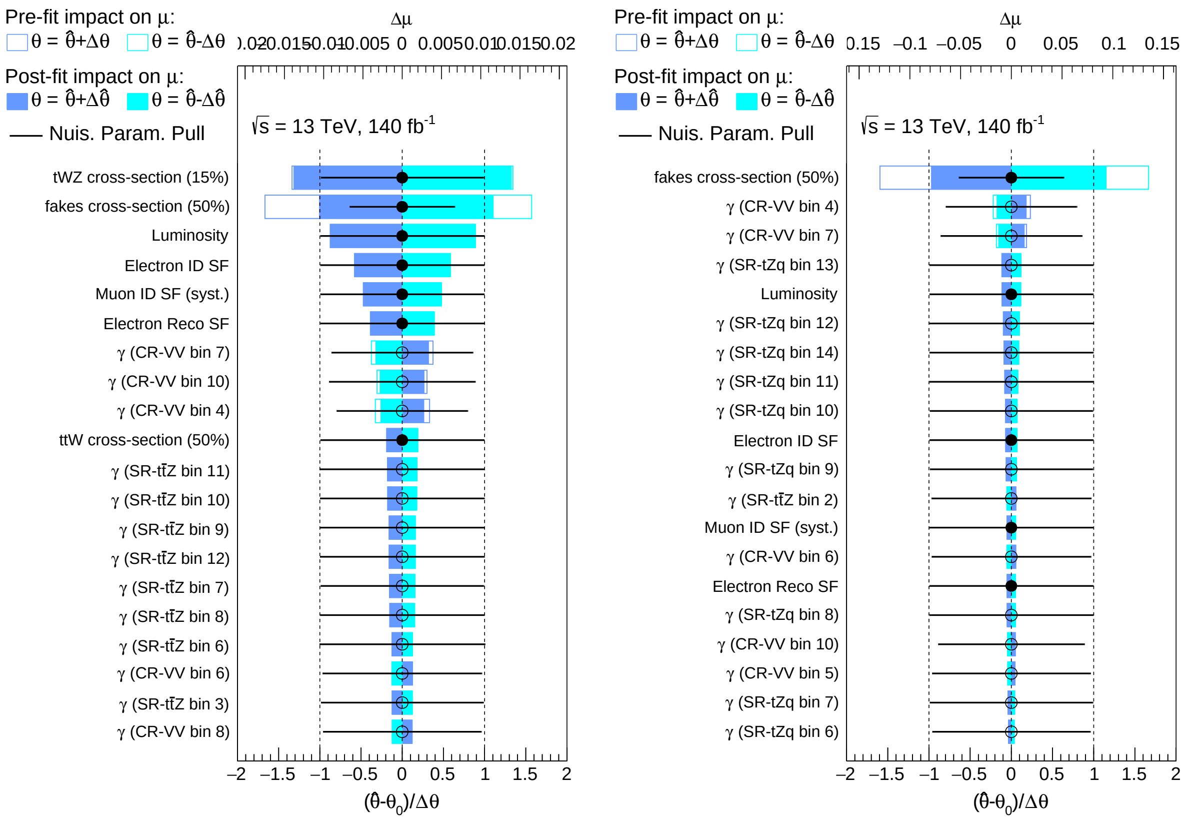


Figure 6.9.: Ranking plots of the uncertainties of the Asimov fit. The uncertainties are ranked by their impact on $\mu_{t\bar{t}Z}$ (left) and μ_{tZq} (right). The ranking for μ_{VV} is shown in Figure E.2. The unfilled blue and turquoise rectangles indicate the prefit, while the filled ones indicate the postfit.

The inclusive $t\bar{t}Z$ cross-section measurements, described in Refs. [81] and [84], have both relative statistical uncertainties of 6 % in the trileptonic channel. This is slightly more accurate than this measurement. In the tZq analysis from Ref. [82] the relative statistical uncertainty is 13.4 %. This is, again, slightly more precise than this analysis. The systematic uncertainty of the tZq analysis of 7.2 % is significantly lower than that of this measurement, although not all systematic uncertainties are applied.
One possible explanation for these differences in precision is the increased complexity introduced by the simultaneous modelling of both $t\bar{t}Z$ and tZq signal processes in a combined fit. In contrast to separate measurements, a joint analysis must disentangle overlapping signal regions, account for correlations between processes, and manage a larger number of nuisance parameters, all of which can lead to reduced sensitivity. Additionally, the combined fit strategy may intentionally trade off statistical power in individual channels to achieve a more global constraint on both processes.

7. Conclusion

In this work, the trileptonic channels of the $t\bar{t}Z$ and tZq processes were analysed in a joint measurement using ATLAS Run 2 data at $\mathcal{L} = 140\ \text{fb}^{-1}$. This allows the study of the direct couplings of top quarks and Z bosons. This chapter summarises the content, and the results are evaluated. Additionally, an outlook is given.

Summary and Evaluation

Chapter 2 introduces the properties of the SM, including the particle spectrum. The elementary particle with the highest mass is the top quark. Top quark pairs can only be produced in hadron colliders with the current accelerator technology. Their decay channels are differentiated between alljets, lepton + jets and dilepton. Of special interest are top quark productions associated with a Z boson because these processes can include a direct tZ coupling. An investigation of this coupling is crucial to test the SM. Two of these processes are $t\bar{t}Z$ and tZq, where the latter is much rarer.

In Chapter 3, the experimental setup is described, including an overview of the ATLAS detectors' layers and the object definitions. The following chapter introduced DNNs. These are complex computational models inspired by a biological nervous system. In particle physics, they are used to identify patterns in simulated data and classify them. Training and regularisation techniques like dropout layers, batch normalisation and k-fold cross-validation are described which improve the DNNs' generalisation and reduce overfitting. DNN models can be evaluated by analysing the separation power, the ROC curve, the AUC, and the permutation importance.

Chapter 5 is a summary of existing $t\bar{t}Z$ and tZq measurements and their limitations. Three analyses and their techniques are presented. The $t\bar{t}Z$ analysis from ATLAS [81] measures the inclusive and differential cross-section of the $t\bar{t}Z$ process in the trileptonic and tetraleptonic channels. Furthermore, the transverse momentum of the Z boson is unfolded. The ATLAS tZq analysis provides the inclusive tZq cross-section [82]. The measurement is performed in the trileptonic channel. In Ref. [83] the processes $t\bar{t}Z$, tZq and, tWZ are investigated jointly in tri- and tetraleptonic final states within a framework of an EFT.

The joint $t\bar{t}Z$ and tZq measurement analysis is outlined in Chapter 6. The Z boson is reconstructed, adding the four-momenta of two OSSF leptons, which lead to the invariant mass closest to m_Z. The reconstruction of the top quarks is more challenging because the neutrinos' four-momentum must be calculated from missing transverse momentum. A distinction is made between hadronically and leptonically decaying top quarks.
To reduce background from events including non-prompt leptons, lower limits on the transverse lepton momenta are set. The selection is optimised by estimating the efficiencies of certain requirements. This method targets to maximise S/F without losing too many signal events. By applying the lepton p_{T} requirements, S/F is increased by 389 % loosing only 12 % of the signal events.
The goal of MVA is to reduce the impact of the diboson background. Therefore, a DNN classifies $t\bar{t}Z$, tZq, and diboson events. It uses MC simulation sets as training data. Multiple methods like applying dropout layers, batch normalisation, and k-folding cross-validation are used to obtain an accurate discrimination power and avoid under- and overfitting. Parameters like the ROC curve and AUC are investigated to ensure appropriate generalisation. The input parameters with the normalised permutation importance are m_t^{lep}, N_{jets} and $p_{\mathrm{T}}(j_i)$. They thus contribute most to the discrimination. The separation power for $t\bar{t}Z$ and tZq is higher than for diboson events.
Two SRs for $t\bar{t}Z$ and tZq, and one CR for the diboson background are set up. The definition of the orthogonal regions is based on the DNN scores. The selection requirements are estimated in a procedure maximising the fraction of $t\bar{t}Z$ events in SR-$t\bar{t}Z$ and the fraction of tZq events in SR-tZq. As in the calculation of the $p_{\mathrm{T}}(\ell_i)$ requirements, signal events are prevented from being lost.
Theoretical and instrumental uncertainties are applied. These include luminosity, lepton reconstruction, lepton identification, lepton trigger efficiency and the cross-sections of the background processes. Uncertainties related to jets and b-tagging are not available. The statistical uncertainties are, therefore, not comparable with other measurements. The binned maximum profile-likelihood fit is performed with an Asimov dataset. The relative statistical uncertainties of $\mu_{t\bar{t}Z}$ and μ_{tZq} are comparable with other measurements in the trileptonic channel [81, 82, 84]. The main systematic uncertainties come from background modelling.

Outlook

While this work focuses on the trileptonic channel, extending the analysis to the dileptonic and tetraleptonic channels would provide valuable insights for $t\bar{t}Z$. This is done in the Ref. [84] and leads to a more accurate measurement of the $t\bar{t}Z$ process. Including this in a joint $t\bar{t}Z$ and tZq measurement is an important step towards comprehensively understanding these processes across different decay channels.

To further reduce the systematic uncertainties, it is necessary to minimise events of non-prompt leptons. The methods used in this work, i.e. applying conditions to the reconstructed Z boson mass and the transverse momentum of the leptons, already significantly reduce this background, but it is still a large source of systematic uncertainties. Here, it is also possible to apply MVA like DNNs. The DNN, which is used for classifying $t\bar{t}Z$, tZq and diboson events, shows a precise separation power. However, other approaches could also be tried here, such as adding more nodes, layers or dropout layers. Optimising the separation power will probably also improve the final results.

Another background that has a high influence on systematic uncertainties is the tWZ process. It has a very similar final state to $t\bar{t}Z$ and involves a direct coupling between a Z boson and a top quark. The analysis of Ref. [83] includes this in a joint $t\bar{t}Z$, tZq and tWZ measurement. This can also be considered here.

The last step is to apply the unfolding method as described in Chapter 5.1 and Ref. [81]. An appropriate unfolding parameter is the transverse momentum of the Z boson because it is the most EFT sensible parameter. Performing an unfolding is crucial since it leads to results on particle level which are directly comparable with theory.

Bibliography

[1] A. Einstein, *Zur Elektrodynamik bewegter Körper*, Annalen der Physik **322(10)**, 891 (1905)

[2] A. Einstein, *Die Grundlage der allgemeinen Relativitätstheorie*, Annalen der Physik **354(7)**, 769 (1916)

[3] M. Curie, *The Atomic Weight of Radium; A new Determination*, School Science and Mathematics **8(2)**, 104 (1908)

[4] E. Rutherford, *LXXIX. The scattering of α and β particles by matter and the structure of the atom*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science **21(125)**, 669 (1911)

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

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

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

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

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

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

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

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

[13] S. Weinberg, *The Making of the Standard Model*, Eur. Phys. J. C **34**, 5 (2004)

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

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

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

[17] F. Zwicky, *Die Rotverschiebung von extragalaktischen Nebeln*, Helvetica Physica Acta **6**, 110 (1933)

[18] WMAP Collaboration, *Nine-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results*, The Astrophysical Journal Supplement Series **208(2)**, 20 (2013)

[19] Super-Kamiokande Collaboration, *Evidence for Oscillation of Atmospheric Neutrinos*, Phys. Rev. Lett. **81**, 1562 (1998)

[20] SAGE Collaboration, *Solar neutrino results from SAGE*, Nuclear Physics B - Proceedings Supplements **91(1)**, 36 (2001)

[21] GNO Collaboration, *Complete results for five years of GNO solar neutrino observations*, Phys. Lett. B **616(3–4)**, 174–190 (2005)

[22] KamLAND Collaboration, *Measurement of Neutrino Oscillation with KamLAND: Evidence of Spectral Distortion*, Phys. Rev. Lett. **94**, 081801 (2005)

[23] SNO Collaboration, *Combined analysis of all three phases of solar neutrino data from the Sudbury Neutrino Observatory*, Phys. Rev. C **88**, 025501 (2013)

[24] Particle Data Group, *Review of Particle Physics*, PTEP **2022**, 083C01 (2022)

[25] C. S. Wu, et al., *Experimental Test of Parity Conservation in Beta Decay*, Phys. Rev. **105**, 1413 (1957)

[26] M. Goldhaber, L. Grodzins, A. W. Sunyar, *Helicity of Neutrinos*, Phys. Rev. **109**, 1015 (1958)

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

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

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

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

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

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

[33] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716(1)**, 30 (2012)

[34] LHCb Collaboration, *Observation of $J/\psi\phi$ structures consistent with exotic states from amplitude analysis of $B^+ \to J/\psi\phi K^+$ decays*, Phys. Rev. Lett. **118(2)** (2017)

[35] LHCb Collaboration, *Amplitude analysis of $B \to J/\psi\phi K^+$ decays*, Phys. Rev. D **95(1)** (2017)

[36] LHCb Collaboration, *Observation of $J/\psi p$ resonances consistent with pentaquark states in $\Lambda_b^0 \to J/\psi K^- p$ decays*, Phys. Rev. Lett. **115(7)** (2015)

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

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

[39] D. Pumplin, et al., *New Generation of Parton Distributions with Uncertainties from Global QCD Analysis*, JHEP **0207**, 012 (2002)

[40] A. D. Martin, et al., *Physical Gluons and High E_T Jets*, Phys. Lett. B **604**, 61 (2004)

[41] A. D. Martin, et al., *Uncertainties of predictions from parton distrbutions, I. Experimental errors*, Eur. Phys. J. C **28**, 455 (2003)

[42] ZEUS Collaboration, *An NLO QCD analysis of inclusive cross-section and jet-production data from the ZEUS experiment*, Eur. Phys. J. C **42**, 1 (2005)

[43] H1 Collaboration, *Deep Inelastic Inclusive eP Scattering at low x and a Determination of α_s*, Eur. Phys. J. C **21**, 33 (2001)

[44] W. T. Giele, S. A. Keller, *Implications of hadron collider observables on parton distribution function uncertainties*, Phys. Rev. D **58**, 094023 (1998)

[45] D. Stump, et al., *Uncertainties of predictions from parton distribution functions, I. The lagrange multiplier method*, Phys. Rev. D **65**, 014012 (2002)

[46] D. Stump, et al., *Uncertainties of predictions from parton distribution functions, II. The hessian method*, Phys. Rev. D **65**, 014013 (2002)

[47] J. Pumplin, et al., *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **0207**, 012 (2002)

[48] M. Botje, *A QCD analysis of HERA and fixed target structure function data*, Eur. Phys. J. C **14**, 285 (2000)

[49] S. I. Alekhin, *Parton Distributions from Deep Inelastic Scattering Data*, Phys. Rev. D **68**, 014002 (2003)

[50] R. D. Ball, et al., *Parton distributions from high-precision collider data*, Eur. Phys. J. C **77(10)** (2017)

[51] S. Bailey, et al., *Parton distributions from LHC, HERA, Tevatron and fixed target data: MSHT20 PDFs*, Eur. Phys. J. C **81(4)**, 341 (2021)

[52] DØ Collaboration, *Observation of Single Top-Quark Production*, Phys. Rev. Lett. **103**, 092001 (2009)

[53] CDF Collaboration, *Observation of Electroweak Single Top-Quark Production*, Phys. Rev. Lett. **103**, 092002 (2009)

[54] ATLAS Collaboration, *Measurement of the t-channel single top-quark production cross section in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, Phys. Lett. B **717(4)**, 330 (2012)

[55] CMS Collaboration, *Measurement of the t-Channel Single Top Quark Production Cross Section in pp Collisions at $\sqrt{s} = 7$ TeV*, Phys. Rev. Lett. **107**, 091802 (2011)

[56] ATLAS Collaboration, *Measurement of the production cross-section of a single top quark in association with a Z boson in proton–proton collisions at 13 TeV with the ATLAS detector*, Phys. Lett. B **780**, 557 (2018)

[57] CMS Collaboration, *Observation of Single Top Quark Production in Association with a Z Boson in Proton-Proton Collisions at* $\sqrt{s} = 13$ TeV, Phys. Rev. Lett. **122**, 132003 (2019)

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

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

[60] CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST **3(08)**, S08004 (2008)

[61] LHCb Collaboration, *The LHCb Detector at the LHC*, JINST **3(08)**, S08005 (2008), also published by CERN Geneva in 2010

[62] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3(08)**, S08002 (2008)

[63] 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)** (2019)

[64] 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)

[65] 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)** (2016)

[66] 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)

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

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

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

[70] ATLAS Collaboration, *Identification and rejection of pile-up jets at high pseudorapidity with the ATLAS detector*, Eur. Phys. J. C **77(9)** (2017)

[71] ATLAS Collaboration, *Neural Network Jet Flavour Tagging with the Upgraded ATLAS Inner Tracker Detector at the High-Luminosity LHC*, Technical Report ATL-PHYS-PUB-2022-047, CERN, Geneva (2022)

[72] ATLAS Collaboration, *Deep Sets based Neural Networks for Impact Parameter Flavour Tagging in ATLAS*, Technical Report ATL-PHYS-PUB-2020-014, CERN, Geneva (2020)

[73] ATLAS Collaboration, *Topological b-hadron decay reconstruction and identification of b-jets with the JetFitter package in the ATLAS experiment at the LHC*, Technical Report ATL-PHYS-PUB-2018-025, CERN, Geneva (2018)

[74] ATLAS Collaboration, *Secondary vertex finding for jet flavour identification with the ATLAS detector*, Technical Report ATL-PHYS-PUB-2017-011, CERN, Geneva (2017)

[75] 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)

[76] W. S. McCulloch, W. Pitts, *A Logical Calculus of the Ideas Immanent in Nervous Activity*, Bull. Math. Biophys. **5(4)**, 115 (1943)

[77] F. Rosenblatt, *The perceptron - A perceiving and recognizing automaton*, Technical Report 85-460-1, Cornell Aeronautical Laboratory, Ithaca, New York (1957)

[78] D. E. Rumelhart, G. E. Hinton, R. J. Williams, *Learning internal representations by error propagation*, page 318–362, MIT Press, Cambridge, MA, USA (1986)

[79] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, *Dropout: a simple way to prevent neural networks from overfitting*, J. Mach. Learn. Res. **15(1)**, 1929–1958 (2014)

[80] S. Ioffe, C. Szegedy, *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*, in F. Bach, D. Blei, editors, *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pages 448–456, PMLR, Lille, France (2015)

[81] 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(8)**, 737 (2021)

[82] ATLAS Collaboration, *Observation of the associated production of a top quark and a Z boson in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **2020(7)** (2020)

[83] CMS Collaboration, *Probing effective field theory operators in the associated production of top quarks with a Z boson in multilepton final states at $\sqrt{s} = 13$ TeV*, JHEP **2021(12)** (2021)

[84] ATLAS Collaboration, *Inclusive and differential cross section measurements of $t\bar{t}Z$ production in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector, including EFT and spin correlations interpretations*, Technical Report ATLAS-CONF-2023-065, CERN, Geneva (2023)

[85] A. Kulesza, et al., *Associated production of a top quark pair with a heavy electroweak gauge boson at NLO + NNLL accuracy*, Eur. Phys. J. C **79(3)** (2019)

[86] G. D'Agostini, *A multidimensional unfolding method based on Bayes' theorem*, Nucl. Instrum. Meth. A **362(2)**, 487 (1995)

[87] ATLAS Collaboration, *Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, Eur. Phys. J. C **83(10)** (2023)

[88] T. Sjöstrand, S. Mrenna, P. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178(11)**, 852 (2008)

[89] T. Sjöstrand, S. Ask, J. R. Christiansen, R. Corke, N. Desai, P. Ilten, S. Mrenna, S. Prestel, C. O. Rasmussen, P. Z. Skands, *An Introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191**, 159 (2015)

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

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

[92] R. Frederix, et al., *The automation of next-to-leading order electroweak calculations*, JHEP **2018(7)**, 185 (2018)

[93] S. Agostinelli, et al., *Geant4—A Simulation Toolkit*, Nucl. Instrum. Meth. A **506(3)**, 250 (2003)

[94] J. Allison, et al., *Geant4 Developments and Applications*, IEEE Trans. Nucl. Sci. **53(1)**, 270 (2006)

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

[96] D. P. Kingma, J. Ba, *Adam: A Method for Stochastic Optimization*, in Y. Bengio, Y. LeCun, editors, *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings* (2015)

[97] Y. E. Nesterov, *A Method for Solving the Convex Programming Problem with Convergence Rate $O(1/k^2)$*, Soviet Mathematics Doklady **27**, 372 (1983)

[98] T. Dozat, *Incorporating Nesterov Momentum into Adam*, in *Proceedings of the 4th International Conference on Learning Representations, Workshop Track*, pages 1–4, San Juan, Puerto Rico (2016)

A. Activation Functions

Binary / Heaviside step function:

$$H(x_j^{(l)}) = \begin{cases} 0 & \text{for } x_j^{(l)} < 0 \\ 1 & \text{for } x_j^{(l)} \geq 0 \; . \end{cases} \tag{A.1}$$

Sigmoid / Logistic function:

$$\sigma_{\text{sig}}(x_j^{(l)}) = \frac{1}{1 + \exp\left(x_j^{(l)}\right)} \tag{A.2}$$

Hyperbolic Tangent:

$$\tanh\,(x_j^{(l)}) = \frac{\exp\left(x_j^{(l)}\right) - \exp\left(-x_j^{(l)}\right)}{\exp\left(x_j^{(l)}\right) + \exp\left(-x_j^{(l)}\right)} \tag{A.3}$$

B. Monte Carlo Samples

Table B.1.: Used MC samples for the MC campaign `MC20a` and processes tZq, $t\bar{t}Z$, diboson, $t\bar{t}W$, tWZ and other. The samples for the campaigns `MC20d` and `MC20e` have the same name but with the `r-tags` `r13144` and `r13145`.

tZq
mc20_13TeV:mc20_13TeV.410560.MadGraphPythia8EvtGen_A14_tZ_4fl_tchan_noAllHad.deriv.DAOD_PHYS.e5803_s3681_r13167_p5855
$t\bar{t}Z$
mc20_13TeV:mc20_13TeV.504330.aMCPy8EG_NNPDF30NLO_A14N23LO_ttee.deriv.DAOD_PHYS.e8255_s3797_r13167_p5855
mc20_13TeV:mc20_13TeV.504334.aMCPy8EG_NNPDF30NLO_A14N23LO_ttmumu.deriv.DAOD_PHYS.e8255_s3797_r13167_p5855
mc20_13TeV:mc20_13TeV.504338.aMCPy8EG_NNPDF30NLO_A14N23LO_ttZqq.deriv.DAOD_PHYS.e8255_s3797_r13167_p5855
mc20_13TeV:mc20_13TeV.504342.aMCPy8EG_NNPDF30NLO_A14N23LO_tttautau.deriv.DAOD_PHYS.e8255_s3797_r13167_p5855
mc20_13TeV:mc20_13TeV.504346.aMCPy8EG_NNPDF30NLO_A14N23LO_ttZnunu.deriv.DAOD_PHYS.e8255_s3797_r13167_p5855
Diboson
mc20_13TeV:mc20_13TeV.364250.Sherpa_222_NNPDF30NNLO_llll.deriv.DAOD_PHYS.e5894_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364253.Sherpa_222_NNPDF30NNLO_lllv.deriv.DAOD_PHYS.e5916_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364284.Sherpa_222_NNPDF30NNLO_lllvjj_EW6.deriv.DAOD_PHYS.e6055_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364288.Sherpa_222_NNPDF30NNLO_llll_lowMllPtComplement.deriv.DAOD_PHYS.e6096_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.345705.Sherpa_222_NNPDF30NNLO_ggllll_0M4l130.deriv.DAOD_PHYS.e6213_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.345706.Sherpa_222_NNPDF30NNLO_ggllll_130M4l.deriv.DAOD_PHYS.e6213_s3681_r13167_p5855
$t\bar{t}W$
mc20_13TeV:mc20_13TeV.700168.Sh_2210_ttW.deriv.DAOD_PHYS.e8273_s3797_r13167_p5855
mc20_13TeV:mc20_13TeV.700205.Sh_2210_ttW_EWK.deriv.DAOD_PHYS.e8307_s3797_r13167_p5855
mc20_13TeV:mc20_13TeV.700000.Sh_228_ttW.deriv.DAOD_PHYS.e7793_s3681_r13167_p5855
tWZ
mc20_13TeV:mc20_13TeV.410408.aMcAtNloPythia8EvtGen_tWZ_Ztoll_minDR1.deriv.DAOD_PHYS.e6423_s3681_r13167_p5855
Other
mc20_13TeV:mc20_13TeV.363358.Sherpa_221_NNPDF30NNLO_WqqZll.deriv.DAOD_PHYS.e5525_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.363359.Sherpa_221_NNPDF30NNLO_WpqqWmlv.deriv.DAOD_PHYS.e5583_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.363360.Sherpa_221_NNPDF30NNLO_WplvWmqq.deriv.DAOD_PHYS.e5983_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364242.Sherpa_222_NNPDF30NNLO_WWW_3l3v_EW6.deriv.DAOD_PHYS.e5887_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364243.Sherpa_222_NNPDF30NNLO_WWZ_4l2v_EW6.deriv.DAOD_PHYS.e5887_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364244.Sherpa_222_NNPDF30NNLO_WWZ_2l4v_EW6.deriv.DAOD_PHYS.e5887_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364245.Sherpa_222_NNPDF30NNLO_WZZ_5l1v_EW6.deriv.DAOD_PHYS.e5887_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364246.Sherpa_222_NNPDF30NNLO_WZZ_3l3v_EW6.deriv.DAOD_PHYS.e5887_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364247.Sherpa_222_NNPDF30NNLO_ZZZ_6l0v_EW6.deriv.DAOD_PHYS.e5887_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364248.Sherpa_222_NNPDF30NNLO_ZZZ_4l2v_EW6.deriv.DAOD_PHYS.e5887_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364249.Sherpa_222_NNPDF30NNLO_ZZZ_2l4v_EW6.deriv.DAOD_PHYS.e5887_s3681_r13167_p5855

Table B.2.: Used MC samples for the MC campaign `MC20a` and processes including only non-prompt leptons. The samples for the campaigns `MC20d` and `MC20e` have the same name but with the `r-tags` `r13144` and `r13145`.

Fakes
mc20_13TeV:mc20_13TeV.410389.MadGraphPythia8EvtGen_A14NNPDF23_ttgamma_nonallhadronic.deriv.DAOD_PHYS.e6155_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364100.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_PHYS.e5271_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364101.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_PHYS.e5271_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364102.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV0_70_BFilter.deriv.DAOD_PHYS.e5271_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364103.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_PHYS.e5271_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364104.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_PHYS.e5271_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364105.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV70_140_BFilter.deriv.DAOD_PHYS.e5271_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364106.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_PHYS.e5271_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364107.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_PHYS.e5271_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364108.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV140_280_BFilter.deriv.DAOD_PHYS.e5271_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364109.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_PHYS.e5271_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364110.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_PHYS.e5271_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364111.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV280_500_BFilter.deriv.DAOD_PHYS.e5271_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364112.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV500_1000.deriv.DAOD_PHYS.e5271_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364113.Sherpa_221_NNPDF30NNLO_Zmumu_MAXHTPTV1000_E_CMS.deriv.DAOD_PHYS.e5271_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364114.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_PHYS.e5299_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364115.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_PHYS.e5299_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364116.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV0_70_BFilter.deriv.DAOD_PHYS.e5299_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364117.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_PHYS.e5299_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364118.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_PHYS.e5299_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364119.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV70_140_BFilter.deriv.DAOD_PHYS.e5299_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364120.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_PHYS.e5299_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364121.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_PHYS.e5299_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364122.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV140_280_BFilter.deriv.DAOD_PHYS.e5299_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364123.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_PHYS.e5299_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364124.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_PHYS.e5299_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364125.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV280_500_BFilter.deriv.DAOD_PHYS.e5299_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364126.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV500_1000.deriv.DAOD_PHYS.e5299_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364127.Sherpa_221_NNPDF30NNLO_Zee_MAXHTPTV1000_E_CMS.deriv.DAOD_PHYS.e5299_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364128.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_CVetoBVeto.deriv.DAOD_PHYS.e5307_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364129.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_CFilterBVeto.deriv.DAOD_PHYS.e5307_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364130.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV0_70_BFilter.deriv.DAOD_PHYS.e5307_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364131.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_CVetoBVeto.deriv.DAOD_PHYS.e5307_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364132.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_CFilterBVeto.deriv.DAOD_PHYS.e5307_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364133.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV70_140_BFilter.deriv.DAOD_PHYS.e5307_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364134.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_CVetoBVeto.deriv.DAOD_PHYS.e5307_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364135.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_CFilterBVeto.deriv.DAOD_PHYS.e5307_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364136.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV140_280_BFilter.deriv.DAOD_PHYS.e5307_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364137.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_CVetoBVeto.deriv.DAOD_PHYS.e5307_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364138.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_CFilterBVeto.deriv.DAOD_PHYS.e5313_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364139.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV280_500_BFilter.deriv.DAOD_PHYS.e5313_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364140.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV500_1000.deriv.DAOD_PHYS.e5307_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.364141.Sherpa_221_NNPDF30NNLO_Ztautau_MAXHTPTV1000_E_CMS.deriv.DAOD_PHYS.e5307_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.410472.PhPy8EG_A14_ttbar_hdamp258p75_dil.deriv.DAOD_PHYS.e6348_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.700011.Sh_228_eegamma_pty7_EnhMaxpTVpTy.deriv.DAOD_PHYS.e7947_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.700012.Sh_228_mmgamma_pty7_EnhMaxpTVpTy.deriv.DAOD_PHYS.e7947_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.700013.Sh_228_ttgamma_pty7_EnhMaxpTVpTy.deriv.DAOD_PHYS.e7947_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.700014.Sh_228_vvgamma_pty7_EnhMaxpTVpTy.deriv.DAOD_PHYS.e7947_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.700015.Sh_228_evgamma_pty7_EnhMaxpTVpTy.deriv.DAOD_PHYS.e7947_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.700016.Sh_228_mvgamma_pty7_EnhMaxpTVpTy.deriv.DAOD_PHYS.e7947_s3681_r13167_p5855
mc20_13TeV:mc20_13TeV.700017.Sh_228_tvgamma_pty7_EnhMaxpTVpTy.deriv.DAOD_PHYS.e7947_s3681_r13167_p5855

C. Requirements on the Lepton Transverse Momenta

Table C.1.: Total yields and several parameters for different combinations of $p_{\mathrm{T}}(\ell_i)$ requirements. The p_{T} requirement of the leading lepton is $p_{\mathrm{T}}(\ell_1) > 27$ GeV. The shown variables are the number of signal event S, of signal events for the loosest $p_{\mathrm{T}}(\ell_i)$ cut S_0, of background events B, of fake events F, of lost signal events L and the efficiencies $Z(B)$, $Z(\sqrt{B})$, $Z(F)$ and $Z(\sqrt{F})$.

$p_{\mathrm{T}}(\ell_1)/\mathrm{GeV}$	27	27	27	27	27	27	27	27	27	27	27	27	27	27	27
$p_{\mathrm{T}}(\ell_2)/\mathrm{GeV}$	25	25	25	25	25	20	20	20	20	15	15	15	10	10	7
$p_{\mathrm{T}}(\ell_3)/\mathrm{GeV}$	25	20	15	10	7	20	15	10	7	15	10	7	10	7	7
$t\bar{t}Z$	510	591	658	706	723	595	666	718	737	669	723	743	724	745	745
tZq	154	187	214	232	238	189	218	239	245	219	241	248	242	249	249
$VV+b$	199	236	267	291	301	238	271	297	308	272	300	312	301	313	313
$VV+c$	355	421	474	513	528	424	482	523	541	483	527	547	528	548	548
$VV+l$	207	249	283	309	320	251	288	317	329	290	320	334	321	335	335
fakes	109	160	264	560	1242	164	284	623	1398	291	673	1497	685	1554	1561
tWZ	91	105	117	126	128	105	118	127	131	119	128	132	129	132	132
$t\bar{t}W$	5	6	7	8	8	6	7	8	8	8	8	8	8	8	8
other	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Total	1631	1956	2284	2745	3490	1972	2336	2854	3698	2351	2922	3822	2938	3885	3892
S	664	778	872	938	961	784	884	957	982	888	964	991	966	994	994
B	967	1178	1412	1807	2529	1188	1452	1897	2716	1463	1958	2831	1972	2891	2898
F	109	160	264	560	1242	164	284	623	1398	291	673	1497	685	1554	1561
L	330	216	122	56	33	210	110	37	12	106	30	3	28	0	0
L/S_0	0.50	0.28	0.14	0.06	0.03	0.27	0.12	0.04	0.01	0.12	0.03	0.003	0.03	0	0
S/B	0.69	0.66	0.62	0.52	0.38	0.66	0.61	0.50	0.36	0.61	0.49	0.35	0.49	0.34	0.34
$S/\sqrt{B}$	21.4	22.7	23.2	22.1	19.1	22.7	23.2	22.0	18.8	23.2	21.8	18.6	21.8	18.5	18.5
S/F	6.09	4.86	3.30	1.68	0.77	4.78	3.11	1.54	0.70	3.05	1.43	0.66	1.41	0.64	0.64
$S/\sqrt{F}$	63.6	61.5	53.7	39.6	27.3	61.2	52.5	38.3	26.3	52.1	37.2	25.6	36.9	25.2	25.2
$Z(B)$	1.01	1.39	1.55	1.42	1.07	1.41	1.55	1.41	1.04	1.56	1.39	1.02	1.39	1.00	1
$Z(\sqrt{B})$	0.58	0.89	1.08	1.12	1.00	0.90	1.10	1.14	1.01	1.11	1.14	1.01	1.14	1.00	1
$Z(F)$	4.81	5.52	4.46	2.47	1.17	5.50	4.28	2.32	1.09	4.22	2.18	1.04	2.15	1.00	1
$Z(\sqrt{F})$	0.64	0.92	1.02	0.98	0.94	0.93	1.03	1.00	0.98	1.03	1.00	0.99	1.00	1.00	1

Table C.2.: Total yields and several parameters for different combinations of $p_{\mathrm{T}}(\ell_i)$ requirements. The p_{T} requirement of the leading lepton is $p_{\mathrm{T}}(\ell_1) > 30$ GeV. The shown variables are the number of signal event S, of signal events for the loosest $p_{\mathrm{T}}(\ell_i)$ cut S_0, of background events B, of fake events F, of lost signal events L and the efficiencies $Z(B)$, $Z(\sqrt{B})$, $Z(F)$ and $Z(\sqrt{F})$.

$p_{\mathrm{T}}(\ell_1)$/GeV	30	30	30	30	30	30	30	30	30	30	30	30	30	30	30
$p_{\mathrm{T}}(\ell_2)$/GeV	25	25	25	25	25	20	20	20	20	15	15	15	10	10	7
$p_{\mathrm{T}}(\ell_3)$/GeV	25	20	15	10	7	20	15	10	7	15	10	7	10	7	7
$t\bar{t}Z$	510	591	658	706	723	594	666	718	736	668	723	743	724	744	744
tZq	154	187	214	232	238	189	218	239	245	219	241	248	241	249	249
$VV+b$	199	236	267	291	301	238	271	297	308	272	300	312	301	313	313
$VV+c$	354	421	474	512	528	424	481	523	541	483	527	546	528	548	548
$VV+l$	207	249	283	309	320	251	288	316	329	289	320	333	321	335	335
fakes	109	160	264	559	1241	164	282	620	1392	290	668	1486	681	1542	1548
tWZ	91	105	117	126	128	105	118	127	131	119	128	132	129	132	132
$t\bar{t}W$	5	6	7	8	8	6	7	8	8	8	8	8	8	8	8
other	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Total	1630	1956	2284	2743	3488	1972	2334	2849	3691	2349	2916	3809	2932	3871	3878
S	664	778	872	938	961	783	884	957	981	887	964	991	965	993	993
B	966	1178	1412	1805	2527	1189	1450	1892	2710	1462	1952	2818	1967	2878	2885
F	109	160	264	559	1241	164	282	620	1392	290	668	1486	681	1542	1548
L	330	216	122	56	33	211	110	37	13	107	30	3	29	1	1
L/S_0	0.33	0.28	0.14	0.06	0.03	0.27	0.12	0.04	0.01	0.12	0.03	0.003	0.03	0.001	0.001
S/B	0.69	0.66	0.62	0.52	0.38	0.66	0.61	0.51	0.36	0.61	0.49	0.35	0.49	0.35	0.34
$S/\sqrt{B}$	21.4	22.7	23.2	22.1	19.1	22.7	23.2	22.0	18.8	23.2	21.8	18.7	21.8	18.5	18.5
S/F	6.09	4.86	3.30	1.68	0.77	4.77	3.13	1.54	0.70	3.06	1.44	0.67	1.42	0.64	0.64
$S/\sqrt{F}$	63.6	61.5	53.7	39.7	27.3	61.1	52.6	38.4	26.3	52.1	37.3	25.7	37.0	25.3	25.2
$Z(B)$	1.34	1.39	1.55	1.42	1.07	1.40	1.56	1.42	1.04	1.56	1.40	1.02	1.39	1.00	1.00
$Z(\sqrt{B})$	0.77	0.89	1.08	1.12	1.00	0.90	1.10	1.15	1.01	1.10	1.14	1.01	1.14	1.00	1.00
$Z(F)$	6.39	5.52	4.46	2.48	1.17	5.48	4.31	2.33	1.09	4.22	2.20	1.04	2.16	1.01	1.01
$Z(\sqrt{F})$	0.84	0.92	1.02	0.98	0.94	0.92	1.03	1.00	0.98	1.03	1.00	0.99	1.00	1.00	1.00

D. DNN Evaluation

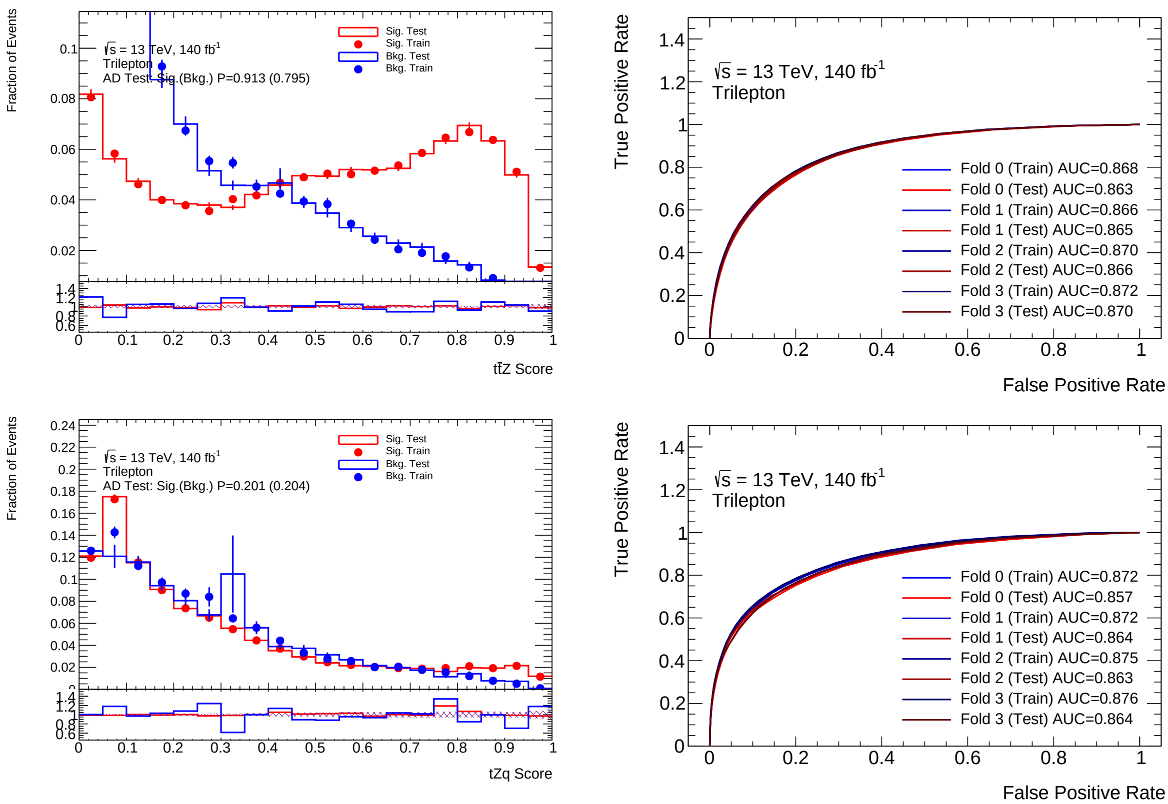


Figure D.1.: Fraction of Signal (blue) and background (red) events for fold 1 (left) and receiver operating characteristic (ROC) curves (right) for the $t\bar{t}Z$ (top), tZq (bottom). All plots compare data and testing. The lower band of the plots to the left shows the training over testing ratio. The shaded error bands are the statistical uncertainty. The comparison of training and data does not show a significant deviation, and the ROC curves form a distinct arc, comparable with the arcs of Figure 4.5. This indicates an accurately generalising DNN model.

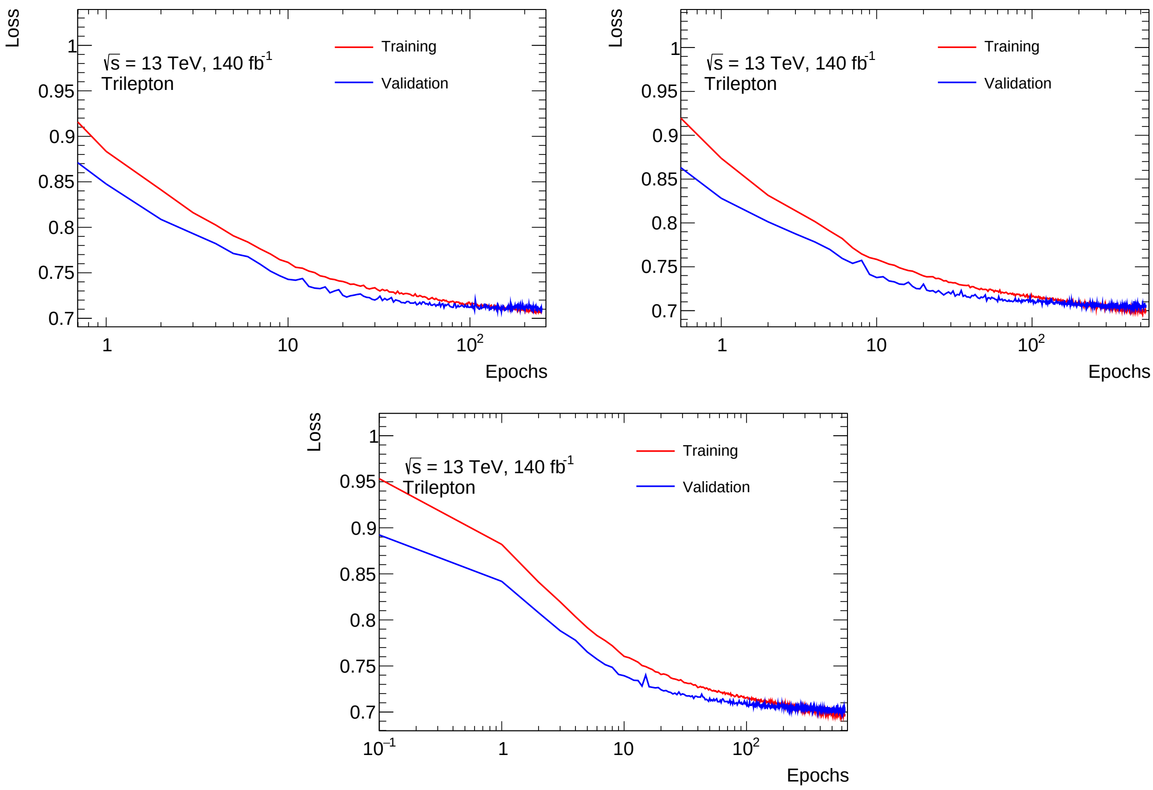


Figure D.2.: Loss functions for folds 0 (top left), 2 (top right) and 3 (bottom). The shapes are comparable with the shapes of Figure 4.3. At the beginning, the training loss (red) is higher than the validation loss (blue) due to dropout features. At the end of the training, the validation loss is higher than the training loss and reaches its minimum.

E. Supplementary Post-Fit Distributions

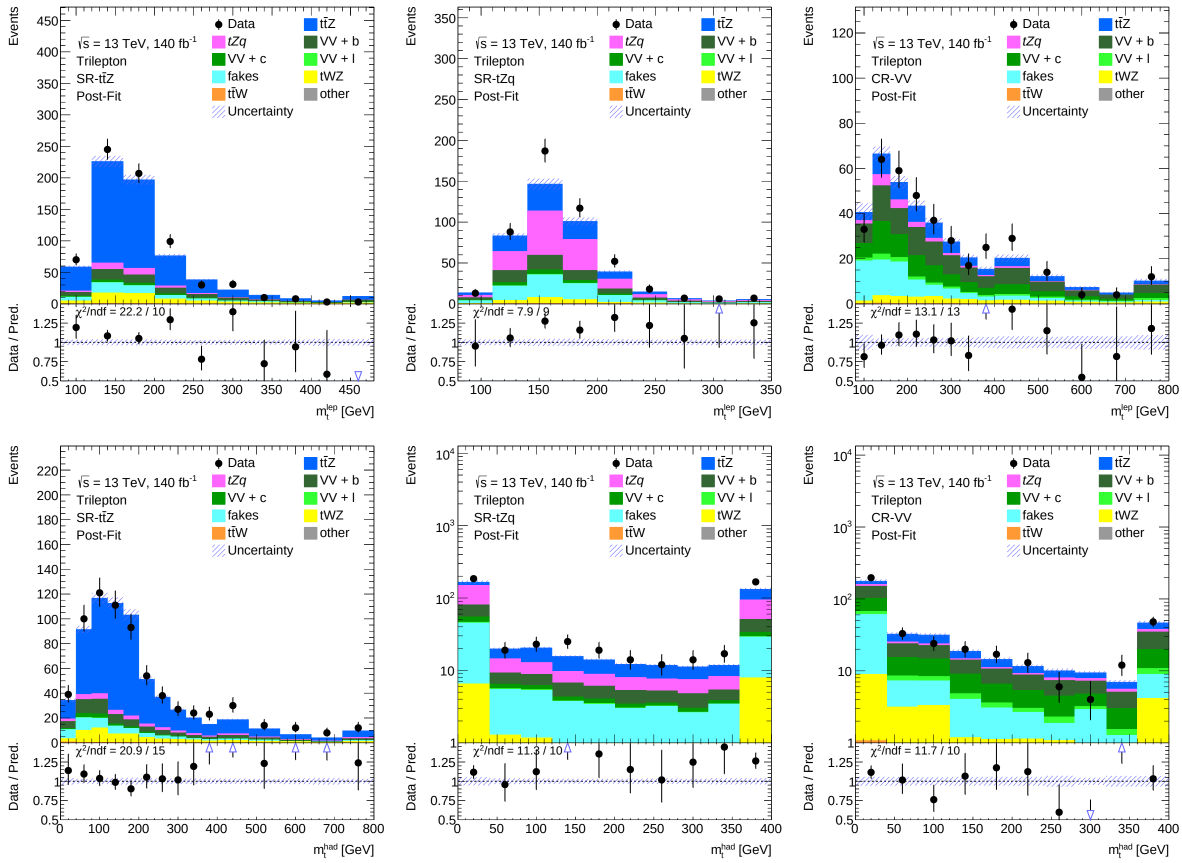


Figure E.1.: Post-Fit distributions of the leptonic top quark mass (top) and the hadronic top quark mass (bottom) in the regions SR-$t\bar{t}Z$, SR-tZq and CR-VV (from left to right). Overflow events are included in the rightmost bin. The lower panels show the data-MC agreement. Arrows indicate data outside of the panels' range. Statistical and systematic uncertainties are included in the uncertainty band.

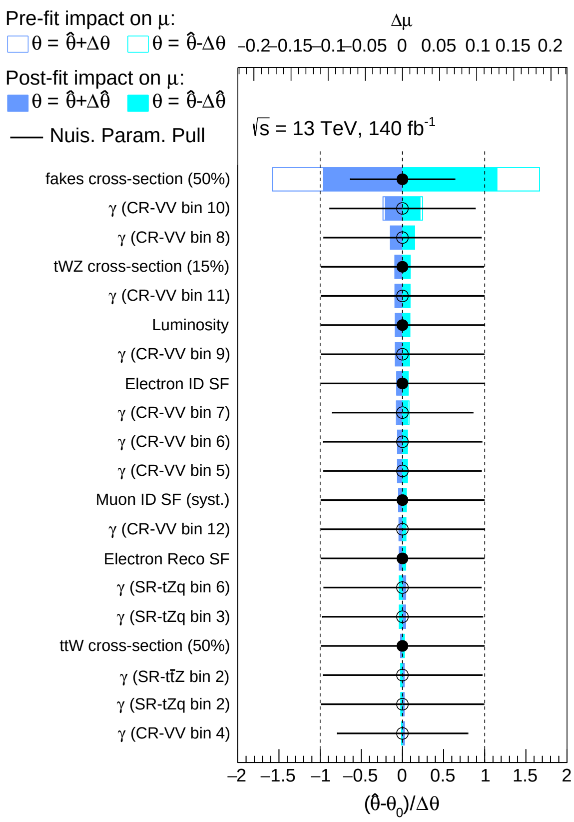


Figure E.2.: Ranking plots of the uncertainties of the Asimov fit. The uncertainties are ranked by their impact on μ_{VV}. The unfilled blue and turquoise rectangles indicate the prefit, while the filled ones indicate the postfit.

Acknowledgements

I would like to take this opportunity to thank everyone who has contributed to the preparation of this Master's thesis, whether through support or motivation. First, I would like to thank Prof. Dr. Arnulf Quadt for giving me the opportunity to write my Master's thesis in particle physics. I truly enjoyed it, and it was great to be part of the ATLAS community and to work for the world's largest particle accelerator. I would also like to thank Prof. Dr. Ariane Frey for reviewing this Master's thesis in her role as second examiner. I would like to thank Dr. Steffen Korn and Dr. Baptiste Ravina for their active supervision. Your support and willingness to answer all my questions in detail made this Master's thesis possible.
In addition to all the official supervisors, I would like to thank all the members of the II Physics Institute. You are a really cool group of people! Finally, I would like to thank my parents for their active support throughout my studies and my girlfriend Pauline for proofreading this Master's thesis.

Erklärung zur Nutzung von ChatGPT und vergleichbaren Werkzeugen im Rahmen von Prüfungen

In der hier vorliegenden Arbeit habe ich ChatGPT oder eine andere KI wie folgt genutzt:

- ☐ gar nicht
- ☐ bei der Ideenfindung
- ☐ bei der Erstellung der Gliederung
- ☐ zum Erstellen einzelner Passagen, insgesamt im Umfang von ... % am gesamten Text
- ☐ zur Entwicklung von Software-Quelltexten
- ☒ zur Optimierung oder Umstrukturierung von Software-Quelltexten
- ☒ zum Korrekturlesen oder Optimieren
- ☐ Weiteres, nämlich: ...

Ich versichere, alle Nutzungen vollständig angegeben zu haben. Fehlende oder fehlerhafte Angaben werden als Täuschungsversuch gewertet.

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 19. Juli 2025

(Sebastian Alexander Preuth)