

Single-Top t -channel Total Production Cross Section Measurement at $\sqrt{s} = 13$ TeV with ATLAS

Simon Thiele


Masterarbeit in Physik
angefertigt im Physikalischen Institut

vorgelegt der
Mathematisch-Naturwissenschaftlichen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität
Bonn

November 2022

I hereby declare that this thesis was formulated by myself and that no sources or tools other than those cited were used.

Bonn, 21st November 2022
Date


Signature

1. Gutachter: Prof. Dr. Ian C. Brock
2. Gutachterin: Dr. Tatjana Lenz

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 1 |
| 2 | The Standard Model of Particle Physics | 2 |
| 2.1 | Force Carrier Bosons | 2 |
| 2.2 | Fermions | 4 |
| 2.3 | The Higgs Boson | 5 |
| 2.4 | Mathematical Formulation | 5 |
| 2.5 | Open Questions | 6 |
| 2.6 | Top-Physics | 6 |
| 2.6.1 | Top Production Modes | 6 |
| 2.6.2 | Top Decay Modes | 7 |
| 3 | Large Hadron Collider | 8 |
| 3.1 | Collider Physics | 9 |
| 3.1.1 | Natural Units | 9 |
| 3.1.2 | Collider Properties | 10 |
| 4 | ATLAS Detector | 12 |
| 4.1 | Detector Physics | 13 |
| 4.1.1 | Semiconductor Detectors | 13 |
| 4.1.2 | Gaseous Detectors | 13 |
| 4.1.3 | Scintillators and Photomultipliers | 14 |
| 4.1.4 | Calorimeters | 15 |
| 4.2 | ATLAS Coordinate System | 15 |
| 4.3 | Detector Parts | 17 |
| 4.3.1 | Inner Detector | 17 |
| 4.3.2 | Calorimeters | 18 |
| 4.3.3 | Muon Spectrometer | 19 |
| 4.4 | Particle Identification | 19 |
| 4.5 | Object Definitions | 19 |
| 4.5.1 | Jets | 19 |
| 4.5.2 | b -tagging | 21 |
| 4.5.3 | Charged Leptons | 21 |
| 4.5.4 | Missing Transverse Energy | 22 |

| | | |
|----------|---|-----------|
| 4.6 | Data and Monte Carlo Samples | 22 |
| 4.6.1 | Grid Computing | 23 |
| 5 | Non-prompt and Fake Charged Leptons | 24 |
| 5.1 | Lepton Terminology | 24 |
| 5.2 | Non-prompt Leptons | 24 |
| 5.3 | Fake Leptons | 25 |
| 5.4 | Matrix-Method | 25 |
| 5.4.1 | Efficiencies | 26 |
| 6 | Analysis | 30 |
| 6.1 | The Single-Top t -channel | 30 |
| 6.2 | Event Selection | 30 |
| 6.3 | Fake Estimation | 32 |
| 6.4 | Signal Background Separation | 34 |
| 6.4.1 | Artificial Neural Networks | 34 |
| 6.4.2 | Neural Network Output | 34 |
| 6.5 | Fit and Cross Section Measurement | 35 |
| 6.5.1 | Maximum Likelihood Fit | 35 |
| 6.5.2 | Statistics Only Fit | 37 |
| 6.5.3 | Cross Section Uncertainty Only Fit | 38 |
| 7 | Summary and Conclusion | 41 |
| A | Full Systematic Fit | 44 |
| A.1 | Systematics | 44 |
| B | Monte Carlo Samples | 46 |
| C | E_T^{miss}- and p_T^ℓ-Distributions in the Control Regions | 47 |
| D | Neural Network Training | 50 |
| E | XRootD for TopLoop derived Code on the ATLAS Grid | 53 |
| E.1 | Instructions | 53 |
| E.2 | HTCondor Batch Jobs | 54 |
| E.3 | Notes, Comments, Tips and Hints | 55 |
| | List of Figures | 62 |
| | List of Tables | 64 |
| | Acknowledgements | 65 |

Introduction

The fundamental building blocks of matter have been a fascination of humans long before the advent of modern science. Even ancient philosophers already thought about this question and the Greek work for indivisible is still used today to label atoms. Modern advances in nuclear physics and the relatively young discipline of particle physics however have found that to be a misnomer. In fact, atoms are made up of protons, neutrons, and electrons, where protons and neutrons are even further divisible.

In the latter half of the 20th century advances in accelerator technology lead to the discovery of hundreds of new particles, believed to be distinct elementary particles. Later the discovery of quarks shed light onto the reality that these particles were in fact all made up from a small set of particles, which nowadays are considered to be elementary. This set of elementary particles is described in the standard model, one of the most successful models in modern physics.

One of these fundamental building blocks is the top quark, the heaviest particle in the standard model. Its high mass means that a lot of energy is necessary to produce it, but also that it is a very interesting probe for current questions in particle physics. For example due to its high mass, the top quark couples strongly to the Higgs, and due to its short life time, it allows precise measurements of the quark-flavour-mixing CKM matrix.

To add to the understanding of the top quark, this thesis aims to measure the cross section of one top quark production mode at the Large Hadron Collider using the ATLAS detector: the single top t -channel. This work will first introduce the standard model of particle physics, including a brief introduction of top physics, the Large Hadron Collider, and the ATLAS detector. Then the concept of fake leptons is introduced and the matrix method to estimate them is explained. Lastly the steps of the analysis are shown: the event selection made to define signal and control regions, the estimation of fakes using the described method, the signal background separation using an artificial neural network, and finally the maximum likelihood fit from which the cross section can be extracted.

The Standard Model of Particle Physics

The Standard Model (SM) of particle physics describes the most fundamental building blocks of matter and three of the four fundamental forces; electromagnetism, as well as the strong and weak nuclear forces. Gravity is so far not explained by the SM. In the following, when referring to forces, gravity will hence be excluded. This chapter will introduce all the particles and forces described by the SM, give a short explanation to some of the mathematical formalisms, and an overview over the physics of top quarks.

The particles in the SM are elementary particles, i.e. particles which cannot be subdivided into smaller constituent particles. These elementary particles have no size, they are pointlike. One can divide these particles by their spin into half-integer spin particles called Fermions and integer spin particles known as Bosons. An overview of the particles the SM describes can be found in figure 2.1. In addition to the particles displayed there every particle has an anti-particle with the same mass and spin but opposite sign for all “charge-like” qualities, including properties like parity, as well as lepton and baryon number.

Often the term particle is used to include both particles and anti-particles, and the same is true for the specific particle names. For example “up quark” or the symbol u often refers to both up and anti-up quarks. Unless otherwise specified, this thesis will use that shortened terminology as well.

2.1 Force Carrier Bosons

The forces are each communicated by spin-1 carrier particles, called gauge bosons. They obey the Bose-Einstein statistics [2]. The gauge bosons are the massless photon γ and gluon g for the electromagnetic and strong force respectively, and the heavy W^\pm and Z^0 for the weak force. The range of the latter is severely limited by the mass of the two bosons. Each of the forces couples to a charge, the electromagnetic and weak force couple to the electromagnetic and weak charge respectively. The former is often shortened to charge, even if that term is technically less precise. During this introduction to the SM, the precise terminology electromagnetic charge is used. The following chapters will then use the less accurate shorthand. Notably, while the photon is electromagnetically neutral, W -bosons carry a weak charge in addition to their electromagnetic charge, making them couple to Z -bosons as well as photons.

The charge that gluons couple to is called colour charge, as it does not simply have a sign and a value, but actually three different charges, called blue, red and green. Due to a process called colour confinement, no free objects with a total non-zero colour charge can exist. Any free object that consists of elementary particles with colour charges has to consist of either the same amount of blue, red,

Standard Model of Elementary Particles

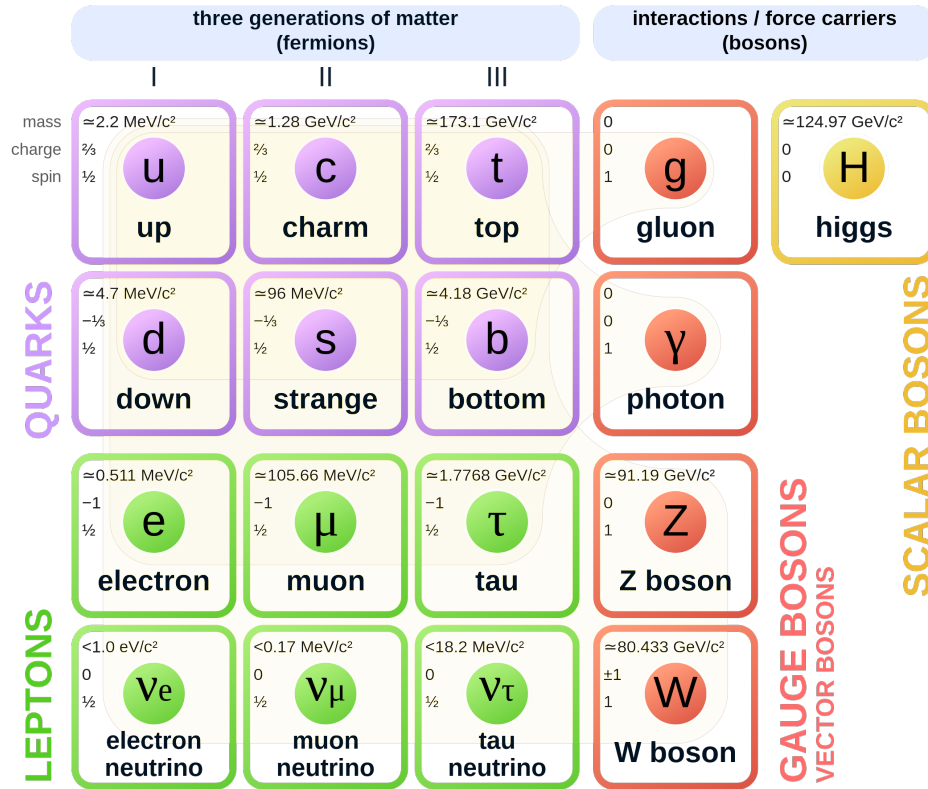


Figure 2.1: Overview of the particles in the Standard Model with their charges, spins, and masses listed [1].

and green particles, anti-blue, anti-red, and anti-green anti-particles or the same amount of particles and anti-particles with corresponding charge-pairs blue and anti-blue, red and anti-red, and green and anti-green. These combinations are called colour-neutral, sometimes also white. Gluons always carry combination of a colour and an anti-colour, e.g. blue and anti-green, making them couple to themselves, strongly limiting the range of the strong force, despite the fact, that the carrier particle is massless. All other bosons are colour-neutral.

As the name suggests, the strong charge has the strongest coupling, being roughly three orders of magnitude stronger than electromagnetism and eight orders of magnitude stronger than the weak force. Gravity has a coupling strength of 37 orders of magnitude below that of the strong charge and therefore is so weak that it has virtually no effect on the predictions of the SM.

Gluons, photons, and Z-bosons are their own anti-particles, while W^+ and W^- are usually just considered the same particle with different charges, similarly to how a red up-quark and a green up-quark are not considered different particles. Hence the term anti-particle typically only refers to anti-fermions, while bosons are considered to not have anti-particles.

2.2 Fermions

Matter itself is constructed from spin- $\frac{1}{2}$ particles, the elementary fermions. They obey the Fermi-Dirac statistics [2]. Each of the following four types of fermions exists in three different versions, which differ in their mass. These groups of four are called families or generations. The fermions can be sorted by the forces they interact with, i.e. by what charges they carry.

Quarks interact with all three forces. All quarks can have any of the three colour charges. One can split quarks into two categories based on their electromagnetic charge. Up-type quarks have an electromagnetic charge of $+\frac{2}{3}$ and down-type quarks of $-\frac{1}{3}$. As described above the fact that quarks carry colour charge means that they cannot exist as free particles but only in bound states. These states of bound quarks are called hadrons. A hadron consisting of three quarks is called a baryon (and one consisting of three anti-quarks an anti-baryon), a hadron consisting of a quark and an anti-quark a meson. States with more quarks than that, like pentaquark states consisting of four quarks and an anti-quark are called exotic states. Some evidence of these has been found, but little is known about their exact substructure [3]. The gluons that bound the hadrons together produce virtual quark-antiquark pairs during the exchanges, so that a hadron consists of a lot more than the two or three quarks listed above at any given time. These additional quarks are known as sea quarks, while the real quarks making up the hadrons are called valence quarks.

The up-type quarks are – in order of increasing mass – the up quark u , the charm quark c , and the particle with the highest mass in the SM, the top quark t . The down-type quarks are the down quark d , the strange quark s , and the bottom quark, sometimes also called beauty quark, b . These different quark types are called quark-flavours.

When quarks couple to a electromagnetically neutral gauge boson (g , γ , Z), they cannot change their flavour. These bosons can be produced from the annihilation of a quark anti-quark pair, if those quarks are the same flavour. And similarly neutral bosons can decay into quark anti-quark pairs of the same flavour. Quarks can also radiate or absorb those particles without changing their flavour. None of these processes can happen for quarks of different flavours, there are no “Flavour Changing Neutral Currents” (FCNC). For example a g cannot decay into a $u\bar{t}$ pair, even though that process would obey all conservation laws, given that the gluon has enough energy.

Flavour Changing Charged Currents on the other hand are possible in the SM. When a W^\pm couples to two quarks, the flavour even needs to change, since electromagnetic charge needs to be conserved. Hence the W can only ever couple to one up-type and one down-type quark. The couplings to quarks of different families are weaker than the couplings to quarks of the same family. The strengths of those couplings are given by the Cabibbo-Kobayashi-Maskawa (CKM) matrix, shown in equation 2.1. For ease of readability, the uncertainties on the magnitudes of the CKM elements are left out, they are all on the order of 10^{-3} or smaller [4].

$$V_{\text{CKM}} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} = \begin{pmatrix} 0.97401 & 0.22650 & 0.00361 \\ 0.22636 & 0.97320 & 0.04053 \\ 0.00854 & 0.03978 & 0.999172 \end{pmatrix} \quad (2.1)$$

The other group of elementary fermions are known as leptons, which do not carry a colour charge. There are the charged leptons with an electromagnetic charge of -1 and the electromagnetically neutral neutrinos, which only interact via the weak force. The charged leptons are the electron e , the muon μ , and the tauon or tau τ . The neutrinos are named after the charged leptons in their family, electron

neutrino ν_e , muon neutrino ν_μ , and tau neutrino ν_τ . In the SM neutrinos are massless. Like the quarks, the different lepton types are referred to as lepton-flavours.

Anti-quarks and anti-neutrinos are usually denoted by a bar over the symbol of the corresponding particle, \bar{t} for example denotes an anti-top quark. Since charged leptons have integer charge, they are often written with their electromagnetic charge as a superscript to show the difference between particles and anti-particles. For example μ^- or μ^+ denote a muon and anti-muon respectively.

2.3 The Higgs Boson

The last particle in the SM is the spin-0 Higgs boson H , which results from the Higgs-field. That field is the mechanism through which the particles get their masses. The boson is an excitation of that field and marked the experimental confirmation of that theory. Like the gauge bosons, the scalar Higgs boson obeys the Bose-Einstein statistics. The coupling strength of the Higgs boson to the other particles of the SM depends on the particles' mass. It increases linearly with the mass of fermions and quadratically with the mass of bosons. Therefore it couples very weakly to light particles like up- or down-quarks and electrons and strongly to heavy particles like the W - and Z -bosons and the top quark.

2.4 Mathematical Formulation

All three forces of the SM can be modelled as a Quantum Field Theory (QFT) in which particles are described as specific excitations of a field.

In the SM the electromagnetic and weak interaction are unified as a $SU(2) \times U(1)$ gauge symmetry and the strong force is described as a $SU(3)$ theory, leading to the total local symmetry of the SM being $SU(3) \times SU(2) \times U(1)$. The QFT describing the electromagnetic force is known as Quantum Electro Dynamics (QED), the QFT for the strong force Quantum Chromo Dynamics (QCD).

The kinematics of the fields can be described with a Lagrange density, consisting of terms describing the propagation of the gauge bosons \mathcal{L}_F , the interaction between the gauge bosons and the fermions $\mathcal{L}_{\text{Dir.}}$, the Yukawa coupling of the Higgs boson to the fermions $\mathcal{L}_{\text{Yukawa}}$, and the Higgs kinematics, describing the coupling of the Higgs to the W - and Z -boson and possibly itself $\mathcal{L}_{\text{Higgs}}$:

$$\begin{aligned}\mathcal{L}_{\text{SM}} &= \mathcal{L}_F + \mathcal{L}_{\text{Dir.}} + \mathcal{L}_{\text{Yukawa}} + \mathcal{L}_{\text{Higgs}} \\ &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ &\quad + i\bar{\psi}\not{D}\psi + h.c. \\ &\quad + \psi_i y_{ij} \psi_j \phi + h.c. \\ &\quad + |D_\mu \phi|^2 - V(\phi),\end{aligned}$$

where F is, ψ are the fermion fields, ϕ the boson fields, y is the Yukawa coupling strength, D is the gauge covariant derivative, and V is the Higgs potential. $h.c.$ is short for the Hermitian conjugate of the previous expression.

2.5 Open Questions

While the SM is successful in explaining many observations and most measurements agree with theoretical expectations, there are still questions that remain unanswered by the model. As mentioned in the opening of this chapter, gravity is not explained by the SM. Similarly the SM does not include explanations for dark matter and dark energy [5], or neutrino masses, which have been observed to be non-zero are not generated by the SM [6]. Furthermore there are parameters of the model, such as the mass of the Higgs Boson, that are not predicted by theory and can therefore only be measured in experiments. There are various different models beyond the SM (BSM) that could explain some of these question.

2.6 Top-Physics

The top quark is the heaviest particle of the SM. It has a mass of (173.34 ± 0.76) GeV [7] and a decay width expected to be on the order of 1 GeV [4]. This corresponds to an average life on the order of 10^{-25} s, which is an order of magnitude shorter than typical hadronisation times. Therefore the t decays as a free particle, despite quarks usually being confined into hadrons. This means that the CKM-Matrix elements V_{td} , V_{ts} , and V_{tb} can be very precisely measured by observing top-decays. Additionally since the coupling strength of the Higgs boson is dependant on a particles' mass, the t being the heaviest particle in the SM makes it an interesting probe for Higgs physics.

2.6.1 Top Production Modes

Due to the high coupling strength of the strong force, the most common production mode of the t at the Large Hadron Collider (see chapter 3) is the production of $t\bar{t}$ -pairs. Examples for leading order $t\bar{t}$ -processes can be seen in figure 2.2.

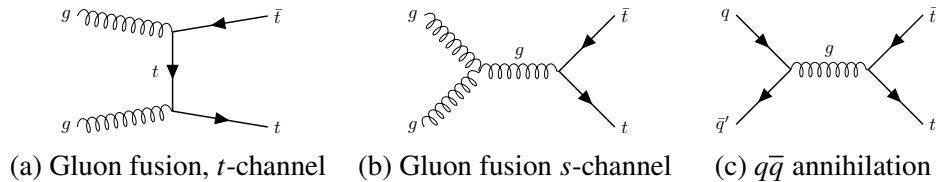


Figure 2.2: Feynman diagrams of some leading order $t\bar{t}$ production processes.

A single t can be produced through electroweak processes. The leading order processes for that are the t -channel, s -channel, and tW , the Feynman diagrams of which are shown in figure 2.3. Due to the particles in the final state, the t -channel is also referred to as tq and the s -channel as tb .

The quarks marked as q can be of any flavour (with the right charge), but are typically the valence quarks of the protons, so u or d . The b quarks in the initial state come from $g \rightarrow b\bar{b}$ splitting.

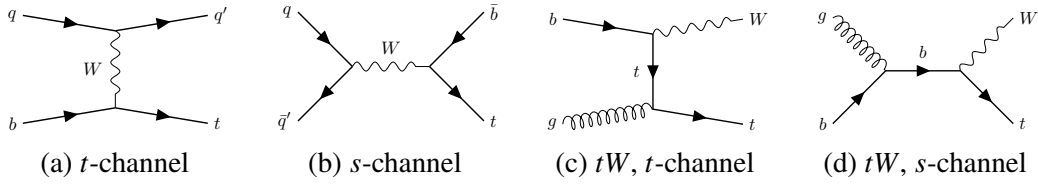
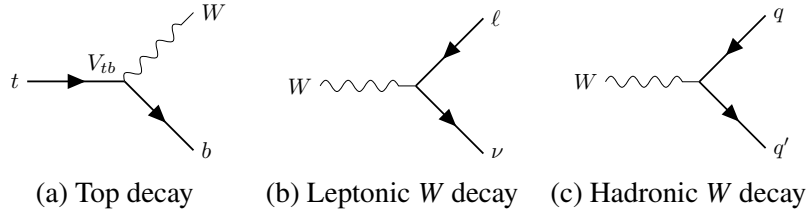


Figure 2.3: Feynman diagrams of the leading order single top production processes.

2.6.2 Top Decay Modes

The top quark can only decay through the weak interaction into a W -boson and a down-type quark. The decay widths of that process scales with the square of corresponding CKM-Matrix element. Since V_{td} and V_{ts} are orders of magnitude below V_{tb} (see 2.1), the top decays almost exclusively into bottom-quarks. The decay into a b -quark and a W -boson is shown in figure 2.4. The W in turn can decay leptonically into a charged lepton and a corresponding neutrino or hadronically into two quarks. The leptonic decay has a branching ratio of approximately 33%, the hadronic decay 67% [4].


 Figure 2.4: Decay of a top into a bottom and W and the two decay modes of the W . Similar decays of the top into down and strange quarks are possible but suppressed by the CKM matrix elements.

Large Hadron Collider

The **Large Hadron Collider (LHC)** is a proton-proton synchrotron at the European Organization for Nuclear Research (CERN). This chapter gives a short overview over some of the technical details of the accelerator and introduces some core concepts of collider physics.

The LHC is installed in a 26.7 km tunnel under the French-Swiss border that was constructed previously for the electron-positron collider LEP. The LHC is designed to accelerate protons up to 7 TeV allowing for a maximum centre of mass energy of $\sqrt{s} = 14$ TeV. Currently the highest energy reached is $\sqrt{s} = 13.6$ TeV. The LHC hosts four main experiments, two general purpose experiments, ATLAS and CMS, ALICE which is a dedicated heavy-ion experiment focussing on Pb-Pb instead of proton-proton collisions, and LHCb with a focus on b -physics.

The storage ring consists of 1232 superconducting main dipoles operating at a temperature of 1.9 K, able to provide a maximal magnetic field of 8.3 T. The particles are accelerated using 8 radio frequency cavities in each direction [8].

As a synchrotron the LHC requires the protons to be pre-accelerated. At CERN several accelerators are operated, most of which serve both as pre-accelerators for higher energy accelerators and also provide beams for lower energy experiments. An overview of the CERN accelerator complex can be seen in figure 3.1. The LHC is “supplied with protons from the injector chain Linac2 – Proton Synchrotron Booster (PSB) – Proton Synchrotron (PS) – Super Proton Synchrotron (SPS)” [8]. The protons are produced by stripping hydrogen atoms of their electron. In 2020 the LINAC2 got replaced by a new linear accelerator, the LINAC4. LINAC4 accelerates negative hydrogen ions, which have both electrons stripped off during injection from LINAC4 to the PS Booster [9].

This thesis analyses data taken during run 2 from 2015 to 2018. A run is a time period in which the LHC is operated without shutdown for maintenance or upgrades. During run 2 the LHC provided pp -collisions at a centre of mass energy of $\sqrt{s} = 13$ TeV. In that time ATLAS recorded an integrated luminosity of 139 fb^{-1} .

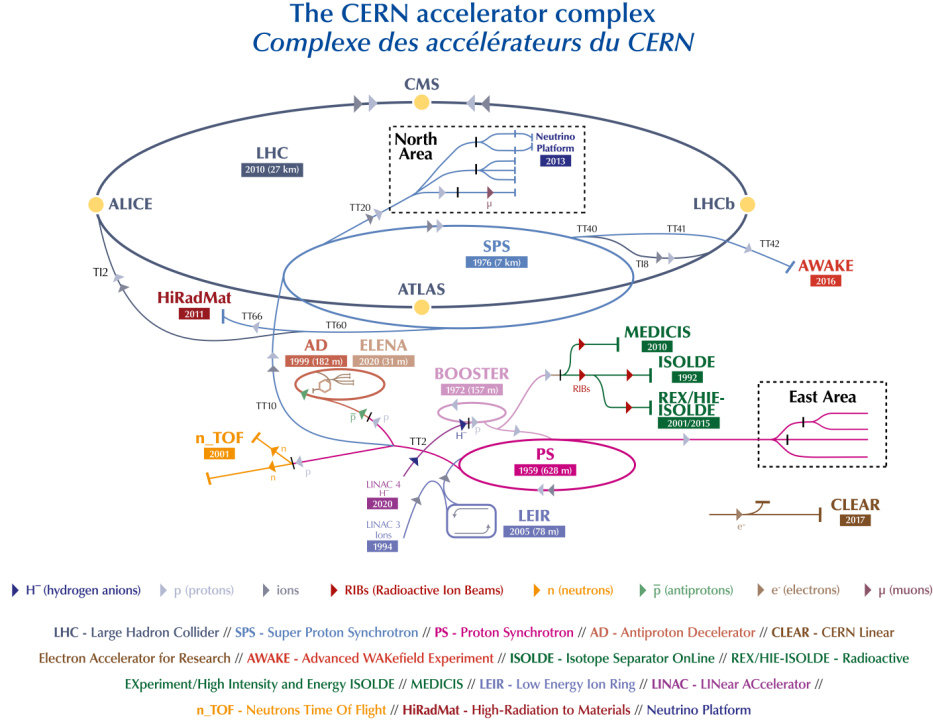


Figure 3.1: Overview of the full accelerator complex at CERN [10].

3.1 Collider Physics

3.1.1 Natural Units

In the context of high energy particle physics it is often not useful to use SI units which were designed to have easy to handle numbers on macroscopic scales. Especially for energy the SI base unit Joule (J) is not useful. Instead the energy of particles is typically given in electronvolt (eV), which is the energy a particle with an electromagnetic charge of 1 has after being accelerated using an electrical field with a strength of 1 V. In SI units this translates to

$$1 \text{ eV} \approx 1.602 \times 10^{-19} \text{ J}. \quad (3.1)$$

The relativistic energy-momentum-relation

$$E^2 = (\vec{p}c)^2 + (mc^2)^2, \quad (3.2)$$

gives an equation to relate the energy E , three-momentum \vec{p} , and rest-mass m of a particle. One can see here, that the dimensionality of E , $|\vec{p}|$, and m only differs in factors of the speed of light in vacuum c . To more easily relate these quantities, three-momentum and mass are typically also not given in SI units, but in eV/c and eV/c^2 respectively. To ease notation high energy physics uses the natural units, in which

c is set to 1.¹ This means, that one can write the energy-momentum-relation as

$$E^2 = |\vec{p}|^2 + m^2. \quad (3.3)$$

Now E , $|\vec{p}|$, and m all have the same dimension and can all be given in eV.

3.1.2 Collider Properties

An important quality for any collider is the centre of mass energy \sqrt{s} of a collision. This is the invariant mass of the system of the two colliding particles, and therefore the energy available to generate new particles in the collision. This is especially relevant when one wants to generate heavy particles and to ensure the produced particles have high energy, making them more reliably detectable. So with equation 3.3 it is given as

$$\sqrt{s} = \sqrt{(E_1 + E_2)^2 - (\vec{p}_1 + \vec{p}_2)^2}. \quad (3.4)$$

Where E_i and \vec{p}_i are the energy and three-momentum of the particles in the colliding beams. For a symmetrical collider like the LHC where two identical particles are accelerated to the same energy in opposite direction this simplifies to

$$\begin{aligned} E_1 &= E_2 =: E \\ \vec{p}_1 &= -\vec{p}_2 \\ \Rightarrow \sqrt{s} &= 2E. \end{aligned} \quad (3.5)$$

Next to \sqrt{s} another important characteristic for a collider is the luminosity L , which is a measure for the number of events produced per time and area. Higher luminosity means more events, reducing statistical uncertainties. This is especially important for precision measurements and for searches for rare processes. The luminosity of a collider can be approximated as

$$L = \frac{f N_1 N_2 n_b}{4\pi \sigma_x \sigma_y}, \quad (3.6)$$

with the revolution frequency f , the number of particles per bunch N_i , the number of bunches n_b , and the beam widths assuming a Gaussian distributed particle density in the beam σ_i . For the LHC these parameters are listed in table 3.1. The two beams of the LHC are symmetrical and identical so that $N_1 = N_2 =: N$ and $\sigma_x = \sigma_y =: \sigma$.

¹ Natural units also set the reduced Planck constant and the vacuum permittivity to one: $c = \hbar = \varepsilon_0 = 1$, however that is not strictly relevant to this discussion.

Table 3.1: Technical parameters defining the luminosity of the LHC [11].

| Parameter | Value |
|-----------|---|
| f | $11.145 \times 10^3 \text{ Hz}$ |
| N | 1.15×10^{11} |
| n_b | 2556 |
| σ | $2.5 \mu\text{m}$ |
| L | $1.0 \times 10^{34} / \text{cm}^2 \text{s}$ |

Every process that can happen at the collider has a cross section σ . The cross section of a process is a measure of the probability that the given process will happen during an interaction. Its dimension is an area. Cross sections are typically not given in SI units, but in barn (symbol: b):

$$1 \text{ b} = 1 \times 10^{-28} \text{ m}^2 . \quad (3.7)$$

Since the luminosity is a measure of events per time and area, one can relate it to the number of expected events N via the cross section of the considered scattering process by integrating over the time t the collider was operated:

$$N = \sigma \int L \, dt =: \sigma \mathcal{L}_{\text{int}} . \quad (3.8)$$

\mathcal{L}_{int} is called the integrated luminosity. It is typically given in b^{-1} .

ATLAS Detector

The data in this analysis has been taken with the ATLAS detector (A Toroidal LHC ApparatuS), one of the four main experiments at the LHC [12]. This chapter will give a short overview over the ATLAS detector by introducing the theory behind the detector types used, as well as explaining the individual subdetectors and how they are used to identify the particles produced in the pp -collisions. This chapter will only serve as a brief introduction to the basic workings of the detector. More details about the individual parts of the detector and the technical performance can be found in the technical design reports [13] and [14].

The ATLAS detector consists of a barrel of layered detectors which surround the beampipe cylindrically, as well as end caps perpendicular to the beampipe to detect particles at a narrow angle relative to the beam. With this setup the detector covers nearly the entire solid angle. A computer generated image of the detector can be seen in figure 4.1.

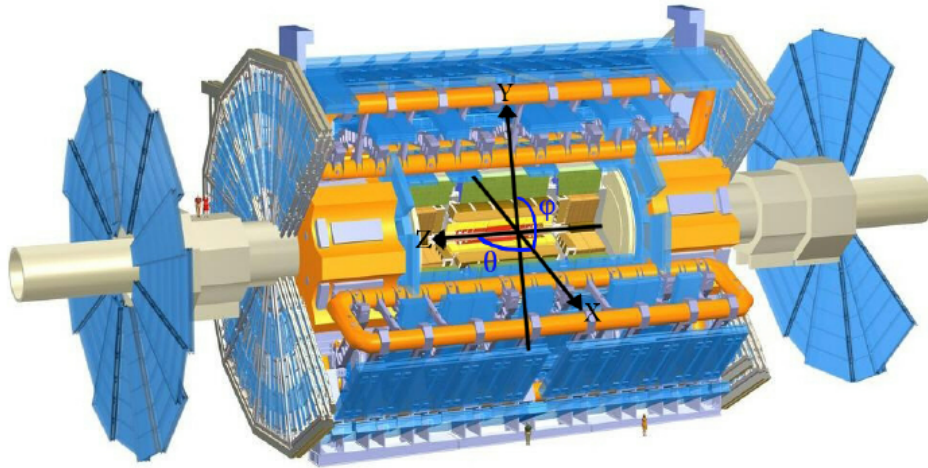


Figure 4.1: Computer generated image of the ATLAS detector with the coordinate system overlaid [15], altered by Sven-Patrik Hallsjö in [16].

4.1 Detector Physics

Since elementary particles are not visible by eye, one needs detectors to measure the particles' presence. In modern particle physics there are several types of detectors available which use different interactions to measure various kinds of particles. This section will discuss the detection mechanisms used in ATLAS, being semiconductor detectors, gaseous detectors, scintillators, and calorimeters.

4.1.1 Semiconductor Detectors

Semiconductors can be used to detect ionising particles. An ionising particle passing through a semiconductor will create electron hole pairs in the material, making a measurable current flow in the semiconductor. By arranging several layers of semiconductors and dividing these layers into pixels, which can be read out individually, it is possible to detect several particles at once and reconstruct their paths through the detector. Typical semiconductors used for these detectors are silicon and germanium.

4.1.2 Gaseous Detectors

Similarly to semiconductors, certain gasses can be ionised by high energy charged particles passing through them. By placing a wire in the gaseous volume and applying a voltage between that wire and the container of the gas, the created ions and free electrons can be measured. The charges will be accelerated to the wire and wall respectively, where they will cause a measurable current to flow.

If the voltage is high enough, the free electrons and ions can be accelerated enough to become ionising themselves, creating more charges and thereby creating a shower of charged particles. This amplifies the signal and can be necessary for the current to be reliably detectable. When an amplification is used, depending on its strength, the strength of the signal current can be proportional to the initial charge created by the primary particle, allowing for a reconstruction of that particles' energy. If the created showers are big enough, the measured signal current will plateau and not allow for that reconstruction. This plateau is called the Geiger-Müller plateau.

Figure 4.2 shows the relationship between the applied voltage and the signal strength. The initial rise in the "Recombination" region is due to the voltage being so low, that some of the created electron-ion-pairs recombine into neutral atoms again before drifting apart. The initial plateau "Saturation of Ion Chamber" happens, when none of the electron ion pairs recombine, but the acceleration also is not strong enough for them to cause further ionisation in the gas. Then the mentioned proportional region and the Geiger-Müller plateau follow, linked by a region of non-linearity. The last rise in signal strength happens, when the voltage is high enough, that a constant current will flow, even without ionising particles passing through.

Generally one separates gaseous detectors into three categories, depending on the region they operate in. Ionisation chambers operate at a voltage high enough to separate the electron-ion-pairs, but without gas amplification, proportional counters operate in the region of proportionality, and Geiger-Müller-counters operate on the high voltage plateau.

Ionisation chambers require very sensitive instruments to measure the low currents and are therefore also strongly affected by operational circumstances like moisture. Geiger-Müller counters, on the other hand, are easier to construct and maintain. But the charged avalanches in the gas mean, that while an avalanche is happening and shortly after, while the ions are recombining with electrons from the mantle of the chamber, the detector cannot measure another particle. This "dead time" does not affect ionisation

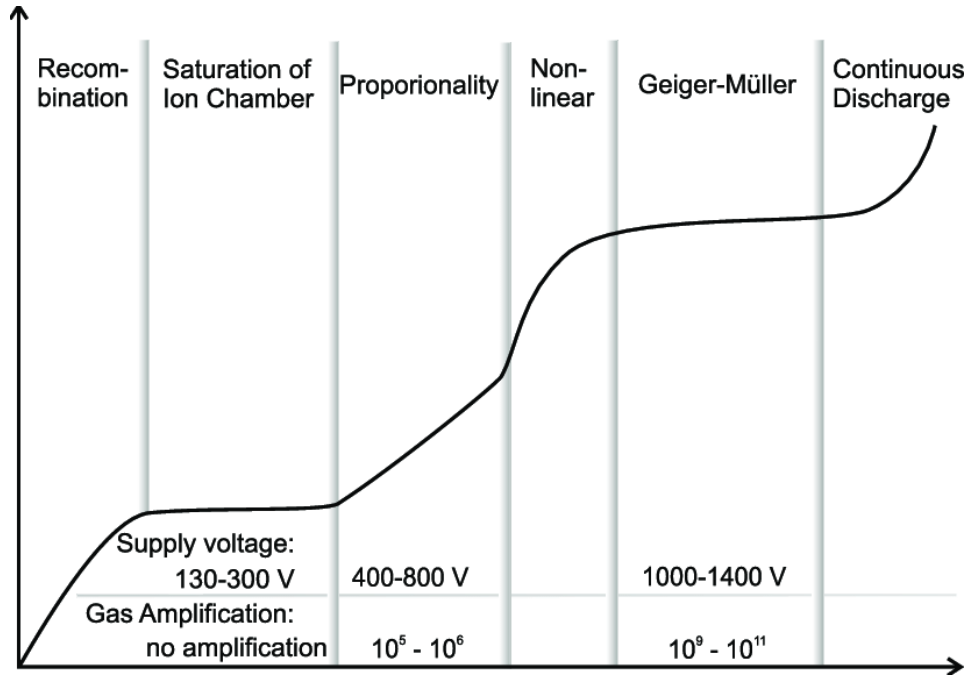


Figure 4.2: Signal strength of a gaseous detector depending on the applied voltage. The values of the supply voltage depend on the exact construction of the detector and just serve as a rough estimate [17].

chambers, which is why they are typically used, when high radiation dose rates need to be measured. The main advantage of proportional counters is naturally their ability to reconstruct the energy of detected particles. However their operation is often quite complex, as they require the properties of the gas and the voltage to be well calibrated.

A special type of gaseous detectors, important for the ATLAS setup, is the straw tube detector. A straw tube detector consists of several thin ionisation chambers called straws. With the straws so thin and densely layered, one can get several hits from a particle transversing the detector volume, similar to the pixels of a semiconductor detector, described in the previous section 4.1.1. The electrons and ions transversing through the gas take a non-zero amount of time to reach the wire or mantle of the straw after being produced. This time information can be used to reach a spatial resolution better than the size of the individual straws.

4.1.3 Scintillators and Photomultipliers

Scintillators are materials which can absorb the energy of incoming particles and then emit it as several photons of lower energy, giving off visible light. This happens when the incoming particle excites the material - for example by lifting electrons from the valence to the conducting band - and the de-excitation can happen over several lower energetic excited states. Then every individual step from one excited state to the next lower one emits a photon of lower energy than the initial particle.

These photons can be detected and turned into a measurable electrical signal by photomultipliers (PMT, short for **Photomultiplier Tube**). PMTs work by having one layer of a material converting the incoming photons into electrons via the photoelectric effect, the photocathode. Since one photon will

only produce a single electron, the electrons get accelerated towards dynodes using high voltage. Upon impact with the dynode they will produce more secondary electrons, which will get accelerated to another dynode, continuing until the signal has been amplified enough to be detected when the electrons reach the final anode. A schematic overview over a setup using a scintillator and a PMT to detect high energy photons can be seen in figure 4.3.

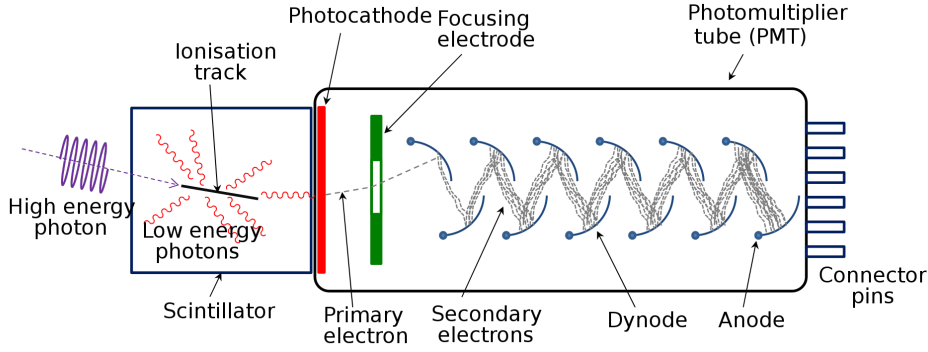


Figure 4.3: Schematic overview of a setup using a scintillator and a PMT to detect high energy photons. The focusing electrode is not strictly necessary for the detection mechanism [18].

The low energy photons in the scintillator are radiated off homogeneously into all directions in space. The total number of low energy photons produced is proportional to the energy of the initial high energy particle. In addition the size of the final electron avalanche in the PMT is also proportional to the energy of the photon hitting the photocathode. Therefore a well calibrated setup can reconstruct the total energy of the initial particle, even though not all low energy photons are detected.

4.1.4 Calorimeters

Calorimeters are detectors which use interactions of particles in matter to fully stop these particles and measure the energy deposited in the detector volume, for example using scintillators. Particles tend to deposit more energy in denser materials with higher proton counts in the atoms. Scintillator materials and other detector components however tend to be made from relatively low density materials. So to fully stop high energetic particles a scintillator would require a prohibitively large volume. Therefore calorimeters typically consist of alternating layers of high density stopping material and volume dedicated to detection. The layers of the stopping material will cause the high energy particles to fragment into lower energy particles, producing a shower of continuously more particles of lower energy, until the particles are fully stopped inside the detector. The energy deposited in the stopping material naturally cannot be detected, so a calorimeter built from layers like that needs to be calibrated using objects with a well known energy.

4.2 ATLAS Coordinate System

The coordinate system used for the ATLAS detector is constructed such that Cartesian coordinates x , y , and z form a right handed system with the z -axis along the beamline, the x -axis pointing towards the centre of the LHC, and the y -axis pointing up. Figure 4.1 shows the coordinate system overlayed

on a computer generated image of the detector. Therefore the x - y -plane, called the transverse plane, is orthogonal to the beamline. The combination of the momenta in x and y direction is called the transverse momentum of a particle:

$$\sqrt{p_x^2 + p_y^2} =: p_T . \quad (4.1)$$

And similarly the momentum along the beam axis is called the longitudinal momentum:

$$p_z =: p_L . \quad (4.2)$$

Using the relativistic relation between energy E , three-momentum $|\vec{p}| =: p$, and invariant mass m , $E^2 = p^2 + m^2$ one can define also define a transverse Energy:

$$E_T^2 := p_T^2 + m^2 . \quad (4.3)$$

Another useful quantity is the so called transverse mass:

$$m_T := E^2 - p_L^2 = m^2 + p_T^2 . \quad (4.4)$$

For vanishing invariant mass $m = 0$ all three of these transverse quantities are the same $p_T = E_T$ ($m = 0$) $= m_T$ ($m = 0$). For most objects except jets, $p_T \gg m$ is given, making $m \approx 0$ a good approximation.

The angle around the z -axis, i.e. in the transverse plane is the azimuthal angle, labelled $\varphi \in [0, 360^\circ[$. This means that the azimuthal angle of an object can be described via its three-momentum in x - and y -direction as

$$\varphi = \begin{cases} \arctan \frac{p_y}{p_x} & p_x \neq 0 \\ \frac{p_y}{|p_y|} \frac{\pi}{2} & p_x = 0 \end{cases} . \quad (4.5)$$

Processes are usually expected to be invariant under rotation of φ . Hence this variable is mostly used to measure angular differences between two objects (see eq. 4.8). The angle between a particle's three-momentum and the z -axis is called $\theta \in [0, 180^\circ]$ with 0° pointing in the positive and 180° the negative z -direction. However since it is not Lorentz invariant under boosts along the z -axis that angle is typically not used to describe the kinematics of a particle, but rather the so called pseudorapidity:

$$\eta := -\ln \left(\tan \left(\frac{\theta}{2} \right) \right) = \frac{1}{2} \ln \left(\frac{|\vec{p}| + p_L}{|\vec{p}| - p_L} \right) = \operatorname{arctanh} \left(\frac{p_L}{|\vec{p}|} \right) . \quad (4.6)$$

This leads to a particle with no longitudinal momentum, so $\theta = 90^\circ$ having a pseudorapidity of $\eta = 0$, while a particle with no transverse momentum, so $\theta = 0^\circ$ or $\theta = 180^\circ$ having an infinite pseudorapidity. The ATLAS detector can measure particles up to $|\eta| < 2.5$ or $|\eta| < 4.5$, depending on the particle type.

When a particles' mass is low compared to its three-momentum, the absolute value of the three momentum is the same as its energy $m \ll |\vec{p}| \Rightarrow |\vec{p}| \approx E$ (see equation 3.3). When we apply that limit to equation 4.6 the pseudorapidity converges to a quantity called rapidity y :

$$\eta \approx \frac{1}{2} \ln \left(\frac{E + p_L}{E - p_L} \right) =: y . \quad (4.7)$$

While the rapidity itself still is not Lorentz invariant under boosts along the z -axis, the *difference* between rapidities is. Accordingly differences in pseudorapidity are invariant for massless particles.

Pseudorapidity often gets used over rapidity, because it is a purely angular property and $m \ll |\vec{p}|$ is given for most objects in a collider experiment like ATLAS. For heavier objects like jets however, rapidity is still used (see for example section 4.5.1).

The angular distance between two objects is then given as

$$\Delta R := \sqrt{\Delta\eta^2 + \Delta\varphi^2}. \quad (4.8)$$

With Δy^2 instead of $\Delta\eta^2$ for massive objects like jets.

4.3 Detector Parts

Detectors at collider experiments are often built from several subdetector parts, as different detector types are needed to measure the position, charge, and energy of all particles that can be produced in the collisions. To cover as much of the solid angle around the interaction point as possible, these subdetectors are often built in layers around each other, so that every individual subdetector covers the full range of the detector. Additionally information from several subdetectors is often combined in the reconstruction of particles, which requires the detectors to be layered, so that a particle can pass through several or all of them, before being stopped or leaving the detector volume.

The ATLAS detector consists of three such layers, the inner detector, the calorimeters, and the muon spectrometer, as well as a magnet system, consisting of a central solenoid and three air-core toroids. The inner detector serves to reconstruct the exact position of the pp -interaction (“interaction vertex”) and to determine the momentum and charge of charged particles, as well as the path they take through the detector. The calorimeters serve to measure the energy of nearly all particles, the only exceptions being muons and neutrinos. The muon spectrometer lastly serves for the detection of muons and the precise measurement of their momentum and charge.

The detector has a cylindrical shape around the beam line and endcaps on either side to cover the forward region. A schematic overview over the detector layers and the particles they are used to detect can be found in figure 4.4.

4.3.1 Inner Detector

The inner detector is the subdetector closest to the beam pipe. It is made up of the pixel detector, the silicon microstrip tracker, and the transition radiation tracker. The pixel detector and the silicon microstrip tracker are semiconductor detectors, the transition radiation tracker is a gaseous straw tube detector. All three of those are contained in the central solenoid magnet, providing a homogenous magnetic field parallel to the beam axis with a flux density of 2 T. The inner detector measures the tracks of charged particles. The magnetic field of the central solenoid causes these tracks to be curved, allowing the sign of the charge to be reconstructed from the direction of the curvature and the momentum from the radius. Both the charge tagging and the momentum reconstruction rely on the curvature of the track and get therefore less precise with higher momenta. The pixel detector provides high resolution detection as close to the interaction point as possible to reconstruct the positions of the event vertices. For that purpose it is located at a distance of only 3.3 cm from the interaction point.

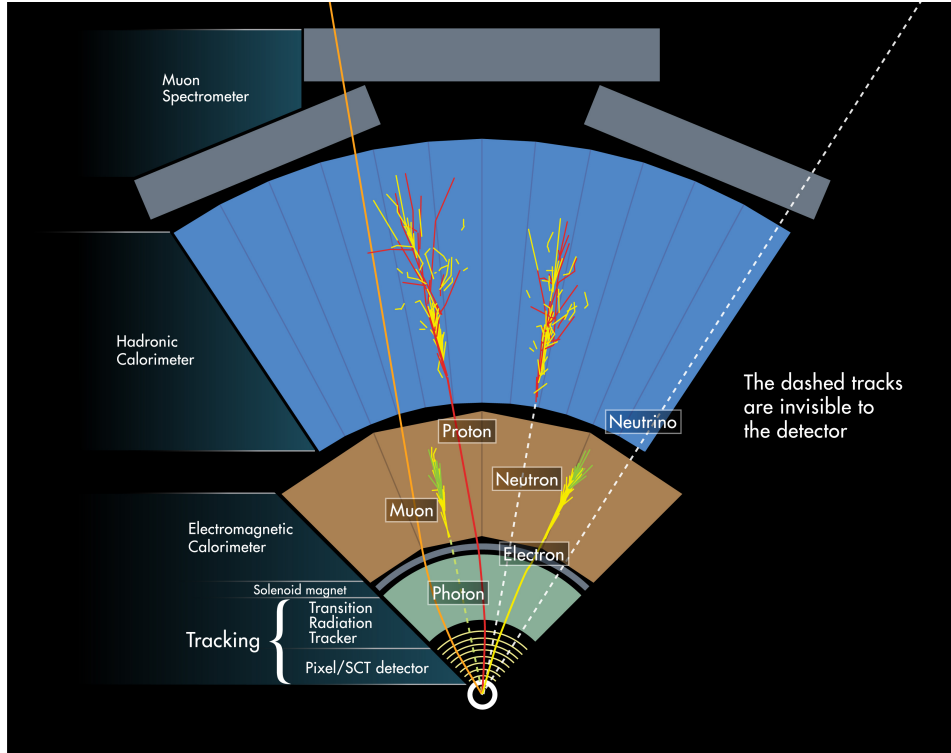


Figure 4.4: Schematic cross section through the layers of the ATLAS detector and the characteristic detector responses caused by the different particles. Proton and Neutron here are stand-ins for any charged and neutral hadrons respectively [19].

4.3.2 Calorimeters

The ATLAS calorimeter system surrounds the inner detector and consists of two components, an electromagnetic and a hadron calorimeter. The former measures the energy of electrons and photons, the latter of hadronic particles. In addition to the energy measurement, some information about the particle track can be reconstructed from the location of the energy deposition. However the spatial resolution is poor compared to that of the inner detector.

The calorimeters are the only part of ATLAS capable of detecting neutral particles, photons and neutral hadrons. If an energy deposition in the electromagnetic calorimeter can be matched to a charged track in the inner detector it can be concluded it belongs to an electron, a deposition that cannot be matched to those tracks stems from a photon. Similarly charged hadrons show a track in the inner detector and deposit energy in the hadronic calorimeter, while neutral hadrons can only be detected in the calorimeter without tracks.

The electromagnetic calorimeter uses lead as a stopping material and liquid argon to measure the deposited energy. It covers the regions up to $|\eta| < 3.2$. The hadron calorimeter uses steel as a stopping material and scintillator tiles to measure the energy in the central barrel up to $|\eta| < 1.7$, and copper and liquid argon in the endcap regions $1.5 < |\eta| < 3.2$. Additionally a calorimeter combining copper-liquid argon and tungsten-liquid argon covers the forward regions $3.1 < |\eta| < 4.9$ for both electromagnetic and hadronic energy measurements [20].

4.3.3 Muon Spectrometer

Muons have a lower chance of interacting with the detector components and can hence leave the entire detector. Therefore their energy cannot be measured in a calorimeter. To account for that the outer-most layer of the detector is the muon spectrometer. It measures tracks behind the calorimeters, which accordingly have to belong to muons. To determine charge and momentum the muon spectrometer is placed within the magnetic field of three air-core toroids, similar to the inner detector in the central solenoid. The muon spectrometer combines various types of gaseous detectors [21].

4.4 Particle Identification

Electrons and photons are identified by energy depositions in the electromagnetic calorimeter. Since electrons are electromagnetically charged, while photons are neutral, they can be differentiated combining the information from the calorimeter to that of the inner detector. When an energy deposition in the calorimeter can be matched to a charged track, it is reconstructed as an electron, with charge tagged by the curvature of the track. When there is no track to match, the cluster is reconstructed as a photon.

Similarly hadronic particles leave energy depositions in the hadron calorimeter. Here clusters are matched to tracks from the inner detector to reconstruct charged hadrons and the remaining clusters are then used to reconstruct neutral hadrons.

Lastly muons are long lived enough to fully transverse the detector. They get reconstructed from hits in the muon spectrometer, which can get matches to tracks in the inner detector. Since the muon spectrometer is contained in a magnetic field like the inner detector it allows for more reliable charge tagging and to more precisely measure the momentum of the muon, compared to electrons, for which momentum measurements and charge tagging rely entirely on the inner detector.

The only SM particles that cannot be detected in ATLAS are neutrinos. However since the two beams collide head on along the beam axis, the total transverse momentum has to be 0. So by reconstructing all other particles in the collision, the presence of neutrinos can indirectly be observed by a lack of transverse momentum/energy in one direction. This does however not give any information about the number of neutrinos produced. Furthermore if more than one neutrino is produced, the total missing transverse energy will not necessarily be equal to that of all neutrinos, since they might cancel that effect partially if there is an angle greater than 90° between them in the transverse plane.

Importantly, there are always several collisions happening close to each other at the same time, or in the same location in short periods of time. Sometimes the products from these events cannot be clearly separated. The influence of these other events on a recorded collision is called pile-up.

4.5 Object Definitions

Based on the particle identification principles described in the previous section, several types of objects get reconstructed from the raw data taken by the detector.

4.5.1 Jets

Jets are sprays of collimated hadrons. They are the result of quarks or gluons in the final state which then fragment and hadronise. Jets of high enough energy can be identified by eye in an event display. But

for the purpose of analysis jets are defined using a jet algorithm, which sorts hadronic particles into jets. Important qualities for a jet algorithm are infrared and collinear (IRC) safety. This means, that the set of high energy jets an algorithm finds should be largely unaffected by adding or removing soft particles from the event (infrared safety) or by a collinear splitting of particles in the jets (collinear safety) [22].

In ATLAS jets are constructed using the IRC safe anti- k_t algorithm. This algorithm produces circular jets around hard objects called seeds in the following steps, which are repeated until all objects are sorted into jets.

1. Calculate a momentum dependent distance measure between all objects as well as between all objects and the beam line (respectively d_{ij} and d_{iB} in equation 4.9).
2. Determine the minimum of all those distances.
3. If that minimum is the distance between two objects (d_{ij}), group these objects together by adding their four-momenta and repeat from the top, treating these objects as one.
4. If instead that minimum is the distance between an object and the beamline (d_{iB}), declare that object a final state jet. Remove it from the event and repeat from the top.

The momentum based distance measure used is

$$d_{ij} = \min \left(p_{Ti}^{-2}, p_{Tj}^{-2} \right) \frac{\Delta R_{ij}^2}{R^2}, \quad \Delta R_{ij}^2 = (y_i - y_j)^2 + (\varphi_i - \varphi_j)^2, \quad (4.9)$$

$$d_{iB} = p_{Ti}^{-2}.$$

Here y , φ and p_T are respectively the rapidity, azimuthal angle, and transverse momentum of the object, d_{ij} is the distance between two objects i and j , and d_{iB} the distance between object i and the beamline. R is a parameter of the algorithm which is typically set to 0.4 in ATLAS analyses¹.

The momentum dependence of this distance measure assures, that jets are formed around hard seeds. This “favours clusterings that involve hard particles rather than clusterings that involve soft particles [...] or energy-independent clusterings” [22]. The disadvantage over alternative jet algorithms like the k_t algorithm or the Cambridge/Aachen algorithm not discussed here is, that the anti- k_t algorithm does not produce a substructure that can meaningfully be related to the QCD branching of the event. “[E]ssentially the anti- k_t recombination sequence will gradually expand through a soft subjet, rather than first constructing the soft subjet and then recombining it with the hard subjet” [22].

The objects that get sorted into jets by the anti- k_t algorithm are constructed from hits in the hadronic calorimeters using topological clusters and track information using particle flow.

A topological cluster is constructed around individual high energy depositions in the hadron calorimeter. When a calorimeter cell measures an energy deposition of four standard deviations (4σ) above the noise threshold it gets used as the seed of a cluster. First all surrounding cells (both in the same and in different layers of the calorimeter, so in all three dimensions) with energy depositions of two standard deviations (2σ) above the noise threshold are added to the cluster and then all the cells immediately surrounding the cluster, independent of their measured energy. This is often shortened to the 4-2-0 rule. At the end a splitting step is performed, where some clusters are split based on local energy maxima [23].

¹ Some analyses use large jets with $R=1.0$

The particle flow algorithm takes tracks from the inner detector and the topological clusters from the calorimeter as input and initially matches clusters to tracks based the direction and momentum of the track as well as the energy and location of the cluster. A single particle can deposit energy in more than one topological cluster. When a track is matched to one or more clusters, the expected energy deposition of the particle that left the track is subtracted from the cells in the cluster, then the next track is matched. Any remaining clusters that cannot be matched to a track are then reconstructed as neutral hadrons [24]. These reconstructed objects are then used at the input to the anti- k_t algorithm.

4.5.2 b -tagging

Hadrons containing b -quarks have higher lifetimes than hadrons only containing light flavour quarks, leading to them transversing some distance from the initial interaction vertex before decaying and forming a jet. This distance is still short enough for them to decay inside the beampipe but the tracking of the inner detector is precise enough to reconstruct the secondary vertex of the heavy flavour decay. Additionally b -jets tend to have higher masses and higher particle multiplicities in the jet than light jets. Based on these parameters, jets can be tagged as b - or light jets². This is done using a boosted decision tree (BDT). For information on BDTs see [25]. Based on that output all jets are then tagged at working points with 60, 70, 77, and 85% efficiency. This means that 60% is the tightest working point and 85% the loosest [26].

The b -jets in the selection of this analysis are tagged at the 60% working point. Every jet not tagged at this working point is considered light.

4.5.3 Charged Leptons

The light charged leptons e and μ leave tracks in the inner detector and respectively either an energy deposition in the electromagnetic calorimeter or further tracks in the muon spectrometer. The heavy τ is too short lived to be detected directly. It decays into a W and a ν_τ , the former can either decay into a light charged lepton and corresponding neutrino or two quarks. A leptonically decaying tauon is not reconstructed as such but seen in the detector as a light lepton. A hadronically decaying tauon however will be seen as a jet, meaning tau-jets need to be differentiated from jets originating from quarks or gluons.

Based on their signatures in the inner detector, electromagnetic calorimeter, and muon spectrometer, light lepton candidates are constructed. There are different likelihood based identification working points called Loose, Medium, and Tight. Prompt electrons with $E_T = 40$ GeV are identified with 93%, 88%, and 80% efficiency respectively [27], muons with $20 < p_T [\text{GeV}] < 100$ have efficiencies of 99%, 97%, and 93% [28]. A variant of the Loose operating point for electrons is the LooseAndBLayer working point. In addition to the same requirements as the Loose working point it requires a hit in the innermost pixel layer of the detector. For more details, including further alternative working points see [27, 28].

Additionally isolation requirements can be applied to the leptons to suppress non-prompt leptons (see section 5.2). One of these possible requirements is the Prompt Lepton Veto (PLV). For the PLV a BDT algorithm is trained to separate prompt and non-prompt leptons. Separate BDTs are trained for electrons and muons. The PLVLoose working point has an efficiency of 95%, PLVTight of 90% [29].

This analysis uses the Single Top Ntuples (see section 4.6) and accordingly shares its definition of loose and tight leptons. An electron is considered loose when it passes the “LooseAndBLayer”

² c -tagging is also possible using similar techniques. However this thesis only differentiates between b - and light jets.

identification cut. A muon is considered loose when it passes the “Loose” identification cut. No isolation cuts are applied to loose leptons. The selection for tight electrons uses the “Tight” identification and “PLVTight” isolation working points. For tight muons the “Medium” identification and “PLVTight” isolation working points are used. These cuts are summarised in table 4.1.

Table 4.1: Overview over the working points used to define “loose” and “tight” leptons in the Single Top Ntuples.

| | | Electrons | Muons |
|-------|----------------|----------------|----------|
| Loose | Identification | LooseAndBLayer | Loose |
| | Isolation | no cut | no cut |
| Tight | Identification | Tight | Medium |
| | Isolation | PLVTight | PLVTight |

4.5.4 Missing Transverse Energy

Particles that do not interact via the strong or electromagnetic force cannot be detected in ATLAS. However since the initial protons collide head on, the total transverse momentum of the event has to be 0. So whenever an imbalance in the total transverse momentum is observed, this “missing transverse energy” E_T^{miss} (sometimes also MET) has to be due to one or several only weakly interacting particles being produced.

$$E_T^{\text{miss}} = - \sum_i |p_T^i|, \text{ for all visible particles } i. \quad (4.10)$$

In the SM the only particles with such a signature would be neutrinos, but E_T^{miss} is also an important quality in many BSM searches.

Important to note is that E_T^{miss} and p_T^{miss} are often used interchangeably, while that is technically only the case if the invisible particle is massless.

4.6 Data and Monte Carlo Samples

The data analysed was taken during run 2 of the LHC, using pp -collisions with $\sqrt{s} = 13$ TeV. In total an integrated luminosity of $138\,965.16 \text{ pb}^{-1}$ with an uncertainty of 1.7% was recorded [11].

Apart from the data recorded by the ATLAS detector this analysis also relies on Monte Carlo (MC) simulations of the events. These model individual processes to provide samples to be compared to data. These simulated samples do not only include the pp -collision, but also further hadronisation and the interactions of the particles with the detector. This can be computationally very expensive. Therefore two types of MC sets exists; those which simulate the full event and detector (“full sim.”, shortened to FS) and those which use a simplified simulation of the detector to speed up computation times (“fast sim” or “ATLAS fast 2”, shortened to AFII). As far as available, all samples used in this analysis are FS. The only exceptions are some rare backgrounds, for which only AFII samples are available. These are tHq , tWH , $tttt$, and $ttbb$. None of these contribute significantly to the signal or control regions.

To reduce statistical fluctuations, one usually simulates a lot more MC events than the amount that would correspond to the given integrated luminosity. As a result more events with different kinematics can be produced. To account for this every event has a weight associated with it, dependant on several factors like the the simulation of the pile-up, the detector triggers, or the cross section of the process.

All of these influences on the MC weight are also sources of systematic uncertainties (see appendix A). Since the samples simulate both the initial interaction at more than leading order and the hadronisation of the particles in the detector volume, sometimes a leading order event with hadronisation and a next to leading order event describe actually the same process. To avoid double counting in these cases, the MC weights can be negative.

Since the MC samples simulate the exact interaction in the initial event MC samples exist as two levels; truth and reconstructed (reco). The latter includes the detector information and takes its uncertainties and limitations into account, while the former contains the exact information of the physical process. This can be useful to gauge the impact, the detector limitations have on the measurement. MC accounts for pile-up by overlaying other collisions from QCD generators.

The samples used in this analysis are the version 34 of the Single Top Ntuples of the ATLAS collaboration. “These are produced with AnalysisTop-21.2.181, using the SingleTopAnalysis package” [30]. Some preselections are made for these samples: The ntuples only contain events with at least one lepton with $p_T > 27$ GeV and at least one central ($|\eta| < 2.5$) jet with $p_T > 25$ GeV. The events must also pass some common quality cuts. Specifically the samples for the single lepton channel were used. A list with all the processes considered can be found in the appendix B. The single top samples (t -channel, s -channel, and tW), $t\bar{t}$, as well as Higgs events and $t\bar{t}H$ are produced with the POWHEG generator [31] using PYTHIA 8 [32] as the pile-up generator. W +jets, Z +jets, and some Diboson events are generated using SHERPA2.2.1 [33] and the other Diboson and the Triboson samples using SHERPA2.2.2 [33]. All the other rare top and Higgs backgrounds use MADGRAPH [34] and for pile-up PYTHIA 8.

One possible source of errors are differences between MC generators. To account for these, AFII samples for the same process are produced using different generators. Comparison between these generators yields uncertainties that can be applied to the nominal FS samples (see appendix A).

For further reading on the theory and history of Monte Carlo simulations see [35]. For an overview over the use of MC in high energy particle physics see [36].

4.6.1 Grid Computing

These data and MC samples are too large to copy and store on every local computing cluster which might need to access them. For that purpose the ATLAS collaboration has a computing grid, a system of interlinked computers at various international locations, each of which stores some of the samples in addition to backups stored locally at CERN.

While the grid also offers computing capabilities one method to access the samples on the grid is to use the XRootD protocol. With that protocol the files containing the samples can be loaded into RAM on local computing clusters by giving a grid-address instead of a local file path as input. This might limit the runtime of the code based on the download speed of the local computing cluster. But typically the computations necessary to process the samples are complicated enough that the runtime is actually limited by the computing time. Fully copying the samples and processing locally appears to not be any faster than using the XRootD protocol but requires a lot of local storage space. Instructions on how to use the XRootD protocol in this fashion with TopLoop based code, including how to set up HTCondor batch jobs with this method can be found in the appendix in section E.

Further information on XRootD and this use case for it can be found under [37, 38].

Non-prompt and Fake Charged Leptons

There are two types of objects that can be wrongly identified as a lepton in the final state of a process, non-prompt and fake leptons. These objects can have a significant background contribution to analyses, especially in single-lepton regions with heavy flavour quarks. This chapter will introduce these types of objects, the processes which cause them, and one method by which their contribution to the signal region can be estimated.

5.1 Lepton Terminology

While the term “lepton” in the SM refers to e , μ , τ , and their corresponding neutrinos, it is often used as shorthand for charged leptons, excluding the neutrinos. Furthermore since taus decay before they can be detected, their identification and accordingly their fake estimation is handled differently. Therefore in the following section, “lepton” only refers to the charged light leptons, e and μ .

The other term often used as technically inaccurate shorthand is fake leptons, to refer to both fake and non-prompt leptons. Technically fake and non-prompt leptons are physically different objects as described in the following sections. However their effect on the analysis is the same: they get reconstructed as real, prompt leptons and therefore their impact on the analysis needs to be estimated. Similarly the terms prompt leptons and real leptons are used interchangeably.

5.2 Non-prompt Leptons

Non-prompt leptons are real leptons but they are not part of the final state of the pp -interaction. Instead they are produced from secondary interactions or come from the interactions of pile-up. Typically this happens in a heavy flavour decay. A quark, often b or c , can decay via a weak interaction, emitting a lepton later. That lepton can then be detected and falsely reconstructed as primary. As a consequence analyses with a heavy flavour jet in the final state tend to have a higher contribution of non-prompt leptons than light flavour final states. However light hadrons can also produce non-prompt leptons. A significant contribution here are muon candidates from in-flight semi-leptonic kaon decays [39].

Another source of non-prompt leptons are photon conversions into e^-e^+ -pairs. When these occur in the beampipe and are asymmetric in the momentum splitting, it may not be possible to reconstruct the conversion vertex, leading to a misidentification of the object as an electron. This is more likely for

photons from sources like initial state radiation or QED processes not involving leptons than when the photon is emitted close to a real electron as bremsstrahlung or a higher-order QED correction. The latter tends to be considered part of the electron, since the energy deposits in the calorimeter usually overlap enough that they are reconstructed as only a single cluster.

5.3 Fake Leptons

Fake leptons are objects that physically are not leptons, but get reconstructed as a lepton anyway. This can happen when a QCD jet deposits a significant fraction of its energy in the electromagnetic calorimeter. That process is rare but due to the high coupling strength of the strong force, the cross section of those QCD multijet events is much higher than the cross section of a desired process like single top production. These jets almost exclusively fakes electrons, since the hadron calorimeter is deep enough, to stop “essentially all pions and kaons before they can reach the [muon spectrometer]” [39]. So even at high p_T there is only a minimal contribution through punch-through hadrons.

One significant contribution to hadronic fakes is the decay of neutral pions into photons ($\pi^0 \rightarrow \gamma\gamma$). The photons of this decay tend to be collimated, producing only a single cluster in the calorimeter, while an associated track might be produced by a conversion of one of the photons in the detector material. This phenomenon is so important that the electron ID criteria specifically attempt to identify cases with two peaks in the cluster associated with an electron [40].

Therefore “non-prompt muons are the only substantial contribution to the fake/non-prompt background, while for electrons there is more variety” [39]. For lower p_T ranges there tend to be more non-prompt electrons from heavy flavour quarks, while at higher p_T hadronic fakes dominate. An overview of the sources of fake and non-prompt leptons and how they effect electrons and muons can be found in table 5.1.

Table 5.1: Overview of processes contributing to the fake/non-prompt background and their effects on electrons and muons.

| Source | Electrons | Muons |
|----------------------------------|----------------------------------|--|
| non-prompt | dominant at low p_T | only significant contribution |
| photon conversions | minor contribution | no contribution |
| $\pi^0 \rightarrow \gamma\gamma$ | mayor contribution at high p_T | no contribution |
| QCD fakes | dominant at high p_T | minimal contribution, only at high p_T |

5.4 Matrix-Method

There are different methods to estimate fake leptons. One of these is the matrix method. This method relies on the rates at which real and fake loose leptons also pass the tight selection criteria. The total number of measured leptons N is the sum of the total number of real and fake leptons N_r and N_f . The number of tight leptons can be written as the number of loose leptons multiplied with an efficiency ϵ for

real and fake leptons respectively:

$$N^l = N_r^l + N_f^l \quad (5.1)$$

$$\begin{aligned} N^t &= N_r^t + N_f^t \\ &= \epsilon_r N_r^l + \epsilon_f N_f^l . \end{aligned} \quad (5.2)$$

The superscript l and t denoted loose and tight. Rearranging these equations we can calculate the number of fake leptons which pass the tight selection as

$$N_f^t = \frac{\epsilon_f}{\epsilon_r - \epsilon_f} (\epsilon_r N^l - N^t) . \quad (5.3)$$

Based on this equation one can calculate a weight for each individual event, by simply setting $N_l = 1$ and $N_t = 1$ or $N_t = 0$, depending on if the lepton in the event passes the tight selection:

$$w = \frac{\epsilon_f}{\epsilon_r - \epsilon_f} (\epsilon_r - \delta) , \quad (5.4)$$

$$\text{with } \delta = \begin{cases} 1 & \text{if tight} \\ 0 & \text{else} \end{cases} . \quad (5.5)$$

By applying these weights to the data samples one then gets a set of samples estimating the fake contribution to the regions.

The titular matrix can be seen in equation 5.6. It contains the real and fake efficiencies and relates a vector with the number of real and fake leptons with a vector containing the number of tight and non-tight leptons. Note that non-tight and loose are not the same, as loose also contains the tight leptons, while non-tight do not: $N^{\bar{t}} = N^l - N^t$.

$$\begin{pmatrix} N^t \\ N^{\bar{t}} \end{pmatrix} = \begin{pmatrix} \epsilon_r & \epsilon_f \\ 1 - \epsilon_r & 1 - \epsilon_f \end{pmatrix} \begin{pmatrix} N_r \\ N_f \end{pmatrix} . \quad (5.6)$$

Equation 5.3 can also be calculated from the inverse of the matrix. This can be expanded to a 4×4 matrix for dilepton analyses.

5.4.1 Efficiencies

To calculate the weights derived in the matrix method, one needs to extract the real and fake efficiencies. Efficiencies are calculated individually for electrons and muons and binned in different variables, typically p_T and $|\eta|$ of the lepton. The binning in p_T is done, because as described above in table 5.1, different fake sources dominate in different p_T regions. Binning in $|\eta|$ is done because ATLAS is non homogenous in η . As mentioned above, the rate of non-prompt leptons increases with the amount of heavy flavour quarks in the final state. So to accurately estimate the fake lepton contribution in the signal region, the efficiency estimation is often done in regions with the same jet and b -jet multiplicity.

Real Efficiencies from Data

The real efficiencies are extracted from the $Z \rightarrow \ell\ell$ region. This is done, because this region has a clean signature with minimal fake contribution. So the number of measured leptons is set to be the same as the number of real leptons $N \approx N_r$. Therefore the efficiency is the fraction of tight and loose leptons:

$$\epsilon_r = \frac{N_r^t}{N_r^l} \approx \frac{N^t}{N^l}. \quad (5.7)$$

Events with two leptons of the same flavour and opposite sign are selected and a region around the Z-peak is chosen. The background is modelled as linear and subtracted based on the side-band method. For that the centre of the Z-peak is chosen as the signal and two “bands” of similar width to that signal region are chosen as the sidebands. The linear background gets estimated in the sidebands, and then interpolated to the signal region. The corresponding plots of the invariant mass of the two leptons $m_{\ell\ell}$ can be seen in figure 5.1. These plots use the 2015 dataset and have the sideband and signal regions marked. The opposite sign events are also shown in grey, however they are not relevant to the extraction of the efficiencies. [41] performs the background for three different side-band and signal region definitions and calculates ϵ_r as the mean value of the three results. They consider this spread as a potential intrinsic uncertainty of the measurement, but in the end find it to be negligible.

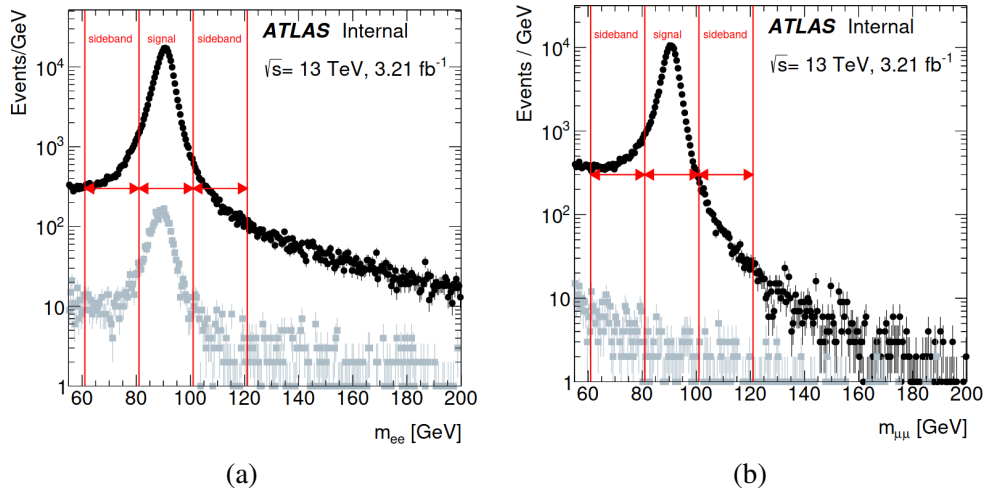


Figure 5.1: “Distribution of the invariant mass of opposite- (black) and same- (grey) sign charge loose (a) electron and (b) muon pairs. Lines show the signal and sideband regions where the yields are calculated” [41].

In this region the real efficiencies get measured using the tag-and-probe method. Here one lepton (tag) is required to pass the tight selection criteria, while the second lepton (probe) selected only needs to pass the loose criteria. The efficiency can then be calculated as the fraction of probes that also pass the tight selection criteria:

$$\epsilon_r^{Z \rightarrow \ell\ell \text{ data}} = \frac{N_{\text{probe}}^t}{N_{\text{probe}}^l}. \quad (5.8)$$

The same analysis is then performed on Monte Carlo simulation. When comparing the results it can be observed, that efficiencies change based on the different processes. The biggest difference can be seen

between Z +jets and $t\bar{t}$ events. To account for this difference a scale factor is derived from the average efficiency of these two events:

$$\epsilon_r = \left| \frac{\epsilon_r(\text{average})}{\epsilon_r(Z + \text{jets})} \right|_{\text{MC}} \epsilon_r^{Z \rightarrow \ell\ell \text{ data}}. \quad (5.9)$$

“In this way, a sort of extrapolation from a Z +jets region to an ‘unknown real lepton source composition’ region is performed” [41].

Real Efficiencies from Monte Carlo

Since prompt leptons are well modelled in the Monte Carlo, the real efficiencies can alternatively also be extracted from simulated samples on truth level. With the same motivation as before a same flavour, opposite sign, dilepton selection is made and efficiencies derived from Z +jets and $t\bar{t}$ samples using the tag-and-probe method.

Since these efficiencies are derived using truth information, scale factors need to be applied to account for the difference in truth and reconstructed level. These are provided by the ATLAS ElectronGamma and Muon Combined Performance Working Groups for electrons and muons respectively.

Fake Efficiencies

Fake leptons are not modelled in Monte Carlo, so these efficiencies have to be extracted from data. For that purpose a fake enriched region is chosen. However no kinematic region consists only of fakes, so an approximation like in equation 5.7 fails for the fake efficiency. Instead the prompt leptons need to be subtracted using simulation. There are two common approaches to identify a region with large fake contribution. One option is to make a dilepton same charge selection. Same charge lepton pairs are rarely produced at the LHC so it is likely that one of them is a fake. By requiring one of the two leptons to be tight, the chance of the other lepton being fake is enhanced. Alternatively single lepton events are selected with kinematic cuts that enhance fake contribution, for example limiting E_T^{miss} or $m_T^{\ell, E_T^{\text{miss}}}$ or using the track impact parameters to ensure the lepton does not originate from the primary vertex [39].

An example for this can be seen in figure 5.2. Like in figure 5.1 these plots show data from 2015, but here a single lepton selection was made. The plots show E_T^{miss} and $m_T^{\ell, E_T^{\text{miss}}}$ for e +jets¹ events. To select the region with the highest fake contribution two cuts $m_T^{\ell, E_T^{\text{miss}}} < 20 \text{ GeV}$ and $m_T^{\ell, E_T^{\text{miss}}} + E_T^{\text{miss}} < 60 \text{ GeV}$ were applied. The remaining real leptons in that region are then subtracted using Monte Carlo leading to the fake efficiency

$$\epsilon_f = \frac{N_f^t}{N_f^l} = \frac{N^t - N_{r, \text{MC}}^t}{N^l - N_{r, \text{MC}}^l}. \quad (5.10)$$

The use of Monte Carlo simulation introduces a systematic uncertainty to ϵ_f .

¹ [41] also contains similar plots for μ +jets events.

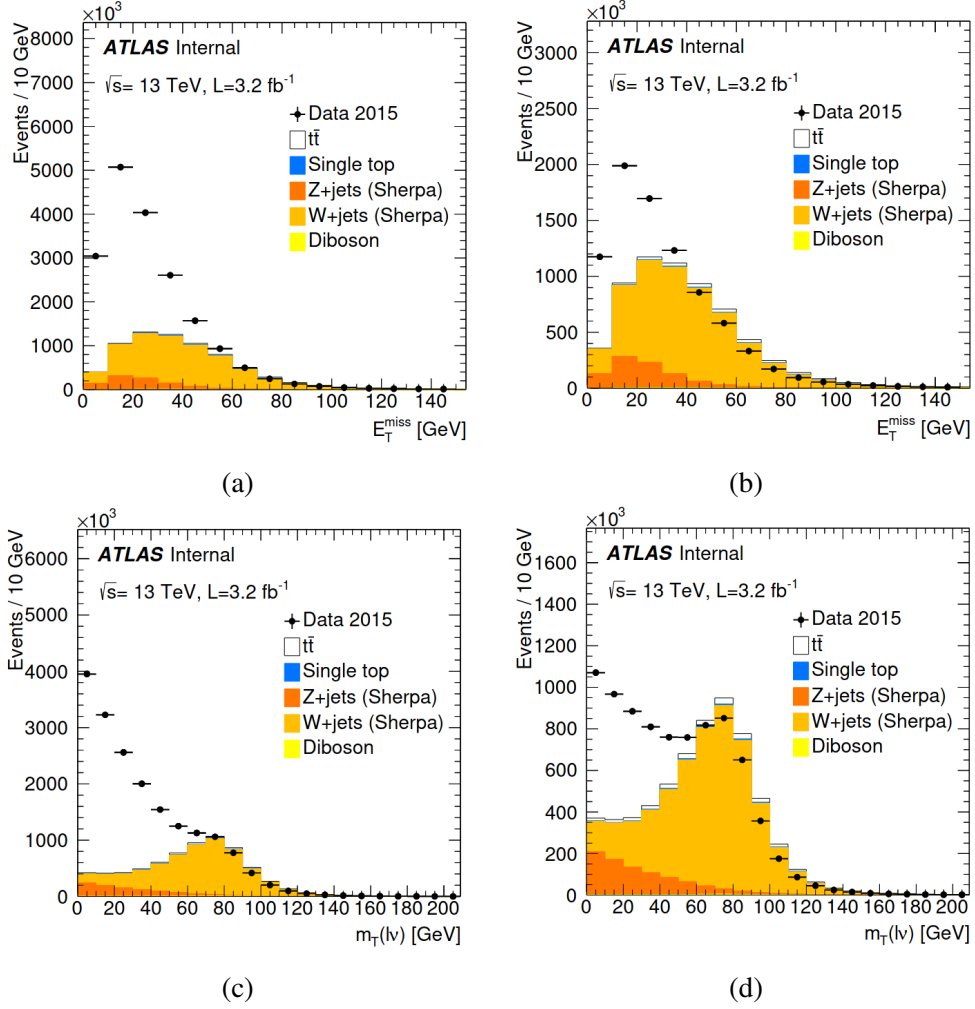


Figure 5.2: “Distribution of (a,b) the missing transverse energy E_T^{miss} and (c,d) the $m_{T(\ell\nu)}^{\ell, E_T^{\text{miss}}}$ for data and the different MC contributions at (a,c) loose and (b,d) tight levels for e +jets events with one electron and at least one jet” [41].

Analysis

This chapter shows the steps for the analysis of the cross section of single top production in the t -channel. First the cuts made to select events for the signal and control regions are described, then the fake estimation using the matrix method discussed in chapter 5. A short introduction to artificial neural networks and their use in signal background separation is given. The maximum likelihood fit is introduced and its application to the analysis presented.

6.1 The Single-Top t -channel

As discussed in section 2.6 one possible production mode of a single top quark at the LHC is the single top t -channel (figure 2.3 (a)). To analyse that channel, the final state of that process needs to be understood. In addition to the production of the single top, one also needs to consider its decay into a W -boson and a bottom quark. This analysis only considers the W -boson decay into light leptons, so an electron or muon and a corresponding neutrino, or into a leptonically decaying tau. The leptonic W -decay is selected since this final state is a lot cleaner than the hadronic decay: There are fewer jets in the final state and the lepton can be triggered on. Additionally the W can be easier reconstructed from the lepton and missing energy, rather than a combination of two jets, which would have to be matched to the W first.

The Feynman diagram of that process for both tq and $\bar{t}q$, including the initial gluon splitting and the decay of the top is shown in figure 6.1. The second b -quark produced by the gluon splitting is not considered part of the final state, as it tends to be low momentum and very forward. As a consequence it is not visible in the detector.

6.2 Event Selection

In addition to the preselection made for the Single Top Ntuples described in section 4.6, additional cuts are applied to generate a region enriched in signal events.

As discussed in the previous section the final state of the single-top t -channel consists of the scattered spectator quark and the titular top quark, decaying into a b -quark and a W -boson. The latter in turn decays into a lepton and a neutrino (see figure 6.1). Hence in the detector the event will have a signature of two jets (one from the spectator quark and one from the b -quark), one of which will be b -tagged, a

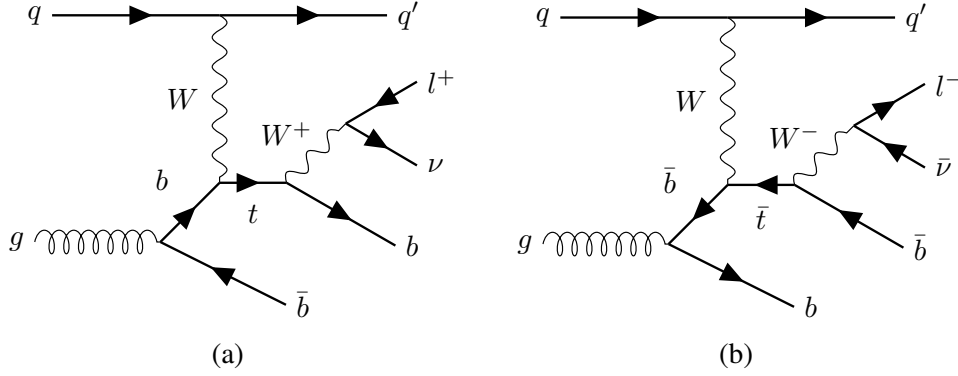


Figure 6.1: Leading order Feynman diagrams for (a) tq and (b) $\bar{t}q$, including the initial $g \rightarrow b\bar{b}$ splitting, the top-decay and the leptonic W -decay. The second b -quark produced by the gluon splitting is typically not visible in the detector.

charged lepton, and missing transverse energy from the neutrino. The jet multiplicity is abbreviated as 2j1b, denoting **2 jets**, **1** of which is to be **b -tagged**. The jets are required to have $p_T > 30$ GeV and the b -jet is required to be tagged at the 60% working point. For the transition region between the endcap and the forward calorimeter $2.7 < |\eta| < 3.5$ the requirement is raised to $p_T > 35$ GeV, because in that area the detector response is worse than in the rest of the acceptance. However an event with a third jet in that transition region and $p_T > 30$ GeV will still be rejected to reduce contamination from events with three jets¹. The lepton has to be tight (see table 4.1) and have $p_T > 28$ GeV. Events with a second loose lepton with $p_T > 10$ GeV get vetoed. Due to the detector limitations, the light jet is required to have $|\eta| < 4.5$, and the b -jet and the lepton $|\eta| < 2.5$.

To account for the neutrino the event is required to have $E_T^{\text{miss}} > 30$ GeV. This also reduces the amount of fake and non-prompt leptons, especially electrons, since the requirement of a neutrino in the event reduces the impact of QCD fakes. Additionally the impact of fake and non-prompt leptons is reduced, by requiring the transverse mass of the W -boson, reconstructed as the transverse mass of the lepton- E_T^{miss} -system to be greater than 50 GeV. Furthermore a triangular isolation cut is applied to the transverse momentum of the lepton: $p_T^\ell > 40 \cdot \left(1 - \frac{\pi - |\Delta\varphi(\ell, j_1)|}{\pi - 1}\right)$ GeV, where $\Delta\varphi(\ell, j_1)$ is the difference in the azimuthal angle between the lepton and the highest p_T jet.

Finally the transverse mass of the b -jet-lepton-system is required to be below 160 GeV. This serves as a proxy of the transverse top-mass, assuring that the t -quark is real and not a virtual particle with mass greater than that of the top. The limit is set slightly below the nominal top mass of approximately 173 GeV because the used proxy-mass of the b -jet and lepton does not account for the neutrino.

For the final fit two control regions are defined, one enriched in W +jets events and one in $\bar{t}\bar{t}$ events. Both these regions differ from the signal region in their jet multiplicity. The W +jets control region requires three jets, none of which are to be b -tagged at the 60% working point (3j0b) and the $\bar{t}\bar{t}$ region requires four jets, two of which are to be b -tagged (4j2b). These are chosen, because they are relatively clean regions, which consist almost entirely of W +jets and $\bar{t}\bar{t}$ events respectively. Those are the main backgrounds in the signal region of the analysis, as can be seen in figure 6.2. Like the signal regions, the control regions are also split by lepton charge. Their E_T^{miss} - and p_T^ℓ -distributions can be found in the

¹ The same is true for the forth and fifth jet in the 3j0b and 4j2b control regions respectively.

appendix in figures C.1 and C.2.

An overview of the selection for the signal and control regions can be found in table 6.1.

Table 6.1: Selection for the signal and control regions. j_1 denotes the jet with the highest p_T . All three regions are split by the lepton charge into two regions, denoted with a + or - denoting the sign of the charge.

| Region | Selection |
|---------------|---|
| All Regions | $p_T^{\text{jet}} > 30 \text{ GeV}$, raised to 35 GeV for $2.7 < \eta < 3.5$, still veto additional jets with p_T between 30 and 35 GeV in that region. 1 tight lepton with $p_T^\ell > 28 \text{ GeV}$ veto 2nd loose lepton with $p_T^\ell > 10 \text{ GeV}$ $E_T^{\text{miss}} > 30 \text{ GeV}$ $m_T^W = m_T^{\ell, E_T^{\text{miss}}} > 50 \text{ GeV}$ $p_T^\ell > 40 \cdot \left(1 - \frac{\pi - \Delta\varphi(\ell, j_1) }{\pi - 1}\right) \text{ GeV}$ $m_T^{b, \ell} < 160 \text{ GeV}$ |
| SR | 2 jets, 1 b -tagged (60% working point) |
| $t\bar{t}$ CR | 4 jets, 2 b -tagged (60% working point) |
| W +jets CR | 3 jets, 0 b -tagged (60% working point) |

To reduce the contribution of events with hadronic W 's to the background, a cut on the invariant masses of the two jets and the two-jet system, rejecting events with jet masses close to the W -mass was considered. However to maintain comparability to the results of [42] this cut was not applied.

6.3 Fake Estimation

The contribution of fake leptons was estimated using the matrix method discussed in section 5. The efficiencies were not derived specifically for this analysis, instead older efficiencies from a single top polarisation analysis were used. The files containing the efficiencies can be found at [43]. These were derived in an inclusive region in terms of jet multiplicity, only requiring at least one jet in the final state and not putting any constraints on the amount of b -tagged jets. The real efficiencies were derived from data, not MC. Both real and fake efficiencies were binned in $|\eta|$ and p_T of the lepton. Since the kinematic region for which the efficiencies were derived does not require a b -jet, it is expected, that non-prompt leptons will not be well modelled in the signal region. This will probably lead to an overall underestimation of the fake background. To account for this mismodelling a conservative uncertainty of 50% is applied to the background (see 6.5.3).

The E_T^{miss} - and p_T^ℓ -distributions of the signal regions can be seen in figure 6.2. The main backgrounds are $t\bar{t}$, W +jets, and fakes. tW contributes less than those three, but still has a contribution impact to the region. As expected, the ratio between data and the prediction is slightly above one, getting worse with low E_T^{miss} and especially low p_T^ℓ . That difference is likely due to the underestimation of non-prompt leptons.

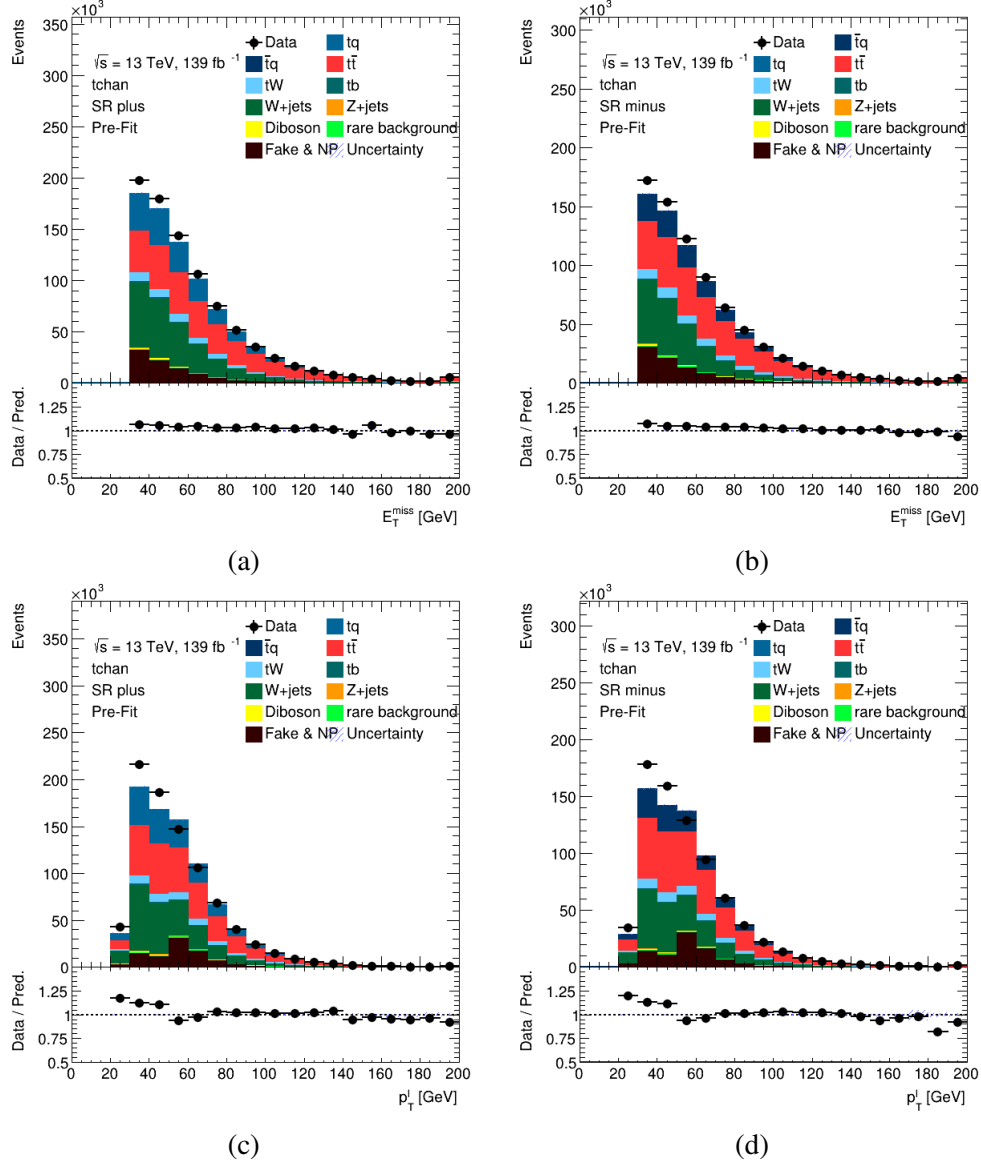


Figure 6.2: (a,b) E_T^{miss} - and (c,d) p_T^l -distributions of the signal regions for leptons with (a,c) positive and (b,d) negative charge. tq and $t\bar{q}$ are the signal in the respective regions.

6.4 Signal Background Separation

To additionally enrich the signal region in single top events an artificial neural network (NN) is trained to separate the signal and background events in the signal region.

6.4.1 Artificial Neural Networks

An artificial neural network is an algorithm “trained” to be able to solve a specific problem. This section will specifically focus on binary classifiers without going into detail on how the algorithm might be implemented. The NN takes a set of input parameters and through a set of nonlinear calculations transforms it into a set of output parameters. For a binary classifier that output is a single number, typically between 0 and 1. In this case that input is a set of kinematic variables describing an event and the output is a measure of “signalness”. Ideally the network will output 0 when receiving the kinematics of a background event as input and 1 when receiving those of a signal event.

To that end the network is “trained” using simulated events. A set of events, labelled as background or signal is provided to the network, and the output is calculated. Then a measure of the difference between that output and the ideal output called loss is calculated and the individual calculations the network does are slightly changed. This process is repeated, where the changes applied to the calculations are dependant on if the previous changes reduced the loss. A sufficient repetition of these steps should then “teach” the network the difference between signal and background events.

The network has several parameters, such as the number of repetitions done and the size of the changes per repetition, which do not get changed throughout the training. These parameters are known as hyperparameters. The ideal set of hyperparameters is highly problem dependant.

After that training the network is provided with the kinematic variables of the recorded data events, thereby sorting the data by its “signalness”. This will provide a region in which the shapes of signal and background events are very different, allowing for isolating the signal contribution for the further fit (see section 6.5).

The field of artificial neural networks and machine learning in general is of course much wider. For a more general introduction to machine learning see for example [44].

6.4.2 Neural Network Output

The NN for this analysis was set up and trained by fellow master student Richard Baumann. The hyperparameters were tuned using a grid search. A table with the hyperparameters, and plots showing the quality of the training can be found in the appendix D.

The output of the trained network can be seen in figure 6.3. The backgrounds are clearly shifted towards 0 and signal to 1. Notably the difference between data and predictions is also larger for lower values of NN_{out} . That difference between data and predictions is likely largely due to an underestimation of non-prompt leptons and the NN is shifting the predicted fakes to the left as well. Therefore it seems like the network is able to separate fakes from the signal, even though they have not been modelled well and were not included in the training samples.

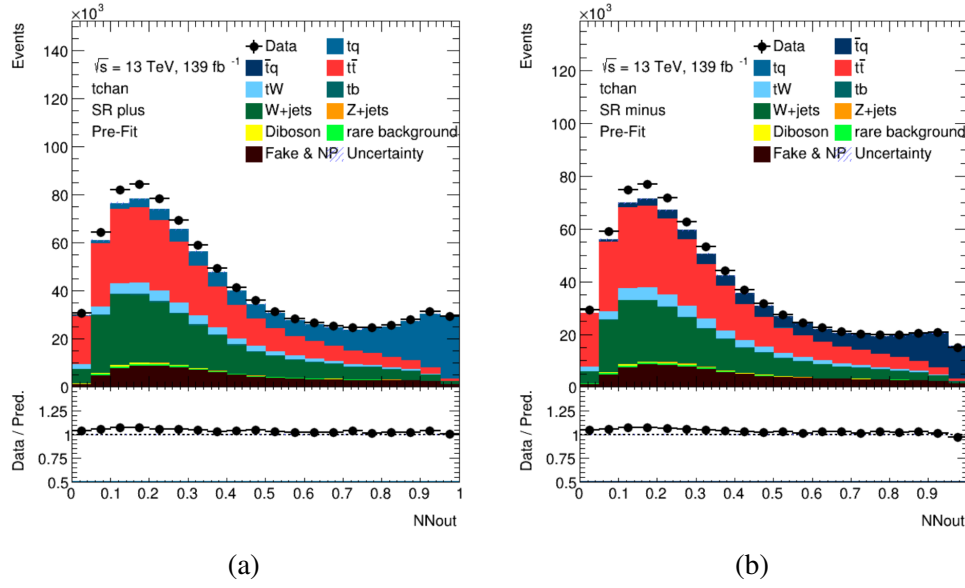


Figure 6.3: Distribution of the neural network output NN_{out} for (a) signal region plus and (b) signal region minus. In both regions all backgrounds, including fakes are visibly shifted towards lower NN_{out} , while the signal is shifted towards higher values. The disagreement between data and prediction is also more pronounced for lower values of NN_{out} .

6.5 Fit and Cross Section Measurement

Using the signal-background-separated signal region a maximum likelihood fit is used to generate normalisation factors to apply to the theoretical t -channel cross sections.

6.5.1 Maximum Likelihood Fit

As discussed in 3.8 the expected number of events N of a certain process depends on the integrated luminosity \mathcal{L}_{int} and the cross section of that process σ . However not all events that occur at the collider can be observed. Rather the number of observed events n depends on the event selection and the quality of the event reconstruction. Dependant on these an efficiency ϵ is introduced:

$$n = \epsilon N = \epsilon \sigma \mathcal{L}_{\text{int}}. \quad (6.1)$$

In the signal region, events of different processes can be observed, giving the total number of observed events as

$$n = \epsilon_s \sigma_s \mathcal{L}_{\text{int}} + \mathcal{L}_{\text{int}} \sum_i^{\text{bkg.}} \epsilon_i \sigma_i. \quad (6.2)$$

Here the subscript s denotes the signal (tq or $\bar{t}q$) and i the different background processes. To extract the signal cross section from the experiment, the theoretical cross sections are used and a scaling parameter

μ is introduced to the signal term:

$$n(\mu) = \epsilon_s \sigma_s \mathcal{L}_{\text{int}} \mu + \mathcal{L}_{\text{int}} \sum_i^{\text{bkg.}} \epsilon_i \sigma_i . \quad (6.3)$$

The probability of observing a number of events x in a certain interval with a mean number of events \bar{x} is given by the Poisson distribution:

$$P(x) = e^{-\bar{x}} \frac{\bar{x}^x}{x!} . \quad (6.4)$$

Using the expected number of events derived above $\bar{x} = n(\mu)$, the probability of observing x events is given by

$$P(x, n) = e^{-n(\mu)} \frac{n(\mu)^x}{x!} . \quad (6.5)$$

Since this gives the probability of observing x events in a certain area, one gets the overall likelihood of a model describing the data by forming the product of these probabilities in each bin:

$$L(\vec{x}, \vec{n}) = \prod_j^{\text{bins}} P(x_j, n_j) , \quad (6.6)$$

where \vec{x} and \vec{n} are vectors containing the entries of x and n in every bin and the subscript j denotes the bin number.

To replace that product with a sum, it is common to take the logarithm of the likelihood. The advantage is of that is that calculating sums is computationally less expensive. By convention a factor -2 is introduced, giving the quantity known as “log-likelihood”. The logarithm is strictly monotone rising and reaches its maximum at the same point as its argument, allowing for this transformation.

$$\begin{aligned} \Lambda(\vec{x}, \vec{n}) &:= -2 \ln(L(\vec{x}, \vec{n})) \\ &= -2 \sum_j^{\text{bins}} \ln(P(x_j, n_j)) . \end{aligned} \quad (6.7)$$

In the fitting process μ serves as the parameter of interest, which is varied to either maximise the likelihood L or minimise the log-likelihood Λ . The expected number of events for each process including the signal the full term $\mathcal{L}_{\text{int}} \epsilon_i \sigma_i$ is provided by Monte Carlo simulation.

In addition to allowing for changes in the parameter of interest μ , the number of background events are allowed to fluctuate within their uncertainties. These factors are called nuisance parameters. While the backgrounds will only vary within their uncertainty, it is still possible that they will be reduced or increased by a larger factor, when the fit parameters are correlated. So to ensure, that the fitting algorithm can separate the impact a change of μ and a change of the number of background events has on the quality of the fit, a region has to be chosen, in which the shapes of the background and signal are different from each other. Therefore the NN_{out} -distribution provided by the network is a good variable to perform the fit in.

To additionally constrain the nuisance parameters and thereby improve the quality of the fit the control regions are used (see section 6.2). These regions are fitted simultaneously to the signal region, while being dominated by one background with significant contribution to the signal region ($t\bar{t}$ and W +jets

respectively).

6.5.2 Statistics Only Fit

As a first step to gauge if the signal background separation provides sufficient shape difference between background and signal, a statistics only fit is performed. For this fit, no nuisance parameter based on the systematic uncertainties is introduced. Instead the only parameter that can vary in this fit is the parameter of interest μ . Looking at the uncertainties of that fit is a first hint, whether the precision of the analysis will be limited by statistical or systematic uncertainties. The resulting post fit plots can be seen in figure 6.4.

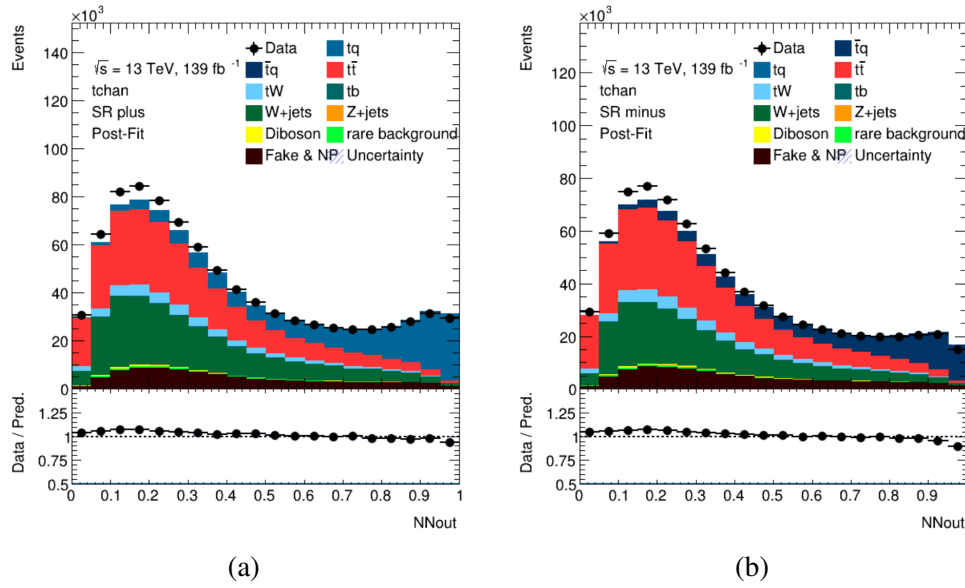


Figure 6.4: NN_{out} -distributions after performance of the statistics only fit for (a) signal region plus and (b) signal region minus.

For the statistics only fit, the parameters of interest are determined to be

$$\begin{aligned}\mu_{t\bar{t}q}^{\text{stat.}} &= 1.0823 \pm 0.0038, \\ \mu_{t\bar{t}t}^{\text{stat.}} &= 1.1025 \pm 0.0055.\end{aligned}\tag{6.8}$$

These values are reasonably close to 1 which is the expected value when not observing new physics and have uncertainties below 1%. So one can conclude, that a fit taking systematic uncertainties into account will likely be able to determine an accurate result with a precision limited by systematics, not a lack of statistics.

One metric to check the quality of the fit is a likelihood scan. These scans work by fixing the parameter of interest μ to a certain value and then running the fit 30 times, only varying the nuisance parameters, and recording the obtained likelihood. One expects to observe a smooth curve with a minimum (for the log-likelihood) around the value of μ the fit determined. If the curve has more than one minimum or any kinks in it, that can be a sign of a problem with the fit. These curves can also be used to check for problems if a fit does not converge. The likelihood scans for the statistics only fit are shown in figure 6.5.

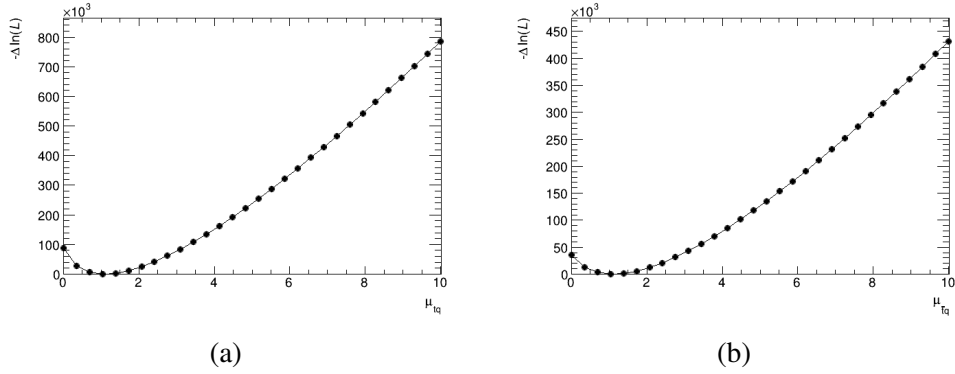


Figure 6.5: Likelihood scans of the statistics only fit for the signal region (a) plus and (b) minus.

6.5.3 Cross Section Uncertainty Only Fit

As a next step to gauge the potential quality of the final fit, the three main backgrounds $t\bar{t}$, W +jets, and tW have the uncertainties of their cross sections applied to them (see table 6.2). Additionally the fake background has a conservative uncertainty of 50% applied to it to account for the fact that the efficiencies were derived for a different kinematic region. This serves largely as a test of the quality of the two control regions. These regions should constrain the possible variation of these backgrounds, mainly $t\bar{t}$ and W +jets. In addition to the parameter of interest μ , the fit also varies these four backgrounds within

Table 6.2: Uncertainties of the cross sections of the three main backgrounds and the uncertainty applied to the fake background.

| Process | Uncertainty |
|------------|-------------|
| $t\bar{t}$ | 6% |
| W +jets | 30% |
| tW | 5% |
| Fakes | 50% |

their uncertainty, by introducing nuisance parameters as described above. The post fit plots can be seen in figure 6.6.

For this fit with the cross section systematics, the parameters of interest are determined to be

$$\begin{aligned}
 \mu_{tq}^{\text{crosssec. sys.}} &= 1.0091 \pm 0.0048, \\
 \mu_{\bar{t}q}^{\text{crosssec. sys.}} &= 1.0315 \pm 0.0072.
 \end{aligned}
 \tag{6.9}$$

These values are again close to the expected value of 1. Notably the introduction of only these four systematics already increased the uncertainties of the results by roughly a quarter. This is another hint that the final fit will be limited by systematics, not statistics. In general one can expect similar systematic uncertainties to the analysis measuring the tW cross section [45]. Especially the introduction of uncertainties on the modelling of tW and $t\bar{t}$ and the jet energy scales are expected to affect the uncertainties significantly. The jet energy scale is used for the calibration of the calorimeters to account for the energy lost in the stopping materials (see section 4.1.4).

More interesting for this fit however are the four nuisance parameters. These can be seen in figure

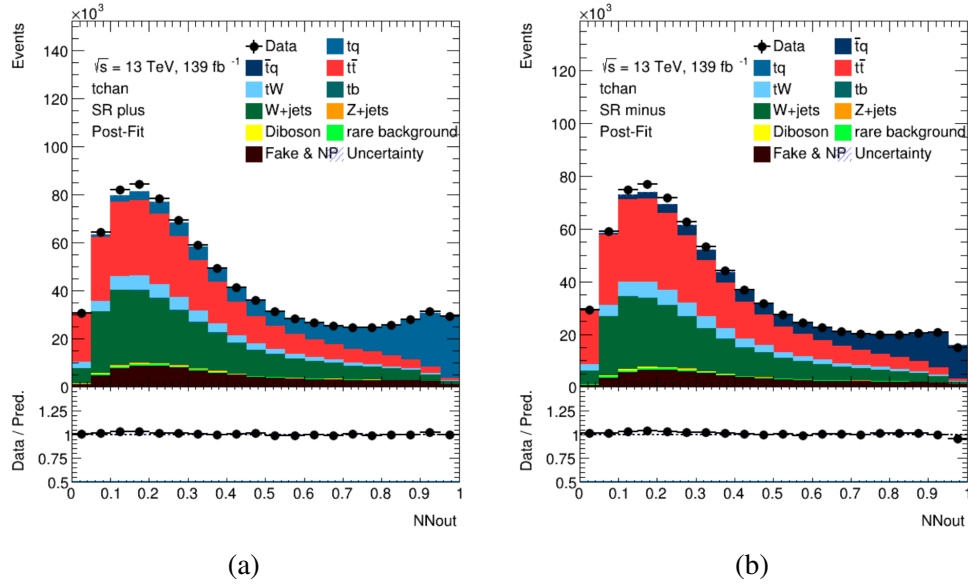


Figure 6.6: NN_{out} -distributions after performance of the fit using only the systematic uncertainties for the fake estimation and the cross sections of the main backgrounds for (a) signal region plus and (b) signal region minus.

6.7. These plots show the difference between the initial theory value for the cross section and the value determined by the fit, normalised to the uncertainty. So if the fit is in full agreement with the theory input, the corresponding point in the plot will be located at 0 and have an uncertainty of ± 1 . The points for $t\bar{t}$, W +jets, and Fakes are all close to 0, however for tW the difference between the initial input and the fit result is off by a factor of 5, meaning the point cannot be seen in the figure. All points have uncertainties below 4%, making their error bars not visible in the plot. These uncertainties as well as the high tW cross section are unphysical. The most likely explanation for this behaviour is that the low number of parameters means that there is not enough possible variation for the fit to reach more reasonable results. This is especially the case in combination with the control regions. In a sense the control regions constrain the nuisance parameters “too much”.

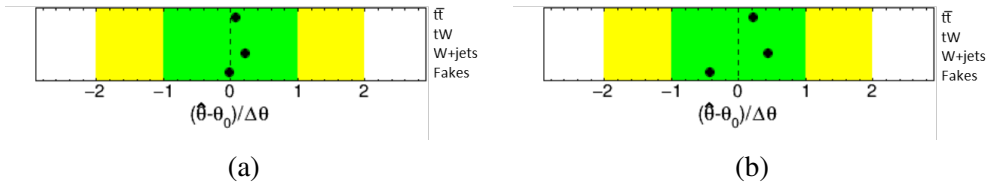


Figure 6.7: Difference between the nominal values for the four varied backgrounds and the fit results, normalised to the uncertainty for the fit to (a) signal region plus and (b) signal region minus. tW is off to the right of the plot at approximately 16.7. All the uncertainties are too small for the error bars to be visible in the plot.

The likelihood scans for this fit are shown in figure 6.8.

A fit taking into account more systematics uncertainties is briefly described in appendix A. This fit however did not converge. Likely the larger amount of uncertainties produces too large of a phase space for the fit to find a clear minimum.

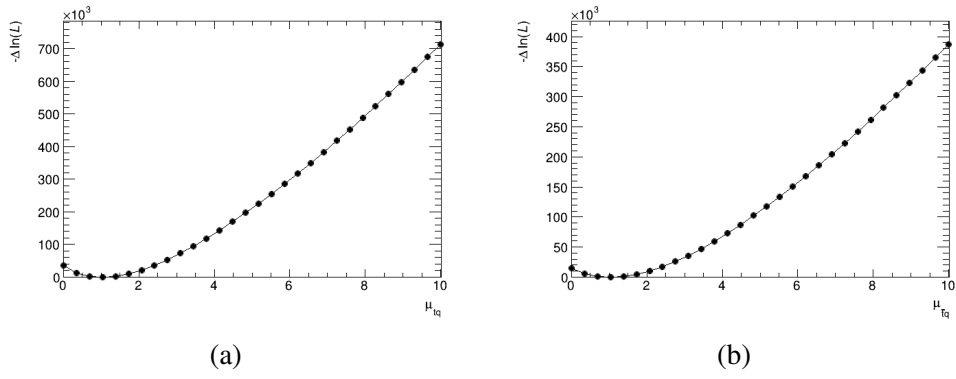


Figure 6.8: Likelihood scans of the cross section uncertainty only fit for the signal region (a) plus and (b) minus.

Summary and Conclusion

To summarise the top quark is an interesting probe for different aspects of the standard model. Notably the CKM matrix elements V_{td} , V_{ts} , and V_{tb} and Higgs physics, making it an interesting particle to measure and understand in detail. One kind of these measurements is the determination of the cross section of various top quark production processes. In this thesis the steps for a measurement of the production cross section of a single top quark in the t -channel were presented. To that end the data the ATLAS detector recorded from 2015 to 2018 was analysed. Said dataset contains 139 fb^{-1} of pp -collisions at $\sqrt{s} = 13 \text{ TeV}$.

For the analysis the dominant decay channel of the top into a bottom quark was selected, and further the leptonic decay channel of the resulting W . Accordingly a single lepton selection with one light and one b -jet was made. The contribution of fake leptons to the signal region was estimated using the matrix method and a separation of signal and background was applied using an artificial neural network. Finally a maximum likelihood fit was applied, first using only statistical uncertainty and then using limited systematic uncertainties.

The performance of the statistics only fit and of a fit with very limited uncertainties revealed results that are broadly in agreement with the SM expectation, however no exact determination of the cross section was possible. Nevertheless it can be concluded, that the results of the final analysis will likely be limited by systematic uncertainties, not statistics, as the statistics only fit had uncertainties below 1% and the introduction of only cross section uncertainties for the main backgrounds already raised the uncertainty by about a quarter. Large contributions to the systematic uncertainty are expected from the jet energy scales and the modelling of tW and $t\bar{t}$. The next step for a full systematics fit would be to investigate, why the fit did not converge. Reducing the number of bins or grouping backgrounds together might yield better results.

Apart from finalising the fit taking all systematics into account, some improvements to the analysis could be made. A derivation of the fake and real lepton efficiencies for the signal regions of this analysis would likely yield a better fake estimation and thereby reduce a large source of uncertainty. Additionally it might be interesting to consider a signal region more inclusive in its number of light jets. The selection made in this thesis is very much a leading order selection. So to get a phase space more representative of the entire process, one might want to consider an analysis of the 2-3j1b or even 2-4j1b region. To avoid domination by $t\bar{t}$ and tW events, one might need to require the additional jets to be close to one of the higher p_T jets. This should increase the probability, they result from final state QCD radiation. However even then, including higher order diagrams will increase the impact, the interference between

tW and $t\bar{t}$ has on the analysis. Engaging with these processes is a non-trivial problem.

Appendix

Full Systematic Fit

The last fit performed includes all the systematic uncertainties listed in the following section A.1 to provide final results with full uncertainties. However that fit did not converge. Figure A.1 shows the likelihood scans. At least the scan in the signal region minus does show a visible minimum close to 1, while the scan in the signal region plus at has two minima in that area. But even after limiting the range in which μ could vary to the interval $[0.5, 1.5]$ the fits did not converge. This is likely a problem with the systematics. There are too many systematics that can vary too greatly for the fit to find a clear minimum. Further investigation, which systematics cause this issue is needed. Possible future steps to produce a converging fit might be combining backgrounds or reducing the number of bins, to limit correlations between the nuisance parameters.

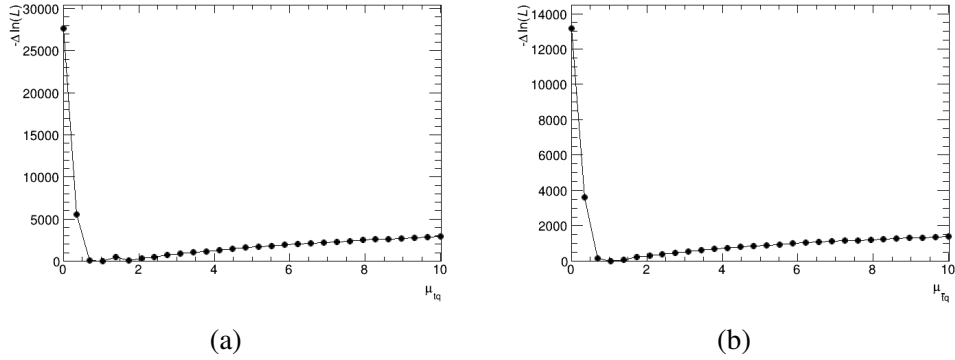


Figure A.1: Likelihood scans of the full systematic fit for the signal region (a) plus and (b) minus.

A.1 Systematics

This analysis like any other is affected by several systematic uncertainties. As already mentioned in section 6.5.3 the cross sections used in the generation of the MC samples (see section 4.6) have uncertainties applied to them. The impact of these is negligible for processes with small contributions to the signal regions, but for the main backgrounds $t\bar{t}$, tW , and W +jets they have to be accounted for. As already shown in table 6.2, 6% uncertainty is applied to $t\bar{t}$, 5% to tW , and 30% to W +jets. Also

discussed in section 6.5.3 is the uncertainty applied to the fake lepton background. The estimation of that background is dependant on the real and fake lepton efficiencies, which were derived for a kinematic region, more inclusive in its jet multiplicity than the signal regions (see section 6.3). Most notably did the derivation of the efficiencies not require a b -jet in the final state, which likely means an underestimation of non-prompt leptons in the signal region of this analysis. The exact impact of that underestimation is difficult to gauge, so a conservative uncertainty of 50% is applied to the fake background.

All samples are affected by the uncertainty of the luminosity of 1.7% [11]. For some important processes, samples were generated using different MC generators. The differences between these generators are used as systematic uncertainties in the modelling of these processes. To save computing power, the systematic samples are only generated using the simplified detector simulation AFII (see section 4.6). For comparability between these samples, nominal AFII samples with the same generators as the FS samples were also generated. The uncertainties resulting from the comparison of the AFII samples are then applied to the FS samples. These modelling uncertainties are applied to the single top t - and s -channel, $t\bar{t}$, and tW . It is expected that especially the $t\bar{t}$ modelling systematics have a significant impact on the quality of the fit.

In next to leading order processes the tW process can have a second on-shell t -quark and therefore have an interference with the $t\bar{t}$ -channel. There are two approaches to this interference called Diagram Removal (DR) and Diagram Subtraction (DS). The nominal samples in this analysis use the DR scheme. Hence the DS samples are used to gauge the modelling uncertainty from this interference. More details on the tW - $t\bar{t}$ -interference and the two schemes can be found in [45].

Additionally weight systematics for the jet vertex trigger (JVT) and pile-up are applied, as well as for the triggers, reconstruction, identification, and isolation of electrons and leptons respectively. Lastly weight systematics for the b -tagging are applied. The weight systematics are applied to all FS samples.

Another weight systematic that would likely be interesting for a lepton+jets analysis would be the jet energy scale. However this was not yet taken into account for the non converging fit described above.

Monte Carlo Samples

As described in section 4.6 the samples used in this analysis are version 34 of the single-lepton channel of the Single Top Ntuples. Further details on these samples can be found at [30, 46]. The nominal MC samples used are listed in table B.1.

Table B.1: All MC Samples used for the simulated prediction with the label they have in the plots in section 6.

| Sample | Label in Plots |
|-------------------------|-----------------|
| single top t -channel | tq |
| | $\bar{t}q$ |
| $t\bar{t}$ | $t\bar{t}$ |
| tW | tW |
| single top s -channel | tb |
| W +jets | W +jets |
| Z +jets | Z +jets |
| Diboson | Diboson |
| Triboson | rare background |
| Higgs | rare background |
| $t\bar{t}V$ | rare background |
| $t\bar{t}H$ | rare background |
| tZq | rare background |
| tHq | rare background |
| tWZ | rare background |
| tWH | rare background |
| $t\bar{t}t$ | rare background |
| $t\bar{t}t\bar{t}$ | rare background |
| $t\bar{t}b\bar{b}$ | rare background |

E_T^{miss} - and p_T^ℓ -Distributions in the Control Regions

This section contains the E_T^{miss} - and p_T^ℓ -distributions of the $t\bar{t}$ (figure C.1) and W +jets (figure C.2) control regions, analogous to figure 6.2 for the signal regions. As described in section 6.2 the control regions differ from the signal region only in the jet multiplicity. While the signal region has a 2j1b selection, the $t\bar{t}$ region selects events with 4j2b, and the W +jets region with 3j0b. This results in regions almost entirely consisting of the titular process.

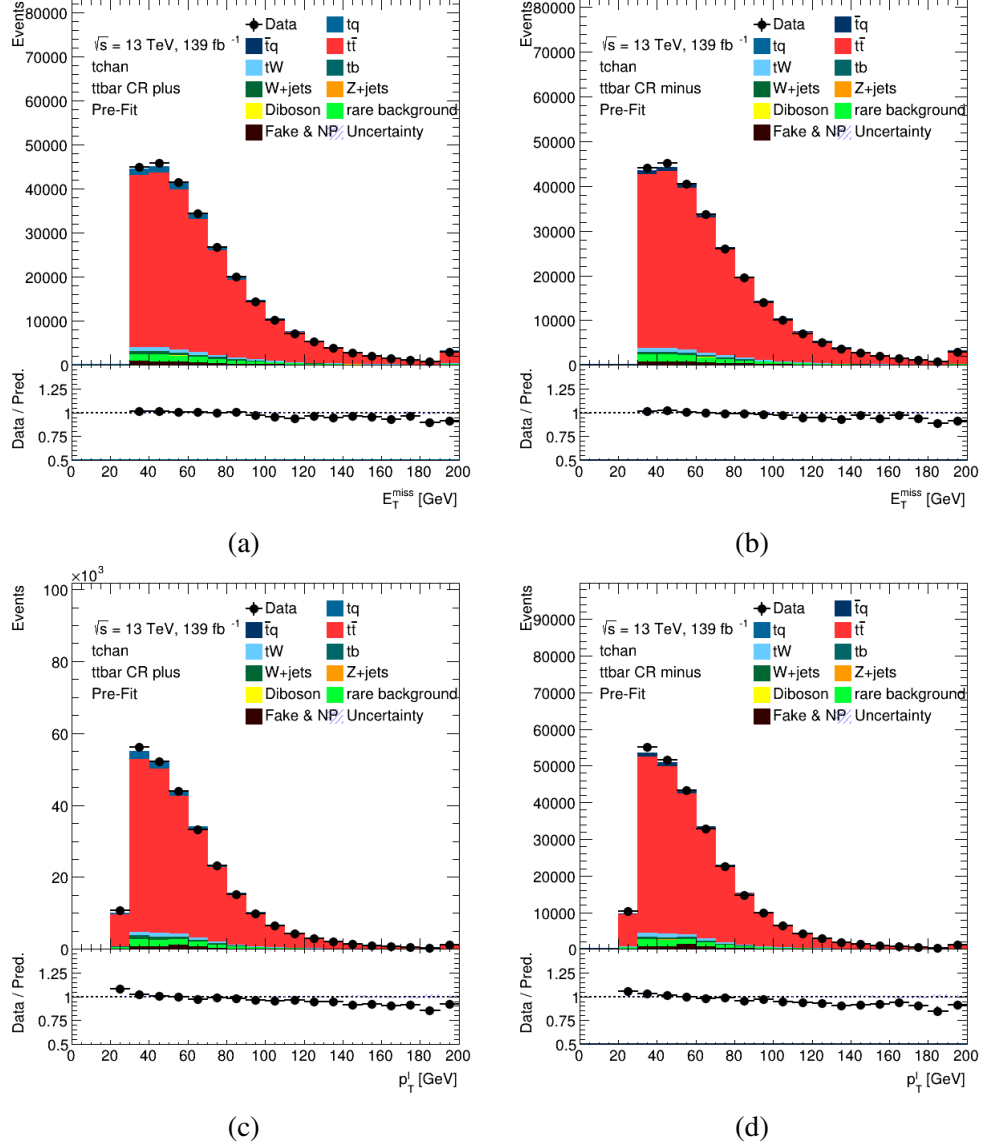


Figure C.1: (a,b) E_T^{miss} - and (c,d) p_T^ℓ -distributions and of the $t\bar{t}$ control regions for leptons with (a,c) positive and (b,d) negative charge.

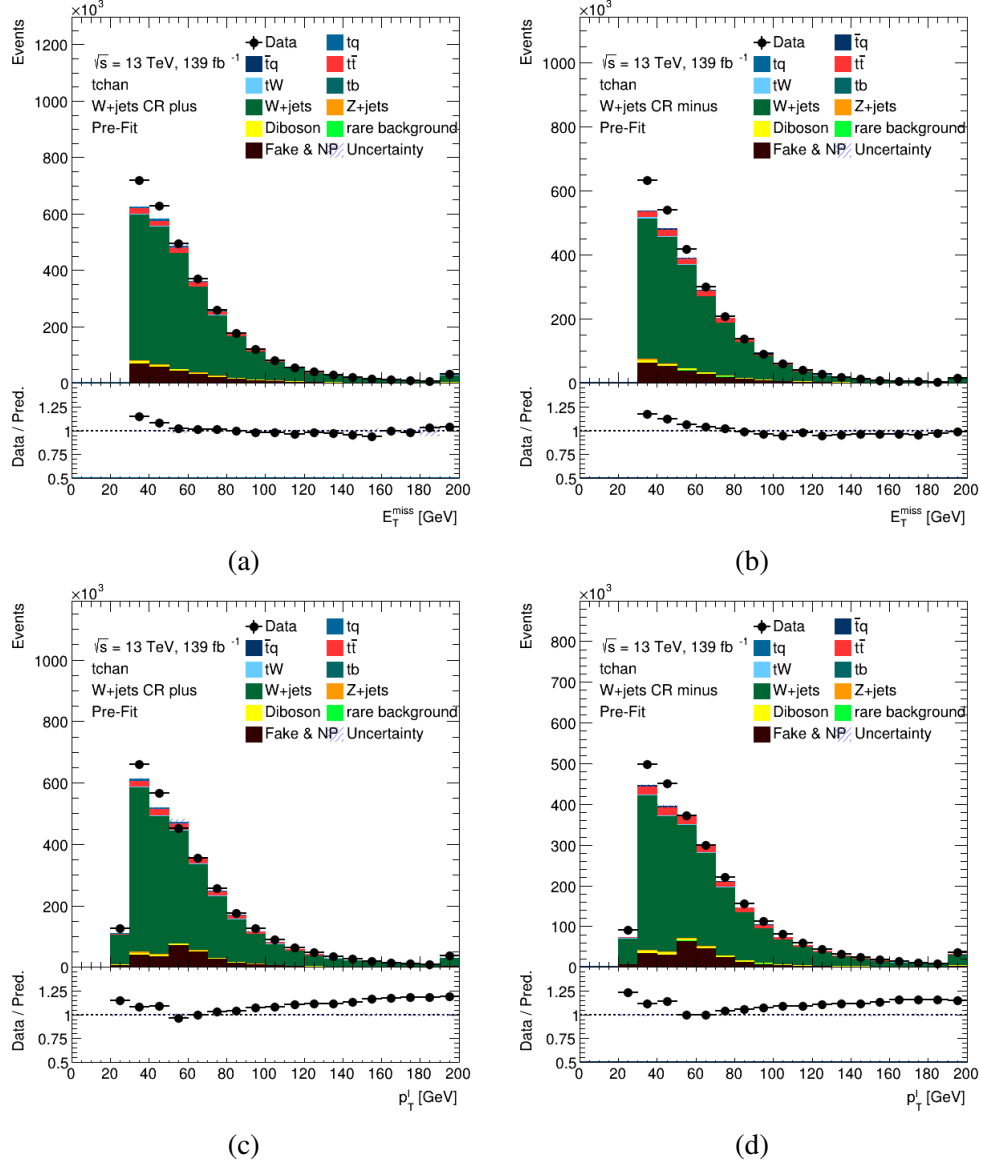


Figure C.2: (a,b) E_T^{miss} - and (c,d) p_T^ℓ -distributions and of the W +jets control regions for leptons with (a,c) positive and (b,d) negative charge.

Neural Network Training

This appendix gives a few more technical details on the NN trained by Richard Baumann. This is intended purely as documentation of these values and as information for the interested reader. No detailed explanation is given for these terms, only a rough overview:

The epochs are the number of optimisation steps performed in the training. Hidden layers and nodes per layer give the network “architecture”, they are basically a measure of how complicated the calculation performed by the network is. Dropout randomly selects parts of the network every epoch, that will be ignored during that one epoch. This is intended to prevent the network hyperfocussing on individual details. The learning rate determines the magnitude of the changes applied every epoch and the momentum modifies those changes based on the changes from the previous epochs, which speeds up the training and prevents the network to get stuck in local minima. The values of these hyperparameters are given in table D.1.

The optimiser used was Adam [47].

Table D.1: Hyperparameters of the NN trained by Richard Baumann. These were optimised using a grid search.

| Parameter | Value |
|-----------------|-------|
| Epochs | 1000 |
| Hidden Layers | 3 |
| Nodes per Layer | 180 |
| Dropout | 0.25 |
| Learning Rate | 0.002 |
| Momentum | 0.02 |

Two important plots to validate the quality of the network are the loss function and the receive operating characteristic (ROC) curve. As described in section 6.4.1 the loss is the measure of the difference between the actual and ideal output of the network. As long as the loss is decreasing, the network is learning information about the training data provided. Whenever one trains a network there is always a chance of it learning incidental details about the training set that cannot be generalised (“Overtraining”). To be able to check for overtraining and thereby avoid it the NN is usually not trained on all simulated events available. Instead they are split into a training and a testing set. The network is trained on the former and then also evaluated on the latter. When the network is overtrained, the training

loss keeps decreasing, while the test loss stays stagnant or even increases again. As can be seen in figure D.1, this is not the case here.

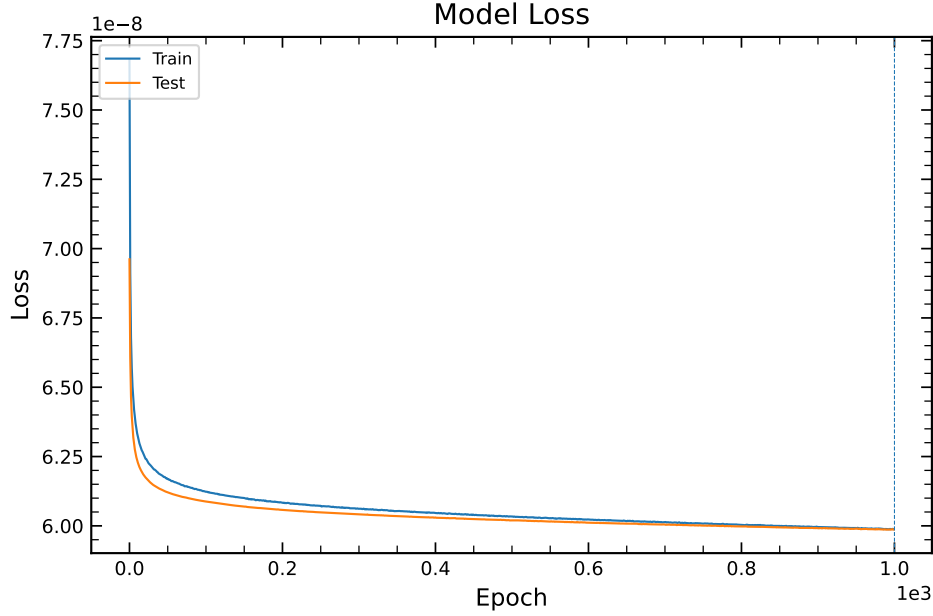


Figure D.1: Development of the loss of the neural network over the epochs.

The ROC-curve plots the true positive rate, i.e. the fraction of signal events correctly identified as such against the false positive rate, i.e. the fraction of background events falsely identified as signal for different cut off points for NN_{out} . The curve will always contain the points (0,0) and (1,1), $NN_{\text{out}} > 1$ will be true for no event and $NN_{\text{out}} > 0$ will be true for all. Figure D.2 shows the ROC-curve for the training and testing samples. It also marks the angle bisector as a dotted red line. This is the ROC-curve one would expect when simply sorting the events randomly. To evaluate the quality of a ROC-curve one often uses the area under the curve (AUC). A random assignment has an AUC of 0.5, a hypothetical network which correctly assigns all background events the value 0 and all signal events the value 1 would have an AUC of 1.0. The AUC of the used network has a value of 0.84. This is identical for the training and testing samples. A large difference in AUC or the shape of the ROC-curves between these two would be another sign of overtraining.

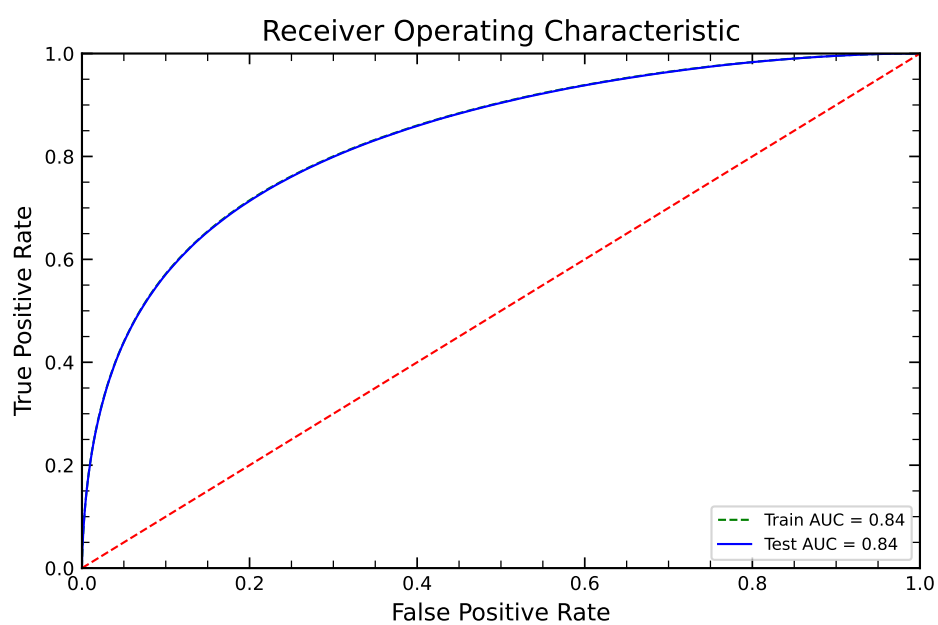


Figure D.2: Receiver Operating Characteristic for both training and testing samples. The dotted red line shows the ROC-curve one would expect, when sorting the events randomly.

XRootD for TopLoop derived Code on the ATLAS Grid

An integral tool to this thesis was XRootD. It enables the user to access files that are saved on the ATLAS computing grid without copying those files to the local computing cluster. Brief instructions on how to use XRootD for this purpose are displayed in this appendix. They have been written as a step by step explanation for someone familiar with TopLoop derived code. “TopLoop provides an [Application Programming Interface] for processing single top ntuples, dealing with metadata describing the samples, and constructing simple physical objects to simplify common calculations” [48]. Parts of these instructions also specifically reference the examples tHqLoop [49] and tchLoop [50], both of which are TopLoop derived projects, and the Bonn Analysis Facility (BAF) cluster. This is done without much introduction since the main target audience for these instructions are the fellow members of the Brock ATLAS working group at the university of Bonn. Therefore the tone is also somewhat more conversational and less formal than the rest of this thesis.

These instructions require a valid grid certificate.

E.1 Instructions

1. Setup rucio (setupATLAS, lsetup rucio, voms-proxy-init -voms atlas).
2. Create a directory `samples` in the run directory and cd into it.
3. Copy the `../condor/samples_AP/Remotfile_rucio.py` into your current directory.
4. Make a file called `datasamples.txt` and list the required samples in it (e.g.
`user.cescobar.mc16_13TeV.412063.aMCPy8EG_t11q_nf4.SGTOP1.e7054`
`_s3126_r9364_p4346.11.v34,`
 for Single Top Ntuples you can look these up at
<https://twiki.cern.ch/twiki/bin/view/AtlasProtected/SgTopRun2NtuplesSamples>)
 - Note however, that this page often takes a few days to update after any changes.
 - If you want to make sure, the information is correct, you can run the command `rucio list-file-replicas user.cescobar.mc16_13TeV.412063.aMCPy8EG_t11q_nf4`

- `.SGTOP1.e7054_s3126_r9364_p4346.ll.v34`, which will produce a list of all the replicas of all the files.
 - If you want to check, that the samples are at a certain grid site, you can add the `-rse`-flag, e.g. `-rse UNI-BONN_LOCALGROUPDISK`.
 - This is also the same command the `Remotefile_rucio.py`-script uses.
5. Check in the drop down menu on the twiki for a sample where the data set is replicated (usually always the same for files run by the same user) e.g. for the file above the RSE is `IFIC-LCG2_LOCALGROUPDISK` and update the line in the `Remotefile_rucio.py` where it has the `-rse` with the one needed.
 - I would recommend to group the containers by RSE, that way you only have to run the script once per grid site.
 - If you often need to access samples from several sites, you might want to slightly modify the script so that it takes in the RSE as a parameter.
 6. Run `python Remotefile_rucio.py datasamples.txt` which will create a directory called `inTxt` which will contain one file per data sample with the xrootd paths to the files related to it.
 7. Start a new session, setup your TopLoop derived algorithm like normal (`setupATLAS`, `asetup AnalysisBase` or `-restore`, etc.).
 8. Setup rucio using `lsetup "emi -w"` (and `voms-proxy-init -voms atlas`), rather than `lsetup rucio`.
 9. `cd` to the `run` directory and execute you TopLoop derived algorithm like normal, but instead of providing a directory with root files using the `-d` command, use the `-T` or `-x` (depending on your TopLoop derived algorithm) option with the path to `samples/inTxt/<name_of_sample>.txt`
 10. To run this over several files it is probably easier to create a script to run over all the files on the batch system with one per job.

E.2 HTCondor Batch Jobs

In order to create batch jobs of this algorithm on BAF, you will need a valid rucio proxy (this assumes that you do already have a grid certificate).

1. Run `setupATLAS -c centos7` on the desktop machine (not in an interactive job).
2. Then execute `lsetup "emi -w"` and `voms-proxy-init -voms atlas` (since we are not doing anything else here, you can also use `lsetup rucio` instead of `lsetup "emi -w"`, does not make a difference in this case).
 - You may want to use `voms-proxy-init -valid 96:0 -voms atlas` instead to increase the validity of the proxy from the default 24 to 96 hours. You cannot create a proxy that is valid longer than that. If you try higher values, you will still only get a proxy that is valid for 96 hours.

- If you need to run a job longer than that, you need to fetch a new proxy. Just repeat these steps, while the job is running. HTCondor will pick it up and ship it into the compute jobs (it checks for updates of the proxy file every 10 minutes by default).
- 3. This will create a file called something like `x509up_u210176` in the `/tmp` directory. Copy that file to a folder of your choice (easiest is the one, that also contains your `.jdl` file and wrapper).
- 4. Add a line to the `.jdl` file telling HTCondor, where to find the proxy:
`x509userproxy = PathToCopiedFile/x509up_u210176.`
Note that HTCondor does not know, how to expand the `~` character. You can either use a relative path, a fully absolute path, or `x509userproxy = $ENV(HOME)/PathToCopiedFile/x509up_u210176.`
- The easiest thing should be to just copy the file to the directory in which you also have and call the `.jdl` file and accordingly use `x509userproxy = ./x509up_u210176.`

E.3 Notes, Comments, Tips and Hints

- Do not use `lsetup rucio` to setup rucio the second time, but `lsetup "emi -w" ! lsetup rucio` adds some libraries that might be incompatible with AnalysisBase and should in general only be used on its own. (Note that \LaTeX does not necessarily produce the right quotation marks in the pdf, so if you copy paste this command, you might have to replace them by hand.) However `lsetup "emi -w"` does not allow for usage of the `Remotefile_rucio.py` which is why creating a new session to run the code after creating the `.txt` files is necessary. I suppose, this is somewhat jury-rigged, but at time of writing I have not found a more elegant way to do this.
- If you do not have the `Remotefile_rucio.py` in your TopLoop derived algorithm, you can copy it from another user. The script is only used to create the `.txt`-files so as far as I can tell it should be compatible with any TopLoop derived code. If you are in Bonn, you can for example copy it from my `tchLoop`:
`/cephfs/user/s6sithie/work/area/run/samples/Remotefile_rucio.py.`
- The `rucio list-file-replicas-command` in `Remotefile_rucio.py` might have the flag `-protocol root`. However some of the RSEs do not support the root, but only the `davs` protocol. You can get the script to work on those as well, by changing the flag to `-protocol root,davs` (One should also be able to just cut the flag entirely, but I did not try that). However you then need to assure, that your code does not rely on the root protocol. As long as you use `TFile::Open("root://somewhere/somefile.root");` or `TFile::Open("davs://somewhere/somefile.root");` that should be fine.
 - Some grid sites support both protocols. I have not tried what happens yet in that case, but I think, there is a possibility, that the files get added to the `.txt`-file twice and you end up double counting in you code. If you end up using that, make sure that does not happen. If it does, just cut one of the flags when accessing those RSEs.
- The option to supply the the path to the `.txt`-files might be called `-T` or `-x` (or anything else for all I know, but these are the two I have encountered). Check the `run-algorithm` (e.g. `runHqLoop.cxx` or `runTchLoop.cxx`) for a command like

```
app.add_option("-T,-inTxt", inTxt, "Input SgTop ntuple dataset txt")
->check(CLI::ExistingFile); or
auto txtOpt = app.add_option("-x,-in-txt", inTxt, "Text file listing files
to process")
->check(CLI::ExistingFile); .
```

The examples are taken from runHqLoop.cxx and runtchLoop.cxx respectively.

Bibliography

- [1] wikipedia users MissMJ and Cush, 2019, URL: https://en.wikipedia.org/wiki/File:Standard_Model_of_Elementary_Particles.svg (visited on 13/10/2022) (cit. on p. 3).
- [2] V. I. Yukalov, *Theory of cold atoms: basics of quantum statistics*, *Laser Physics* **23** (2013) 062001, URL: <https://doi.org/10.1088%2F1054-660x%2F23%2F6%2F062001> (visited on 19/11/2022) (cit. on pp. 2, 4).
- [3] J.-M. Richard, *Exotic Hadrons: Review and Perspectives*, *Few-Body Systems* **57** (2016) 1185, URL: <https://doi.org/10.1007%2Fs00601-016-1159-0> (visited on 19/11/2022) (cit. on p. 4).
- [4] R. L. Workman et al., *Review of Particle Physics*, *PTEP* **2022** (2022) 083C01, URL: https://pdg.lbl.gov/2022/reviews/contents_sports.html (visited on 19/11/2022) (cit. on pp. 4, 6, 7).
- [5] V. Sahni, “5 Dark Matter and Dark Energy”, *The Physics of the Early Universe*, Springer Berlin Heidelberg, 2004 141, URL: https://doi.org/10.1007%2F978-3-540-31535-3_5 (visited on 19/11/2022) (cit. on p. 6).
- [6] C. Weinheimer and K. Zuber, *Neutrino masses*, *Annalen der Physik* **525** (2013) 565, URL: <https://doi.org/10.1002%2Fandp.201300063> (visited on 19/11/2022) (cit. on p. 6).
- [7] T. ATLAS, CDF, CMS and D. Collaborations, *First combination of Tevatron and LHC measurements of the top-quark mass*, 2014, URL: <https://arxiv.org/abs/1403.4427> (visited on 19/11/2022) (cit. on p. 6).
- [8] L. Evans and P. Bryant, *LHC Machine*, 2008, URL: <https://iopscience.iop.org/article/10.1088/1748-0221/3/08/S08001/pdf> (visited on 30/09/2022) (cit. on p. 8).
- [9] CERN, *CERN’s accelerator complex*, 2022, URL: <https://home.cern/science/accelerators/accelerator-complex> (visited on 11/10/2022) (cit. on p. 8).
- [10] E. Mobson behalf of CERN, *The CERN accelerator complex in 2019*, 2019, URL: <https://cds.cern.ch/record/2684277> (visited on 20/11/2022) (cit. on p. 9).

- [11] ATLAS Collaboration, *ATLAS Note: Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC*, 2019,
URL: <https://cdsweb.cern.ch/record/2677054> (visited on 19/11/2022)
(cit. on pp. 11, 22, 45).
- [12] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008) S08003 (cit. on p. 12).
- [13] ATLAS Collaboration,
ATLAS Detector and Physics Performance: Technical Design Report, Volume 1,
ATLAS-TDR-14; CERN-LHCC-99-014, 1999,
URL: <https://cds.cern.ch/record/391176> (visited on 27/09/2022) (cit. on p. 12).
- [14] ATLAS Collaboration,
ATLAS Detector and Physics Performance: Technical Design Report, Volume 2,
ATLAS-TDR-15; CERN-LHCC-99-015, 1999,
URL: <https://cds.cern.ch/record/391177> (visited on 27/09/2022) (cit. on p. 12).
- [15] J. Pequeno on behalf of the ATLAS Collaboration,
Computer generated image of the whole ATLAS detector, 2008,
URL: <https://cds.cern.ch/record/1095924> (visited on 15/11/2022) (cit. on p. 12).
- [16] S.-P. Hallsjö, *Search for Dark Matter in the Upgraded High Luminosity LHC at CERN: Sensitivity of ATLAS phase II upgrade to dark matter production*, 2014 (cit. on p. 12).
- [17] D. Makowski and A. Napieralski, *The Impact on Electronic Devices with the Special Consideration of Neutron and Gamma Radiation Monitoring*,
Zesz. Nauk. Elektr./Politech. Łódz. (2007) (cit. on p. 14).
- [18] wikipedia user Qwerty123uiop, 2013, URL: <https://commons.wikimedia.org/wiki/File:PhotoMultiplierTubeAndScintillator.svg> (visited on 27/10/2022) (cit. on p. 15).
- [19] J. Pequeno, P. Schaffner on behalf of the ATLAS Collaboration,
How ATLAS detects particles: diagram of particle paths in the detector, 2013,
URL: <https://cds.cern.ch/record/1505342> (visited on 13/10/2022) (cit. on p. 18).
- [20] ATLAS Collaboration, *ATLAS Calorimeter Performance: Technical Design Report*,
ATLAS-TDR-1; CERN-LHCC-96-040, 1996,
URL: <https://cds.cern.ch/record/331059> (visited on 18/11/2022) (cit. on p. 18).
- [21] L. Pontecorvo on behalf of the ATLAS Collaboration, *The ATLAS Muon Spectrometer*, 2003,
URL: <https://cds.cern.ch/record/676896/files/sn-atlas-2003-030.pdf> (visited on 18/11/2022) (cit. on p. 19).
- [22] G. P. Salam, *Towards Jetography*, (2010),
URL: <https://arxiv.org/pdf/0906.1833.pdf> (visited on 29/09/2022) (cit. on p. 20).
- [23] ATLAS Collaboration,
Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1,
Eur. Phys. J. C **77** (2017) 490, arXiv: 1603.02934 [hep-ex] (cit. on p. 20).
- [24] ATLAS Collaboration,
Jet reconstruction and performance using particle flow with the ATLAS Detector,
Eur. Phys. J. C **77** (2017) 466, arXiv: 1703.10485 [hep-ex] (cit. on p. 21).

- [25] Y. Coadou, *Boosted Decision Trees and Applications*, *EPJ Web of Conferences* **55** (2013) 02004, URL: <https://doi.org/10.1051/epjconf/20135502004> (visited on 20/11/2022) (cit. on p. 21).
- [26] ATLAS Collaboration, *Measurements of b -jet tagging efficiency with the ATLAS detector using $t\bar{t}$ events at $\sqrt{s} = 13$ TeV*, *JHEP* **08** (2018) 089, arXiv: 1805.01845 [hep-ex] (cit. on p. 21).
- [27] 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** (2019) 639, arXiv: 1902.04655 [hep-ex] (cit. on p. 21).
- [28] 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** (2021) 578, arXiv: 2012.00578 [hep-ex] (cit. on p. 21).
- [29] F. He, K. Liu, R. Ospanov and R. Roberts, *ATLAS Note: Tagging non-prompt electrons and muons using lifetime and isolation information*, 2018, URL: <https://cds.cern.ch/record/2632152/files/ATL-COM-PHYS-2018-1083.pdf> (visited on 17/11/2022) (cit. on p. 21).
- [30] ATLAS Collaboration, *Single-top ntuple production (v34) for Run-2 samples (single-lepton channel with taus)*, *TWiki*, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/SingleTopProductionV34lj> (visited on 19/11/2022) (cit. on pp. 23, 46).
- [31] P. Nason, *A New method for combining NLO QCD with shower Monte Carlo algorithms*, *JHEP* **11** (2004) 040, arXiv: hep-ph/0409146 (cit. on p. 23).
- [32] T. Sjöstrand, S. Mrenna and P. Skands, *A brief introduction to PYTHIA 8.1*, *Computer Physics Communications* **178** (2008) 852, URL: <https://doi.org/10.1016%2Fj.cpc.2008.01.036> (visited on 19/11/2022) (cit. on p. 23).
- [33] E. Bothmann et al., *Event generation with Sherpa 2.2*, *SciPost Physics* **7** (2019), URL: <https://doi.org/10.21468%2Fscipostphys.7.3.034> (cit. on p. 23).
- [34] 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*, *Journal of High Energy Physics* **2014** (2014), URL: <https://doi.org/10.1007%2Fjhep07%282014%29079> (cit. on p. 23).
- [35] D. P. Kroese, T. Brereton, T. Taimre and Z. I. Botev, *Why the Monte Carlo method is so important today*, *WIREs Computational Statistics* **6** (2014) 386, eprint: <https://wires.onlinelibrary.wiley.com/doi/pdf/10.1002/wics.1314>, URL: <https://wires.onlinelibrary.wiley.com/doi/abs/10.1002/wics.1314> (visited on 20/11/2022) (cit. on p. 23).
- [36] A. Buckley et al., *Monte Carlo event generators for high energy particle physics event simulation*, 2019, URL: <https://arxiv.org/abs/1902.01674> (visited on 20/11/2022) (cit. on p. 23).

- [37] ATLAS Collaboration, *Using Xrootd Service (AAA) or Remote Data Access*, 2014, URL: <https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookXrootdService> (visited on 18/11/2022) (cit. on p. 23).
- [38] O. Freyermuth, *User Documentation, BAF, XRootD*, 2020, URL: <https://confluence.team.uni-bonn.de/display/PHYIT/XRootD> (visited on 18/11/2022) (cit. on p. 23).
- [39] F. Cardillo et al., *ATLAS Note: Tools for estimating fake/non-prompt lepton backgrounds with the ATLAS detector at the LHC*, 2022, URL: <https://cds.cern.ch/record/2686919/> (visited on 04/10/2022) (cit. on pp. 24, 25, 28).
- [40] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, *JINST* **14** (2019) P12006, arXiv: 1908.00005 [hep-ex] (cit. on p. 25).
- [41] F. Derue, *ATLAS Note: Estimation of fake lepton background for top analyses using the Matrix Method with the 2015 dataset at $\sqrt{s} = 13$ TeV with AnalysisTop-2.3.41*, 2016, URL: <https://cds.cern.ch/record/2135116/> (visited on 04/10/2022) (cit. on pp. 27–29).
- [42] W. Wagner, J. A. Reidelsturz, M. R. Estabragh, D. Hirschbuehl and O. B. Bylund, *ATLAS Note: Measurement of the total single-top-quark t -channel production cross section at 13 TeV using the full Run 2 dataset*, (2022), URL: <https://cds.cern.ch/record/2677346> (visited on 18/11/2022) (cit. on p. 32).
- [43] N. Bruscinio et al. on behalf of the ATLAS Collaboration, *atlasHTop GitLab*, URL: <https://gitlab.cern.ch/atlasHTop/thbbskimmer/-/tree/master/skimNtuple/data/matrixmethod> (visited on 21/11/2022) (cit. on p. 32).
- [44] I. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*, <http://www.deeplearningbook.org>, MIT Press, 2016 (cit. on p. 34).
- [45] ATLAS Collaboration, *Measurement of differential cross-sections of a single top quark produced in association with a W boson at $\sqrt{s} = 13$ TeV with ATLAS*, *Eur. Phys. J. C* **78** (2018) 186, arXiv: 1712.01602 [hep-ex] (cit. on pp. 38, 45).
- [46] ATLAS Collaboration, *Single-top Run2 samples*, *TWiki*, URL: <https://twiki.cern.ch/twiki/bin/view/AtlasProtected/SgTopRun2NtuplesSamples> (visited on 27/09/2022) (cit. on p. 46).
- [47] I. K. M. Jais, A. R. Ismail and S. Q. Nisa, *Adam Optimization Algorithm for Wide and Deep Neural Network*, *Knowledge Engineering and Data Science* **2** (2019) 41, ISSN: 2597-4637, URL: <http://journal2.um.ac.id/index.php/keds/article/view/6775> (visited on 19/11/2022) (cit. on p. 50).
- [48] D. Davis on behalf of the ATLAS Collaboration, *TopLoop Documentation*, URL: <https://atlas-toploop.web.cern.ch/atlas-toploop/docs/> (visited on 21/11/2022) (cit. on p. 53).

- [49] *tHqLoop GitLab*, URL: https://gitlab.cern.ch/atlasphys-top/singletop/thq_13tev_rel21/thloop (visited on 21/11/2022) (cit. on p. 53).
- [50] *tchLoop GitLab*, URL: <https://gitlab.cern.ch/sithiele/tchLoop/-/tree/tchLoop> (visited on 21/11/2022) (cit. on p. 53).

List of Figures

| | | |
|-----|---|----|
| 2.1 | Overview of the particles in the Standard Model with their charges, spins, and masses listed [1]. | 3 |
| 2.2 | Feynman diagrams of some leading order $t\bar{t}$ production processes. | 6 |
| 2.3 | Feynman diagrams of the leading order single top production processes. | 7 |
| 2.4 | Decay of a top into a bottom and W and the two decay modes of the W | 7 |
| 3.1 | Overview of the full accelerator complex at CERN [10]. | 9 |
| 4.1 | Computer generated image of the ATLAS detector with the coordinate system overlayed [15], altered by Sven-Patrik Hallsjö in [16]. | 12 |
| 4.2 | Signal strength of a gaseous detector depending on the applied voltage [17]. | 14 |
| 4.3 | Schematic overview of a setup using a scintillator and a PMT to detect high energy photons [18]. | 15 |
| 4.4 | Schematic cross section through the layers of the ATLAS detector and the characteristic detector responses caused by the different particles [19]. | 18 |
| 5.1 | Distribution of the invariant mass of loose electron and muon pairs [41]. | 27 |
| 5.2 | Distribution of E_T^{miss} and $m_T^{\ell, E_T^{\text{miss}}}$ for data and the different MC contributions at loose and tight levels for e +jets events with one electron and at least one jet [41]. | 29 |
| 6.1 | Leading order Feynman diagrams for tq and $\bar{t}q$ | 31 |
| 6.2 | E_T^{miss} - and p_T^ℓ -distributions of the signal regions. | 33 |
| 6.3 | Distribution of the neural network output NN_{out} for the signal regions. | 35 |
| 6.4 | NN_{out} -distributions after performance of the statistics only fit. | 37 |
| 6.5 | Likelihood scans of the statistics only fit for the signal regions. | 38 |
| 6.6 | NN_{out} -distributions after performance of the fit using only the systematic uncertainties for the fake estimation and the cross sections of the main backgrounds. | 39 |
| 6.7 | Difference between the nominal values for the four varied backgrounds and the fit results, normalised to the uncertainty. | 39 |
| 6.8 | Likelihood scans of the cross section uncertainty only fit for the signal regions. | 40 |
| A.1 | Likelihood scans of the full systematics fit for the signal regions. | 44 |
| C.1 | E_T^{miss} - and p_T^ℓ -distributions and of the $t\bar{t}$ control regions. | 48 |
| C.2 | E_T^{miss} - and p_T^ℓ -distributions and of the W +jets control regions. | 49 |

List of Figures

| | | |
|-----|--|----|
| D.1 | Development of the loss of the neural network over the epochs. | 51 |
| D.2 | Receiver Operating Characteristic for both training and testing samples. | 52 |

List of Tables

| | | |
|-----|---|----|
| 3.1 | Technical parameters defining the luminosity of the LHC [11]. | 11 |
| 4.1 | Overview over the working points used to define “loose” and “tight” leptons in the Single Top Ntuples. | 22 |
| 5.1 | Overview of processes contributing to the fake/non-prompt background and their effects on electrons and muons. | 25 |
| 6.1 | Selection for the signal and control regions. | 32 |
| 6.2 | Uncertainties of the cross sections of the three main backgrounds and the uncertainty applied to the fake background. | 38 |
| B.1 | All MC Samples used for the simulated prediction with the label they have in the plots in section 6. | 46 |
| D.1 | Hyperparameters of the NN trained by Richard Baumann. These were optimised using a grid search. | 50 |

Acknowledgements

I would like to thank Prof. Dr. Ian C. Brock for the opportunity to work in his group and the guidance throughout this year. Dr. Tatjana Lenz for reviewing this thesis as the second referee. All the other members of the Brock group, especially my supervisor Federico “Rico” Guillermo Diaz Capriles, for all the help, support, and incredibly fun times. My family and friends for always having my back, especially my parents and my girlfriend. Finally Oliver Freyermuth from the IT support of the physical institute for the patience and many, many quick replies.