Bachelor's Thesis

Studien zu Top-Quark und Photon Kopplung mit dem ATLAS-Experiment

Studies of top quark and photon coupling with the ATLAS experiment

prepared by

Joey Kalis

from Mülheim a. d. Ruhr

at the II. Physikalischen Institut

Zusammenfassung

In der Teilchen- und der Hochenergiephysik finden Neuronale Netze und andere Machine Learning Algorithmen stetig wachsende Anwendung. Hauptsächlich zur Diskriminierung zwischen Signal und Background werden eben jene als binäre Klassifizierungsmethoden eingesetzt.

Die vorliegende Thesis handelt von der Anwendung eines Neuronalen Netzwerks zur Unterscheidung zwischen prompten Photonen, die von Top-Quark Paaren ausgesandt werden, und hadronischen "fake" Photonen, die einen dominanten Beitrag zu der Untergrundverteilung ausmachen, um die elektromagnetische Kopplung des Top-Quarks weitergehend für die Schwerpunktsenergie $\sqrt{s} = 13\,\text{TeV}$ am CERN zu untersuchen. Für diese Untersuchung ist insbesondere der Aufbau des ATLAS Detektors von tragender Bedeutung.

So genannte Shower Shapes werden in das Neuronale Netz implementiert, um bestmöglich systematische Unsicherheiten zu eliminieren. Wie sich herausstellen wird, sind diese bestens geeignet, um die gewünschte Separierung zwischen prompten Photonen und hadronischen Fake Photonen hervorzubringen.

Abstract

Neural Networks and other machine learning algorithms enjoy an ever-expanding usage in particle and high energy physics. Primarily, they are used as binary classifiers to discriminate between signal and background contributions.

The thesis deals with the usage of a Neural Network to discriminate between prompt photons, which will be radiated off top quark pairs, and the contribution of hadron fakes being a dominant background contribution to get sensitive to the $t\gamma$ coupling for a centre-of-mass energy $\sqrt{s} = 13\,\text{TeV}$ at CERN. The structure of the ATLAS detector is important for this study.

So-called shower shapes will be implemented in the used Neural Network to eliminate systematic uncertainties in a best possible way. As it turns out, they will be used to get the desired separation between prompt photons and hadron fakes.

Contents

Contents

1. Introduction

"If it disagrees with experiment, it is wrong. In that simple statement is the key to science. It does not make any difference how beautiful your guess is, it does not matter how smart you are who made the guess, or what his name is... If it disagrees with experiment, it is wrong. That is all there is to it." Richard Feynman said these words about scientific methods in one of his famous lectures in the 1960s. Particle physicists study the nature of the particles that constitute matter and radiation. For this purpose, particle colliders like the LHC and important detectors like ATLAS were built. The Standard Model is one of these 'beautiful guesses', which is not in disagreement with the experiments until today. Scientific theories and hypothesis base on the principle of falsificationism. A famous example is the statement: "all swans are white", which is falsifiable since there is another statement that contains observations of black swans anywhere in the world and contradicts it [1]. Such a statement can not be proven, but it was considered right until it was proven wrong. Scientific theories are treated the same way, so scientists appreciate a theory like the Standard Model until it would be proven wrong.

The Standard Model is one of the most successful theories of contemporary physics, but it is certainly not perfect. Down to the present day, the Standard Model predicts one observation after another right, which lead to its enormous triumph nowadays. But there are numerous effects that were not experimentally observed (limitations) and this is why scientists try to observe further particles or effects today at CERN.

One of these effects is the electromagnetic coupling of the heaviest known particle in the Standard Model: the top quark. This thesis deals with the improvement of techniques like a Neural Network to get sensitive to this kind of top quark coupling. Mainly, the discrimination between prompt photons and hadron fakes, which are hadrons misidentified as photons and photons originating from hadronic decays, will be discussed in the analysis. Version 21.0 of the Neural Network will be implemented for analysing the top quark pair production in association with a photon at centre-of-mass energy $\sqrt{s} = 13\,\text{TeV}$.

The following chapter deals with the theoretical background, which includes a brief introduction to the Standard Model, its limitations and a section about top quark physics. The latter section contains besides its properties and interactions previous results at

$\sqrt{s} = 7\,\text{TeV}$ and $\sqrt{s} = 8\,\text{TeV}$. Besides the physical foundations, the experimental ones, so the setup of the Large Hadron Collider and the ATLAS detector, will be presented afterwards. Chapter 4 is about the photon measurements at ATLAS, which includes the difference between prompt photons and hadron fakes, photon reconstruction and identification. Then, Monte Carlo truth studies are presented to analyse the behaviour of the photon candidates for different cuts and types. After giving a short introduction to machine learning in Chapter 6, the used Neural Network, the Prompt Photon Tagger, will be introduced and in Section 6.1 finally in the $t\bar{t}\gamma$ analysis implemented. Chapter 7 presents a short summary of all results and achievements.

2. Theoretical Background

This chapter deals with the theoretical background, which is necessary for the analysis of the $t\bar{t}\gamma$ process the thesis deals with.

For this purpose, the Standard Model of elementary particle physics with its properties and limitations will be introduced and, after that, features and fundamental interactions of the top quark and previous $t\bar{t}\gamma$ analyses will be presented.

2.1. The Standard Model of elementary particle physics

In the following lines, a brief description of the Standard Model (SM) of elementary particle physics and a short summary about the most recent discoveries at the Large Hadron Collider (LHC) will be given. After that the limitations of the SM will be discussed.

2.1.1. Brief description of the SM

The Standard Model [2] describes the fundamental particles and their interactions. Figure 2.1 shows a sketch of the fundamental particles in the SM.

There are 12 fermions (and also 12 antifermions) with spin-$\frac{1}{2}$, six quarks and six leptons, grouped into three generations. The mass of these fermions increases from the first to the third one. Quarks and Leptons are even more divided into up-type and down-type fermions. Up-type quarks (up, charm, top) all have an electric charge[1] of $Q = +\frac{2}{3}$ (which has to be measured for the top quark), whereas the down-type (down, strange, bottom) quarks have $Q = -\frac{1}{3}$. The charged leptons (e^-, μ^-, τ^-) have an electric charge of $Q = -1$, and their corresponding neutrino (ν_e, ν_μ, ν_τ) is uncharged.

Thus, the difference in units of electric charge Q between down-type and up-type fermions is always 1.

Another difference between the up-type and down-type fermions is the property of the third component of the weak isospin I_3. For up-types it is $I_3 = +1/2$ and for down-types

---

[1]in elementary charge $e = 1.602 \times 10^{-19}$ C.
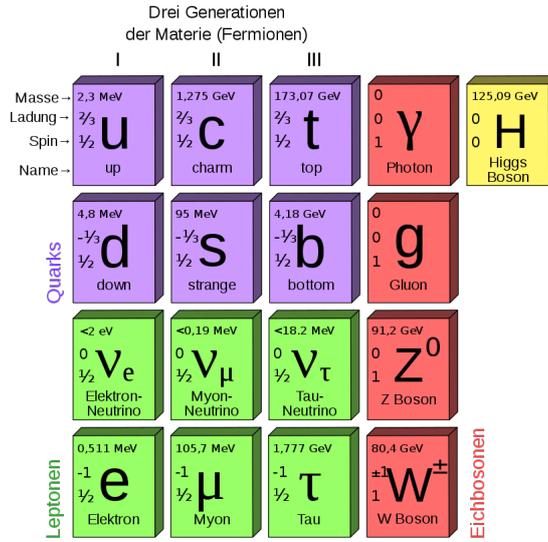
Figure 2.1.: Schematic depiction of the Standard Model of elementary particles with the three generations of matter in the first three columns, the gauge bosons in the fourth and the Higgs boson in the fifth column.

$I_3 = -1/2$.

The SM describes three out of four known fundamental interactions: weak, strong and electromagnetic interaction. For each of these interactions there are force carriers (known as gauge bosons) with spin 1.

The bosons of the weak interaction are the $W^\pm$ (charged weak interaction) and the Z^0 (neutral weak interaction) [3]. Whereas the Z boson couples to both left- and right-handed fermions, only left-handed fermions (or right-handed anti-fermions) can couple with the $W^\pm$ because of the W-Vertex factor of the charged electric current [4]

$$-i\frac{g_W}{2\sqrt{2}}\gamma^\mu \underbrace{(1-\gamma^5)}_{=2P_L}, \tag{2.1}$$

where P_L is the left-handed projection operator. Neutrinos just interact weakly and they appear only as left-handed fermions.

The photon γ is the gauge boson of the electromagnetic interaction. All charged fermions and bosons can couple with it.

Quarks are the participating particles in the strong interaction. They carry besides of electric and weak charge a colour charge (red, green, blue). This is the charge of the Quantum Chromodynamics (QCD), describing the strong interaction. In nature there are only colourless particles, so quarks only exist in bound colour-neutral states. They build colourless hadrons in the process of hadronisation. Because of the underlying group

theory of QCD, there are eight gluons, which carry one colour and one anticolour. So they can couple to themselves via triple or quartic gauge couplings.

With the discovery of the Higgs boson [5, 6] with spin-0 the SM was completed. In 2012 ATLAS and CMS discovered the Higgs boson in diphoton-channel. The Higgs boson couples only to massive particles. So it couples to the W and Z bosons, but not to gluons and photons.

Mathematically, the SM is described as a local gauge and Lorentz invariant Lagrange density function with the underlying gauge group $U(1)_Y \times SU(2)_L \times SU(3)_C$, which is fully renormalisable [7].

2.1.2. Limitations of the SM

As it was already mentioned, the SM only describes three out of the four fundamental interactions, excluding gravity. Besides the missing hypothetical spin-2 graviton, the SM is widely considered to be incompatible with general relativity.

Another problem is the dark matter, which is a common concept for explaining the Galaxy rotation curve [8]. A particle with the properties, that are needed to describe this kind of matter, is still missing today.

Noteworthy is also the search for heavy resonances (W', Z') or supersymmetric particles [9–11]. If their existence would be confirmed, theories beyond the SM could be examined more.

Lastly, the neutrinos are supposed to be massless, because they do not interact with the Higgs boson, which is inconsistent with neutrino flavour oscillations [12].

As it is shown, there are some problems the SM can not explain, which means that the search for these (until now) hypothetical particles is really important. This is one of the aims of the LHC: to find physics beyond the SM.

2.2. Top quark physics

The top quark [13–15] is the heaviest known fermion in the SM (see Section 2.1.1). After discovering the bottom quark [16] in 1977 at FERMILAB, it was suggested that there will be also a third up-type quark, completing the third generation. In the end, it took over 18 years to discover the top quark. Finally it was observed in 1995 at the TEVATRON by the DØ [17] and CDF [18] experiments.

2.2.1. Properties

As it is shown in Figure 2.1 the top quark has an electric charge of $Q = +\frac{2}{3}$ and a spin of 1/2.

One of the top quark's peculiarities is its enormous mass. With a value of $m_t = 172.44 \pm 0.13(\text{stat}) \pm 0.47(\text{syst})\text{GeV}$ [19] the top quark is the heaviest known particle, which made the search for it challenging.

Another property is the really short lifetime. Theoretically, the width of the top quark is predicted to $\Gamma_t \approx 1.33\,\text{GeV}$ [20], which corresponds to a lifetime of $\tau_t \approx 5 \times 10^{-25}\,\text{s}$. The time scale of hadronisation [21] processes is two orders of magnitude larger, which means that the top quark decays before forming a hadronic bound state. Thus, further studies onto the decay products and the reconstruction of its properties are important and allow accessing the quantities of this quark.

2.2.2. Interactions of the top quark

According to the SM (see Section 2.1.1), the top quark interacts weakly, strongly and electromagnetically. These interactions are described individually in the following paragraphs.

Weak interaction of the top quark

In the weak interaction, there is also single top quark production possible [22]. A top quark decays mainly via $t \to W^+ b$. Furthermore, the W^+ boson decays into a quark pair $q\bar{q}'$ ($\approx 67\%$) or leptonically $\ell\nu_\ell$ ($\approx 33\%$).

In the case of $t\bar{t}$ production, there are three final state topologies (with decreasing probability): full-hadronic (alljets), single lepton (ℓ + jets) and dilepton (2 leptons + jets). The branching ratios of these final state topologies are depicted in Figure 2.2.

Strong interaction of the top quark

Due to flavour conservation in the strong interaction, there is only $t\bar{t}$ pair production possible. Different from the production at the TEVATRON [23], where $q\bar{q}$ annihilation is the dominant mechanism of $t\bar{t}$ pair production, gluon fusion is the main production at the LHC ($\approx 90\%$). The reason for this is that $q\bar{q}$ annihilation of valence quarks is not possible because of the usage of a *pp* collider and due to parton distribution functions (PDF) [24], but gluon fusion is available.

Top Pair Branching Fractions
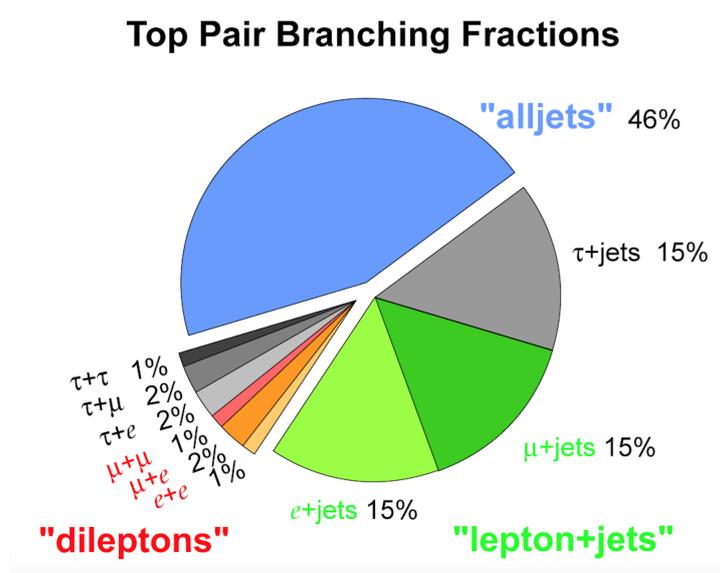
Figure 2.2.: Final state topologies for $t\bar{t}$ production and their branching ratios.

The top quark in the electromagnetic interaction

Top quarks can also couple to a photon (electromagnetic) via $pp \to t\bar{t}\gamma$. In the SM this coupling is expected to be a vector coupling. Examples of the electromagnetic interaction of the top quark are shown in the case of $t\bar{t}$ production because of $q\bar{q}$ annihilation in Figure 2.3 and because of gluon fusion in Figure 2.4.
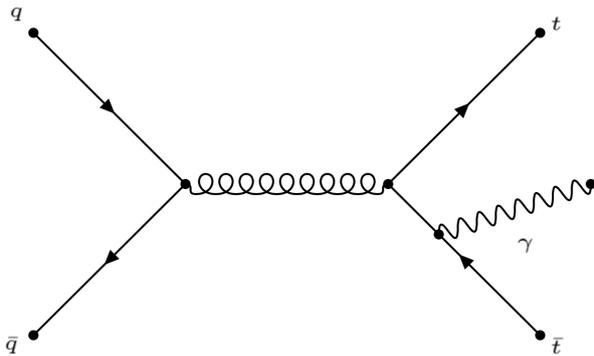
Figure 2.3.: Possible photon radiation of $t\bar{t}$ system for quark annihilation.
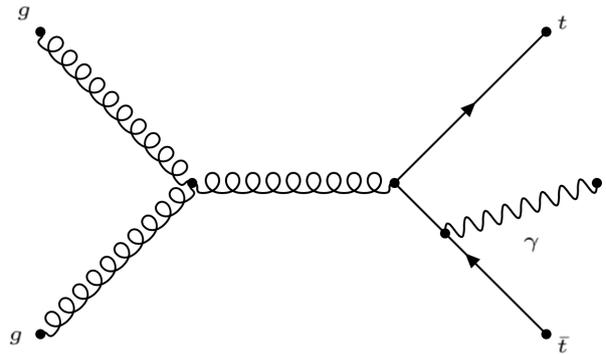
Figure 2.4.: Possible photon radiation of $t\bar{t}$ system for gluon fusion.

To confirm or to get sensitive to possible deviations of the $t\gamma$ coupling, detailed measurements are necessary.

2.2.3. Results of $t\bar{t}\gamma$ analysis at centre-of-mass energies of $7\,\mathrm{TeV}$ and $8\,\mathrm{TeV}$

At a centre-of-mass energy of $\sqrt{s} = 7\,\mathrm{TeV}$ and an integrated luminosity of $\mathcal{L} = 4.59\,\mathrm{fb}^{-1}$ the cross section measurements for the electromagnetic coupling of the top quark ($pp \to t\bar{t}\gamma$) were analysed in the single lepton channel and in a kinematic fiducial region above the photon's transverse energy of $E_T > 20\,\mathrm{GeV}$ at ATLAS. This measurement is mentionable because it was the first time the observations are at a significance of 5.3 standard deviations.

The fiducial cross section is defined as

$$\sigma_{t\bar{t}\gamma}^{fid} = \frac{N_S}{\varepsilon\mathcal{L}},$$

where $N_S = N - N_b$ is the corrected number of $t\bar{t}\gamma$ events. The number of estimated background events N_b is subtracted from potential $t\bar{t}\gamma$ candidates N. ε is the efficiency, which is determined from Monte Carlo simulations as the ratio of all $t\bar{t}\gamma$ events passing the event selection to the events generated in the fiducial region.

The analysis showed that the fiducial cross section at $\sqrt{s} = 7\,\mathrm{TeV}$ is found to be $\sigma_{t\bar{t}\gamma}^{fid} = 63 \pm 8(\mathrm{stat.})_{-13}^{+17}(\mathrm{syst.}) \pm 1(\mathrm{lumi.})\mathrm{fb}$, which is in good agreement to the theoretical result, predicted in QCD NLO calculations, of $\sigma_{theo.} = (48 \pm 10)\,\mathrm{fb}$.

At $\sqrt{s} = 8\,\mathrm{TeV}$ and an integrated luminosity of $\mathcal{L} = 20.2\,\mathrm{fb}^{-1}$ a measurement of photons with a transverse momentum above $p_T > 15\,\mathrm{GeV}$ was performed.

In this case the theoretical prediction at NLO calculations yielded a value of $\sigma_{theo.} = (151 \pm 24)\,\mathrm{fb}$. The experimental value was $\sigma_{t\bar{t}\gamma}^{fid} = 139 \pm 7(\mathrm{stat.}) \pm 17(\mathrm{syst.})\mathrm{fb}$ [25].

3. Experimental Setup - LHC & the ATLAS detector

In the following section the experimental setup will be described. First, there will be a brief description of the Large Hadron Collider at CERN. Then, the ATLAS detector will be introduced.

3.1. Hadron Collider physics at the Large Hadron Collider

The Large Hadron Collider [26] (LHC) is a proton-proton (pp) collider. In 2009, it started operating. The LHC has a circumference of about 27 km and is placed 100 m under ground near Geneva, Switzerland. Currently, the centre-of-mass energy is $\sqrt{s} =$ 13 TeV.

Two beams of protons are accelerated in the opposite direction. At the LHC there are four points, where the beams cross surrounded by one of the detectors (ATLAS [27], CMS [28], ALICE [29], LHCb [30]). ATLAS and CMS are supposed to examine collisions getting SM measurements and hints for physics beyond the SM, whereas ALICE is specialised on heavy-ions to produce a quark-gluon plasma and LHCb studies the CP violation. A sketch of the CERN complex is shown in Figure 3.1. The LHC is a circular accelerator. To hold these high-energetic protons on a circular orbit, a magnetic field strength of $|B| =$ 8.33 T is necessary. Technically, the LHC is limited by this high, but necessary, magnetic field strength. This strength is reached using liquid Helium to get a temperature of 1.9 K and using superconducting NbTi magnets.

In Figure 3.1 the accelerator chain of the CERN complex is presented ending at the LHC. First of all, the protons are collected into bunches, containing 1.15×10^{11} protons per bunch. Each proton beam consists of 2808 bunches. Every 25 ns two bunches cross.

The protons start at the linear accelerator LINAC2. Then they reach the BOOSTER accelerator, where the protons already have an energy of about 1.4 GeV. After leaving the Proton Synchrotron (PS) with an energy of 26 GeV, the protons are injected to the Super Proton Synchrotron (SPS). There, the protons reach an energy of about 450 GeV
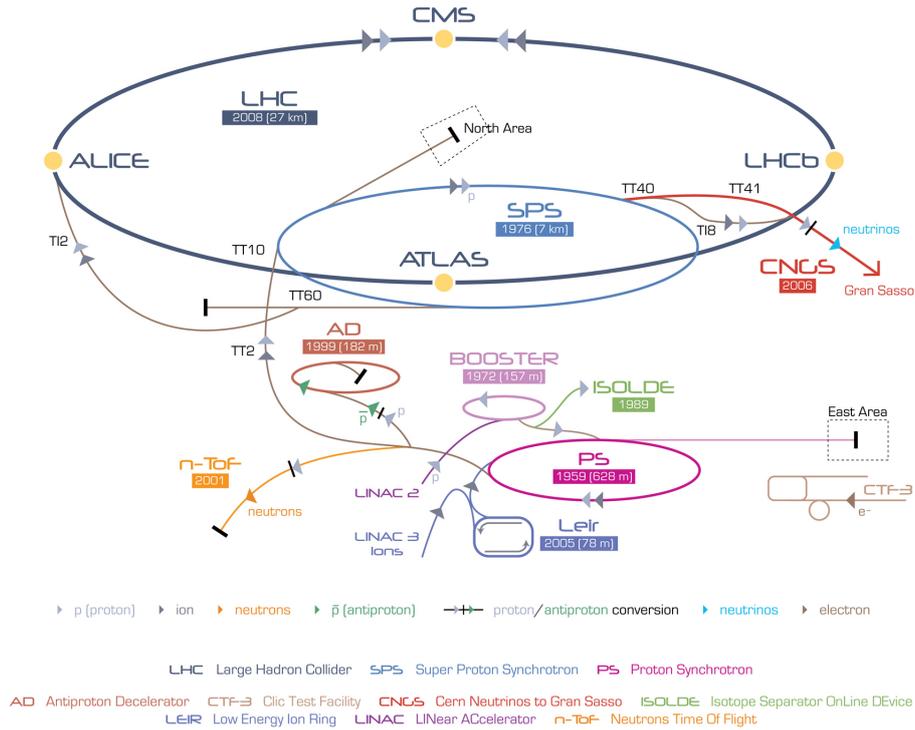
Figure 3.1.: Representation of the accelerator chain at CERN (©CERN, Geneva, Switzerland).

and are injected to the LHC until they reach the desired energy.

In contrast to the static quark model, which was the first idea of a proton and assumes that the proton consists of three quarks (uud), QCD describes the behaviour of the proton and its quarks much better. According to QCD, a proton consists of three valence quarks and, because of the interaction of these valence quarks, gluons can form further quark-antiquark pairs. These are described as sea quarks. All these components of the proton are called partons.

The issue of a high luminosity and therefore a high cross section σ is solved using pp collisions. The current luminosity of the LHC goes up to a maximum of $L = 2 \times 10^{34} \, \mathrm{cm}^{-2} \, \mathrm{s}^{-1}$. An advantage of using protons, with beam energy E, instead of $e^- e^+$ pairs as LEP did, is the non limiting effect of energy loss E_X by synchrotron radiation. The reason for this is the following proportionality

$$E_X \propto \frac{E^4}{m^4}. \tag{3.1}$$

Due to the much larger mass m of the proton, compared to the electron (≈ 2000 times), E_X becomes small.

It is planned that in the next years the centre-of-mass energy ($\sqrt{s}$) will reach its maximum of 14 TeV. From the year 2026 onwards, larger luminosity values are also planned.

3.2. The ATLAS detector

The ATLAS detector [27] is designed for many purposes, for example to find hints for physics beyond the SM or to identify particles. Figure 3.2 shows a schematic picture of the ATLAS detector and its important components.
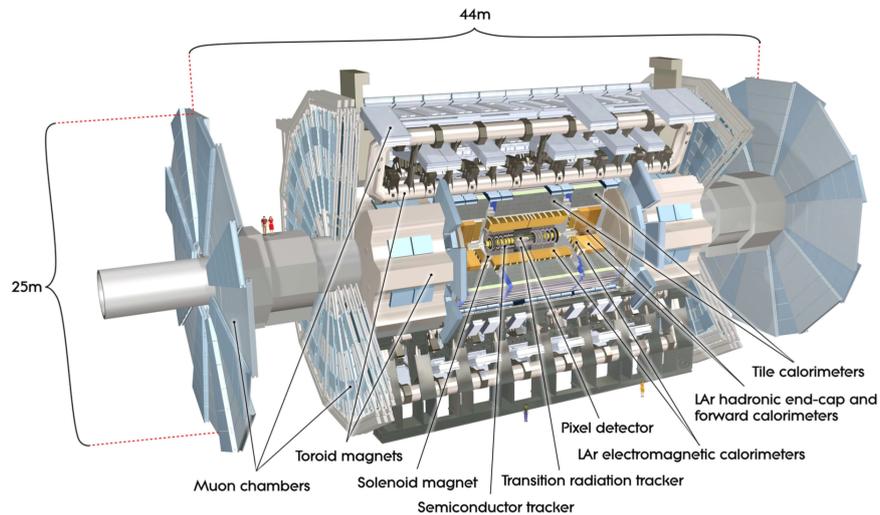
Figure 3.2.: Sketch of the ATLAS detector with its most important sub-components.

In Figure 3.2 one can see that the ATLAS detector is a cylindrical one and therefore a special coordinate system is required (see Figure 3.3). Obviously, the z-axis matches the beam line, Θ is the polar and Φ is the azimuthal angle.
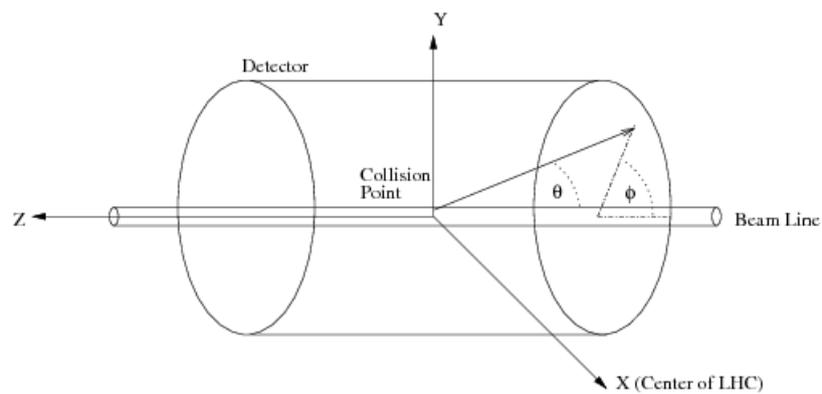
Figure 3.3.: Sketch of the cylindrical coordinate system of the ATLAS detector.

For the observed particles at ATLAS, so E $\gg$ m, where m describes the object's mass and E the total energy, the pseudorapidity η

$$\eta = -\ln \tan \frac{\Theta}{2} \tag{3.2}$$

is a really important quantity, because it depends only on the polar angle Θ, which can be measured directly. Also, η is preferred to Θ, because differences in pseudorapidity are Lorentz invariant under boosts along the longitudinal axis. Lorentz invariant variables are independent of the reference frame, which is also helpful. The following Lorentz invariant variable is used due to distance measurements of two objects in the η - Φ plane

$$\Delta R = \sqrt{(\Delta\Phi)^2 + (\Delta\eta)^2}. \tag{3.3}$$

The Inner Detector

The Inner Detector of ATLAS consists of three components: Pixel detector, Semiconductor Tracker (SCT) and Transition Radiation Tracker (TRT). The whole Inner Detector is placed in a homogenous magnetic field, with a strength of $|B| = 2\,\text{T}$ and field lines parallel to the beam. Because of the Lorentz force, charged particles move onto circular orbits with a radius r, which can be measured. By equating centripetal force and Lorentz force and knowing the charge q of the particle, the absolute value of the transverse momentum p_T of the particle can be calculated with $|p_T| = q|B|r$.

Ultrarelativistic particles are hard to identify in this tracking system, because they travel so fast through this magnetic field, that the radius r becomes nearly immeasurable. The resolution of the reconstructed transverse momentum gets worse with higher transverse momentum, thus

$$\frac{\sigma(p_T)}{p_T} \propto p_T. \tag{3.4}$$

The innermost part of the Inner Detector, the Pixel Detector, is made of Si-diodes and the pixels have an extent over $50\,\mu\text{m} \times 400\,\mu\text{m}$. If an ionizing particle travels through the depletion zone and therefore creates electron-hole pairs, a pulse is measurable. With this information the position, and so the track of the particle is known with a good resolution. Next, the Semiconductor Tracker (SCT) achieves a nearly equivalent resolution of a particle's track. Here, Si strips are used. Financially, this part of the detector is more inexpensive than the Pixel Detector.

To identify a massive particle, one can use the equation $p_T = m\gamma\beta$, where m is the particle's mass, γ is the Lorentz factor and β is the particle's velocity. With the Pixel Detector and the SCT, p_T can be measured, which leads to the need of the particle's velocity to measure the particle's mass.

The outermost part of the Inner Detector is the Transition Radiation Tracker (TRT), which is used for further tracking resolution and to identify particles. This kind of radiation occurs whenever a particle passes the boundary with two different refraction indices

and is released in a cone under an angle ϑ, dependent on the Lorentz factor γ, so

$$\vartheta \propto \gamma^{-1}. \tag{3.5}$$

By measuring the angle ϑ with the calorimeter system, γ can be determined.

The TRT is composed of straw tubes to be conductive to the detection of the particle's tracks. It can detect tracks up to a pseudorapidity of $0 \leq |\eta| \leq 2$. The *pp* collisions create charged particles that ionize the mixture of gases (70% Xe, 27% CO_2, 3% O_2) the straw tubes are filled with. This ionization triggers a measurable pulse. Due to their small masses, mostly electrons can be identified with the TRT. Another important feature of this tracker is to distinguish electrons and pions (different Lorentz factors) with high accuracy.

A profile of the different components of the ATLAS detector is shown in Figure 3.4.

Calorimeter system

First, there is the Electromagnetic Calorimeter (EMC), which is based on the principle of Bremsstrahlung. The EMC is composed of three layers and is shown in Figure 3.5. The first layer is the strip layer, which is used to discriminate single photons and two overlapping photon showers coming from the decays of neutral mesons (mostly $\pi^0 \to \gamma\gamma$ and η-mesons). The second layer collects most of the energy deposited in the EMC by photon and electron showers. Lastly, the third layer corrects the leakage beyond the EMC of high-energy showers.
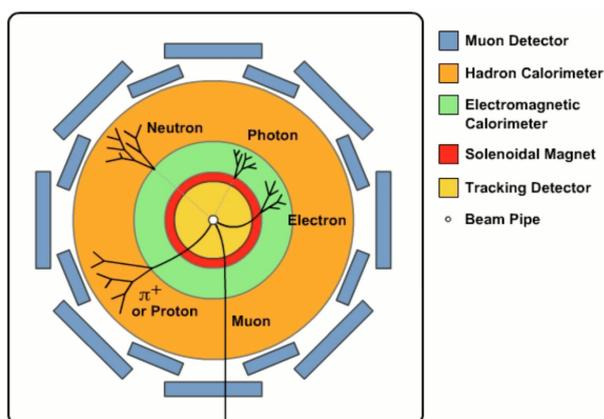
Figure 3.4.: Prediction of the detectable particles and their behaviour in the different components of the ATLAS detector to identify them.
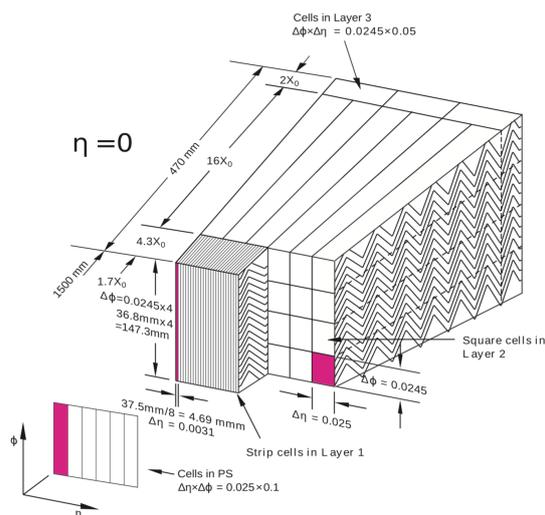
Figure 3.5.: Sketch of a part of ATLAS EMC around $\eta = 0$.

Charged particles mostly lose their energy due to Bremsstrahlung or ionisation (Bethe-Bloch formula). The EM calorimeter is designed for absorbing electrons/positrons and photons completely. If highly energetic and low mass particles like e^- or e^+ with energy E passes the passive material (lead), they will emit highly energetic photons under an opening angle

$$\xi \propto E^{-1} \tag{3.6}$$

due to their deceleration. The photons are so energetic, they can create $e^- e^+$-pairs, which can also emit photons. This chain occurs until the emitted photons do not have enough energy for further pair production anymore. By measuring the length of the electromagnetic shower (active material (liquid argon)), the energy of the incident particle E can be calculated with

$$N_{max} = \frac{E}{E_C}. \tag{3.7}$$

E_C is the critical energy, defined as the energy value, the particle loses more energy because of ionisation than Bremsstrahlung and N_{max} is the number of radiation lengths X_0 for a maximal shower length.

Strongly interacting particles like hadrons lose their energy in the hadronic calorimeter (HaC). Hadrons stop in the hadronic calorimeter due to a hadronic shower by passing the passive material (steel). Hadrons often decay into pions, which again decay into photons evoking an electromagnetic shower in the HaC as well. The characteristic length for the hadronic shower is the so called interaction length λ, which is similar defined as X_0. By knowing both X_0 and λ, the particle's shower can be measured (active material (scintillating tiles)).

Muon Detector

The outermost part of the ATLAS detector is the Muon Detector. Because muons are much heavier than electrons ($\approx$ 200 times), they only leave a small, nearly not noteable curvature in the inner detector system. Energy loss by Bremsstrahlung is proportional to m^{-2}, so this is a negligible effect. They only leave a light track in the EM calorimeter by ionisation. Moreover, muons do not interact strongly, so they can traverse the hadronic calorimeter. Therefore the Muon Detector was built. Only muons can leave a track in the Muon detector.

Other particles, leaving no tracks in any detector, are neutrinos or unidentified particles, which are only detectable by calculating the missing transverse momentum.

4. Photon measurements at ATLAS

In the following chapter the procedure of photon reconstruction and identification [31] at ATLAS will be presented. Photons can be reconstructed as two different types, which are essential for the $t\bar{t}\gamma$ analysis. These different reconstructed types will be discussed at first.

4.1. Prompt photons and hadron fakes

Many charged particles like electrons or charged hadrons leave similar signatures like the photons in the EMC, which compounds the ability for a good identification. This thesis mainly deals with the discrimination between two reconstructed types of photons: prompt photons and hadron fakes.

Prompt photons are real photons in the final state, not originating from hadronic decays (e.g. see Figure 2.4). For studies concerning the electromagnetic coupling of the top quark these photons are crucial, because they can be radiated off a top quark.

The set of hadron fakes contains hadrons misidentified as photons (EMC) and real photons originating from hadronic decays (e.g. $\pi^0 \rightarrow \gamma\gamma$). For the $pp \rightarrow t\bar{t}\gamma$ analysis the estimation of the hadron fakes contribution is important as it is a dominant background.

In simulation, one can discriminate between prompt photons and hadron fakes using the truth information of the photons presented in Figures 4.1 and 4.2. The different values on the photon truth-information-axis correspond to different photon origins, which will be apparent from Table A.3 in Appendix A.3.

Truth origin number 37 is used in the jet $+$ γ sample (see Figure 4.1) to get prompt photons. Numbers 23-35 for photons originating from different mesons and baryons and particularly number 42 for the $\pi^0 \rightarrow \gamma\gamma$ decay in the dijet sample (see Figure 4.2) are used to get the hadron fakes contribution.
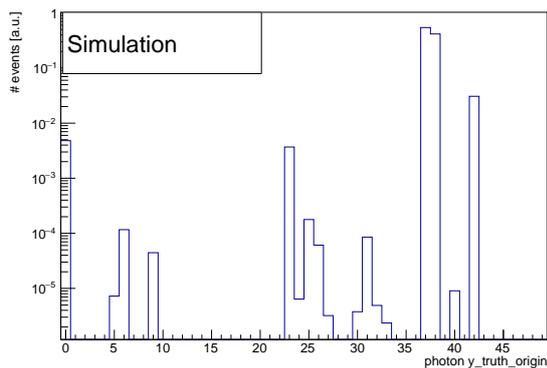
Figure 4.1.: Depiction of the true origin of γ in the jet $+ \gamma$ sample containing prompt photons.
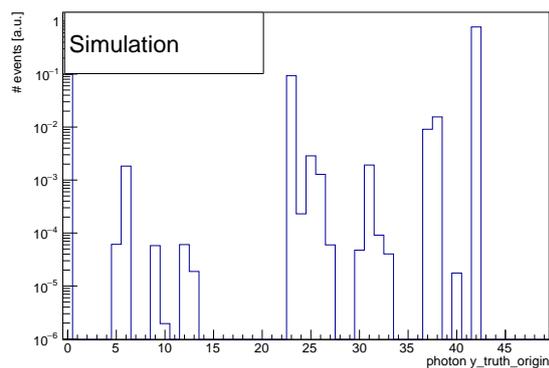
Figure 4.2.: Depiction of the true origin of γ in the dijet sample containing hadron fakes.

4.2. Photon reconstruction

Electrons and photons leave similar signatures in the EMC, and so their reconstruction proceeds in parallel.

A window with an extent of 3×5 cells in the $\eta \times \Phi$ plane searches for electromagnetic cluster seeds with a total transverse energy over $E_T > 2.5\,\text{GeV}$. For photons with $E_T > 20\,\text{GeV}$ the simulation efficiency of cluster search by the usage of a special algorithm [32] exceeds 99%.

Additionally, photons are reconstructed as converted or unconverted with the corresponding variable of the *convType* (see Table A.4 in Appendix A.3). The state of converted means the conversion of the photon into an $e^- e^+$ pair via Bremsstrahlung, which implies one or two tracks with Si hits (diodes or stripes) or without those hits. Only the TRT gets the signal in the latter case.

Probable tracks are matched to the clusters and the appropriate algorithm can distinguish between electrons and photons.

4.3. Photon identification

For a reliable identification of photon candidates variables, delivering good separation between prompt photons and hadron fakes are needed. For this issue, shower shape variables are used, which only depend on the photon candidate's reconstructed energy.

Additionally, a rectangular cut according to pseudorapidity η is given: $|\eta| < 2.37$ excluding the crack region of the EMC $1.37 < |\eta| < 1.52$. This region, where the barrel accordion meets the endcap accordion will not be considered in the analysis further.

In Figure 4.3 shower shape variables, describing various kinds of energy ratios are presented. Their definitions are depicted in Table 4.1.

Figures 4.3 and 4.5 contain the expression of $E_P^{S_n}$, which indicates the electromagnetic energy collected in the n-th longitudinal layer of the EMC in a cluster of properties P to identify the number or properties of selected cells. In Figure 4.4 E_i is the energy and η_i is the pseudorapidity in the i-th cell.

$$f_1 = \frac{E_{S1}}{E_{\text{Tot.}}} \quad f_{\text{side}} = \frac{E_7^{S1} - E_3^{S1}}{E_3^{S1}}$$

$$R_\eta = \frac{E_{3\times7}^{S2}}{E_{7\times7}^{S2}} \quad R_\phi = \frac{E_{3\times3}^{S2}}{E_{3\times7}^{S2}}$$

$$R_{\text{Had}} = \frac{E_T^{\text{Had}}}{E_T}$$

Hadronic
Second Layer
Strips
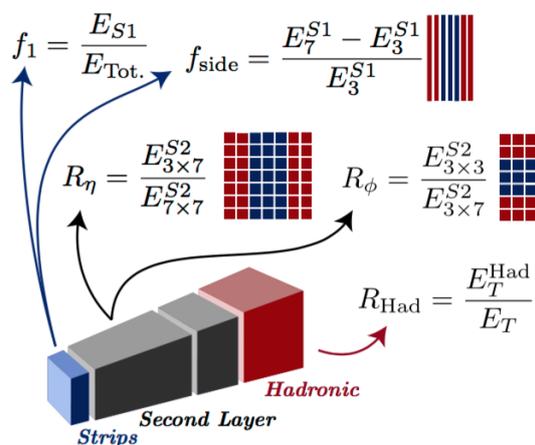
Figure 4.3.: Visualization of shower shape variables (energy ratios).

Quantity	Definition
R_η	Energy ratio of 3×7 to 7×7 cells in the $\eta \times \Phi$ plane.
R_Φ	Energy ratio of 3×3 to 3×7 cells in the $\eta \times \Phi$ plane.
R_{had}	Transverse energy leakage in the HaC normalised to the transverse energy of the γ candidate in the EMC.
f_1	Ratio of the energy measured in the first sampling of the electromagnetic calorimeter to the total energy of the EM cluster.
f_{side}	Energy within 7 strips without 3 central strips normalised to energy in 3 central strips.

Table 4.1.: Description of different shower shape variables (energy ratios).

R_η and R_Φ are measured within the second layer (see Figure 3.5) of the EMC, whereas R_{had} describes the hadronic leakage. f_1 and f_{side} characterise the energy ratios in the strip layer (first layer) of the EMC.

The shower shape variables, presented in Figure 4.4 and defined in Table 4.2, describe widths in the strip layer ($w_{\eta,1}$ and $w_{tot,s1}$) and in the second layer ($w_{\eta,2}$) of the EMC. The third representation of shower shape variables shows energy ratios and differences of energy maxima and minima in the strip layer of the EMC (see Figure 4.5 and Table 4.3).

All variables only depend on their reconstructed energy, which means that they can be used in any analysis with photons. Furthermore, they are all ratios (except ΔE), which

$$w_{\eta,2} = \sqrt{\frac{\sum E_i \eta_i^2}{\sum E_i} - \left(\frac{\sum E_i \eta_i}{\sum E_i}\right)^2}$$

Width in a 3×5 (Δη×Δφ) region
of cells in the second layer.

$$w_s = \sqrt{\frac{\sum E_i (i - i_{\max})^2}{\sum E_i}}$$

ws3 = w1 uses 3×2 strips (η×φ);
wstot is defined similarly, but
uses 20×2 strips.
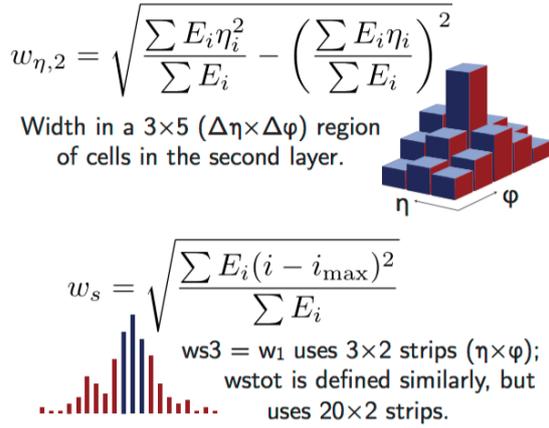
Figure 4.4.: Visualization of shower shape variables (widths in strip and second layer).

Quantity	Definition
$w_{\eta,1}$ (w_s)	Energy weighted width in units of the number of strips using 3 strips around the maximum.
$w_{\eta,2}$	Lateral width of cluster in $\eta \times \Phi = 3 \times 5$.
$w_{tot,s1}$ (w_{s3})	Energy weighted width using 20 strips around the maximum, see $w_{\eta,1}$.

Table 4.2.: Description of different shower shape variables (widths in strip and second layer).

$$\Delta E = E_{\max,2}^{S1} - E_{\min}^{S1}$$

$$E_{\text{ratio}} = \frac{E_{\max,1}^{S1} - E_{\max,2}^{S1}}{E_{\max,1}^{S1} + E_{\max,2}^{S1}}$$
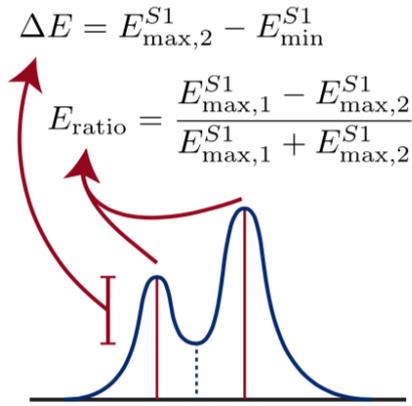
Figure 4.5.: Description of different shower shape variables (energy ratios in strip layer).

Quantity	Definition
E_{ratio}	Ratio of the energy difference of largest and second largest energy maxima to the sum of these energies.
ΔE	Difference between the energy of the second largest maximum and the energy of the minimum value found between the first and second maxima in the strip layer.

Table 4.3.: Description of different shower shape variables (energy ratios in strip layer).

minimises the impact of systematic uncertainties.

Moreover, two reference sets of cuts are defined: looseID and tightID. The selection of the looseID obtains only R_{had}, $w_{\eta 2}$ and R_η, thus shower shapes in the second layer of the EMC and the HaC, whereas tightID also adds selection criteria to the other shower shapes, which are illustrated in Figures 4.3-4.5 and Tables 4.1-4.3 [31].

5. Monte-Carlo truth studies of shower shapes

In the following chapter shower shapes of prompt photons (prompts) and hadron fakes[1] (hfakes) in simulation are analysed and presented. Therefore, two different Monte-Carlo (MC) samples are used: single photon and dijet. In Table 5.1 the corresponding ROOT files for the two MC samples and in Figures 5.1 and 5.2 LO Feynman diagrams are shown.
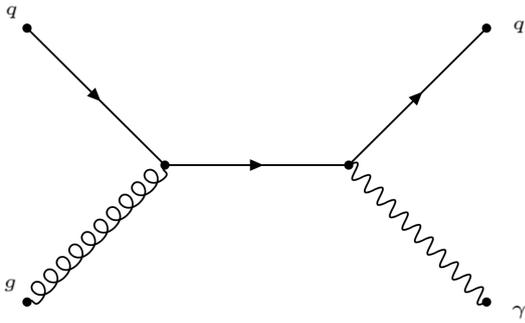
Figure 5.1.: LO Feynman diagram of the single photon sample (it is known as the "QCD Compton process").
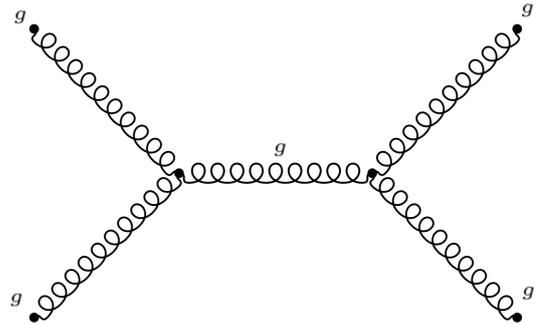
Figure 5.2.: LO Feynman diagram of the dijet sample via triple gauge coupling.

For these two MC samples, reconstructed photons have a transverse momentum above $p_T > 20\,\text{GeV}$ so that the electronic noise [27] and low energetic particles can be removed. Additionally, the photons have to pass a pseudorapidity cut of $|\eta| < 2.37$, excluding the crack region of $1.37 < |\eta| < 1.52$. If it is not particularly mentioned, both converted and unconverted photons will be taken into account. As a last selection cut, they have to pass the looseID criteria.

First, there will be a comparison between the behaviour of prompt photons and hadron fakes in different shower shape variables. Subsequently, the behaviour of the shower shapes of prompts and hfakes in three different p_T and *convType* bins will be presented to assess

---

[1]Both will be discriminated using the MCTruthClassifier, as it was mentioned in Section 4.1.

the impact of different p_T regions or conversion status of the photons on their behaviour.

Table 5.1.: Used ROOT-files for two different MC samples for 21.0 version.

Sample	ROOT-File
$\gamma + jets$	PyPt17_inf_mc16c_v18.root
$dijet$	Py8_jetjet_mc16c_v18.root

5.1. Comparison between prompt photons and hadron fakes

The $\gamma + jets$ and the $dijet$ samples contain approximately 3.97×10^9 prompt and 2.28×10^{12} hadronic fake photon events.

For a good selection of prompt photons in the NN, which is used for the $t\bar{t}\gamma$ analysis in Section 6.1, a good separation between prompt and hadronic fake photons is necessary. Separation S is defined as

$$S = \frac{1}{2} \sum_{i \in \text{bins}} \frac{(p_i - h_i)^2}{p_i + h_i},$$
(5.1)

where p_i is the prompt photon and h_i is the hadron fake contribution in bin i.

Table 5.2.: Representation of separation values for shower shapes.

Variable	Separation S in %	# bins
R_{had}	11.75	50
R_η	45.96	50
R_Φ	35.07	50
$w_{\eta 1}$	10.76	50
$w_{\eta 2}$	21.39	50
f_{side}	24.32	40
$w_{tot,s1}$	22.18	50
ΔE	9.97	50
E_{ratio}	18.04	40

All separation values for the different shower shapes and binning numbers are listed in Table 5.2. The values vary between 10% and 46%.

Different normalised distributions of R_{had}, R_η, R_Φ, $w_{\eta 1}$, $w_{\eta 2}$ and f_{side} are illustrated in Figure 5.3, whereas $w_{tot,s1}$, ΔE and E_{ratio} are shown in Figure A.1 in Appendix A.1. At

the bottom, the ratio between the number of prompts and the number of hfakes per bin was calculated and assigned with error bars.
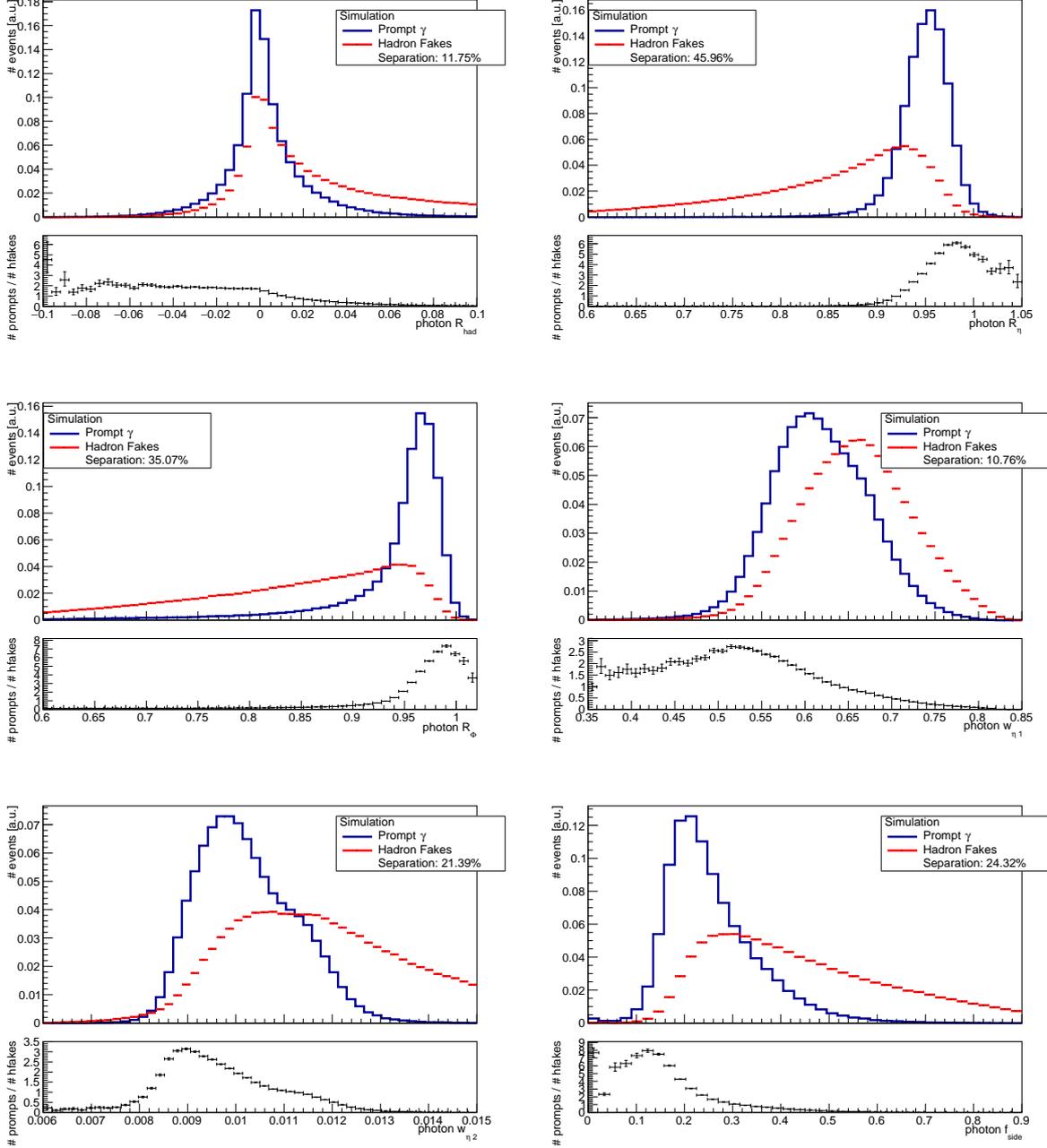
Figure 5.3.: Six shower shape variables with ratio plots.

Although all shower shapes (except ΔE) are ratios, negative values appear in the hadronic leakage distribution R_{had}. The reason is the issue of electronic noise, which is not negligible and has to be subtracted. The same argument applies to the values

above 1 for the distributions of R_Φ and R_η.

5.2. Sensitivity of prompt photons and hadron fakes according to kinematic variables

The interaction between photons and the material of the ATLAS detector is energy dependent, which means that the shape of the shower shapes might differ for different energetic photons. For this purpose, it is important to study the impact of different kinematic variables like transverse momentum p_T and conversion state *convType*. Additionally, the influence of these variables on prompt and hadronic fake photons will be analysed separately to compare their behaviour for different p_T and *convType* cuts.

Primarily, shower shapes of prompts and hfakes with a transverse momentum between $7\,\text{GeV} < p_T \leq 50\,\text{GeV}$ (low p_T), $50\,\text{GeV} < p_T \leq 100\,\text{GeV}$ (medium p_T), and photons with $p_T > 100\,\text{GeV}$ (high p_T) will be studied. Furthermore, the different behaviour of unconverted photons and converted photons will be presented.

5.2.1. Shower shapes for different p_T cuts

The number of prompt and hadronic fake photons in the different p_T regions decrease with increasing p_T. There are approximately 92.7% prompt and 93.0% hadronic fake ones with low p_T, 6.8% prompt and 6.6% hadronic fake ones with medium p_T and 0.5% prompt and 0.4% hadronic fake photons with high p_T.

In Figure 5.4 R_{had}, R_η and R_Φ in the three p_T bins are illustrated. The residual six shower shapes are shown in Figure A.2 in Appendix A.1. On the left are the distributions of prompts and on the right of hfakes.

R_{had} describes the transverse energy leakage in the HaC. In the case of prompt photons, there is a distinct peak for a value of nearly 0. Prompt photons are mostly absorbed in the EMC (see Section 3.2), so they do not lose energy in the HaC causing the peaks around 0. Photons with high p_T display the strongest peak, whereas the medium and low p_T distributions are wider and more symmetric. Hadron fakes are partially photons originating from hadronic decays, which explains why there is more energy leakage in the HaC. The distribution does not approach 0 (until $R_{had} = 0.1$) and is broader and more asymmetric.

The energy fraction in the η direction is measured by R_η. It can be observed that the peak for prompt photons is shifted to the left for decreasing p_T. A decreasing p_T
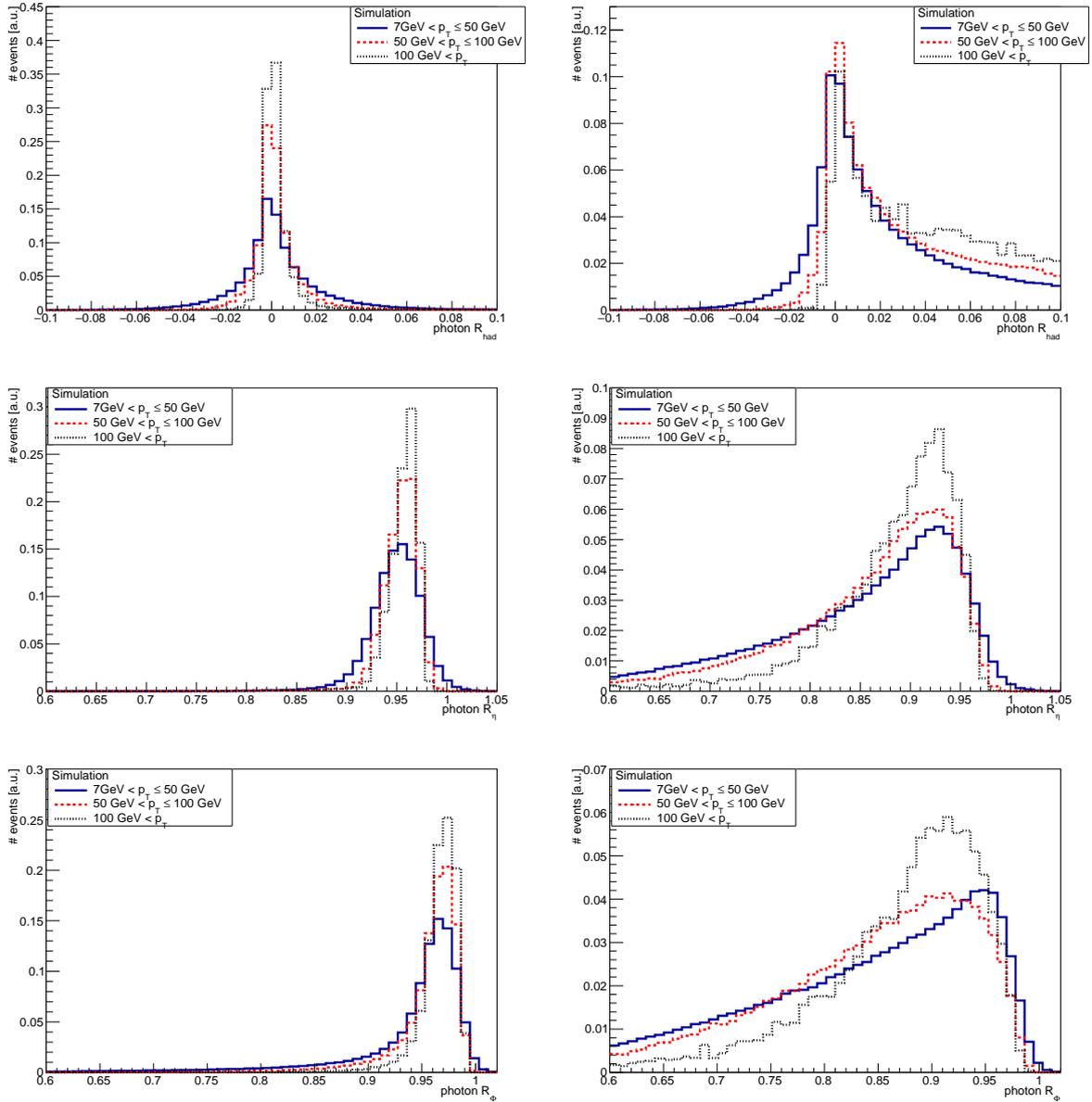
Figure 5.4.: Shower shapes in 3 different p_T bins. Distributions of prompt on the left and hadronic fake photons on the right.

implies an increasing longitudinal energy[2], which again implies a larger energy leakage in η direction. The more highly energetic a photon is, the less the leakage of energy, because they lose less energy due to their travel. In general, prompt photons are more energetic than hadronic fake ones. As a result, the distribution of R_η for hadronic fake photons is wider than that for prompt ones, which leads to a more centered contribution of prompt photons.

---

[2]For constant total energy E_{tot}.

The same argument applies to the energy fraction in Φ direction. Once again, the hadronic fake photon curves are broader than those of the prompt photons. Additionally, prompt ones with low p_T induce a broader electromagnetic shower in the EMC (see Equation 3.6), which leads to a more right shifted peak for high p_T.

Since $w_{\eta 2}$ describes the shower width in a $\eta \times \Phi = 3 \times 5$ region of cells in the second layer of the EMC, the same reasons are valid for the behaviour of the $w_{\eta 2}$ distribution, shown in Figure A.2 in Appendix A.1. If p_T for the prompts rises, the maximum for $w_{\eta 2}$ distributions is shifted to the left.

As $w_{\eta 1}$ and $w_{tot,s1}$ both describe the energy weighted width in units of the number of strips using 3 or 20 strips around the maximum, their distributions (Figure A.2 in Appendix A.1) can be explained simultaneously. Besides the known left shift of the maxima for both prompts distribution, one can observe that the curve of hfakes $w_{\eta 1}$ looks smoother than that of hfakes $w_{tot,s1}$. This is due to the usage of more strips in the latter distribution, which leads to a more chaotic allocation of the photons originating from hadronic decays. As it was defined in Section 4.3, f_{side} is the photon energy in 4 outer strips normalised to the energy in 3 central strips. According to Equation 3.6, a narrower shower is induced by high p_T prompts, which causes more energy in the 3 central strips relating to the 4 outer strips than the medium and low p_T prompts, and therefore again the left shift for increasing p_T. This is illustrated in Figure A.1 in Appendix A.1. Vice versa a wider distribution can be observed for the hfakes distribution.

Commonly, hfakes do not induce two energy maxima[3] in the strip layer. This circumstance justifies the reason for the distinct maximum for $E_{ratio} = 1$ and $\Delta E = 0$, shown in Figure A.2 in Appendix A.1. A high p_T prompt photon also does not induce a second maximum or even a small one, which causes the right shift for E_{ratio} and the left shift for ΔE for increasing p_T.

5.2.2. Shower shapes for different conversion status

The batch of prompt photons contains approximately 66.0% unconverted and 34.0% converted photons, whereas the cluster of hadron fakes includes 40.8% unconverted and 59.2% converted photons.

$w_{\eta 1}$, $w_{\eta 2}$ and f_{side} are shown in Figure 5.5, whereas R_{had}, R_η, R_Φ, E_{ratio}, ΔE and $w_{tot,s1}$ are covered in Figure A.3 in Appendix A.1.

It is more likely that hfakes are reconstructed as converted photons, which can be seen in Figure 5.6 (meaning of different bin numbers in Table A.4 in Appendix A.3), because

---

[3] As a result, there is also no minimum (see ΔE in section 4.3).
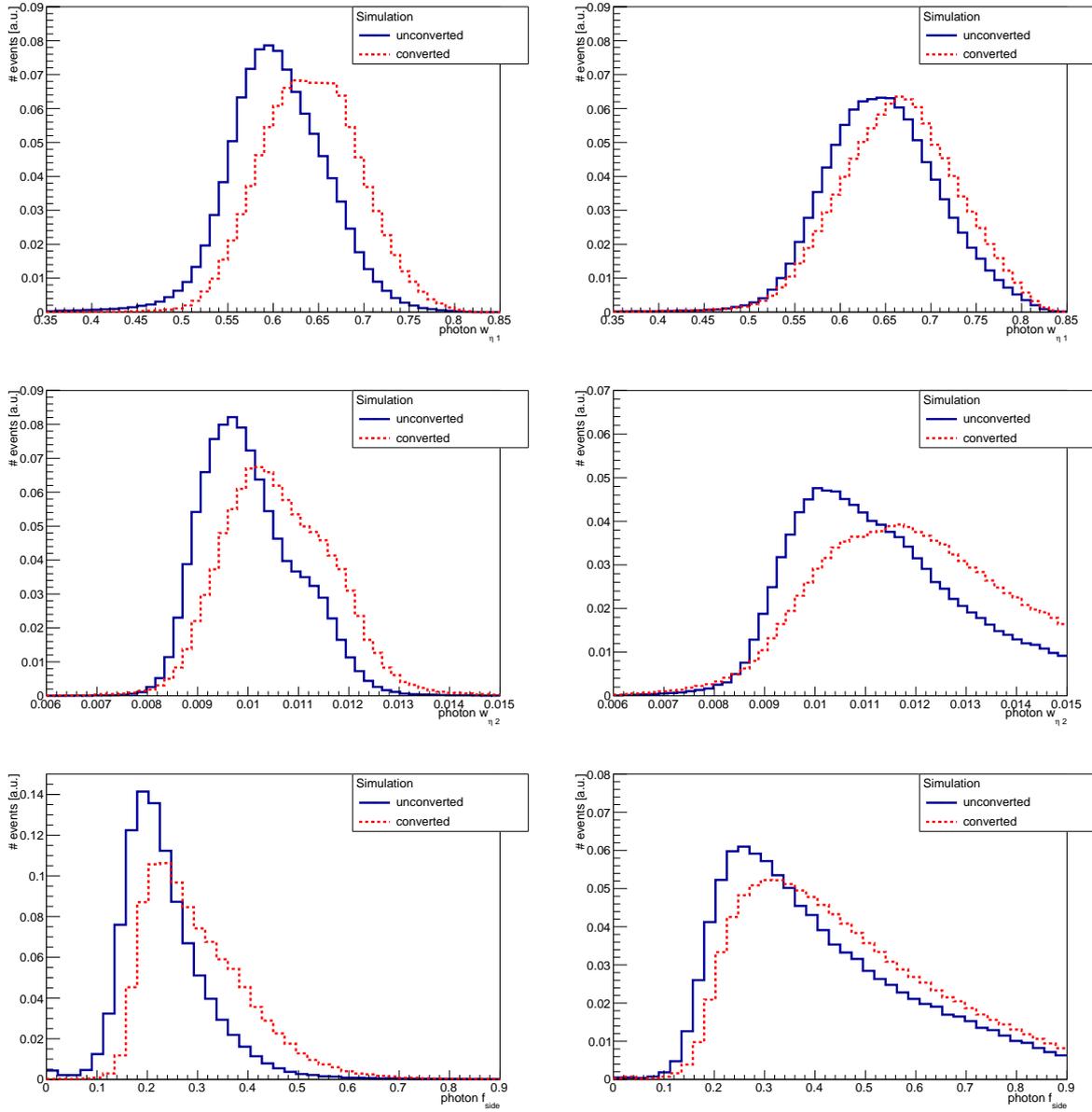
Figure 5.5.: Shower shapes for different conversion types. Distributions of prompts on the left and hfakes on the right.

they have a greater tendency to leave Si hits or to evoke a TRT signal.

In the case of a converted photon, the energy of the photon is split into an $e^- e^+$ pair. This implies that an unconverted photon is more energetic than a converted one, thus a converted one can be compared with a photon with low and medium p_T, whereas a unconverted one can be compared with a high p_T photon. This can also be seen in Figure 5.7. Therefore the arguments, presented in Section 5.2.1, are valid to explain the shape of the 9 different shower shapes.
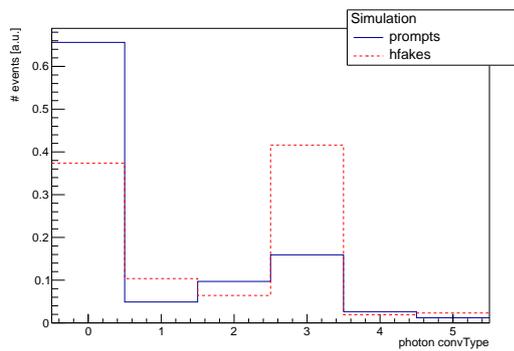
Figure 5.6.: *convType* distribution for prompts and hfakes.
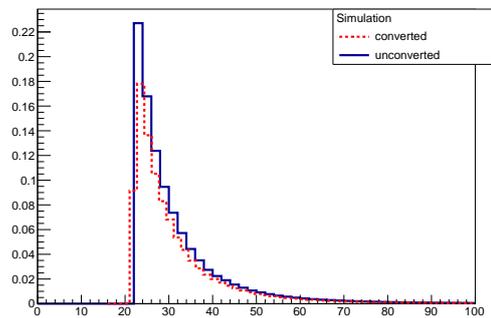
Figure 5.7.: p_T distribution for prompt photons as converted/unconverted.

6. The Prompt Photon Tagger

For the following $t\bar{t}\gamma$ analysis a Neural Network [33] (NN) is used to improve the classification between prompt photons and hadron fakes. NNs are inspired by biological neural networks, which constitute animal brains for example. It is a nonlinear transformation, which depends on adaptable parameters, on a set of input variables to a set of output variables. The adaptable parameters are the features, selected to get an optimal result. To optimise its classification regarding selection criteria, the features have to be trained. After a brief introduction to the used NN with its training, feature selection, architectures and sensitivities, its usage and results according to the $t\bar{t}\gamma$ analysis will be discussed. In Figure 6.1 the basic components of a NN are presented. Neurons (circles) are composed in layers. Each neuron is connected to each in the previous and succeeding layer. Neurons[1] are connected with activation functions [34]. For a more reliable selection of the NN, continuous activation functions are used.

An important aspect is a special type of layer: the batch normalisation layer (bn). Its usage leads to a higher learning rate, because bn ascertains that there is no activation that is gone really high or really low, and a higher independence towards other layers. Additionally, it reduces largely the internal covariate shift. The bn layer whitens the inputs (zero mean, unit variances) and makes them uncorrelated, which leads to a faster convergence of the NN.
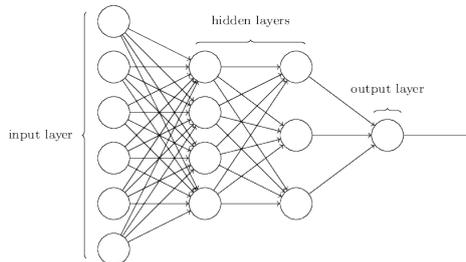
Figure 6.1.: The basic building blocks of a NN.

The Prompt Photon Tagger [35] (PPT) is a NN for a better selection of prompt photons and rejection of hadron fakes. For this purpose it labels looseID photon candidates

---

[1]or perceptrons, a special neuron model.

between a value of 0 and 1. A value of 0 suggests a hadronic fake photon, whereas a value of 1 indicates a prompt photon. It is construed to be analysis independent, thus it can be used in any photon analysis.

First, there will be information on feature selection and training. Then, different architectures (different numbers and types of layers, activation functions and neurons) will be compared to use the best possible one for the $t\bar{t}\gamma$ analysis. Lastly, the sensitivity of the PPT regarding p_T, η and *convType* is examined. For these purposes the root2kerasPipeline is used per training and the LightWeight Trained Neural Network [36] (LWTNN) to operate on the PPT outputs and evaluate the PPT.

Feature selection and training of the PPT

To ensure its analysis independence, it is necessary to deliver analysis independent features to the PPT. As it was already mentioned in Section 4.3, shower shapes only depend on the reconstructed photon energy. Additionally, the separation between prompt photons and hadronic fake ones is large enough (see Table 5.2) for the PPT, thus it will identify them separately. The performance of the PPT is expected to depend on its selection criteria according to p_T, *convType* and even η.

The same selection cuts according to η and truth information (see Section 4.1) as in Chapter 5 are used. Also the photon's transversal momentum p_T is demanded to surpass $p_T > 20\,\text{GeV}$ to remove a huge amount of background. Then, 348870 prompt and 365416 hadronic fake photons can be measured. Neither prompts nor hfakes are weighted. For a more reliable result, the training samples are divided into 2 sets: train and test. Randomly selected, the train set contains 80% of the events. The remaining 20% of the events are contained in the test set.

To reduce the loss of the training[2] the NN was trained over 100 epochs.

Choice of used model

There are different possibilities (models) for the PPT according to the number of neurons in each layer, number of layers, types of layers, feature selection and activation functions. Multiple scenarios are tested and the best model is chosen.

For this purpose, Receiver Operating Characteristics (ROC) curves are used. A ROC curve illustrates the diagnostic ability of a binary classifier system. This curve is used to suggest, which model should be chosen since it shows the classification and generalisation performance of the NN.

---

[2]For more information, see [35].

A NN will have a good generalisation capability if its Roc curves for the test and train set are in agreement.

The Area Under Curve (Auc) value represents its classification performance, whereas a value of 1 matches a perfect classifier and a value of 0.5 indicates that the classifier is not better than flipping a coin.

In Table 6.1 one can see the Auc values for different models. The corresponding Roc curves are presented in Figure A.4 in Appendix A.2. Dense denotes the number of neurons per layer. Also, two different training sets are compared. The first one involves 6 variables: R_{had}, R_η, R_Φ, $w_{\eta 1}$, $w_{\eta 2}$ and f_{side}, whereas the second set includes the additional variables $w_{tot,s1}$, E_{ratio}, p_T, $convType$ and η. ΔE was not considered, because this variable is not a ratio (see Section 4.3) and therefore can be a source of systematic uncertainties.

Table 6.1.: Comparison of different architectures with different number of trained variables and the corresponding Auc values.

Architecture	# variables	AUC (train)	AUC (test)
1 layer (dense: 64)	6	0.9194	0.9192
	11	0.9485	0.9475
2 layer (dense: 20, 20)	6	0.9216	0.9215
with bn	11	0.9448	0.9449
3 layers (dense: 64, 40,	6	0.9246	0.9242
52) with bn, softmax	11	0.9516	0.9504
5 layers (dense: 64, 152,	6	0.9219	0.9210
60, 50, 60) with bn	11	0.9546	0.9516

The highest percentage deviation of the train and test set is approximately 0.3% (5 layers, 11 variables), which suggests a suitable generalisation performance for all architectures. The architecture with 3 layers, batch normalisation and softmax activation function [33] shows the highest Auc = 0.9377 on average. Despite the fact that all architectures shown would be suitable for $t\bar{t}\gamma$ analysis, the latter is the reason why the model with 3 layers was chosen for further studies.

Sensitivity of the PPT according to p_T, η and $convType$

Before implementing the Ppt to analyse the $t\bar{t}\gamma$ process, the influence of the variables p_T, η and $convType$ on the Ppt output will be studied. For this purpose two sets with different features are trained within the 3 layer model. The first set contains the six shower shapes used before, whereas the second one includes in addition p_T, η and $convType$. Figure A.5 in Appendix A.3 shows the shapes for the two sets of prompts and hfakes in different bins

of p_T, η and *convType*. The same regions for p_T and *convType* as in Section 5.2 are used. The different η bins are adjusted to the detector assembly. They cover the barrel part ($0 \leq |\eta| < 0.6$ and $0.6 \leq |\eta| < 1.37$) and the endcap ($1.52 < |\eta| < 2.37$).

To get sensitive to their correlation, all correlation values among the variables of the two different sets are listed in Table A.1 (first set) and Table A.2 (second set) in Appendix A.3. As expected, the first row/column makes the difference between those tables.

At first, the correlation of the NN output to p_T, η and *convType* will be expected to become larger, if they are included in the training, because the NN gets accustomed and becomes therefore more sensitive to those.

By comparing the two Tables A.1 and A.2 in Appendix A.3 one notices that the absolute value of the correlation value between the NN output and p_T and η gets smaller as expected, but the correlation between the NN output and *convType* gets larger.

This unexpected behaviour might be explained by the fact that the *convType* variable is highly correlated to the shower shapes in comparison with η or p_T. As an example R_Φ is more than -50% correlated to *convType*, whereas it is less than 3% correlated to p_T or η. Because of these inter-correlations[3] it is a complex[4] issue to predict the change of the correlation value.

6.1. Analysis of $t\bar{t}\gamma$ with the PPT

To be able to analyse the electromagnetic coupling of the top quark (possible LO Feynman Diagrams are shown in Figures 2.3 and 2.4) at a centre-of-mass energy $\sqrt{s} = 13\,\text{TeV}$, the PPT (3 layer model) was implemented in the $\ell +\ jets$ final state topology. As can be seen in Figure 2.2 this channel is divided into $e +\ jets$ and $\mu +\ jets$.

For the analysis there are exactly one lepton (e or μ) and exactly one photon passing the looseID criteria. Additionally, the photon has to have a transverse momentum above $p_T > 20\,\text{GeV}$ for two reasons. First, a highly energetic photon for $pp \to t\bar{t}\gamma$ is demanded and therefore a huge amount of low energetic background can be removed. Second, this is the same cut used in the training for labelling photons with the PPT, thus the PPT would not be good in labelling lower energetic ones. At last, the number of jets is expected to be ≥ 4 of which at least one is b-tagged with the Mv2c10 tagging algorithm [37].

Figure 6.2 presents the number of $t\bar{t}\gamma$ events and hfakes, which contain $t\bar{t}$, $Z +\ jets$, $W +\ jets$ and diboson processes, in the $\ell +\ jets$ channel. A special weighting[5] is used to achieve a more precise result. To get sensitive to the performance of the PPT the

---

[3]Correlation between shower shapes and *convType*.

[4]For the scope of this Thesis to go into this in detail would take us too far afield.

[5]E.g. for the b-tagger, which is currently used for the ntuples.

corresponding training graphic is shown in Figure 6.3.
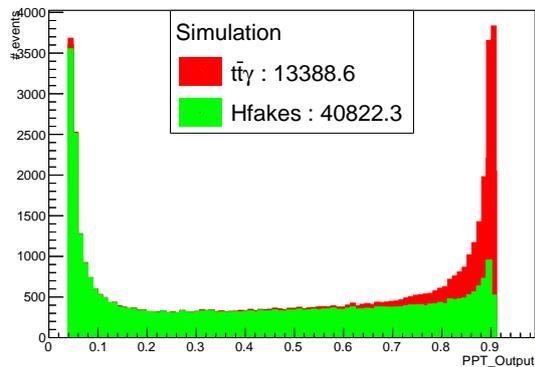
Figure 6.2.: PPT distribution containing hadronic fake photons and $t\bar{t}\gamma$ without an isolation cut.
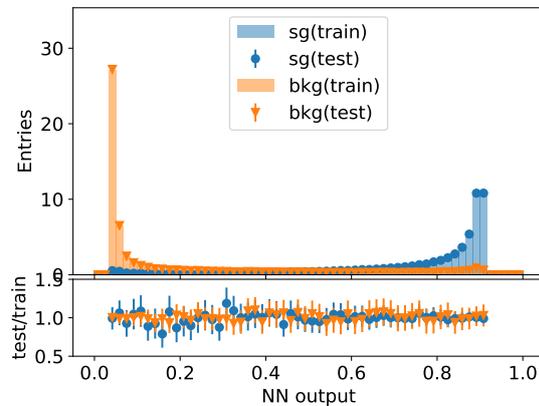
Figure 6.3.: Corresponding NN output plot.

The difference between the shapes of those plots occurs, because the left one presents a stackplot, whereby the peak near 0.9 becomes that high.

Due to the large number of hfakes, another selection cut on photons can be applied. Additionally, photons have to pass the so-called FixedCutTight (FCT) isolation [31], because hadronic fake photons are expected to be less isolated.
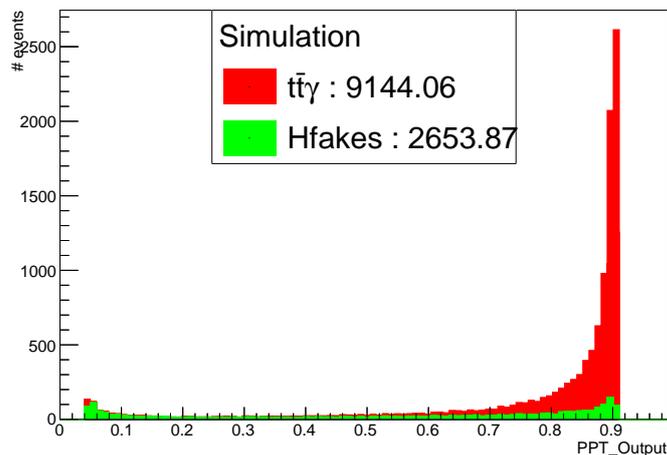
Figure 6.4.: PPT distribution containing hfakes and $t\bar{t}\gamma$ with FCT isolation.

Besides the FCT cut, the same selection cuts are used in Figure 6.2 and 6.4. The number of $t\bar{t}\gamma$ events drops by 31.7 % and hadronic fake ones strongly by 93.5 %.

Lastly, Figure 6.5 represents the MC prediction and the official data from 2015 and 2016.

To achieve a better prediction and due to the fact that also electrons are able to fake a photon signal in the detector (e-Fakes) as a large fraction of the background contribution, these ones are also included in the plot. The red contribution represents the total MC prediction with version 21.0.
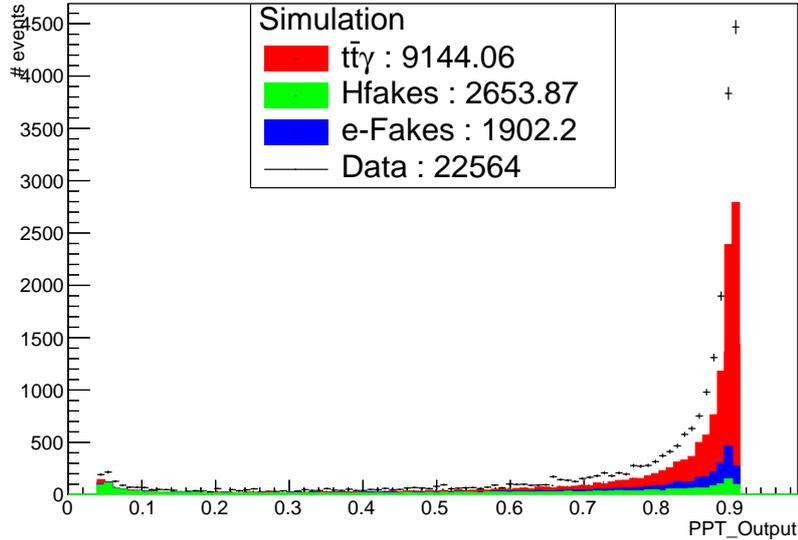
Figure 6.5.: PPT distribution of $t\bar{t}\gamma$ signal in the single lepton channel.

In Figure 6.5 the data points are on average 1.53 times higher than the MC prediction. There are different reasons for this disagreement between data and MC prediction in the single lepton channel.

First, the MC prediction in Figure 6.5 does not include every possible process, that contributes to the background. Mainly, the $W\gamma$ process (e.g. [35]) and other prompts can fake the signal in the detector causing the missing event numbers for the prediction.

Moreover, the fake processes (e-Fakes, hfakes) are only weighted with the MC weighting, which is not perfectly precise according to the probability of appearance.

7. Summary

In the following section, a short summary of the results achieved and presented in this thesis, and a brief outlook will be given.

The Monte-Carlo truth studies for looseID photons of different shower shape variables for $\gamma + jets$ and $dijet$ samples revealed that the separation S between prompt photons and photons originating from a hadronic decay or hadrons faking a photon signal (hfakes) varies between 10% and 46%, which is enough for training them in a NN. Moreover, it was shown that shower shapes depend on the photon kinematics like p_T and $convType$, which also corresponds to the performance of the used NN: version 21.0 of the PPT in different p_T, η and $convType$ bins.

The PPT was successfully implemented for the $t\bar{t}\gamma$ analysis at ATLAS at a centre-of-mass energy $\sqrt{s} = 13\,\text{TeV}$ for looseID photons in the single lepton channel. In Section 6.1 it was presented that the PPT is a tool, which can be used for the discrimination between prompt photons and hadron fakes. Different architectures were tested before, to find the best possible one. As it turns out, each tested model could be used for further analysis, but finally the architecture with 3 layers, a batch normalisation layer and softmax activation function with 6 features was chosen.

In the end, the measured data from 2015 and 2016 was compared to the MC prediction, containing $t\bar{t}\gamma$ events, hfakes ($t\bar{t}$, $Z + jets$, $W + jets$ and diboson processes) and electrons faking photon signals. It can be observed that the MC prediction contains in total approximately 9000 less events than the real data. It turns out, not all processes, contributing to the total background, like $W\gamma$ or other prompts were included. Additionally, the MC weighting can make for this high difference of event numbers.

A. Appendix

A.1. Ratio and sensitivity plots
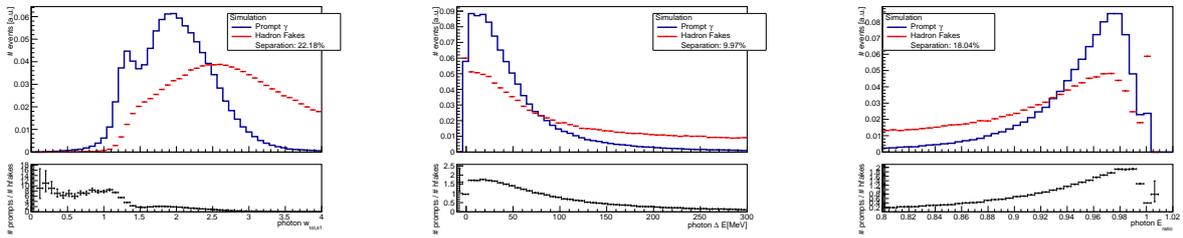
Figure A.1.: Three shower shape variables with ratio plots..
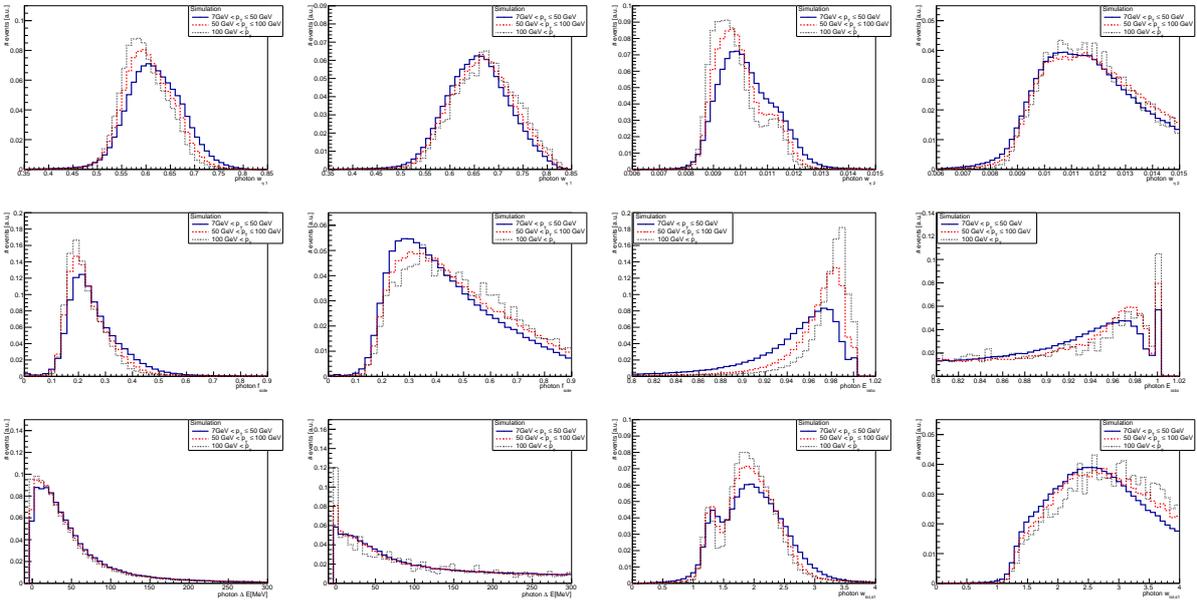
Figure A.2.: Shower shapes in 3 different p_T bins. Distributions of prompts on the left and hfakes on the right.
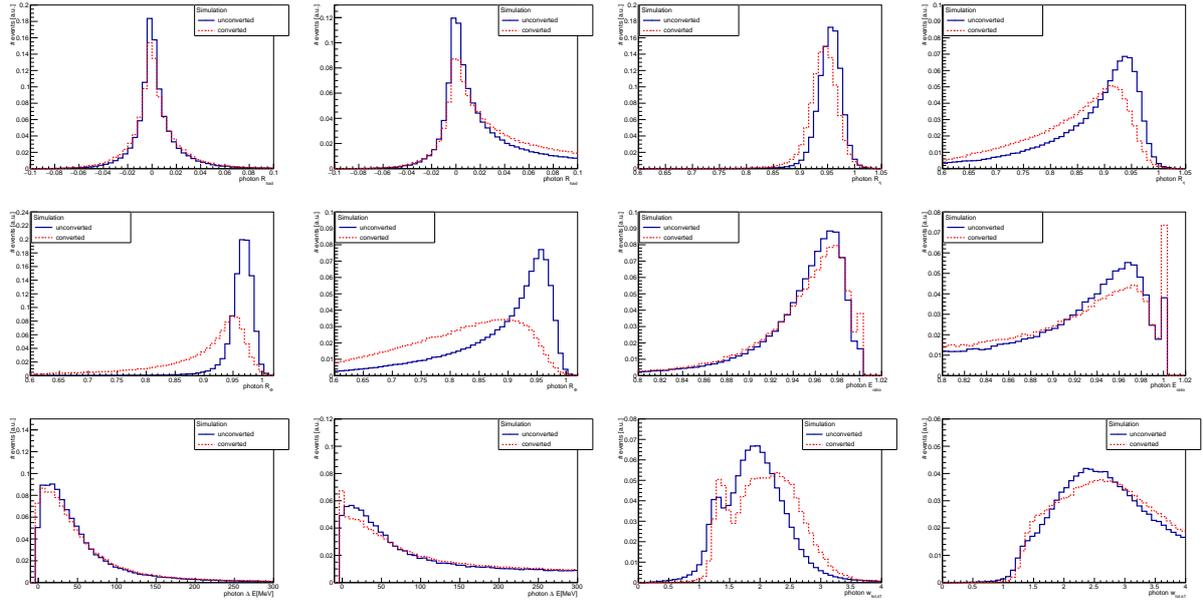
Figure A.3.: Shower shapes for different conversion types. Distributions of prompts on the left and hfakes on the right.

A.2. ROC-Curves
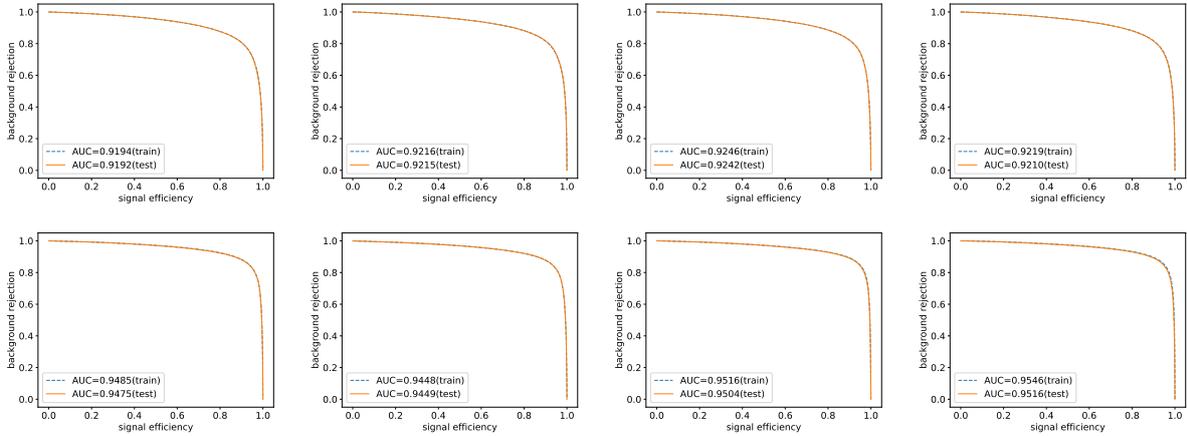
Figure A.4.: Roc curves for 8 different architectures.

A.3. Correlations, PPT sensitivity and Truth information

Table A.1.: Correlation values without including p_T, η and $convType$ in training.

Name	NN	R_{had}	R_η	R_Φ	$w_{\eta,1}$	$w_{\eta,2}$	f_{side}	p_T	η	$convType$
NN	100	-40.09	58.92	33.13	2.44	-32.70	-31.59	3.45	-0.21	-25.42
R_{had}	-40.09	100	-18.50	-4.54	-0.26	4.85	1.27	2.89	0.04	0.23
R_η	58.92	-18.50	100	23.26	2.95	-37.05	-7.98	-0.64	0.16	-17.57
R_Φ	33.13	-4.54	23.26	100	-0.24	-27.05	-23.53	2.76	0.14	-51.86
$w_{\eta,1}$	2.44	-0.26	2.95	-0.24	100	5.70	3.89	0.94	-0.08	2.37
$w_{\eta,2}$	-32.70	4.85	-37.05	-27.05	5.70	100	11.59	-2.74	-0.73	24.36
f_{side}	-31.59	1.27	-7.98	-23.53	3.89	11.59	100	-1.02	-0.14	19.92
p_T	3.45	2.89	-0.64	2.76	0.94	-2.74	-1.02	100	0.21	1.98
η	-0.21	0.04	0.16	0.14	-0.08	-0.73	-0.14	0.21	100	-0.20
$convType$	-25.42	0.23	-17.57	-51.86	2.37	24.36	19.92	1.98	-0.20	100

Table A.2.: Correlation values with including p_T, η and $convType$ in training.

Name	NN	R_{had}	R_η	R_Φ	$w_{\eta,1}$	$w_{\eta,2}$	f_{side}	p_T	η	$convType$
NN	100	-32.26	47.82	24.36	6.56	-27.08	-18.74	-22.52	-9.98	-16.91
R_{had}	-32.26	100	-18.50	-4.54	-0.26	4.85	1.27	2.89	0.04	0.23
R_η	47.82	-18.50	100	23.26	2.95	-37.05	-7.98	-0.64	0.16	-17.57
R_Φ	24.36	-4.54	23.26	100	-0.24	-27.05	-23.53	2.76	0.14	-51.86
$w_{\eta,1}$	6.56	-0.26	2.95	-0.24	100	5.70	3.89	0.94	-0.08	2.37
$w_{\eta,2}$	-27.08	4.85	-37.05	-27.05	5.70	100	11.59	-2.74	-0.73	24.36
f_{side}	-18.74	1.27	-7.98	-23.53	3.89	11.59	100	-1.02	-0.14	19.92
p_T	-22.52	2.89	-0.64	2.76	0.94	-2.74	-1.02	100	0.21	1.98
η	-9.98	0.04	0.16	0.14	-0.08	-0.73	-0.14	0.21	100	-0.20
$convType$	-16.91	0.23	-17.57	-51.86	2.37	24.36	19.92	1.98	-0.20	100
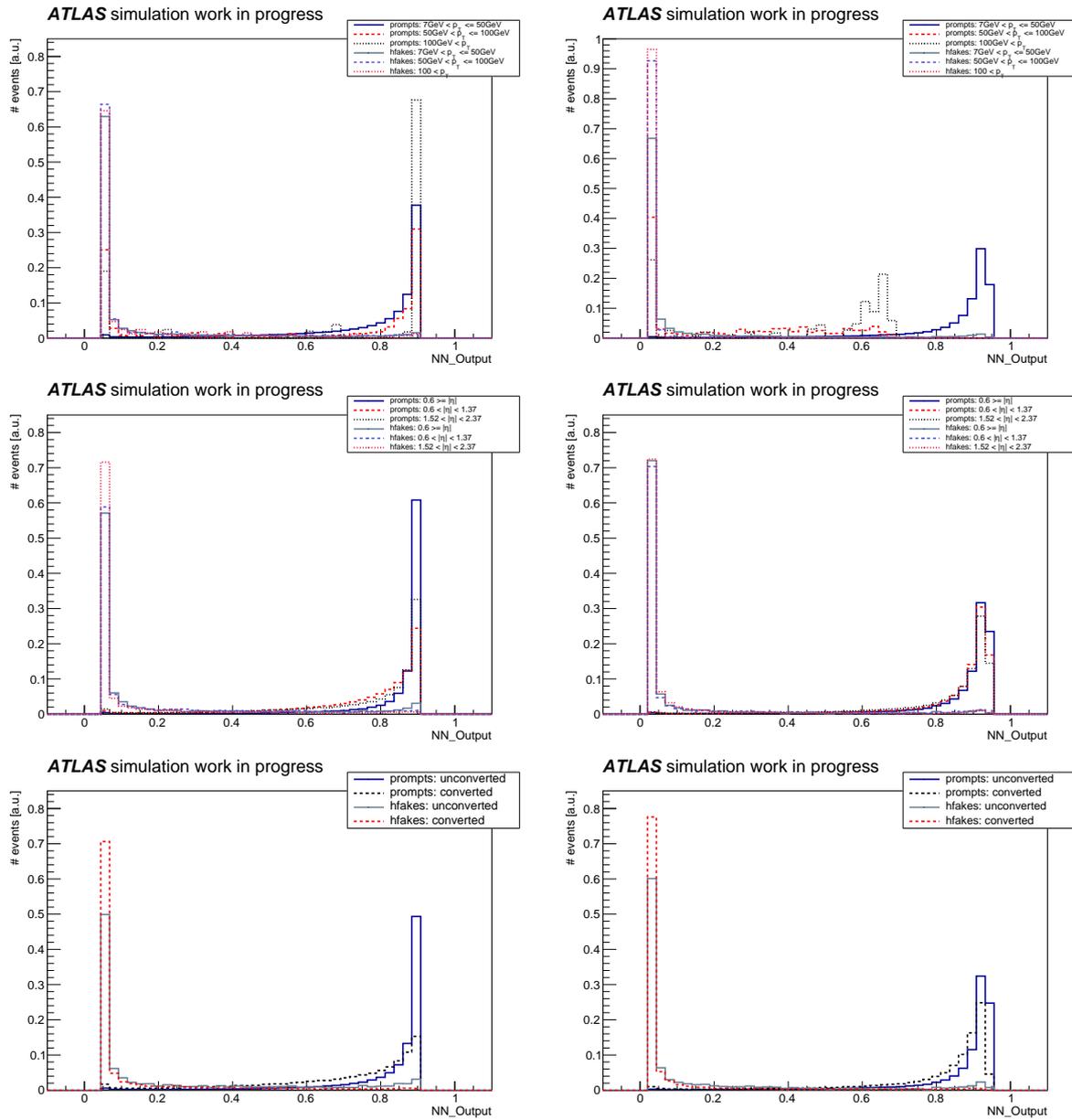
Figure A.5.: Sensitivity plots of the Ppt according to p_T, η and *convType*.

Table A.3.: Number of the truth information variable and the corresponding photon origin.

Origin	Number	Origin	Number	Origin	Number
NonDefined	0	WBosonLRSM	17	BottomBaryon	33
SingleElec	1	NuREle	18	PionDecay	34
SingleMuon	2	NuRMu	19	KaonDecay	35
SinglePhot	3	NuRTau	20	BremPhot	36
SingleTau	4	LQ	21	PromptPhot	37
PhotonConv	5	SUSY	22	UndrPhot	38
DalitzDec	6	OtherBSM	46	ISRPhot	39
ElMagProc	7	LightMeson	23	FSRPhot	40
Mu	8	StrangeMeson	24	NucReact	41
TauLep	9	CharmedMeson	25	PiZero	42
top	10	BottomMeson	26	DiBoson	43
QuarkWeakDec	11	CCbarMeson	27	ZorHeavyBoson	44
WBoson	12	JPsi	28	QCD	45
ZBoson	13	BBbarMeson	29	MultiBoson	47
Higgs	14	LightBaryon	30		
HiggsMSSM	15	StrangeBaryon	31		
HeavyBoson	16	CharmedBaryon	32		

Table A.4.: Number of the *convType* variable and its meaning.

Conversion	Number
unconverted	0
one track only, with Si hits	1
one track only, no Si hits	2
two tracks, both with Si hits	3
two tracks, none with Si hits	4
two tracks, only one with Si hits	5

Bibliography

[1] N. N. Taleb, *Der Schwarze Schwan: Die Macht höchst unwahrscheinlicher Ereignisse*, Carl Hanser Verlag GmbH Co. KG (2008)

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

[3] P. M. Watkins, *Discovery of the W and Z bosons*, Contemp. Phys. **27**, 291 (1986)

[4] S. M. Bilenky, J. Hosek, *Glashow-Weinberg-Salam Theory Of Electroweak Interactions And The Neutral Currents*, Phys. Rept. **90**, 73 (1982)

[5] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716**, 30 (2012)

[6] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716**, 1 (2012)

[7] G. 't Hooft, M. J. G. Veltman, *Regularization and Renormalization of Gauge Fields*, Nucl. Phys. **B44**, 189 (1972)

[8] Y. Sofue, *The Mass Distribution and Rotation Curve in the Galaxy*, PSSS **5**, 1 (2013)

[9] J. Wess, B. Zumino, *Supergauge Transformations in Four-Dimensions*, Nucl. Phys. **B70**, 39 (1974)

[10] J. Wess, B. Zumino, *Supergauge Invariant Extension of Quantum Electrodynamics*, Nucl. Phys. **B78**, 1 (1974)

[11] S. Dimopoulos, S. Raby, F. Wilczek, *Supersymmetry and the Scale of Unification*, Phys. Rev. **D24**, 1681 (1981)

[12] Y. Ashie, et al., *Evidence for an oscillatory signature in atmospheric neutrino oscillation*, Phys. Rev. Lett. **93**, 101801 (2004)

[13] D. Chakraborty, J. Konigsberg, D. Rainwater, *Top-Quark Physics*, Annu. Rev. Nucl. Part. Sci. **53**, 301 (2003)

Bibliography

[14] W. Wagner, *Top Quark Physics in Hadron Collisions*, Rept. Prog. Phys. **68**, 2409 (2005)

[15] A. Quadt, *Top quark physics at hadron colliders*, Eur. Phys. J. **C48**, 835 (2006)

[16] L. M. Lederman, *The discovery of the Upsilon, bottom quark, and B mesons*, in *The Rise of the standard model: Particle physics in the 1960s and 1970s. Proceedings, Conference, Stanford, USA, June 24-27, 1992*, pages 101–113 (1992)

[17] S. Abachi, et al (D0), *Observation of the Top Quark*, Phys. Rev. Lett. **74**, 2632 (1995)

[18] F. Abe, et al. (CDF), *Observation of top quark production in $\bar{p}p$ collisions*, Phys. Rev. Lett. **74**, 2626 (1995)

[19] CMS Collaboration, *Measurement of the top quark mass using proton-proton data at $\sqrt{(s)}$ = 7 and 8 TeV*, Phys. Rev. **D93(7)** (2016)

[20] S. M. Oliveira, et al., *Electroweak corrections to the top quark decay*, Phys. Rev. **D64**, 017301 (2001)

[21] E. Norrbin, T. Sjostrand, *Production and hadronization of heavy quarks*, Eur. Phys. J. **C17**, 137 (2000)

[22] CDF Collaboration, *Observation of electroweak single top-quark production with the CDF experiment*, PoS **HCP2009**, 083 (2009)

[23] R. R. Wilson, *The Tevatron*, Phys. Today **30N10**, 23 (1977)

[24] M. Breidenbach, et al., *Observed Behavior of Highly Inelastic electron-Proton Scattering*, Phys. Rev. Lett. **23**, 935 (1969)

[25] ATLAS Collaboration, *Observation of top-quark pair production in association with a photon and measurement of the $t\bar{t}\gamma$ production cross section in pp collisions at $\sqrt{s}$ = 7 TeV using the ATLAS detector*, Phys. Rev. **D91(7)** (2015)

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

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

[28] CMS Collaboration, *The CMS Experiment at the CERN LHC*, JINST **3**, S08004 (2008)

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

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

[31] M. Delmastro, et al., *Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run-2 data collected in 2015 and 2016*, Technical Report URL https://cds.cern.ch/record/2290912

[32] W. Lampl, et al., *Calorimeter clustering algorithms: Description and performance* (2008)

[33] B. Müller, J. Reinhardt, M. T. Strickland, *Neural Networks, An Introduction*, Springer-Verlag Berlin Heidelberg (1995)

[34] F. Rosenblatt, *The Perceptron: A Probabilistic Model for Information Storage and Organization in The Brain*, Psychol. Rev. **56(6)**, 65 (1958)

[35] B. Völkel, *Studies of the discrimination between prompt photons and hadron fakes using neural networks*, Master Thesis, II.Physik-UniGö-MSc-2017/07 (2017)

[36] *LightWeight Trained Neural Network*, https://github.com/lwtnn/lwtnn, visited on 12.07.2018

[37] ATLAS Collaboration, *Performance of b-Jet Identification in the ATLAS Experiment*, JINST **11** (2016)

Acknowledgment

Due to writing this thesis, I have learned numerous skills, which will be useful in my further life. Especially, the strategy and method how to work with data was fundamentally for writing this thesis.

First of all, I want to give thanks to Prof. Dr. Arnulf Quadt, who gave me the chance working in the Göttingen ATLAS group. The organisation and the whole environment in this group are really amazing (and also funny). I felt really nice incorporated and I want to recommend working in this group to everybody.

I would have never been able to write this thesis without the huge amount of help I got from Josh and Knut. Anytime I was in trouble and did not understand a special process or my laptop issued an error and I did not understand what was going on there (which happened really often), you two, Josh and Knut, helped me and explained me everything there was to know, although you were often really busy. Thank you really much, I owed you a lot of time and I could not imagine having better 'caretakers'.

Additionally, I have to say 'thank you' to Andreas, who helped me first understanding the basics of Neural Networks. This was really helpful.

Moreover I want to thank you, Royer, for giving me helpful comments and proofreading everything.

Thanks, to all institution members for the pleasant time and helpful and also funny conversations, we have had in the monday top quark meeting.

I want to thank Timo and Timon for making my days funnier. We have had nice conversations and a lot to laugh about.

At last, I want to thank my family and all my friends for giving me support and a nice time over the last years.

Erklärung nach §13(9) der Prüfungsordnung für den Bachelor-Studiengang Physik und den Master-Studiengang Physik an der Universität Göttingen:

Hiermit erkläre ich, dass ich diese Abschlussarbeit selbständig verfasst habe, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe und alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen wurden, als solche kenntlich gemacht habe.

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

Göttingen, den 28. August 2018

(Joey Kalis)