UNIVERSITY OF BELGRADE

FACULTY OF PHYSICS

Evelin Bakos

RADIATIVE W BOSON DECAY STUDIES AND THE UPGRADE OF THE ATLAS MUON SPECTROMETER READOUT SYSTEM

Doctoral Dissertation

Belgrade, 2024

UNIVERZITET U BEOGRADU

FIZIČKI FAKULTET

Evelin Bakoš

IZUČAVANJE RADIJATIVNIH RASPADA W BOZONA I UNAPREDJENJE SISTEMA ZA OČITAVANJE MIONSKOG SPEKTROMETRA DETEKTORA ATLAS

Doktorska disertacija

Beograd, 2024

Committee members:

prof. dr. Maja Burić
Full Professor
Faculty of Physics, University of Belgrade

dr. Duško Latas
Docent
Faculty of Physics, University of Belgrade

dr. Marija Vranješ Milosavljević
Research Associate
Institute of Physics, Belgrade

prof. dr. Wim Beenakker
Full Professor
Radboud University, Nijmegen

prof. dr. Antonio Pellegrino
Full Professor
University of Groningen

dr. Mengqing Wu
Docent
Radboud University, Nijmegen

prof. dr. Gerco Onderwater
Full Professor
Maastricht University

Abstract

This thesis incorporates three research topics, all related to the ATLAS detector at the Large Hadron Collider (LHC) at CERN. The first topic covers the contribution to the Data Acquisition (DAQ) of ATLAS, precisely the implementation of the MDT Read Out Driver (MROD) functionality to the Front-End Link eXchange (FELIX) readout system of the ATLAS detector. The second topic represents the search for the rare radiative decay of W boson to meson and a photon, analyzing data already collected with ATLAS during Run 2. The third topic outlines two machine learning algorithms to identify D mesons from radiative W boson decays at the LHC.

ATLAS is one of the general purpose detector placed on the LHC. ATLAS consist of many layers of sub-detectors which are designed to detect the different particles. The outermost layer of ATLAS, the muon spectrometer is made from four detection systems, each one exploiting different technologies. The major part of the muon spectrometer consist of Monitored Drift Tube (MDT) chambers which measures the properties (momentum, angular distributions, electric charge) of muons.

Within the DAQ system, during the previous data-taking periods the MDT chambers were readout via the MDT Read Out Driver (MROD). This readout system is dependent on MROD cards, which will reach end of their lifetime during Run 3. Since repairing the broken MROD cards or ordering new ones is not possible, a new implementation is needed which will use the new FELIX system. As part of the system, the swROD implements data fragment building and formatting. An additional module of the swROD, the `CSMProcessor` has been developed, which mimics the MROD building mechanism and incorporated the sub-detector specific settings which allows the configuration of the different MDT chambers, stores information about them, like their number or position. Multiple tests have been performed using the `CSMProcessor`, and the results show that the modified FELIX readout chain will be able to process the increased data rates in Run 3.

Using ATLAS, many experimental validations have been performed to investigate the accuracy and applicability of the Standard Model (SM). To test the SM as thoroughly as possible, searches for not (yet) detected decays and more precise measurements are necessary. Radiative decays of the W boson to $W \to M\gamma$, where M is a meson, are sensitive to the coupling of the W boson with the photon and, more importantly, probe the strongly coupled Quantum Chromodynamics (QCD) regime. The search for the $W \to \rho\gamma$, where a W boson decaying to a ρ meson and a photon, and other rare decays are experimentally challenging, due to the multijet background, which consists of hadronized quarks and gluons. No search for the $W \to \rho\gamma$ has been performed so far, so no previous bounds exist.

During the decay the ρ meson and the γ are decaying "back-to-back", sharing the momentum of the W boson. Since the lifetime of the ρ meson is short, it decays further to charged and a neutral π meson. Due to the nature of the ρ meson, isolated prompt ρ and a τ lepton decaying hadronically with exactly one charged and one neutral pion are indistinguishable within the detector, therefore the algorithm used to detect the τ candidate can be used to identify the ρ meson without any modification. After the initial trigger and object selection, processes with the same or similar final states are considered as background. The main sources of background are events involving inclusive photon + jet and dijet processes, where a track is reconstructed within a hadronic jet. This background cannot be reliably modelled with Monte Carlo (MC) simulation due to the complicated mixture of contributing processes. Instead, this contribution to the total background is modelled with a data-driven non-parametric approach. To extract the limit on the branching fractions of the $\mathscr{B}(W \to \rho\gamma)$ a binned maximum-likelihood is performed to the selected events. The search provides a limit of $\mathscr{B}(W \to \rho\gamma) < 6.29 \times 10^{-6}$ at 95% confidence level. This limit can be further improve if we include the search in the track-plus-photon final state where the ρ meson is identified as a track. Due to the different triggers and selection criteria, the tau-plus-photon and track-plus-photon final state are orthogonal, therefore the selected events can be combined in the final fit. This approach provides, for the first time, a limit of $\mathscr{B}(W \to \rho\gamma) < 5.17 \times 10^{-6}$ at 95% confidence level.

Reducing the background processes is also possible using the meson properties. However, it is possible to construct many of different variables, which makes it difficult to make a proper selection and measure their usefulness. To make things easier a machine learning algorithm can be developed specifically for meson tagging. One way to do this is to list many high-level variables, and using deep neural network create a classification algorithm, which is able to distinguish between signal and background mesons originating from quarks and gluons. An other approach is to develop a convolutional neural network using low level variables, such as the momentum of the particle and the energy deposited in the calorimeter. However, we can reach the best results if we combine the two exploiting the advantage of both models.The developed algorithm based on these models is able to identify jets originating from D_s mesons in radiative W decays and shows a good efficiency of 47% for signal with a 100 times rejection of jets from quarks and gluons. This presents an opportunity to improve measurements related to D_s mesons, particularly in the context of the rare decays.

Keywords: ATLAS detector; muon spectrometer; data acquisition system; standard model QCD processes; W boson rare decays; jet tagging; machine learning
Scientific field: Physics
Research area: High energy physics

Sažetak

Ova teza obuhvata tri teme koje se odnose na istraživanja na detektoru ATLAS na Velikom sudaraču hadrona (LHC) u CERN-u. Prva tema pokriva doprinos sistemu za prikupljanje podataka (Data Acquisition (DAQ)) detektora ATLAS, konkretno implementaciju MDT Read Out Driver (MROD) funkcionalnosti u Front-End Link eXchange (FELIX) sistem očitavanja. Druga tema predstavlja potragu za retkim radijativnim raspadom W bozona na mezon i foton. Analizirani su podaci iz proton-proton sudara na energiji $\sqrt{s} =$ 13 TeV prikupljenih pomoću detektora ATLAS tokom Run 2 perioda (2015-2018). Treća tema opisuje dva algoritma mašinskog učenja za identifikaciju D mezona nastalih u radijativnim raspadima W bozona na LHC-u.

ATLAS je jedan od dva detektora opšte namene na LHC-u. Sastoji se od više pod-detektora koji su dizajnirani da detektuju tragove naelektrisanih čestica, elektrone, fotone i džetove, kao i mione. Kombinujući navedene podatke mogu se meriti energije i uglovne raspodele navedenih objekata i odrediti znak naelektrisanja naelektrisanih čestica. Takodje, moguće je identifikovati hadronske raspade tau leptona i džetove koji potiču od b- ili c-kvarkova, a može se meriti nedostajuća energija. Najudaljeniji pod-detektor ATLAS-a, mionski spektrometar, napravljen je od četiri sistema za detekciju, od kojih svaki koristi različite tehnologije. Najveći deo mionskog spektrometra čine *Monitored Drift Tube* (MDT) komore.

U okviru sistema DAQ, tokom prethodnih perioda uzimanja podataka, MDT komore su očitavane preko *MDT Read Out Driver-a* (MROD). Ovaj sistem očitavanja zavisi od MROD kartica, koje će dostići kraj svog životnog veka tokom Run 3. Pošto popravka MROD kartica ili naručivanje novih nije moguća, potrebna je implementacija koja će koristiti novodizajnirani sistem za očitavanje podataka FELIX. Kao deo sistema, software ROD (swROD) implementira izgradnju i formatiranje fragmenata podataka. Razvijen je modul swROD-a, `CSMProcessor`, koji oponaša mehanizam izgradnje MROD-a i ugradjuje specifična podešavanja pod-detektora koja omogućavaju konfiguraciju različitih MDT komora, čuvanje informacije o njima, kao što su njihov broj ili položaj. Više testova je obavljeno korišćenjem `CSMProcessor`-a, a rezultati pokazuju da će modifikovani sistem za očitanja FELIX može da obradi povećane brzine prenosa podataka tokom Run 3.

Jedna od osnovnih tema istraživanja na eksperimentu ATLAS je testiranje Standardnog modela (SM). Jedan od mogućih testova SM-a čine potrage za (još) neotkrivenim raspadima gradijentnih (kao i Higsovog) bozona. Radijativni raspad W bozona $W \to M\gamma$ (gde je M mezon) osetljiv je na sprezanje W bozona sa fotonima i, što je još važnije, omogućava ispitivanje režima jako spregnute kvantne hromodinamike. Potraga za $W \to \rho\gamma$, gde se W bozon raspada na ρ mezon i foton, i drugi retki raspadi su eksperimentalno izazovni, zbog velikog fona koji postiče iz produkcije džetova. Na hadronskim sudaračima do sada ovaj kanal raspada nije izučavan, tako da ne postoje prethodne granice za faktor grananja $W \to \rho\gamma$.

Tokom raspada, ρ mezon i γ se raspadaju *"back-to-back"*, deleći impuls W bozona. Pošto je vreme života ρ mezona kratko, on se dalje raspada na naelektrisane i neutralne π mezone. Zbog prirode ρ mezona, izolovani ρ i tau lepton (koji se raspada hadronski sa jednim naelektrisanim i jednim neutralnim pionom) se ne mogu eksperimentalno razlikovati, pa se algoritam koji se koristi za detekciju tau može se koristiti za identifikaciju ρ mezona bez bilo kakvih modifikacija. Nakon inicijalnog trigera i selekcije čestica, procesi sa istim ili sličnim konačnim stanjima se smatraju za fon. Glavni izvori fona su dogadjaji koji uključuju inkluzivne procese foton + džet i didžet, gde se tragovi čestica rekonstruišu unutar hadronskog džeta. Ovaj fon se ne može pouzdano modelovati Monte Carlo (MC) simulacijama, već se koriste data-driven tehnike. Da bi se izračunao limit na faktor grananja $\mathscr{B}(W \to \rho\gamma)$ korišćen je pristup zasnovan na metodu maksimalne verodostojnosti. Dobijen je limit od $\mathscr{B}(W \to \rho\gamma) < 6.29 \times 10^{-6}$ na nivou poverenja 95% sa uračunatim sistematskim greškama. Ovo ograničenje se može dodatno poboljšava ako se u analizu uključi konačno stanje trag-plus-foton gde je ρ mezon identifikovan kao trag unutar detektora. Zbog različitih kriterijuma selekcije dogadjaji iz dva navedena konačna stanja su statistički nezavisni, pa se izabrani dogadjaji mogu kombinovati. Konačni dobijeni $\mathscr{B}(W \to \rho\gamma) < 5.17 \times 10^{-6}$ na nivou poverenja 95%.

Suzbijanje fonskih procesa u identifikaciji mezona koji sadrže charm kvark se može ostvariti korišćenjem različitih svojstava ovih mezona. Moguće je konstruisati veliki broj varijabli, što otežava pravi izbor i merenje njihove efikasnosti. Stoga je razvijen algoritam zasnovan na tehnici mašinskog učenja. Prvi izučavani pristup je primena duboke neuronske mreže za kreiranje klasifikacionog algoritma koji je u stanju da razlikuje D_s od mezona koji potiču od kvarkova i gluona. Drugi pristup je da se razvije konvoluciona neuronska mreža koristeći varijable niskog nivoa, kao što su impuls čestice i deponovana energija u kalorimetrima. Optimalan rezultat se postiže kombinovanjem dve navedene neuronske mreže. Dobijeni algoritam je u stanju da identifikuje džetove koji potiču iz D_s mezona u radijativnim W raspadima sa efikasnošću 47% za signal i faktorom potiskivanja džetova iz kvarkova i gluona od 100. Pokazano je da se na taj način na LHC-u mogu poboljšati merenja vezana za D_s mezone, posebno u kontekstu retkih raspada.

Ključne reči: ATLAS detektor; mionski spektrometar; sistem za prikupljanje podataka; QCD procesa standardnog modela; retkih raspada W bozona; džet tagging; mašinsko učenje
Naučna oblast: Fizika
Oblast istraživanja: Fizika visokih energija

Table of contents

Introduction

In 2020, when the COVID pandemic hit Europe and the borders got locked down, I found myself in the same small village in north Serbia where I grew up. To maintain my social life and my passion for physics, I decided to teach in my old elementary school, and to help students discover the world of physics for the first time.

One of topic most talked about by the students (besides black holes - they are always popular) was particle physics. Most of them were interested in what matter is made of, what the smallest particles are that we, particle physicists, study, and what the fundamental interactions between them are. They found it unbelievable that the proton is not an elementary particle and that the strong interaction prevents the nucleus from collapsing. I often told them about the Standard Model (SM), the theory that summarizes our current knowledge of particle physics, and the questions that even the SM could not yet answer. But I also often mentioned the measurements and searches that we carry out at CERN on the LHC, and that they often support the theory.

During my PhD work, I also performed such a search. The main topic of my dissertation is one of the rare decays of the W boson and it represents the first ever measured upper limit on the $W \to \rho\gamma$ branching fraction. These results improve our knowledge of the SM and expand our understanding of the properties of the W boson. This topic is further expanded with a development of a meson tagger based on machine learning algorithms.

In addition, this thesis also describes my contribution to the Data Acquisition (DAQ) of ATLAS, precisely the implementation of the MROD functionality to the Front-End Link eXchange (FELIX) readout system of the ATLAS detector. This enables to potentially replace MROD readout cards with the FELIX cards, while maintaining the required performance of the readout system.

Based on these topics, the thesis is organized as follows:

Chapter 1 introduces the foundations of the SM and summarises the properties of the W boson and its phenomenology at proton colliders with a particular attention to the exclusive W boson hadronic decays.

Chapter 2 contains the detailed description of experimental setup, which consists of the Large Hadron Collider (LHC) and the ATLAS detector at the CERN accelerator complex. A detailed explanation is given to the Muon Spectrometer (MS) and the ATLAS Trigger and Data Acquisition system as they present the basis of the MS current readout system. The LHC timeline and ATLAS detector upgrades are also outlined.

Chapter 3 presents the detailed description of the current MS readout system together with an explanation of the MDT Read Out Driver (MROD) data format. In addition it introduces a new solution for replacing this readout system, which combines the FELIX card together with a software based component, the swROD. In this context, I implemented a new capability to the swROD which mimics the MROD functionality, while still meeting the performance requirements. In addition, I was also actively participating in testing this new setup.

Chapter 4 describes the techniques used to reconstruct the various particles created in the proton-proton collisions within the ATLAS detector. A more detailed description is given to photons, jets and taus since the W boson decay studies outlined in this thesis depend on the precise and correct reconstruction of those objects.

Chapter 5 introduces the search for $W \to \rho\gamma$ decays, including the procedures for modelling ATLAS collision data, the selection, classification of signal and background events, as well as the statistical treatment which led to the obtained results. In this study, my work included developing and implementing a custom data derivation, measuring trigger efficiencies, performing and optimizing the data selection used to obtain the final estimates, as well as constructing the optimal background model in each analysis category. The results have been published in the EPS 2023 conference note [1] and in a paper [2].

Chapter 6 introduces a neural network based D_s meson tagger for the ATLAS detector. Besides the theoretical introduction to machine learning algorithms, this chapter describes the three different algorithms used to detect D_s mesons, their optimisation and performance, and the estimated upper limit for the $W \to D_s\gamma$ decay using this technique. In the context of this study, my work included the simulation and validation of the signal and background samples and the development and optimisation of the networks. In addition I've tested the network on other signal-like samples, estimated its performance and derived the simulation uncertainties. The results of the study have been published in a paper [3].

1 Theoretical overview

1.1 The Standard Model of particle physics

The Standard Model (SM) [4, 5] describes all known fundamental elements of the universe and the interactions between them. It has been formulated in the 20th century, based on the principles of Quantum Field Theory (QFT) [6] in order to provide a consistent and unified explanation to the increasing amount of experimental discoveries in the field of particle physics. The SM has been incredibly successful at predicting wide array of physics processes and surviving many high-precision tests. However, despite the agreement between the model and the experimental observations SM is considered incomplete. Most notably, gravity is entirely left out of the theory and SM provides no explanation of dark matter either. Furthermore, more recently experiments have found hints of new physics in the form of possible violations of lepton flavor universality [7]. Therefore, several models for physics Beyond Standard Model (BSM) were developed during the years, and many of them are being researched.

1.1.1 Particles of the Standard Model

Already during our first encounter with physics we learn, that there are four known fundamental forces in our universe: electromagnetism, the weak interaction, the strong interaction, and gravity. The most well known is electromagnetism, which occurs between particles with electric charge and it involves all physical phenomena related to electricity, magnetism, electromagnetic fields, light, and atoms. While the weak force is responsible for the weak decays of particles and weak neutral currents, the strong force is what holds protons and neutrons together, meaning that it is responsible for the formation of the nuclei. The SM describe almost all of these forces, except for gravity, which currently is only defined by the classical theory, through General Relativity. Although scientist crave for a single unified theory, the addition of gravity to the SM has been a long standing theoretical problem, due to its incompatibility with QFT. Fortunately, gravity is weak at the atomic scales, so its effects can be considered negligible in high energy physics experiments.

According to the SM, forces are carried by gauge bosons. These particles are the photon (γ), which mediates the electromagnetic force, the gauge bosons of the weak interaction (W^+, W^-, and Z^0), and finally the carrier of the strong force, the gluon (g). In addition, the SM has a single boson with spin value of 0, the Higgs boson. While it is not connected to any fundamental force, the Higgs field is responsible for the Brout-Englert-Higgs (BEH) mechanism [8] for electroweak (EW) symmetry breaking.

Besides forces, matter particles are also present around us, which we can detect via the forces interacting between them. These particles are called fermions. They can be further split into two categories, depending on whether they interact with the strong force. The first group consist of quarks, which interact with the strong force via their colour charge, which can take 6 discrete values (red, green, blue and anti-red, -green and -blue). The up-type quarks carry an electric charge of $+\frac{2}{3}|e|$ and are grouped in three flavours: up (u), charm (c), top (t). Similarly, the down-type quarks, which carry a negative charge of $-\frac{1}{3}|e|$ identified with three flavours: down (d), strange (s), and bottom (b). The second group of particles, referred as leptons, do not have a colour charge and therefore do not interact with the strong force. Leptons, with unitary electric charge ($|e|$) include three flavours: the electron (e), muon (μ) and tau (τ). They are paired with neutral leptons, referred to as neutrinos, and divided into ν_e, ν_μ, and ν_τ. Since neutrinos are electrically neutral, they only interact via the weak force.

The same division can be made according to the spin of the particles: bosons are associated with integer spin values $(S[0, 1, 2, ...])$, while fermions with half-integer spin $(S[1/2, 3/2, ...])$ values. Spin is a property of each particle and it describes its intrinsic angular momentum. Next to spin and other already mentioned properties just as mass, electric or colour-charge, some other properties which describe particles are their hyper-charge, their lepton and baryon number. Each particle has a corresponding antiparticle, which has the same mass and spin but opposite electrical charge, lepton and baryon number. However, it can also happen that a particle is its own antiparticle. A summary of the particle content in the SM can be seen in Figure 1.1.

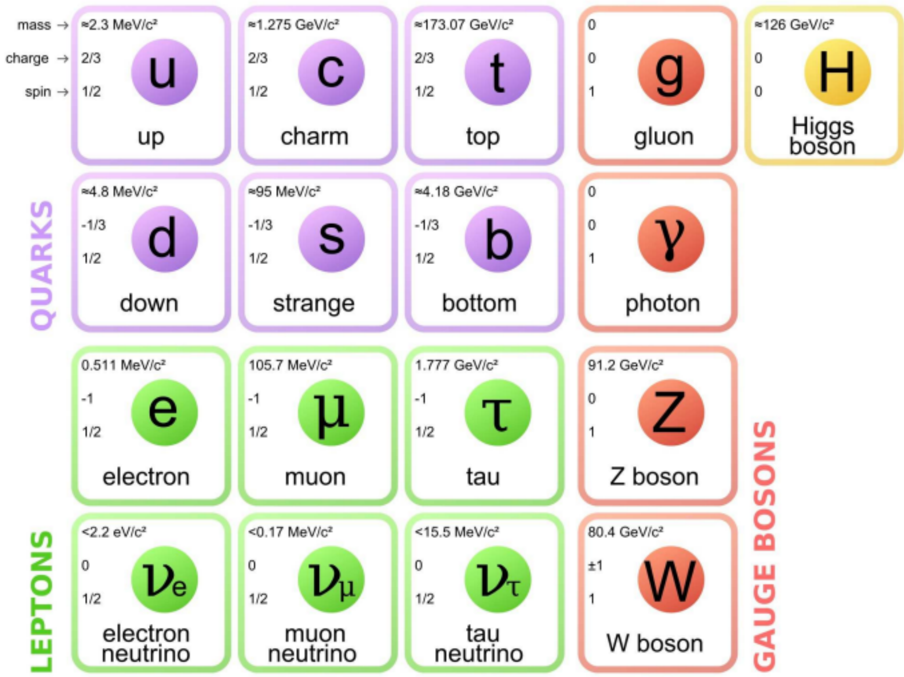


Figure 1.1: Summary of the SM elementary particles and some of their properties. For each particle, charge, mass, and spin are indicated. The 12 fermions are presented on the left side, while the 5 fundamental bosons are shown on the right side of the table [9].

1.1.2 The SM Lagrangian

The mathematical expression which describes the SM is provided by the Lagrangian formalism of the particle fields. The Lagrangian specifies the dynamics and kinematics of the model with the help of symmetries: the transformations under which the system is left invariant. The SM Lagrangian is invariant under the local symmetries of the $SU(3)_C \times SU(2)_L \times U(1)_Y$ group. Here the C defines colour charge, L is the weak isospin, Y defines hypercharge. Within the SM these parameters are conserved.

The SM Lagrangian is constructed as a generalization of Quantum Electrodynamics (QED). The dynamics of the free fermions are described by the Dirac equation of relativistic quantum mechanics, and the Lagrangian density associated to this equation is:

$$\mathscr{L}_{Dirac} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi \tag{1.1}$$

where $\psi = \psi(x)$ defines the spinor of a spin 1/2 fermion, γ^μ represents the Dirac γ-matrices and m is the mass of the particle. It is compulsory that the Lagrangian for a Dirac field is invariant under the $U(1)$ local transformation, corresponding to a rotation of the field phase by an angle $\theta(x)$:

$$\psi(x) \to \psi'(x) = e^{iq\theta(x)}\psi(x) \tag{1.2}$$

This is achieved by replacing the derivative ∂_μ with the covariant derivative D_μ:

$$\partial_\mu \to D_\mu = \partial_\mu + iqA_\mu \tag{1.3}$$

where A_μ is a vector field and is identified with the massless photon. This field is invariant under the gauge transformation

$$A_\mu \to A'_\mu = A_\mu - \partial_\mu \theta(x) \tag{1.4}$$

This allows to define the QED Lagrangian density the following way:

$$\mathscr{L}_{QED} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi + e\bar{\psi}\gamma^\mu\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{1.5}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the electromagnetic field tensor. The first term represents the fermion field, the second determines the interactions between fermions and photons, while the third term defines the photon field.

While QED expressed with the $U(1)$ local gauge symmetry, the symmetry associated to Quantum Chromodynamics (QCD) is the invariance under the $SU(3)$ local phase transformations. QCD is mediated by eight massless gluons, corresponding to the generators of $SU(3)$. Instead of a single charge, like in QED, QCD is characterised by the three colour and the three anti-colour charges, while the mediator of the strong force is the gluon field G_μ. The covariant derivative can be expressed as:

$$D_\mu = \partial_\mu - ig\frac{\lambda_a}{2}G^a_\mu \tag{1.6}$$

where where λ_a are the 8 generator Gell-Mann matrices and g is a constant related to the strong coupling constant, $\alpha_S = g^2/4\pi$. The QCD Lagrangian can be written in the form:

$$\mathscr{L}_{QCD} = \bar{\psi}(i\gamma^\mu \partial_\mu - m\delta_{ij})\psi - \frac{1}{4}G^a_{\mu\nu}G^{\mu\nu}_a \tag{1.7}$$

where m is the mass of the quark.

As experiments show, colour charged particles can only be found in a bound state, meaning that quarks always form composite particles. This phenomenon is known as colour confinement and affects the all particles with colour charge. The composite particles are called hadrons, which can be divided into two groups. Baryons are formed by three quarks and mesons made out of one quark and one antiquark. The mechanism by which quarks and gluons form the hadrons is called hadronisation.

Another important property of QCD is asymptotic freedom, which is related to the significant variation of the coupling constant α_s of the strong interaction, over the range of energies relevant to particle physics. In particular, α_s has large values for small energy scales, which correspond to large distances, making the theory non-perturbative. For higher energies, above ~100 GeV the coupling strength decreases, allowing perturbative methods to be used for QCD calculations. However, even in this range, the convergence of the perturbative expansion is very slow and requires higher-order corrections. A specific class of corrections, the quark self-energy corrections, could lead to infinities, known as ultraviolet divergencies, which can be eliminated by introducing a redefinition of fields or parameters to absorb the divergent term. This method is called renormalisation, which mitigates the problem by introducing a renormalisation scale μ_R and a factorisation scale μ_F. For all these reasons, QCD calculations for processes at the LHC are very challenging.

The SM unifies the weak and electromagnetic interactions into a single electroweak force. The weak charged-current interaction is associated with an SU(2)$_L$ gauge symmetry, whose generators are the weak isospin. The subscript L of the SU(2)$_L$ group indicates that the weak charged-current interaction only couples with left-handed particles and right-handed antiparticles. In this description, the left-handed particles and right-handed antiparticles are represented by weak isospin doublets, with total weak isospin $I_W = \frac{1}{2}$, while the right-handed particle and left-handed antiparticle states are described by weak isospin singlets where $I_W = 0$. The local gauge invariance can be satisfied by introducing three gauge fields, W^a_μ, with $a = 1, 2, 3$ corresponding to three gauge bosons. Linear combinations of these fields give rise to weak charged-currents, corresponding to the physical W bosons:

$$W^{\pm}_\mu = \frac{1}{\sqrt{2}}(W^{(1)}_\mu \mp iW^{(2)}_\mu) \tag{1.8}$$

An additional symmetry U(1)$_Y$ introduces the weak hypercharge $Y = 2(Q - I_W)$, where Q is the electromagnetic charge. A new gauge field B_μ couples to Y. In this model (named GSW after Glashow, Salam and Weinberg) the physical photon and Z boson result from linear combinations of the fields B_μ and $W^{(3)}_\mu$:

$$A_\mu = +B_\mu \cos\theta_W + W^{(3)}_\mu \sin\theta_W Z_\mu = -B_\mu \sin\theta_W + W^{(3)}_\mu \cos\theta_W \tag{1.9}$$

where θ_W is the weak mixing angle. Within the GSW model, the coupling constants of the weak and electromagnetic interactions, g and g', are related as follows:

$$e = g\sin\theta_W = g'\cos\theta_W \tag{1.10}$$

The EW Lagrangian is:

$$\mathscr{L}_{EW} = \bar{\psi}(i\gamma^\mu D_\mu)\psi - \frac{1}{4}W^a_{\mu\nu}W^{\mu\nu}_a - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \tag{1.11}$$

where D_μ is the covariant derivative for the electroweak interaction. It is defined as:

$$D_\mu = \partial_\mu + ig'\frac{Y}{2}B_\mu(x) + ig\frac{\tau_a}{2}W_\mu^a(x) \tag{1.12}$$

The SM description unifies the EW and QCD theories under the symmetry group $SU(3)_C \times SU(2)_L \times U(1)_Y$. The required local gauge symmetry can only be satisfied if all the particles are massless as the introduction of mass terms in the SM Lagrangian would break the local gauge invariance. However, experimental evidence shows that many particles have mass. It is then necessary to include a mechanism in the EW sector of the SM that spontaneously breaks the symmetry and gives rise to the mass of the particles. According to the BEH mechanism, the gauge bosons acquire mass upon their interaction with the Higgs field. The new field is identified with a weak isospin doublet of complex scalar fields with four degrees of freedom:

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^- \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \tag{1.13}$$

where ϕ^0 and ϕ^+ are a neutral field and a positively charged field, respectively. The corresponding Lagrangian is

$$\mathscr{L} = (D_\mu\phi)^\dagger(D^\mu\phi) - V(\phi) \tag{1.14}$$

where the first term describes the coupling of the gauge bosons to the Higgs field and $V(\phi)$ represents the Higgs potential.

1.2 Phenomenology of pp collisions

Protons are composite subatomic particles made of three, two up type and one down type valence quarks between other virtual quark-antiquark pairs and gluons. This means that in the pp collisions not the proton itself but its components, the partons actually interact. This makes the collision of two protons, also called and event, a complex process which consists of several processes as shown in Figure 1.2.

The two partons with the highest momentum transmission belong to the hard scatter (HS). This is the most important part of the event, since here heavy particles can form due to the large momentum transfer. These heavy particles then decay within the detector and their decay products can be detected. Interactions caused by the other partons are forming the underlying event. Particles originating from the underlying events also can be detected. In addition, since most of the partons are coloured and/or charged, they also emit additional gluons and/or photons. This can occur both before and after the HS, thus defined as Initial State Radiation (ISR) and Final State Radiation (FSR) respectively. Within the event, every generated gluon is able to emit additional gluons or generate quark-antiquark pairs. This process describes the Parton Shower (PS), where the partons lose their initial energy. Due to the color confinenement, after the partons lose enough energy, they hadronize into colourless states which can decay further. The decay products of these hadrons are also detected by the detector. All of the detected particles originating from the different processes provide a complete picture of an event making the reconstruction possible.

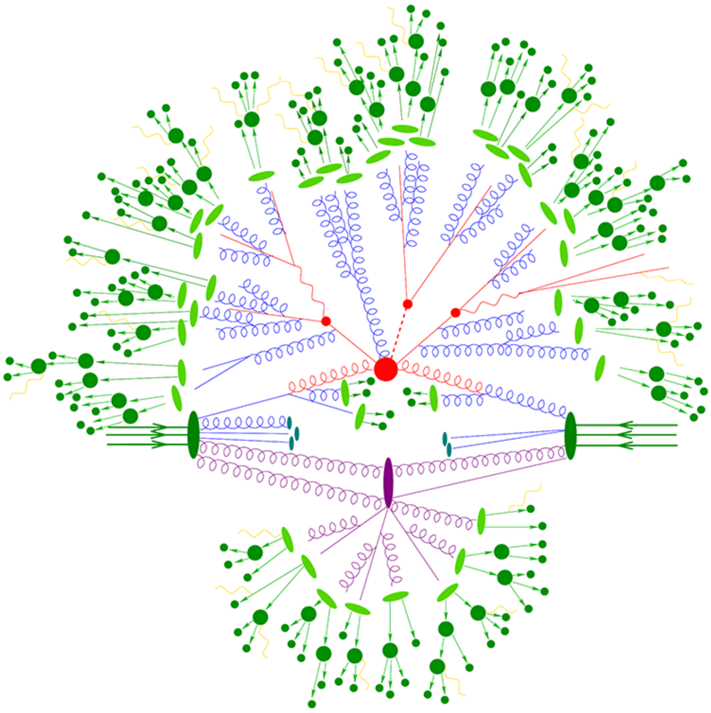


Figure 1.2: Pictorial representation of pp collision. The HS depicted by the red circles. Beside the HS, QCD radiation, shown by the light green circles, is produced, before the final-state partons hadronize. These hadrons also decay, depicted by the dark green circles. The underlying event is also shown on the picture, coloured with purple [10].

The description of the particle production at the LHC relies on the definition of cross section in hadronic collisions. The factorization theorem, first formulated by Drell and Yan [11], states that in hadronic collisions the cross section of a hard scattering process can be separated into two parts. The first one is the partonic cross section, while the second is an universal part corresponding to the distribution of partons inside the colliding hadrons, described by the Parton Distribution Function (PDF). If hadrons p_1 and p_2 interact to produce X, the cross section for the process $\sigma_{p_1 p_2 \to X}$ can be written as the convolution of the cross section of the intervening partons i and j, $\hat{\sigma}_{ij \to X}$, and the PDFs of the hadron, $f_{i,p_1}(x_i)$, for parton i in hadron p_1, and similarly $f_{j,p_2}(x_j)$ for parton j in hadron p_2.

The partonic cross section $\hat{\sigma}_{ij \to X}$ can be written as a power series expansion in the coupling α_s, where each order of the strong coupling corresponds to contributions from higher order emissions in the perturbation theory. It may thus be written in terms of leading order (LO), next-to-leading order (NLO), next-to-next-to-leading order (NNLO) and so on. Terms beyond leading order, the higher order corrections, give rise to logarithmic divergences because of the soft and collinear gluon emissions. These logarithms can be absorbed into the definition of the PDFs, and the factorized cross section can be re-written in terms of renormalized PDFs, depending on μ_F, as:

$$\sigma_{p_1p_2\to X}(s,\mu_F^2,\mu_R^2) = PDF \times \hat{\sigma}_{ij\to X}(\hat{s},\mu_F^2,\mu_R^2)$$
$$= \sum_{ij}\int dx_1dx_2 f_{i,p_1}(x_i,\mu_F^2)f_{j,p_2}(x_j,\mu_F^2)\hat{\sigma}_{ij\to X}(\hat{s},\mu_F^2,\mu_R^2), \quad (1.15)$$

where x_i and x_j are the momentum fractions of hadrons p_1 and p_2 carried by partons i and j respectively. s denotes the squared centre of mass energy of the collision, while $\hat{s} = sx_1x_2$ is the fraction of it that is used in the HS. μ_F defines the scale separating the long and short-distance physics, where perturbative and nonperturbative calculations apply. Similarly the scale dependence of α_s can be explicitly expressed via μ_R.

As an example, at the LHC W bosons at leading order are produced from the annihilation of a quark-antiquark pair. Charge conservation requires an up-type and a down-type quark to interact. In the case of resonant scattering, the scale of the process is directly related to the momenta of the incoming partons and the W boson mass:

$$\mu_R^2 = \mu_F^2 = sx_ix_j = M_W^2. \quad (1.16)$$

Since QCD becomes unperturbative for low energies, the PDFs cannot be calculated analytically. Thus to get some information about the PDF at some well understood energy scales experimental data from precision measurements, such as the deep inelastic scattering measurements are used. At other energy scales an extrapolation is used. An example of a PDF can be found in Figure 1.3, what is measured from high-precision collider data described in [12].

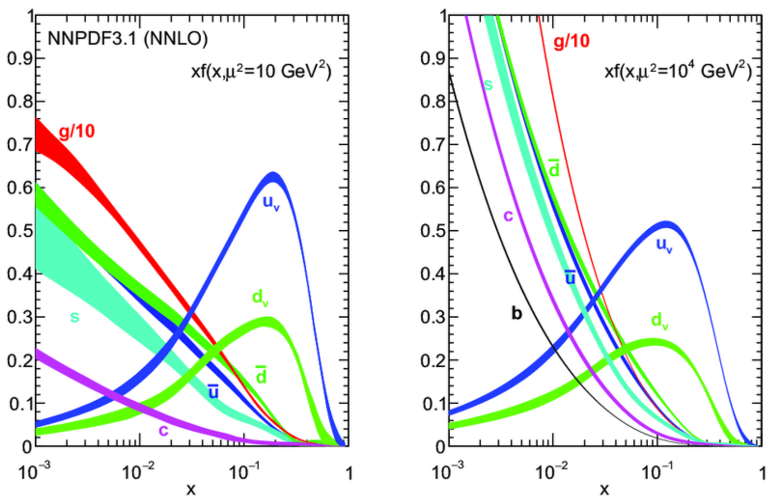


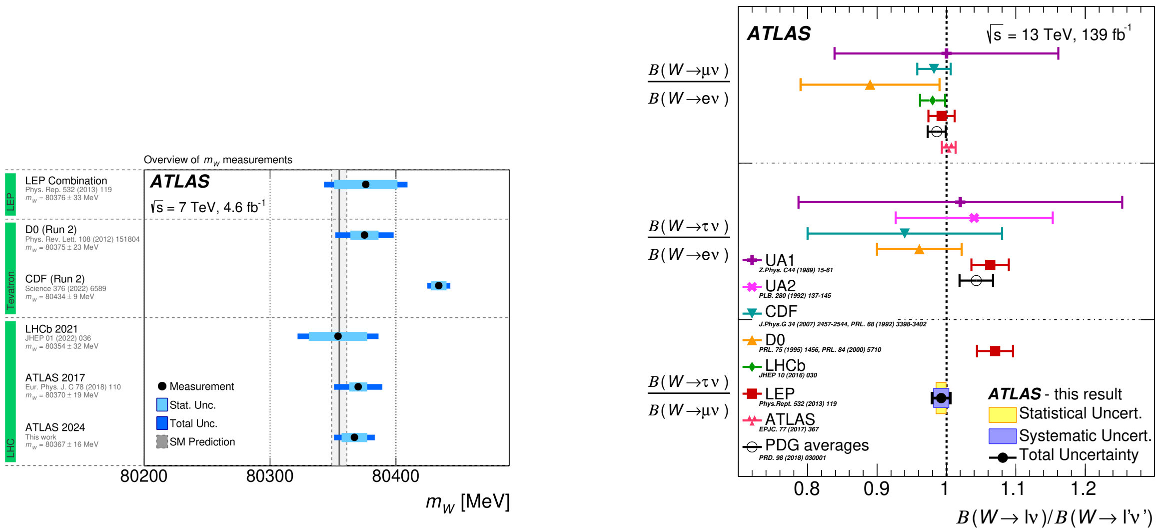
Figure 1.3: Example of an NNLO PDFs, evaluated at $\mu_F^2 = 10$ GeV^2 and $\mu_F^2 = 10^4$ GeV^2. Here, the PDF is shown on the y-axis, the momentum fraction on the x-axis and the width of the bars indicate the errors [12].

In the case of our previous example, Figure 1.3, shows that at leading order in α_S for W boson production is $u + \bar{d} \to W^+$ and $d + \bar{u} \to W^-$. Since in the lowest order there are 2 u quarks for 1 d quark in the proton, the u quarks together carry more momentum than the single d quark, $x(u) > x(d)$, it is kinematically more likely for a $u\bar{d}$ combination to satisfy the equation above, than a $d\bar{u}$ combination. As a result more W^+ relative to W^- are produced in pp collisions.

1.3 W bosons at the LHC

The W boson was discovered at the CERN SPS collider in 1983 [13], which led to a physics Nobel prize in 1984. The properties of W bosons have been studied for over 30 years and many of them have been measured with high precision.

The W boson mass had been measured at electron-positron and proton-antiproton colliders, as well as at the LHC, yielding a combined world average of $m_W = 80379 \pm 15$ MeV consistent with the SM constraints of 80356 ± 8 MeV [14]. The latest measurement is set by the CDF collaboration. The W boson mass is measured to be 80433 ± 9 MeV with a precision of 0.01% [15]. This means a significantly higher mass than the SM predicts as it is shown on Figure 1.4a.



(a) **(b)**

Figure 1.4: (a) Summary of the W boson mass measurements where the grey line presents the SM prediction [16]. (b) Measured value of the ratio of the different decay rates. A vertical dashed line indicates the SM prediction of equal branching ratios to different lepton flavours [17].

The W boson decays almost immediately, in 3×10^{-25} s with a total width measuring $\Gamma_W = 2.085 \pm 0.042$ GeV [5]. This means that the W boson can never be observed directly, and must be reconstructed from its decay products: lepton-neutrino or to quark anti-quark pairs. The total hadronic branching fraction is approximately two-thirds, (67.41 $\pm$ 0.27 %) dominated by $u\bar{d}$ and $c\bar{s}$ decays. These diquark decays lead to a final state consisting of two hadronic jets, which are the signatures of the quarks and gluons produced in the collision. At hadron colliders it is extremely difficult to separate this final state from the vastly larger multijet background, which consist of any number of jets and other hadronic activity, therefore, the leptonic decays are analysed instead. This mainly refers to measurements involving only electrons and muons since these particles can be directly detected in the detector. Due to its short lifetime, only the decay products of the tau can be detected, but despite this, ATLAS performed a cross section measurement involving taus as well, which found to be consistent with the $e\nu$ and $\mu\nu$ results [18]. The leptonic decays, to $\ell\nu$ have almost equal branching fractions in the electron, muon and tau channels which equals to 10.86 $\pm$ 0.09 % [5].

1.3.1 Radiative hadronic decays

Although quite rare, the exclusive hadronic W boson decay allows the quark-anti-quark pair to turn into one or more mesons. In some cases, the formation of the meson is only allowed by the emission of photon by the W boson or one of the two quarks. These decays referred as radiative decays and their representation is displayed in Figure 1.5. These radiative decays of $W \to M\gamma$, where M is a meson, are sensitive to the coupling of the W boson with the photon and, more importantly, probe the strongly coupled QCD regime.

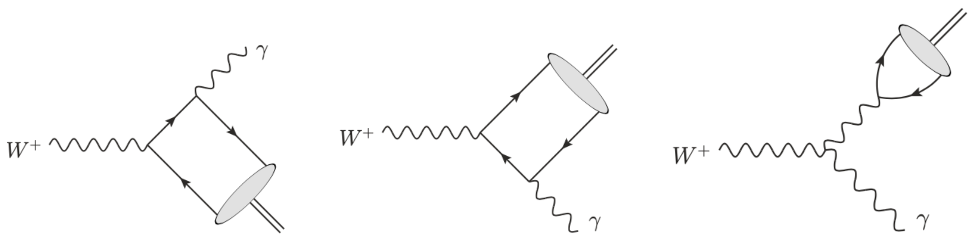


Figure 1.5: Representation of the hadronic W boson decay [19].

Mesons produced in these decays can be both pseudoscalar or vector. A pseudoscalar meson is a meson with 0 total spin and odd parity. Pseudoscalar mesons are, for example, the pion (π), kaon (K) and the D mesons. In contrast, vector mesons have total spin of 1 and also odd parity. An example of the vector meson is the rho (ρ) meson.

The accurate theoretical description of QCD in the transition from perturbative to non-perturbative, is crucial for the study of various processes ranging from $B \to X\ell\nu$ decays to the exclusive Higgs boson decays, which gives enough motivation to study them. The observation of these decays would improve the QCD factorization formalism, and enhance the possibility to perform precise calculations using such an approach. The basis of the most accurate to date prediction for the radiative decay branching fractions is a factorization theorem derived in soft-collinear effective theory. This approach expresses the decay amplitudes as convolutions of calculable hard-scattering kernels with light-cone distribution amplitudes (LCDAs), in a systematic expansion in powers of $(\Lambda_{QCD}/m_W)^2$ and $(m_M/m_W)^2$, where Λ_{QCD} represents the strong coupling constant [19]. The theory also includes the complete set of one-loop QCD radiative corrections. The leading power amplitudes for $W \to M\gamma$ process for final-state meson with a 4-momentum k and photon with a 4-momentum q is given:

$$i\mathscr{A}(W^+ \to M^+\gamma) = \pm\frac{egf_M}{4\sqrt{2}}V_{ij}\left(i\epsilon_{\mu\nu\alpha\beta}\frac{k^\mu q^\nu \varepsilon_W^\alpha \varepsilon_\gamma^{*\beta}}{k \cdot q}F_1^M - \varepsilon_W^\perp \cdot \varepsilon_\gamma^{\perp *} F_2^M\right). \tag{1.17}$$

Here the W polarisation is ε_W and the photon polarisation is noted with ε_γ. g is the coupling constant of the weak interaction and V_{ij} represents the Cabibbo-Kobayashi-Maskawa (CKM) matrix which describes the probability of a transition from one flavour quark to another. f_M is the decay constant of the meson and F_1^M, F_2^M are the form factors defined in Ref. [19]. The first term in the brackets describe the decays originating from the first two Feynman diagram in Figure 1.5, while the second term describing the third additional process, which is possible because the W boson has a direct coupling to the photon. Here the final-state meson is produced by the conversion of an off-shell W boson. The upper sign in Equation 1.17 refers to the case when the meson is pseudoscalar, while the lower sign refers when the meson is a vector meson. Averaging over the polarization states of the W boson, and summing over the photon polarization, the corresponding decay rate is given:

$$\Gamma(W^+ \to M^+\gamma) = \frac{\alpha m_W f_M^2}{48 v^2} |V_{ij}|^2 (|F_1^M|^2 + |F_2^M|^2). \tag{1.18}$$

Here α = 1/137.036 is the fine-structure constant evaluated at $q^2 = 0$, and v denotes the Higgs vacuum expectation value, which enters through the relation $(g/\cos\theta_W)^2 = 4m_W^2/v^2$, which can be solved with

$$v = v(m_W) = m_W \frac{\sin\theta_W \cos\theta_W}{\pi\alpha(m_W)}. \tag{1.19}$$

The form factors F_i^M are given in terms of overlap integrals of calculable hard-scattering coefficients with LCDAs. Solving Equation 1.3.1, it can be seen that the branching fraction is suppressed by a factor $(f_M/m_W)^2$ resulting in values of $\mathscr{B}(W \to M\gamma) \approx 10^{-9}$. The predicted branching fractions for various $W \to M\gamma$ decays are presented in Table 1.1. Significant uncertainties in the calculation arise from the hadronic input parameters, in particular from the meson decay constants and from the intrinsic limitations of the calculation. In addition, the calculations in Ref. [19] also shows, that the decay to a transversely polarized vector meson is strongly suppressed by a factor of $(m_M/m_W)^2$.

$W^\pm \to \pi^\pm\gamma$	$(4.00 \pm 0.83) \times 10^{-9}$
$W^\pm \to \rho^\pm\gamma$	$(8.74 \pm 1.91) \times 10^{-9}$
$W^\pm \to K^\pm\gamma$	$(3.25 \pm 0.69) \times 10^{-10}$
$W^\pm \to K^{*\pm}\gamma$	$(4.78 \pm 1.15) \times 10^{-10}$
$W^\pm \to D_S^\pm\gamma$	$(3.66 \pm 1.49) \times 10^{-8}$
$W^\pm \to D^\pm\gamma$	$(1.38 \pm 0.51) \times 10^{-9}$
$W^\pm \to B^\pm\gamma$	$(1.55 \pm 0.79) \times 10^{-12}$

Table 1.1: Predicted branching fractions for various $W \to M\gamma$ decays taken from [19].

The expected cross section for the $W \to M\gamma$ process at the LHC can be estimated as:

$$\sigma(pp \to W \to M\gamma) = \frac{\sigma(pp \to W \to \ell\nu)}{\mathscr{B}(W \to \ell\nu)} \times \mathscr{B}(W \to M\gamma), \tag{1.20}$$

In case of the W boson cross sections measured by ATLAS at $\sqrt{s} = 13$ TeV [20], the value for $\mathscr{B}(W \to \ell\nu) = 0.1086$ and for $\mathscr{B}(W \to M\gamma)$ is taken from Table 1.1, the predicted rate is $\sigma(W^- \to \ell\nu) = 8540$ pb and $\sigma(W^+ \to \ell\nu) = 11540$ pb.

Although, approximately 3·10^{10} W bosons have been produced at the Large Hadron Collider (LHC), the background originating from various multijet processes and the trigger challenges makes many precision studies of W decays extremely difficult. With the data collected from pp collisions during so-called Run 2 (2015-2018) of the LHC data taking the amount of expected events is roughly 100 for $W \to \pi\gamma$, 10 for $W \to K\gamma$ and 200 for $W \to \rho\gamma$. This could be doubled for the end of the following Run 3, meaning around 500 expected events for $W \to \rho\gamma$, 250 events for $W \to \pi\gamma$ and 24 events for $W \to K\gamma$. For the high luminosity LHC the expected amount of events could increase ten times. The most promising case is the decay $W \to D_s\gamma$, of which over 10k events should be produced on the high luminosity LHC.

The most straightforward way is to study these decays directly through they decay products. The $W \to M\gamma$ decay follows the $A \to B + C$ decay type, called the two-body decay. The general kinematics of two-body decay are best described in the centre-of-mass frame, where the decaying particle A is at rest. Conservation of 4-momentum implies that particle B and C are emitted back-to-back, with their 3-momenta being equal and opposite. Since the W boson is produced on the kinematic threshold, the behaviour of the decay products are approximately the same as in the centre-of-mass frame of the W boson. Given that the produced meson is light, and the photon is massless, the momenta of the decay products are approximately the half of the W boson mass. In cases when the produced meson decays further, due to the large momentum of the meson, the decay products are collimated, sharing the momentum of the parent meson. This means that in the case of $W \to \pi\gamma$ and $W \to K\gamma$ one needs to detect the photon and the pion or kaon directly. In case of the $W \to \rho\gamma$, the rho meson can be detected via it's decay products, most commonly via $\pi^{\pm}\pi^0$.

With this approach one can select all the events with the correct final states and kinematic signature. The main difficulty arises from the trigger: single photon triggers either have too high threshold or large prescales (thus not suitable fo searches), while the photon+meson triggers are too generic for such a study, allowing too much background process to pass the trigger selection. It is possible to develop dedicated triggers for such a processes but often these triggers will not allow the study of the $W \to \rho\gamma$ final state due to the excess π^0 in the final state. However, even with a correct trigger selection, one needs to deal with the large amount of multijet background. These processes are not precisely simulated with the most used methods, so a more advanced background modelling approach is also in need.

An other approach is to exploit the large cross section for $t\bar{t}$ production at the LHC. This process offers a relatively clean environment for a study the W boson properties: after identifying the leptonic decay of one of the W and the two b jets, an additional W still remains in the event. Top quark pair production also comes with a limited trigger bias. Triggering on two b-jets and the leptonic decay of one W boson suppresses the majority multijet background. The disadvantage is that the amount of W bosons created through the $t\bar{t}$ production is much smaller making the result more statistically limited. Ref [21] shows that it is unlikely to perform such a study even with the HL-LHC. This makes the previous approach more probable and favourable.

Due to the difficulties already mentioned, no exclusive hadronic decays have been observed yet. Direct experimental bounds exist for three exclusive hadronic decays: $W \to \pi\gamma$, $W \to \pi\pi\pi$ and $W \to D_s\gamma$. The upper limit for $\mathscr{B}(W \to \pi\gamma) < 7.0 \times 10^{-6}$ [22] is set by the CDF collaboration at the Fermilab Tevatron collider while the $\mathscr{B}(W \to \pi\pi\pi) < 1.01 \times 10^{-6}$ set recently by the LHC CMS collaboration, at 95%

CL [23]. The best upper limit for $W \to D_s\gamma$ is set by the LHCb collaboration with the value $\mathscr{B}(W \to D_s\gamma) < 6.4 \times 10^{-4}$ at 95% confidence level [24]. The limit is obtained analysing $K^+K^-\pi^+$ final states, which make up 5.4% of the D_s decays. This improves on an earlier limit of $\mathscr{B}(W \to D_s\gamma) < 1.3 \times 10^{-3}$ set by the CDF collaboration [25], using only $\phi(K^+K^-)\pi^+$ and $K^{*0}K^+$ final states, which comprise 3.9% of all D_s decays.

2 The LHC and the ATLAS experiment

The search for the W boson hadronic decay described in Chapter 5 uses data collected by the ATLAS experiment, while the D_s tagger introduced in Chapter 6 is developed for the ATLAS detector as well. In addition, the upgrade project described in the next chapter also take place on the ATLAS detector. Hence, in this chapter a short introduction of the LHC and the ATLAS experiment is provided.

2.1 Large Hadron Collider

The Large Hadron Collider (LHC) [26] is located at the European Organisation for Nuclear Research (CERN) near Geneva. Is is a circular hadron accelerator with a circumference of 27 km. LHC is designed to collide proton beams at a centre of mass energy up to 14 TeV. Proton beams are brought to collide head-on at four points along the ring, where the four particle physics experiment are located. With the LHC it is also possible to accelerate beams with heavy ions. Both lead-lead and proton-lead collisions are performed on a regular basis, but LHC also performed xenon-xenon collisions as well.

2.1.1 The accelerator complex

At the beginning of the LHC chain hydrogen atoms are exposed to an electric field, which strips of the electrons yielding bare protons. The Linear Accelerator (LINAC2) accelerates these protons to 50 MeV using Radiofrequency (RF) cavities. From the LINAC2 protons are transferred to the Proton Syncrotron Booster (PSB), where they are accelerated from 50 MeV to 1.4 GeV. The protons are then injected into the Proton Synchotron (PS) which accelerates them up to 25 GeV. Following the PS, protons are further accelerated in the Super Proton Synchotron (SPS). SPS is 7 kilometres in circumference, and accelerates protons to 450 GeV. Finally, SPS injects particles into the LHC which accelerates them further to their collision energy. This is achieved with eight RF cavities and over 8000 superconducting magnets. The magnetic field, created by the magnets, curve the particles trajectories and force them onto a circular orbit via the Lorentz force. To make a collision occur between particles with the same electric charge, two different magnetic fields are used to circulate the beams in opposite directions.

The beams can collide in four points of the LHC ring, which is where the particle detectors are located. ATLAS [27] and CMS [28] are general purpose experiments designed to primarily study and measure the SM parameters and look for new physics beyond the SM [29, 30]. Two further dedicated experiments, ALICE [31] and LHC-b [32] are designed and specially equipped to study heavy ion collisions and physics related to the b-quark respectively. In addition, five smaller experiments on the LHC: TOTEM, LHCf, MoEDAL, FASER and SND@LHC. TOTEM measures the total (in)elastic pp cross section, LHCf studies collision remnants very close to the beam-pipe, MoEDAL is searching for magnetic monopoles, FASER is designed to search for light and extremely weakly interacting particles, while SND@LHC detects and studies neutrinos. A schema of the accelerator complex is shown in Figure 2.1.

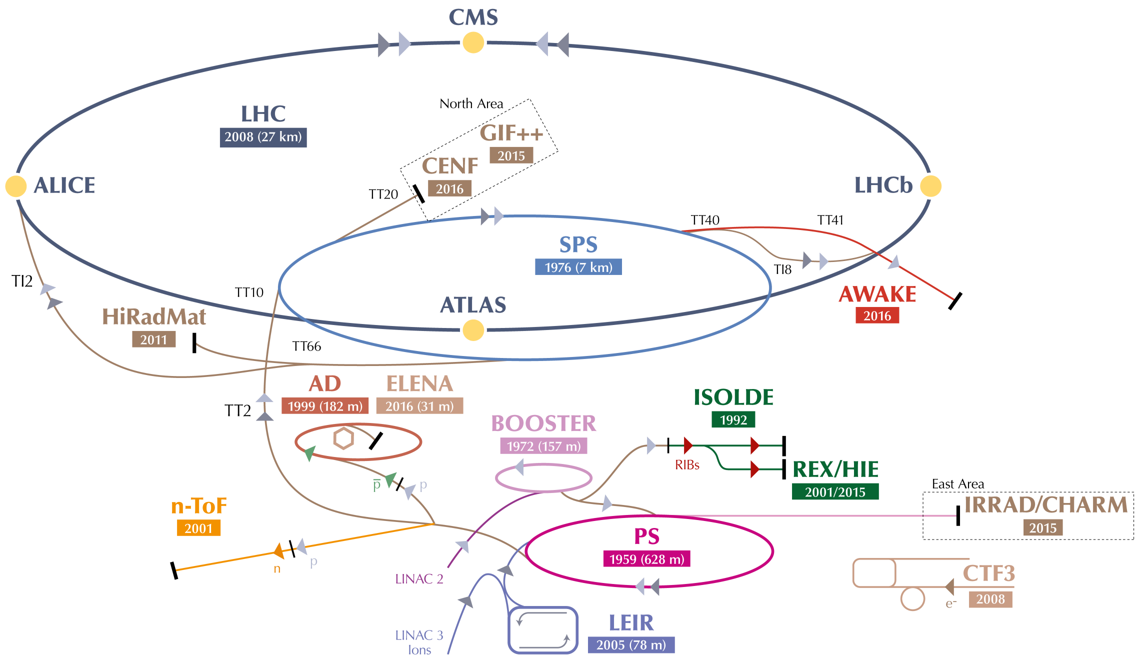


Figure 2.1: Schematic representation of the CERN accelerator complex [33].

2.1.2 Luminosity and pile-up

One of the main tasks of the LHC is to discover rare processes and particle decays. To be able calculate the event rate at which a given process occurs, one needs to determine the the delivered instantaneous luminosity $\mathscr{L}$, which is equal:

$$\mathscr{L} = \frac{N_b N_p^2 f_{rev}}{4\pi\sigma_T^2}, \tag{2.1}$$

where N_b denotes the number of bunches, N_p is the number of protons per bunch, f_{rev} is the revolution frequency and σ_T is the transverse beam size at the interaction point.

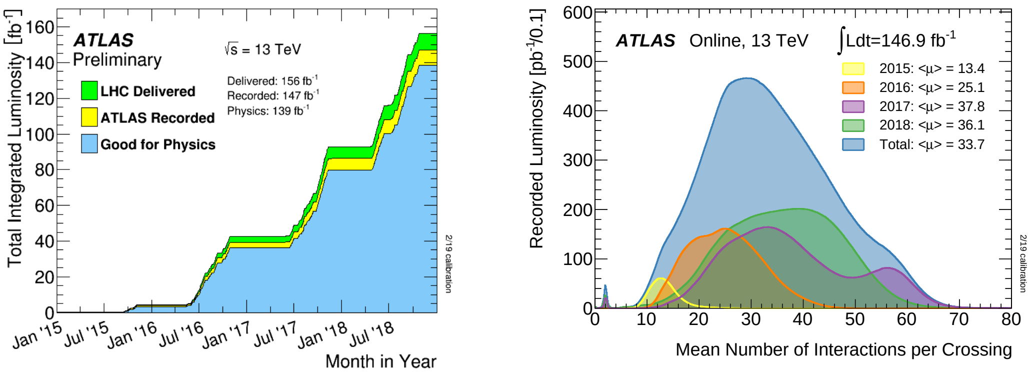
Accurate measurement of the luminosity is essential for the performance of the experiments. This is especially the case in precision measurements of the cross sections where the luminosity plays a major role. In order to accomplish this, the LHC experiments use dedicated detectors for measuring luminosity. ATLAS uses multiple of these dedicated detectors with different technologies [34]. In Run 2, the primary luminosity measurement in ATLAS is provided by the LUCID 2 Cherenkov detector [35]. The beam conditions monitor (BCM) [36] diamond detectors also measure luminosity at the bunch-crossing level, providing complementary information. The amount of data delivered by the LHC and recorded by ATLAS is expressed as luminosity integrated over time:

$$L = \int \mathscr{L} dt. \tag{2.2}$$

From here, the number of produced particles in a given process can be calculated as the product of the integrated luminosity, L, and the cross section of the process, σ, which depends on the centre of mass energy of the collider, $\sqrt{s}$:

$$N = L\sigma(\sqrt{s}). \tag{2.3}$$

Figure 2.2a shows the integrated luminosity delivered by the LHC over the Run 2 data taking period, recorded by the ATLAS detector.



(a) **(b)**

Figure 2.2: (a) Luminosity delivered to ATLAS (green), recorded by ATLAS (yellow), and validated good quality data (blue) during Run 2 pp collisions at 13 TeV centre of mass energy. (b) Luminosity-weighted distribution of the mean number of interactions per crossing for Run 2 data taking period of the LHC [37].

An increase in instantaneous luminosity implies an increase of the number of interactions that occur in a given bunch crossing. Particles produced in additional interactions outside of interest are considered as pile-up. Pile-up events can occur together with the interaction of interest or from interactions between subsequent or preceding bunches. The presence of pile-up deteriorates the performance of the detectors, since, particles originating from pile-up interactions can contaminate hard-scatter events. The number of pile-up interactions per bunch-crossing, μ, can be expressed with the following formula:

$$\mu = \frac{\mathscr{L}\sigma_{inel}}{N_b f}, \tag{2.4}$$

where σ_{inel} is the proton-proton inelastic cross section. Figure 2.2b shows the distribution of $<\mu>$ for each year of the Run 2 pp collisions. The average pile-up for the entire run was approximately 34 interactions per bunch crossing.

2.1.3 LHC timeline

The current schedule for the LHC is shown in Figure 2.3. LHC started its full operation in 2009. The first data taking period lasted from 2010 to 2012 (Run 1), during which the LHC produced pp collisions at 7 and 8 TeV centre of mass energy with a total integrated luminosity of $30\ fb^{-1}$ [38, 39]. Following the first long shutdown the LHC resumed pp collisions in 2015. Between 2015 and 2018, referred to as Run 2 of the LHC, it provided pp collisions at 13 TeV centre of mass energy and delivered $156 fb^{-1}$ of integrated luminosity to the ATLAS detector [34]. The time spacing between two consecutive bunches is 25 ns ensuring bunch collision rate of 40 MHz.

After the recent Phase-I upgrade it started with Run 3 data-taking, which will last from 2022 to 2025. During the Run 3 period LHC should deliver proton collisions with 13.6 TeV centre of mass energy and collect data corresponding to an integrated luminosity of 300 fb^{-1}. A Phase-II upgrade is scheduled following Run 3 to further develop the LHC into the High Luminosity LHC (HL-LHC). HL-LHC will start with its operation in 2027. During its run it should deliver pp collisions at a centre of mass energy of 14 TeV with a baseline luminosity of 5$\times$10^{34} cm^{-2}s^{-1} and an achievable peak luminosity of 7.5$\times$10^{34} cm^{-2}s^{-1}. The HL-LHC will enable the ATLAS experiment to increase the collected integrated luminosity of about 3000 fb^{-1} [40].

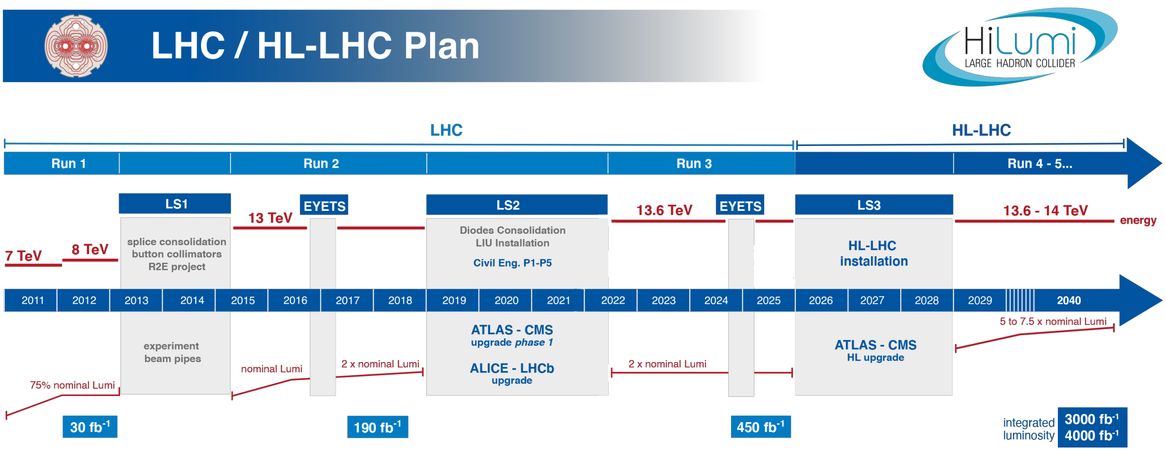


Figure 2.3: The current schedule for the LHC, including the HL-LHC upgrade.

2.2 ATLAS

The ATLAS [27] experiment is a multi-purpose particle detector operating on the LHC ring. Beams of particles from the LHC collide at the centre of the ATLAS detector, forming new particles, which fly out in all directions. For this reason ATLAS has a cylindrical geometry consisting of multiple sub-detectors surrounding the interaction point and covering almost the full solid angle of 4π. One can distinguish two main areas of the detector: the central region (or barrel) and the two forward regions (or end-caps). ATLAS is the biggest detector in the world, 44 m in length and 25 m in height. It is estimated that the detector weighs 7000 tons. An illustration of the ATLAS detector is provided in Figure 2.4.

The detector uses a right-handed coordinate system with its origin at the interaction point. The x-axis points to the centre of the LHC ring, the y-axis points upward

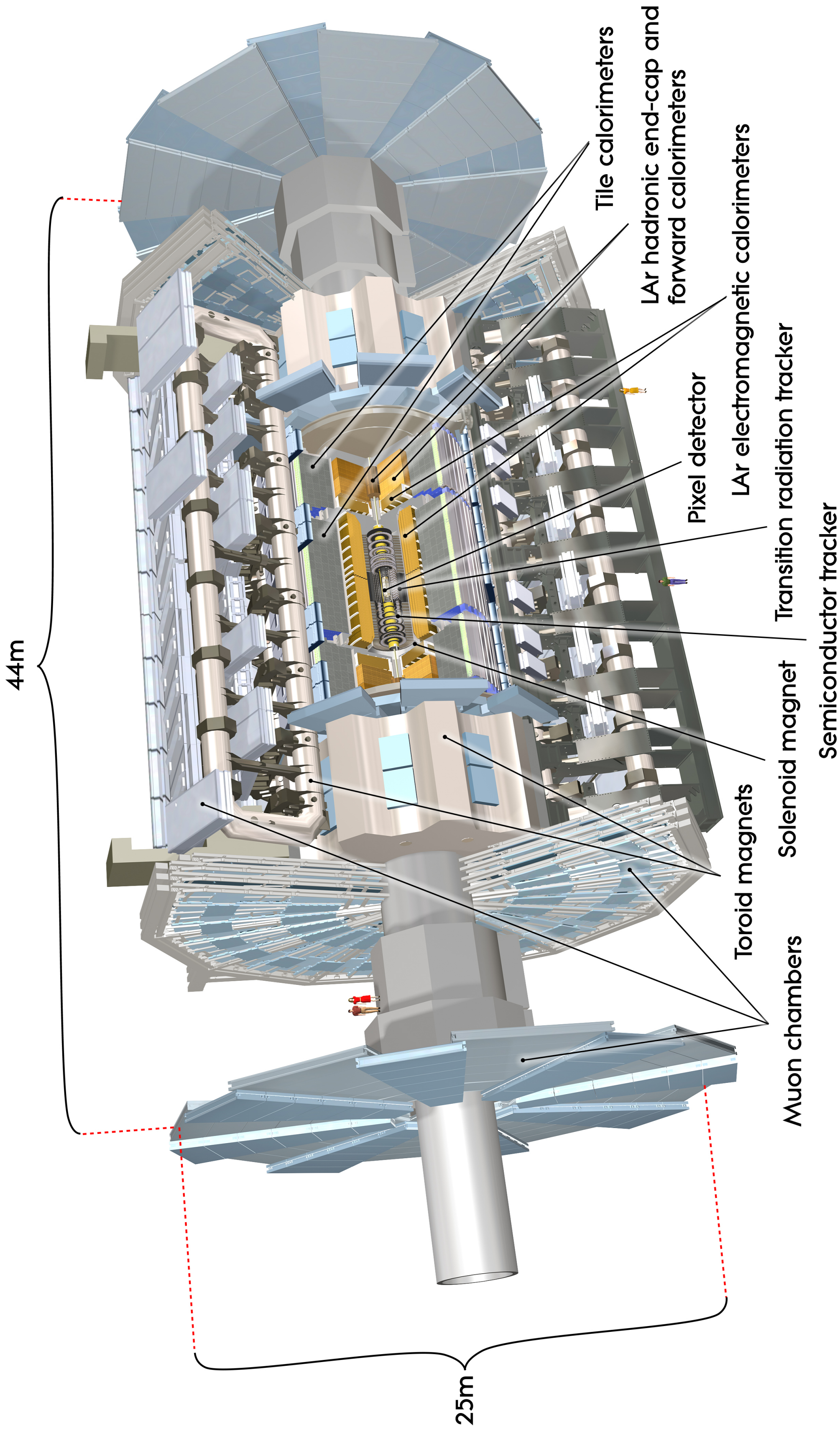


Figure 2.4: Schematic view of the ATLAS detector with the main sub-detectors indicated [27].

and the z-axis is parallel to the beam direction. Often, cylindrical coordinates (z, ϕ, θ) are used instead, ϕ being the azimuthal angle around the beam pipe and θ the polar angle in the transverse $x - y$ plane. The rapidity of the particle is defined as

$$y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) \tag{2.5}$$

where E is the energy of the particle and p_z is the projection of the particle momentum along the z axis. In the relativistic case, $E >> m$ the rapidity becomes equal to the pseudorapidity, which is defined as:

$$\eta = -\ln[\tan(\theta/2)] \tag{2.6}$$

A pseudorapidity of $\eta = 0$ corresponds to the transverse direction within the detector, while $\eta \to \infty$ points to the beam axis. In hadron collider physics, the pseudorapidity is preferred over the polar angle θ since measurements in η are invariant under the longitudinal boost of the reference frame. This is an important feature since the colliding partons carry different longitudinal momentum fractions meaning that the collision centre of mass frame will rarely be coincident with the detector rest frame.

Angular distance is measured in the modified cylindrical coordinate system using η and ϕ as following:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \tag{2.7}$$

For a better description of an event the transverse momentum $p_T = p \cos\theta$ and the transverse energy $E_T = E \cos\theta$, defined on the $x - y$ plane are often used. Moreover the missing transverse energy E_T^{miss} quantifies the energy that is not detected on the transverse plane and mostly corresponds to particles that do not leave signature in the detector, such as neutrinos.

2.2.1 Inner Detector

The ATLAS Inner Detector (ID) [41] is the closest sub-detector to the interaction point, begins a few centimetres from the beam axis and extends to a radius of 1.2 m, covering up to $|\eta| = 2.5$. Its basic functionality is to track charged particles, reconstruct their trajectory and it is responsible of locating the primary vertex of the event. The designed transverse momentum resolution for the ID is σ_{p_T}/p_T = 0.05% p_T [GeV] $\oplus$ 1%.

The ID consist of three different systems of sensors, all placed in a 2 T magnetic field, which is used to curve the trajectory of the charged particles. The sensors are arranged in concentric cylinders around the beam axis in the barrel region while in the end-cap regions they are placed as disks perpendicular to the beam axis. Figure 2.5 shows the schematic illustration of the barrel section of the ID.

Pixel detector

The innermost layer is the Insertable B-Layer (IBL) [43]. The IBL was installed between the beam-pipe and the pixel detector during the shutdown followed Run 1. It is a silicon (Si) semiconductor pixel detector extending up to $|\eta| = 2.9$. The pixel size is 50×250 μm^2 in ϕ and z. The fine pixel size and its close location with respect to the

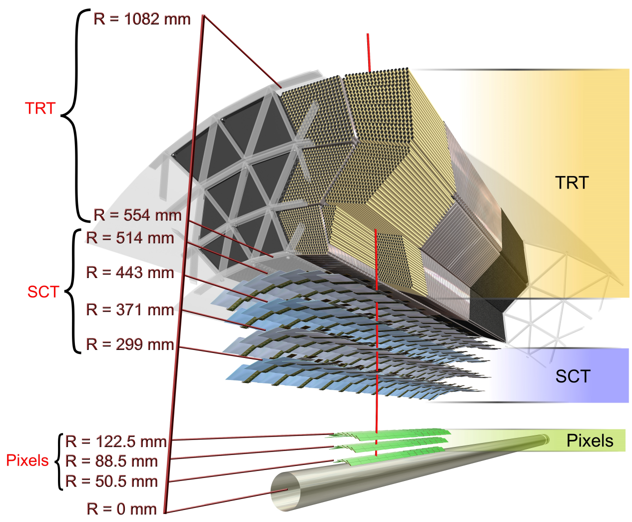


Figure 2.5: Schematic view of the ATLAS Inner Detector in the barrel area [42].

interaction point ensure high quality tracking and vertexing in view of the increase of instantaneous luminosity and radiation damages of the existing systems. It is particularly useful in the reconstruction of secondary vertices and the identification of b-jets.

The IBL is surrounded by three layer of silicon pixel detector. The outermost layer in the barrel region is placed at a distance of 12 cm, while in the end-cap regions, three disks are placed perpendicular to the beam axis. The nominal pixel size is 50×400 μm^2 which ensures high precision measurements in the closest region to the interaction point.

The Semi-Conductor Tracker

The Semi-Conductor Tracker (SCT) [44] is the second sub-detector of the ATLAS ID. It is constructed from silicon microstrips that provides excellent spatial resolution over a large area. The SCT is placed in four concentric layers in the barrel region covering $|\eta| < 1.4$. The sub-detector in the end-cap regions consist of 9 disks on each side in the region of $1.4 < |\eta| < 2.5$. The SCT uses silicon-strip detectors placed at a distance of 80 μm to measure both coordinates, resulting in a precision of 17 μm in the $r - \phi$ plane and 580 μm in the z coordinate.

The Transition Radiation Tracker

The Transition Radiation Tracker (TRT) [45] is the outermost component of the Inner Detector covering up to $|\eta|$=2. It is a constructed from straw trackers and a transition radiation detector. In the barrel there are around 50.000 straws with 144 cm length parallel to the beam axis, while each end-cap is equipped with 125.000 straws of 39 cm length arranged radially in wheels. The straws are drift tubes with a diameter of 4 mm, filled with xenon-based gas, and equipped with a central cathode wire, acting as a proportional counters. The TRT straws allow for continuous tracking, typically providing about 36 space points per charged particle track. The large number of space points compensates for the low resolution compared to the silicon detectors which is equal to 130 μm in the $r - \phi$ direction.

The TRT is able to measure transition-radiation photons. The area between the straws is filled with materials with widely varying refraction indices. Within this material ultra-relativistic charged particles produce transition radiation. The produced photons are absorbed by the gas mixture in the tubes enhancing the readout signal. Via this mechanism, the TRT provides particle-identification which is used primarily to distinguish between electrons and pions.

2.2.2 Calorimeter system

The ATLAS calorimeter system is made of the electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL), covering the pseudo-rapidity range up to $|\eta|$ =4.9. It is placed after the solenoid magnet that surrounds the ID. With the calorimeter system it is possible to measure the energy lost by electrons, photons and hadrons within the detector. Particles that cross the calorimeters initiate a shower of secondary particles through their interaction with a heavy material. The energy of the initial particle can be measured by counting the number of particles constituting the shower. To ensure a precise measurement, it is important for all the particle energy to be deposited within the calorimeter volume.

An useful quantity to measure the size of the electromagnetic calorimeter is the radiation length, X_0, defined as the distance after which the incoming particle retains $1/e$ of its original energy. The total thickness of the ECAL is about 22 radiation lengths in the barrel and about 24 X_0 in the end-cap regions. Similarly, the size of the hadronic calorimeter is given in terms of the interaction length, λ, defined as the mean distance travelled by a hadron in the material before undergoing an inelastic nuclear interaction. The total thickness of the HCAL is 9.7 λ in the barrel and 10 λ in the end-cap regions.

The ATLAS calorimeter system is made from sampling calorimeters, where absorber layers are alternated with layers of active material. This configuration presents the advantage that a fine segmentation can be achieved also in the longitudinal direction, allowing for a precise reconstruction of the shower evolution in all dimensions. This is achieved at the expense of the energy resolution, since part of the energy is deposited in the absorber layers and thus never measured.

Electromagnetic calorimeter

The ATLAS ECAL [46] consists of two parts, the barrel covering a range up to $|\eta|$ =1.47 and the end-cap ranging from 1.37 $< |\eta| <$ 3.2. The electromagnetic calorimeter (ECAL) is designed to match the relative energy resolution of σ_E/E = 10%/$\sqrt{E[GeV]}\oplus$ 0.7 %. The active material is Liquid Argon (LAr) maintained in a temperature of 89 K, while the absorber is lead (Pb). The liquid argon has been chosen due to its linear behaviour, stability of the response over time and its intrinsic radiation-hardness. Each active layer is placed within a gap of 4.5 mm between two lead plates. The absorber thickness varies with pseudorapidity, so that the particles will cross the same amount of material in radiation lengths. Particles initiate electromagnetic showers in the absorber layers and, when the shower products cross the active layers, they cause ionization in the LAr. The resulting ionisation particles drift with the aid of an electric field towards an anode where they are collected. One of the most characteristic properties of the ATLAS barrel ECAL is its accordion shape, which enables for the full ϕ range to be covered.

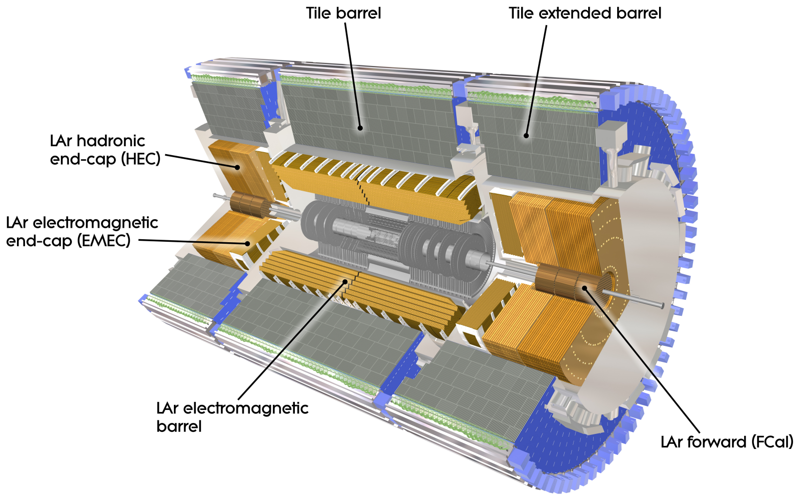


Figure 2.6: Illustration of the ATLAS calorimeter system [27].

The ECAL has three radial layers and an additional thin LAr presampler layer in the front which estimates the energy loss of the particles before they reach the calorimeter. The fine granularity of the first layer allows particle identification tasks, mostly by separating isolated photons from π_0 hadrons. The second layer has moderate η and ϕ granularity capturing most of the energy. The third layer is relatively thin and measures only the tail of high energy showers.

Hadronic calorimeter

The HCAL [46] includes the Tile Calorimeter (TileCal), End-Cap Hadronic Calorimeter and the Forward Calorimeter (FCal). Its main task is to absorb energy of charged and neutral hadrons that pass through the ECAL. The designed relative energy resolution of the hadronic calorimeter is σ_E/E = 50%/$\sqrt{E[GeV]}\oplus$3 % for the barrel and end-cap, and σ_E/E = 100%/$\sqrt{E[GeV]}\oplus$ 10% for the FCal. The TileCal [46] is using steel as an absorber and tiles of scintillating material as an active medium. The end-cap HCAL is made of copper and LAr and it consist of two wheels perpendicular to the beam axis. In addition, the full coverage of the hadronic calorimeter is provided by the FCal. FCal uses LAr as active material, while the absorber is copper in the layers close to the interaction point and tungsten in the following two layers.

2.2.3 The Muon Spectrometer

Unlike electrons and hadrons, muons can travel across the whole detector losing only a small fraction of their energy due to their minimum ionizing nature. For this reason, the Muon Spectrometer (MS) [47] is the outermost sub-detector of ATLAS. The MS has two functions: it provides precise measurements of trajectories of muons and fast signals to trigger events containing muon candidates. It has a coverage up to $|\eta|$ =2.7 consisting of the barrel and two end-cap detectors. It is designed to reach the relative momentum resolutions of σ_{p_T}/p_T = 10% for 1 TeV muons. The MS chambers are arranged in three cylindrical layers around the beam axis in the barrel region, while in the end-cap region the chambers are placed in a form of large wheels perpendicular to the beam axis.

As it can be seen in Figure 2.4, the MS is the biggest sub-detector of ATLAS. It consists of four detection systems relying on different technologies. The major part of MS consist of Monitored Drift Tube (MDT) [48] chambers which ensure precise momentum measurements. In the innermost plane Cathode Strip Chambers (CSC) [49] are used, which are more suitable for regions of very high particle flux. CSCs are multi-wire proportional chambers with cathodes segmented into strips in the orthogonal directions. They are ideal in the $2 < |\eta| < 2.7$ region due to their high radiation resistance, high rate capability and time resolution. Additional subsystems of the MS with lower resolution and faster readout are used for fast muon identification in the trigger system. Resistive Plate Chambers (RPC) [50] are used in the barrel region, as their time resolution is below the bunch crossing time. They are gaseous parallel electrode-plate detectors where the primary ionization particles are multiplied into avalanches by a high electric field. In addition, Thin Gap Chambers (TGC) [51] are used in the end-cap wheels. They are multi-wire proportional chambers with very good time resolution similar to the RPC and a spatial resolution of a few mm.

MDT chambers

The MDT system consists of approximately 1200 chambers containing about 300k drift tubes placed in a toroidal magnetic field. The typical spatial resolution of a single MDT tube is approximately 70 μm while the chambers provide more precise momentum measurements, with a precision of about 35 μm. The main task of the MDT is to measure the momentum and determine the charge of the muon.

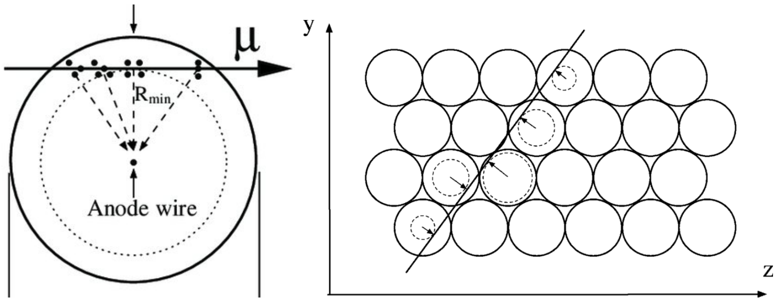


Figure 2.7: On the left side the cross section of a MDT tube is shown, together with the ionization clusters alongside the muon track [52]. On the right side a track fit within a multilayer is illustrated [53].

Similarly to the the other muon sub-detectors, the MDT is a gaseous detector. It is filled with aluminium gas mixture (Ar:CO$_2$ = 93:7). The tubes have a diameter of 30 mm and a wall thickness of 0.4 mm, with a gold-plated tungsten-rhenium anode wire positioned at the centre. The tube wall functions as the cathode. The tube is under a pressure of 3 bar and a voltage of 3080 V. When the muon passes through the chamber it ionises the gas resulting in an ionisation charge, which is collected by the anode. The cross section of a MDT tube is shown in Figure 2.7.

Three or four layers of MDT tubes form a multilayer, and two of these multilayers form a chamber, as it is shown in Figure 2.8. Chambers in the innermost layer has four layers of tubes due to the higher particle rate which ensures improved local pattern recognition.

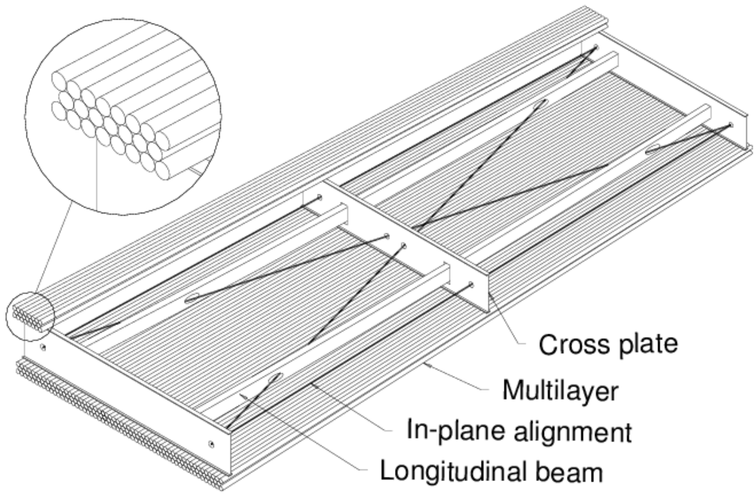


Figure 2.8: Schematic view of a MDT chamber [47].

In the barrel region, the MDTs are placed in three concentric layers around the beam axis, at an approximate radius of 5, 8 and 10 m respectively [54]. In the end-cap region the chambers have a trapezoidal shape and assembled onto three wheels. These wheels are positioned at z = 7.5, 14 and 22.5 m. In both regions the chambers are partly overlapping to prevents holes in detector coverage. Each chamber is classified by a three letter name. The fist letter notes the region of the chamber: barrel (B) or end-cap (E). The second letter identifies the position of the chamber in respect to the beam: inner (I), middle (M) or outer (O). The third letter can be and S or an L, depending if the chamber is small or large. In addition, chambers placed in the low coverage regions have special names, which are not compliant with the naming scheme. The name of some of the barrel region MDT chambers can be seen in Figure 2.9.

2.2.4 Magnet System

The ATLAS magnet system is divided into two parts: the central solenoid [56] and the toroidal magnet system [57, 58]. The central solenoid is a superconducting magnet, 2.3 m in diameter, 5.3 m in length. It is located between the ID and the ECAL. It is cooled to a temperature of 4.5 K providing a magnetic field of 2 T to the ID along the z-axis.

The toroidal magnet system is an air-core system made of a barrel toroid magnet and two end-cap toroids yielding a 0.5 T magnetic field in the barrel region and 1 T in the end-cap region. The barrel toroid has 8 coils, which range from 9.4 m to 20.1 m in diameter and are 25.3 m long. The end-cap magnets also equipped with 8 coils and are 5 m deep with a 10.7 m diameter.

2.2.5 The ATLAS trigger system

Collisions at the LHC occur in every 25 ns, which would correspond to over 60 TB of data recorded by ATLAS written to disk in each second. However it is impossible to process and store this amount of data. Also out of all this information only a few events are of interest for the physics research. For this reason, the ATLAS

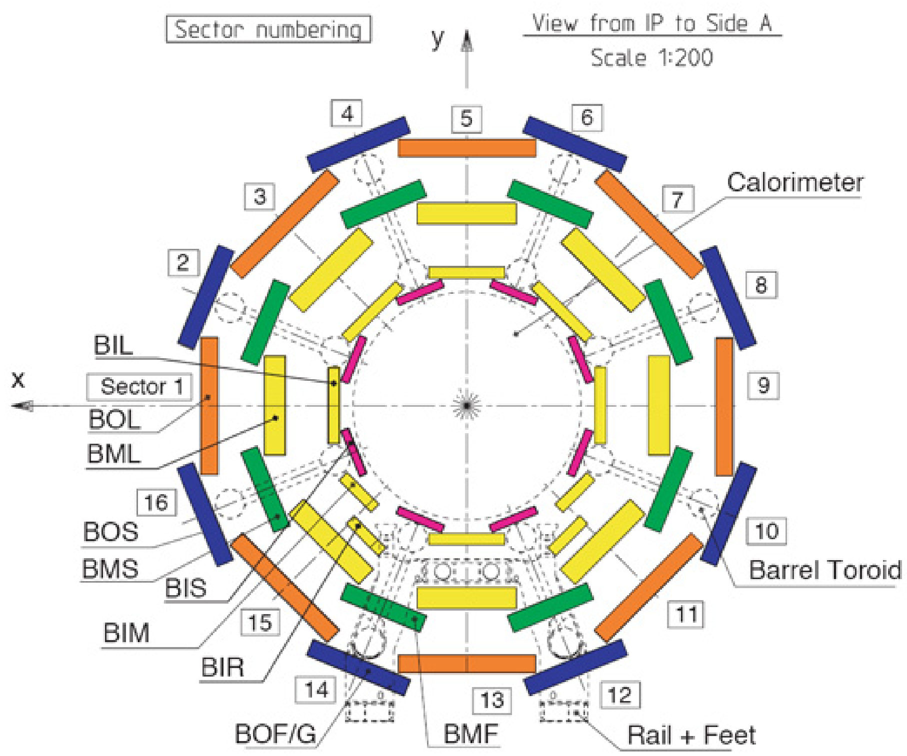


Figure 2.9: Cross-section of the barrel muon system perpendicular to the beam axis [55].

experiment at the LHC uses a Trigger and Data Acquisition (TDAQ) system [59] that includes a hardware Level-1 and a software based high-level trigger. The combination of the two triggers reduces the event rate from the LHC bunch-crossing rate of 40 MHz to an average recording rate of around 1000 Hz. For each bunch crossing, the trigger system uses a fast electronic system and different algorithms based on partial event information to decide whether or not to save an event to disk for offline analysis. The decision is made in three steps with the Level-1 (L1), Level-2 (L2) and Event Filter (EF). L2 and EF are collectively referred to as the High Level Trigger (HLT). The layout of the ATLAS TDAQ in Run 2 is shown in Figure 2.10.

The L1 trigger is divided into two sub-triggers, the L1Calo and L1Muon, that defines the Regions-of-Interest (RoI) in the detector by taking as input low granularity information from all calorimeters and muon detector information. The L1 trigger decision is taken by the Central Trigger Processor (CTP). Here the topological trigger (L1Topo) applies topological selections at the L1 stage, combining kinematic information of the trigger objects received from the L1Calo or L1Muon systems. CTP compares the event with a menu of pre-programmed combinations of objects and thresholds, and determine if the event is accepted or not for further processing. While waiting for the trigger decision, the event information is stored in the front-end memory buffers. If the event is accepted, the CTP sends the L1 accept signal and LHC timing signals to the sub-detector readout systems via the Timing, Trigger and Control (TTC) network. The events are transferred to the Read Out Driver (ROD). ROD processes the events, send them to the Read Out System (ROS), which stores them in the Read Out Buffer (ROB), from where they are further processed by the

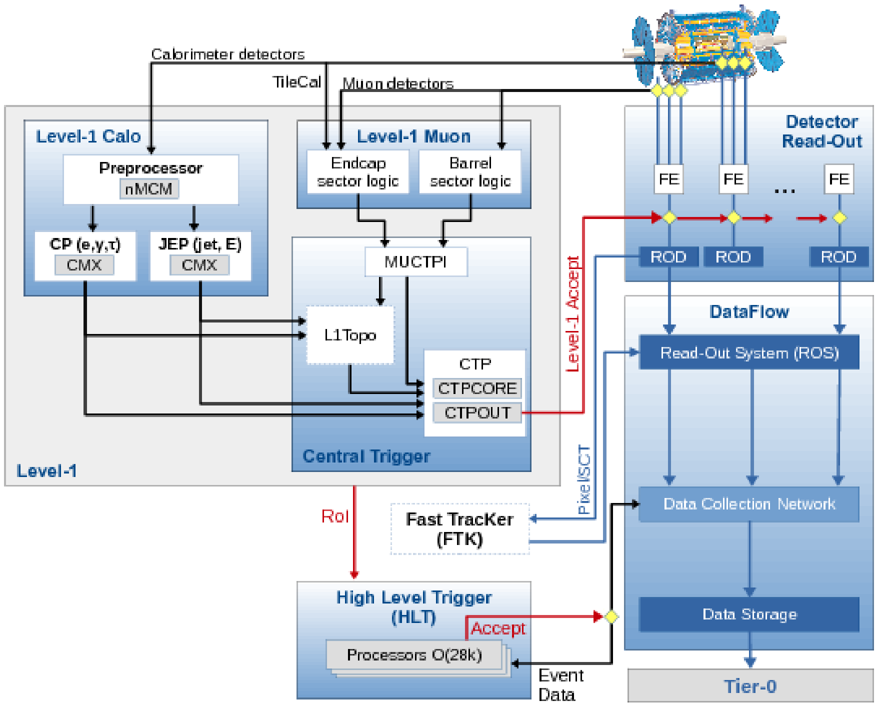


Figure 2.10: The ATLAS TDAQ system in Run 2 with emphasis on the components relevant for triggering [60].

HLT. L1 reduces the event rate from the LHC interaction rate of 40 MHz to approximately 100 kHz [61].

The HLT [62] uses all the sub-detector system to performs a more refined selection of an event. It implements the ID tracking, finer granularity for the calorimeter information, and the precise measurements from the MS. The HLT reduces the rate from the L1 output rate to approximately 1 kHz. After the events are accepted by the HLT, they are transferred to the CERN's computing centre for offline reconstruction.

2.3 ATLAS data and simulation

Figure 2.2a shows that during the Run 2 data taking period LHC delivered a total integrated luminosity of 156 fb^{-1}, out of which 147 fb^{-1} was recorded by the ATLAS detector and 139 fb^{-1} passed the good-physics selection criteria. During data taking, each run is is divided into luminosity blocks, during which the instantaneous luminosity, detector and trigger configuration and data quality conditions are considered constant. It is important that each of these blocks are free from any integrity or detector related issues and that data is available from all the subsystem. To ensure the data quality of ATLAS detector, each sub-detector is constantly monitored and the recorded data is extensively inspected by a team of experts. This process results in the Good Runs List (GRL), a set of XML files that containing the list of luminosity blocks that are certified for use in physics analyses [63]. Data is labeled as "good for physics" if all detector systems are operational (or their defects are tolerable) and all reconstructed physics objects are assessed to be of good data quality. This data is analysed by ATLAS physicist when probing the SM or searching for new physics, so it needs to be processed before it can be used for various searches.

On a parallel chain Monte Carlo (MC) techniques are used to simulate what happens during a hadron collision. They are necessary to develop reconstruction and identification algorithms as well as for the interpretation of physics results. With this technique, we are able to simulate all the details and different steps of the collision together with the final-state particles. During the simulation process, every part of the collision is simulated separately which are include the hard scatter (HS), the underlying event, the parton shower and the hadronisation. All of this processes together form the complete picture of a simulated event.

To be able to compare MC simulations and data collected by the detector, it is important to model what happens when the particles interact with the detector materials. `Geant4` [64] is a simulation framework which contain the model of the ATLAS detector and is able to simulate the sub-detector response. The output of the simulation also has an identical format with the real data recorded by the ATLAS detector so the simulated events resemble the actual data with high precision.

To reconstruct what happened during the collision an extensive software suite [65] is used. Both the real and simulated events are converted into physics objects written into Analysis Object Data (AOD) format. However, since only 1-2% of the full dataset is used for physics analysis, custom derivations have been developed, which follows the Derived AOD (DAOD) format and contains only the relevant information for the specific physics analysis. To be able to efficiently perform the reconstruction, the data reprocessing and analysis the LHC Computing Grid (LCG) is used [66]. LCG incorporates over 170 computing centers in 42 countries in a grid-based computer network, which is specifically designed to be able to handle the enormous data produced by the LHC.

2.4 ATLAS detector upgrades

As it can be seen from Figure 2.3, so far two major update program has been performed on the ATLAS detector. The Phase-I upgrade program at ATLAS was finished in the beginning of 2022, while the Phase-II upgrade is in its final stage of design. The construction and installation of the Phase-II upgrades are taking place after the current Run 3, planned in the 2025-2027 time frame. Phase-II upgrade will prepare the ATLAS detector to take data delivered by the HL-LHC.

Phase-I Upgrade

During Run 3 the ATLAS experiment is required to operate with higher collision rates. During this run, the LHC will deliver luminosities up to three times its design value, with approximately 80 proton-proton collisions per bunch crossing. To maintain physics performance in the new environment, the ATLAS experiment received a series of upgrades during the shutdown. The upgrades include new detector systems, new trigger systems exploiting a more granular readout of the electromagnetic calorimeter and a novel readout chain in the TDAQ system [67].

New L1 calorimeter read out processors have been installed which allow finer granularity data from the Liquid Argon (LAr) calorimeter. This will be used to improve electron, photon and tau selection, while allowing the use of larger-area algorithms to improve jet selection. This finer granularity data will be transmitted from new

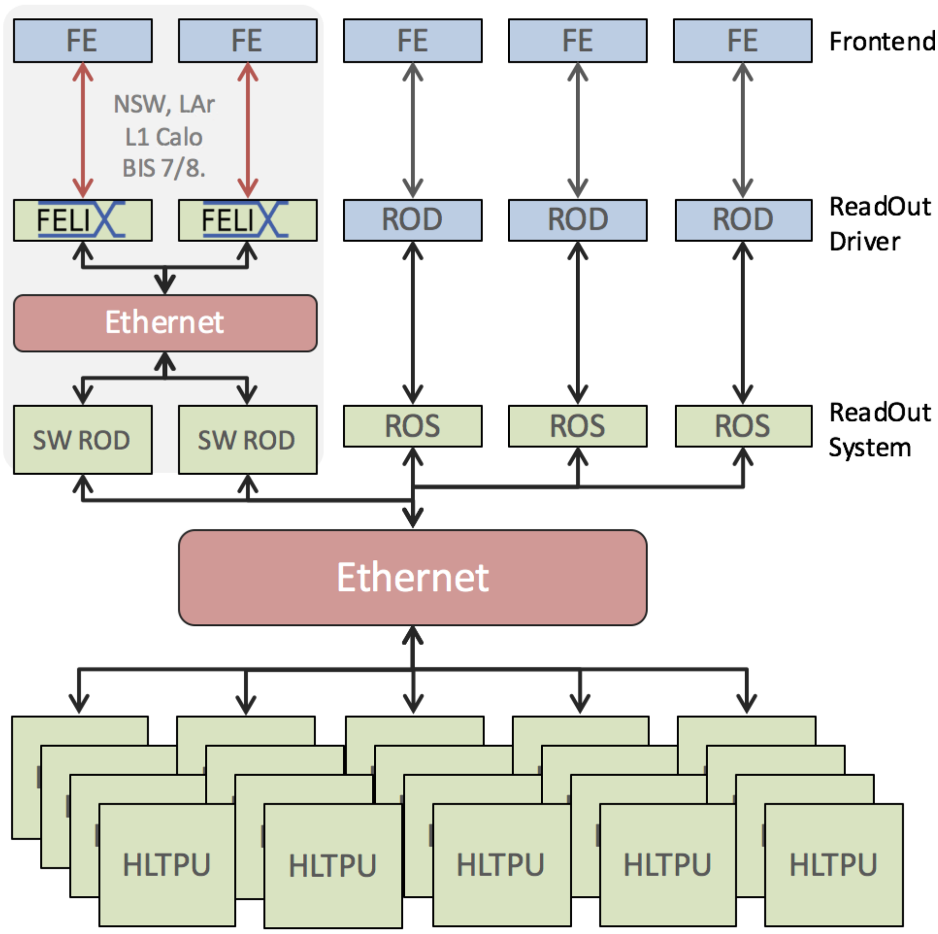


Figure 2.11: Representation of the TDAQ architecture in Run 3, which is consist of the combination of the legacy system and the new FELIX system [68].

dedicated LAr Calorimeter hardware [69]. The Phase-I TDAQ upgrade will also benefit from the construction of the New Small Wheels (NSW) [70]. The signals from the NSW will be included in the Level-1 muon endcap trigger. This will reduce the overall trigger rate by rejecting a large fraction of fake triggers. Besides the NSW project a smaller size project, known as BIS78 [71] (from Barrel Inner Small sectors), is being developed. The BIS78 project will improve the fake muon rejection and the selectivity of the muon trigger in the transition region between the barrel and end-cap regions.

Among the upgraded components there will be the TDAQ, which in the new environment will have to process significantly more complex events while maintaining stable the selection performance. The upgraded ATLAS systems will make use of newer readout link technologies. To connect the new systems, and handle the significantly increased data volumes in a detector agnostic and easily scalable way, a new readout architecture named the Front-End Link eXchange (FELIX) is under development. After the upgrade the NSW and the L1 calorimeter trigger systems will use the new FELIX readout system. FELIX will fit in the existing TDAQ architecture to serve the upgraded detector systems, leaving the rest of the legacy TDAQ system in place, as it is shown in Figure 2.11.

Phase-II Upgrade

The Phase-II upgrade of the ATLAS TDAQ system must satisfy the various ATLAS physics programs planned for the HL-LHC, while coping with HL-LHC conditions [72].

For this purpose the ATLAS Inner Tracker (ITk) will be replaced to provide improved tracking in HL-LHC the high pile-up environment. The new silicon-only design ITk will obtain better momentum resolution for reconstructed tracks with extended $|\eta|$ coverage. The ATLAS LAr Calorimeter and the ATLAS Tile Calorimeter will have entirely new front-end and readout electronics and optical link interface boards optimized to withstand radiation. A large fraction of the ATLAS MS front-end and on- and off-detector readout and trigger electronics will be replaced to enable higher trigger rates and longer latencies. Additional muon chambers will be introduced to manage muon identification and reconstruction performance and increase trigger acceptance. The detector upgrades present new requirements and new opportunities for the TDAQ systems as well. In addition the ATLAS High-Granularity Timing Detector (HGTD) will be installed front of the LAr calorimeter, which will precisely measure the timings of charged particles reducing background from pile-up jets.

3 The upgrade of the MS readout system

As already mentioned, the MDT chambers are the main component of the ATLAS muon tracking system. As part of the Data Acquisition (DAQ) system, during the previous data-taking periods the MDT chambers were read out via the MDT readout chain, consisting of the ATLAS Muon Time to Digital Converter (TDC), the Chamber Service Module (CSM) and the MDT Read Out Driver (MROD) [73]. The main task of the readout electronics is to ensure the measurement accuracy of the tubes and to cope with the rates at high luminosity. To process the high data rates the system is based on proper distribution of the readout processors, large storage capacities and high-speed data links.

The MDT readout chain is dependent on the MROD modules, which may reach end of their lifetime during the Run 3 data taking period. Since repairing the broken MROD cards or ordering new ones is not possible, a new implementation has been developed which is using the new Front-End Link eXchange (FELIX) system [74].

3.1 MDT readout system overview

The readout chain begins with the front-end cards, which are placed directly on the MDT chambers. They are equipped with TDC mezzanine cards to digitize the wire signals [75]. Each front-end card handle maximum of 24 drift tubes and the largest MDT chambers have 18 front-end cards connected to the same amount of TDCs. After receiving a L1 trigger signal, each TDC sends its data to the CSM [76]. The CSM collects data from up to 18 TDC boards, and multiplexes the input data into a single output stream before sending them to the off-chamber DAQ system. The multiplexer operates as a rotating disk that is repeatedly checking each input TDC link for data. When a TDC has forwarded a word[1], it is passed to the output CSM link. When no data was delivered inside the given time frame, a filler ("zero") word is transmitted. A separator word is sent at the start of each rotating scan sequence to guarantee the synchronization The separator word is resetting the multiplexer such that the next word will be always assigned to TDC 0, 1 and so on until the last. The CSM connected to the off-chamber electronics through two fibers, one coming from the TTC distribution box and the other going to the MROD [77].

The main task of the MROD is to receive and demultiplex the data streams from five to eight CSMs. The MROD builds the event fragments from the incoming data and sends them over to the Read Out Buffer (ROB) where the data can be retrieved by the TDAQ system. In addition, the MROD detects and reports errors and inconsistencies in the incoming data streams, it collects statistics and it allows data

[1] In computer architecture, a word is a unit of data of a defined bit length. In the context of this thesis, the length of a word is 32 bits.

observation. The simplified scheme of the full MDT readout chain is showed on Figure 3.1.

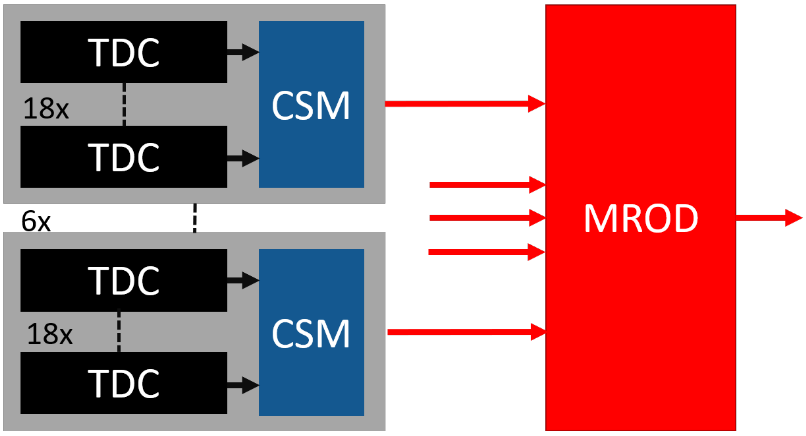


Figure 3.1: Schematic view of the MDT readout system. Each TDC serves 24 drift tubes, each CSM is connected up to 18 TDCs, while the MROD collects data from up to 6 CSMs.

The chambers belonging to the same MROD form a tower covering a predetermined $\Delta\phi\Delta\eta$ slice of the full 4π solid angle. The MDT consists 1150 of these chambers, which form 204 $\Delta\phi\Delta\eta$ towers, means a total of 192 MRODs in need. The MRODs are stored in crates as showed in Figure 3.2. These are placed in the USA15 cavern (USA stands for Underground Service ATLAS) which contains most of the electronics for the experiment. Each crate contains 12 MRODs and a total of 16 MROD crates are in place.

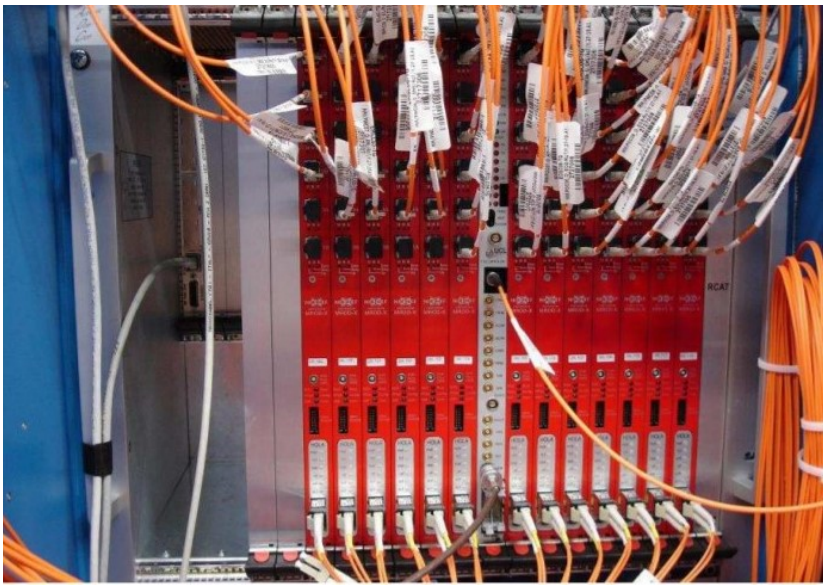


Figure 3.2: Overview of the MROD module connectivity. Twelve or in some case thirteen modules placed in a crate. The CSMs are connected via the optical fibres at the top, while the fibres at the bottom are the ReadOut Links connecting to the ROBs.

3.1.1 MROD data format

The MROD data output consists of three nested levels of fragments [78]. The lowest of the three levels is the TDC level. The next level is formed by the CSM level, which is a collection of the TDC fragments, whereas the highest level corresponds to a group of up to 6 chambers defining the MROD level fragment. Each fragment has the same basic structure: one or more header words followed by a number of data words and completed with a trailer containing the word count for the fragment as a whole. Fragments can be empty, so they contain no data words. At each level the fragments of the next lower level are fully included in the fragment of the current level. The TDC fragment are generated by the TDCs in the form of header and trailer words. In the absence of hits, a TDC sends an empty fragment for each L1 trigger to the CSM. Since the CSM does not add any data word to the data stream, the CSM level header and trailer is generated by the MROD, as is the MROD level ones. The format of the MROD data is visualized in Figure 3.3.

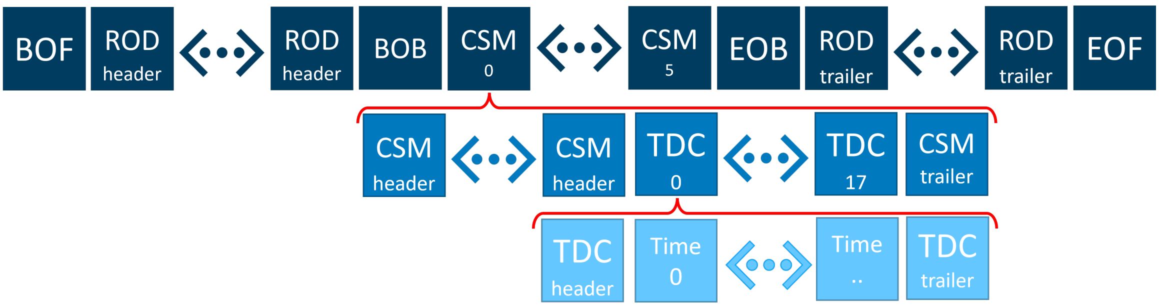


Figure 3.3: MROD data format structure. The top line represents the outermost ROD envelope, the second line shows the CSM data envelope, while the last line depicts the TDC level envelope.

Every event constructed by the MROD has to fulfill the ATLAS TDAQ event format requirements [79]. The full event is built from sub-detector fragments, where each of these fragment is an aggregation of ROS fragments. In turn, each ROS fragment is an aggregation of ROB fragments, which is also built from one or more ROD fragment. Each fragment has a header which contains all the event information and indicates the beginning of the the event, sub-detector, ROS, ROB and ROD fragments. Headers are invariant of sub-detectors, and each fragment header begins with a specific Start of Header Marker. In case of the MDT chambers, the general ROD header and trail words are constructed by the MROD. These words are the following:

- ROD header words:
  - **Start of ROD Header marker**: Indicates the start of a ROD fragment header, which is defined to be `0xEE1234EE` for all RODs.
  - **ROD Header size**: Indicates the size of the header including the header marker. The header contains 9 words.
  - **Format version number**: Indicated the format version of this ROD fragment. The upper two byte state the Major and the lower two byte state the Minor version number.
  - **Source identifier**: Identifies the origin of the ROD fragment. The word consists of a module type, a sub-detector ID and a module ID. The sub-

detector IDs for the MDT detector are in the range `0x61` thru `0x64`.
  - **Run Number**: The highest 8 bits are defined by the type of the Run (physics, calibration, etc.). The low order 24 bits represent the ordered sequence of runs within a type.
  - **Level 1 Identification or Event number (L1ID)**: Contains the 24 bit event identifier generated by the L1 trigger system.
  - **Bunch Crossing Identification (BCID)**: Contains the 12 bit long bunch crossing identifier generated by the L1 trigger system.
  - **L1 Trigger type**: Contains the 8 bit long event trigger type defined by the L1 trigger system.
  - **Detector event type**: Identifies an event which may have been generated by a sub-detector, independently of the ATLAS trigger systems.
- ROD fragment trailer types:
  - **MROD status element (MSE)**: When zero it shows that no known errors are associated with this event fragment.
  - **Number of status elements (NSE)**: This is the total number of words that were inserted in the status block. At least one status word must be present in the event.
  - **Number of data elements (NDE)**: The total number of words in the data block, excluding the 9 words in the MROD header.
  - **Status block position (SBP)**: Defines the relative order of the data and status elements. A value of zero indicates that the status block placed before the data block and a value of one indicates that the status block is placed after the data block.

Besides the general ROD header and trailer words, additional, MROD specific header and trailer words are also included to the data stream. These specific headers, together with the ROD header and trailer words consist the MROD level fragment. The MROD specific header and trailer words are the following:

- **MROD header, BOB (MROD Begin Of Block)**: The BOB word starts with `0x80` and used to store a copy of the L1ID in the lower 24 bits.
- **MROD trailer, EOB (MROD End Of Block)**: The EOB word, starting as `0xF0`, contains 16 bits of word count for this MROD event block, counting from and including the BOB word, up to and including the EOB word.

Besides the trailer and header words the MROD event fragment contains the full CSM data. Just like in the case of the MROD data fragments, the CSM data fragments are also consisting of predefined specific header and trailer words, which are:

- MROD header types:
  - **LWC (MROD Link Word Count)**: The LWC word starts with `0x81`. It indicated the first word of an event fragment coming from one CSM. This word contains 4 bits of the L1ID for debugging purposes and 16 bits word count for the number of words in this CSM fragment of all words including the LWC itself, up to and including the TWC word. The LWC word always followed by the BOL word.
  - **BOL (MROD Begin Of Link)**: The MROD BOL word, starting with `0x18`, indicates which CSM link is giving its data. It contains 4 bits for CSM link number, 12 bits for MROD module serial number, and a number of status bits indicating different working conditions.

  - **TLP (MROD TDC Link Present)**: The MROD TLP word starts with `0x89`. It contains status information about the TDC links connected to this CSM as each bit in the 18 lower position represents a TDC. This header is always present, even if this CSM had no data at all.
- MROD trailer types:
  - **TWC (MROD Trailer Word Count)**: The TWC word starts with `0x8A`. It indicates the last word of an event fragment coming from one CSM link. It contains 12 lower bits of the 24 bit L1ID stored in the MROD. Also contains 12 bits of word count for the number of words coming from this CSM link, starting from (and including) the TLP word and all TDC words up to and including the TWC word itself.

The CSM aggregates data from up 18 TDCs, which means, that within the CSM event fragment, there is up to 18 TDC event fragment placed. Every data fragment sent by the TDC contains the following header and trailer type:

- **TDC data header, BOT (Begin Of TDC)**: The TDC BOT word starts with `0xAt` or `0xBt`. It marks the begin of TDC data for this event. The header starts with 0xA for words from TDC 0-15 and with 0xB for BOT words from TDC 16-17. It also contain 12 bits for event counter, counted by the TDC and 12 bits of BCID.
- **TDC data trailer, EOT (End Of TDC)**: The TDC EOT word starts with `0xCt` where $t =$0,1,2,3. It indicates the end of TDC data fragment for this event. It contains 12 bits of event counter and also 12 bits word count which indicates the number of words in this TDC data block including itself.

In normal running conditions the presence of the following words in each event is expected: the BOB word, followed by a maximum of 6 CSM blocks, each consisting of the LWC, BOL and TLP words, a number of TDC words and a terminating TWC word. The last CSM block is terminated by the EOB word. This can be seen in the example event shown in Listing 3.1. In this example event only one CSM block present. When more CSMs are connected, the block between the MROD LWC and MROD TWC word is present for each CSMs.

To perform consistency check, on all the six CSM links, each TDC will always send both header and trailer words, even if the fragments don't contain data words. The MROD verify the event ID encoded in the BOT and EOT word, and if there is no inconsistency present, both words may be skipped or zero-suppressed, if the zero-suppression feature within the MROD is turned on. If there is real measurement data present in the TDC fragments, the BOT can not be suppressed, since it is the only word containing the full TDC identifier. One other feature of the MROD is the trailer-suppression, when the TDC trailer words are skipped even if that TDC fragment contains measurement data. This will only happen if the EOT word contains the correct event number and word count. If this not the case, the EOT will still be present in the end of the TDC fragment.

3.1.2 Event building mechanism

The MROD uses Xilinx Virtex-II Pro Field-Programmable Gate Array (FPGA)s for data processing, which can be divided into two categories, the `MRODin` and `MRODout`, as it is shown in Figure 3.4. Each `MRODin` FPGA connects to the CSM link and to the MRODout FPGA.

During data processing, in the absence of any errors, the FPGA processes the

```
Start of ROD Header marker: ee1234ee
ROD Header size:            00000009
Format version number:      03000000
Source identifier:          00610023
Run Number:                 00002019
Level 1 ID:                 000150e0 (L1ID)
Bunch Crossing ID:          00000c46 (Bunch crossing ID)
Level 1 Trigger type:       00000001
Detector event type:        00000000

MROD BOB: 800150e0

MROD LWC: 8100002d (wcnt=002d)
MROD BOL: 18000000 (MROD#=00, CSM#=0)
MROD TLP: 8903ffff (Link present=3ffff)
      a00e0c46 c00e0002
      a10e0c46 300400cf 340000eb c10e0004 (General TDC datawords)
      a20e0c46 c20e0002
      a30e0c46 c30e0002
      a40e0c46 c00e0002
      a50e0c46 34000001c10e0003
      a60e0c46 30440171 3440018d c20e0004
      a70e0c46 c30e0002
      a80e0c46 c00e0002
      a90e0c46 c10e0002
      aa0e0c46 c20e0002
      ab0e0c46 c30e0002
      ac0e0c46 c00e0002
      ad0e0c46 c10e0002
      ae0e0c46 c20e0002
      af0e0c46 c30e0002
      b00e0c46 c00e0002
      b10e0c46 c10e0002
MROD TWC: 8a0e002b (wcnt=02b)

MROD EOB: f000002f (nwords: 47.)

MSE: 00000000 (the general ROD status word)
NSE: 00000001 (number of status words: 1.)
NDE: 0000002f (number of data words: 47.)
SBP: 00000001 (status block, at end when 1)
```

Listing 3.1: An example MROD event. The full and partial L1ID is noted with red, the bunch crossing ID is with blue, and the general datawords are with purple.

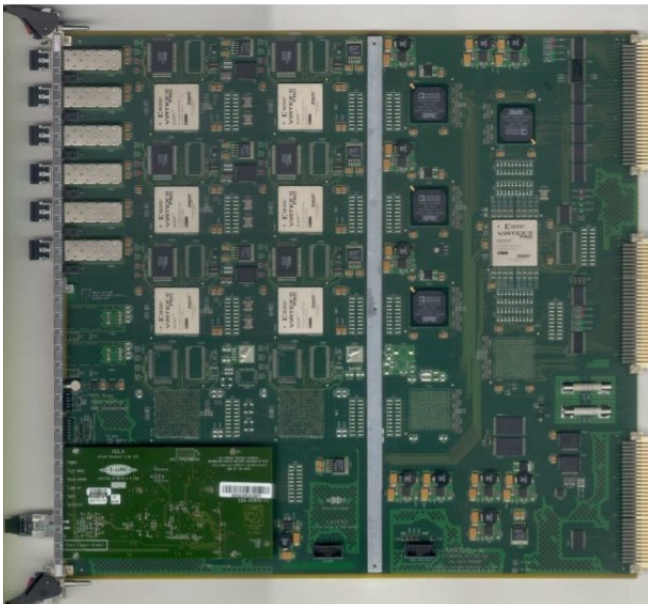


Figure 3.4: The MROD board. The optical cables from up to 6 CSMs are connected at the upper left. The metallic square integrated circuits are the data processing FPGAs, those of the `MRODin` sections can clearly be seen at the upper and middle left, the `MRODout` FPGA is at the middle right [77].

data without intervention. The FPGA demultiplexes the CSM data, removes the empty and separator words, and reconstructs the data streams of the individual TDCs. After reconstruction, data are subsequently stored in the associated memory. The FPGA recognizes the TDC trailer words and records L1 trigger information. Empty TDC envelopes may optionally be zero suppressed. Once the FPGA has recognized a complete event and the trailer words of all active TDCs have been received, the CSM level header and trailer words are generated. The `MRODin` passes the data on to the `MRODout`. The first step of the fragment building in the `MRODout` FPGA consists of generating the correct MROD header words. The MRODout FPGA sends the output streams of the MRODin FPGAs one after the other to the MROD output link. When all the CSM fragments sent out, the data stream is terminated with the trailer words.

During run time, the FPGA checks for a number of error and exception conditions:

- Parity errors on the TDC to CSM link (these errors are encoded in the data by the CSM),
- Link and/or parity errors on the CSM to MROD link
- Absence of data, incorrect or too long event fragments from a TDC
- Absence of expected trailer words or corruption of trailer words

In case an error is detected, the FPGA intervene appropriately. For very serious conditions (too large event fragment or a memory buffer overrun) the FPGA will independently decide to ignore an individual TDC channel.

3.2 Front-End Link eXchange

Front-End Link eXchange (FELIX) [74] is a new detector readout system developed to improve the capacity and flexibility of the current readout chains. FELIX is designed to act as a data router, receiving data packets from detector front-end cards and sending it to the detector software for reconstruction. Whereas previous detector readout systems relied on custom hardware and software solutions, the idea behind FELIX is to unify all readout systems to a well supported and flexible platform. Besides optimizing performance, FELIX reduces the reliance on custom hardware. By maximizing the use of commodity hardware, FELIX is easy to maintain and to upgrade. FELIX is also very modular so different commercial components can be easily implemented to resize the FELIX infrastructure, while keeping the system detector independent. For this reason, FELIX has been selected as the ATLAS system of choice for the HL-LHC detector readout and the system has also been adopted by other non-LHC projects like DUNE [80].

The FELIX system consists of a commodity server, equipped with a Network Interface Card and up to two FPGA based PCIe cards, named FLX-712. The card is responsible for handling data inputs and transferring data packets both to and from a host PC. A FELIX software application processes the data packets and sends them over the network to the software of the ReadOut Driver (swROD). The swROD is designed to act as the data-handling interface between the FELIX readout system and the HLT. The swROD implements data fragment building and formatting, which in the Run 1 and 2 systems were done by the detector specific ROD components, like the MROD.

The FELIX firmware supports multiple working modes. The GBT mode uses the GigaBitTransceiver (GBT) protocol [81] and supports up to 24 links, each aggregating several e-links with configurable bandwidth. The GBT protocol, developed at CERN, provides a high speed radiation-hard optical link for data transmission allowing multiplexing data from several front-ends into a single fiber. The FULL mode uses the GBT logic towards the front-end but implements a simple protocol from the front-end which supports 24 links. While 24 links are supported, 12 are capable of saturating the PCIe bandwidth. The use case of FULL mode is the communication with other devices not required to be radiation-tolerant. In addition, to be able to interpret the CSM data format and re-implement the MROD functionalities a custom ”FELIX-MROD” firmware has been developed which is able read out up to 48 CSM modules and transfer the incoming data on the memory of the host PC.

The FELIX software responsible for routing the data from the FLX-712 card to the network and back is called felix-star. It handles the communication between one or more FELIX cards on one side and a set of network clients on the other side. Felix-star has multiple functions, like packet forwarding from the front-ends to the DAQ system and back, configure the FELIX card, recover from host failures and report operational status information, as well to handle more FELIX card if needed.

Since there is a two way communication between the FELIX card and the host, the data handled by the felix-star application has two kinds of data format. ”To Host” data is read from the card and routed to the network subscribers. ”To Host” data chunks are split, by the FELIX firmware, into sub-chunks to fit in 1024 byte blocks with an e-link and sequence number in the header. Felix-star reads the blocks, sorts them by e-link, checks sequence numbers, and recombines sub-chunks into chunks of data. These chunks of data are subsequently published on the network, inserted with the FELIX header. Within the ”From Host” data handling the network clients may send data to felix-star over the network, which is then copied to the card. Beside data handling, FELIX is also required to interface with the ATLAS TTC system. FELIX must provide TTC information both to the front-ends and to network peers in a reduced form. The propagation of TTC information to the front-end is performed via dedicated e-links.

3.2.1 swROD

The FELIX firmware and software does not perform any data processing, but only provides data routing between detector front-end and the DAQ system. The task of data aggregation and processing is fulfilled by the swROD application [82] before transferring data to the HLT farm. Since swROD is used by multiple subdetectors, the swROD application has been designed to supports high degree of customization.

swROD is running on a set of commodity computers. Given that a single server can serve only a limited amount of input data the software needs to be distributed over multiple computers, in order to satisfy the size of the ATLAS readout system. In the current design this is achieved by splitting the input data channels between a number of software processes, which are referred to as swROD applications. Each instance of the swROD application can run on a separate computer, but it originates from the same binary executable. This executable implements support for detector specific event building and data processing algorithms provided in the form of shared libraries (plugins).

The swROD application is split internally into a number of independent components, with each of them providing a simple interface that defines how other components can interact with it. There are three main components defined by the swROD application architecture:

- `DataInput` interface: receives data from the network. It protects the other components of the swROD application from any changes in the network input protocol, and gives a possibility to use another data source for testing and debugging.
- `ROBFragmentBuilder` interface: the main data aggregation algorithm. It is responsible for aggregating data chunks from individual e-links received via the `DataInput` interface into event fragments according to the given configuration. Such a configuration defines the set of event fragments to be produced as well as a list of input links for each fragment.
- `ROBFragmentConsumer` interface: processing the fully aggregated event fragments. At this step it is also possible to apply a custom subdetector specific processing procedure to the event fragments before passing them to the HLT farm.

Each interface contains a set of default implementations which are supplied along with the swROD application. The swROD package provides two fragment building algorithms that can be used to handle data received from FELIX. Both packages are implementations of the `ROBFragmentBuilder` interface. The GBT mode algorithm builds data chunks from multiple e-links into a single ROD data block using L1IDs for alignment. The FULL mode algorithm treats every individual data chunk from any e-link as a completely built ROD data fragment, that may also optionally contain ROD header and trailer. As the incoming data format is detector specific this feature allows subdetectors to supply custom procedures for these algorithms. If a detector specific processing has to be applied to the ROB fragments produced by the fragment building algorithm a custom plugin library has to be implemented through the CustomProcessor interface. The trigger information extraction procedure extracts the L1ID from a given data chunk and uses it to assign data chunks to a particular event fragment. The Data Integrity Checking procedure is used if the input data chunks could be corrupted. In most cases detector developers have only to define these functions and reuse the data aggregation strategies provided by default. The complete swROD interface implementations are shown in Figure 3.5.

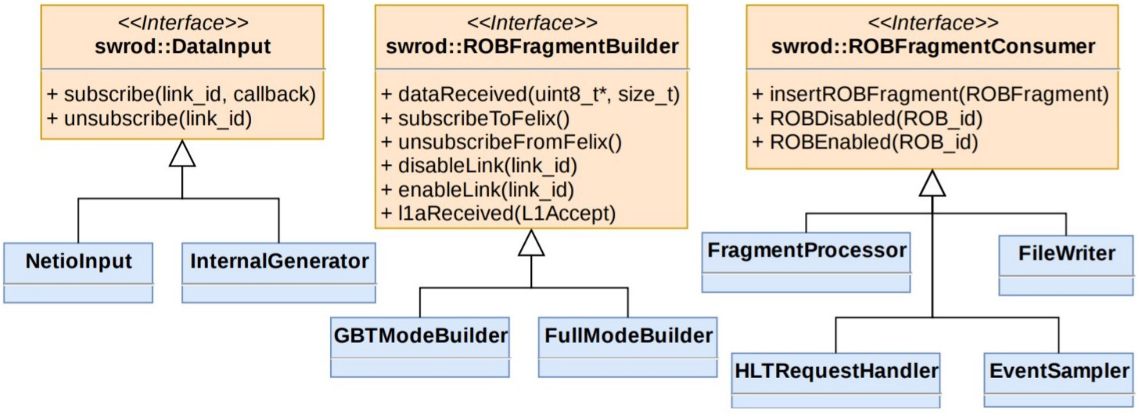


Figure 3.5: Default swROD interface implementations. swROD provides two ROBFragmentBuilder interface: the GBTMode and FullMode builder [82].

3.3 CSMswROD

The goal of the project is to implement the MROD event building functionality within the FELIX readout chain in order to replace MROD readout cards with the FELIX cards. Here the swROD is responsible for event building, therefore it is the suitable platform to implement the MROD functionality.

By default, swROD supplies two event building mechanism, the FULLMode and the GBTMode builder. Both of the implementations expect fully or partially built ROD fragments as an input. Since data arrives in multiplexed from the TDCs in a form of a separator word followed by 18 TDC data words, there is no easily detectable start to an event, nor an easy end. The data first needs to be demultiplexed and slightly annotated before reforming to a ROD fragment is possible. It is apparent that in the case of the MROD none of the default configurations can be used.

To implement the MROD functionality into the swROD a custom ROBFragmentBuilder and processing plugin were developed, using the programming language C++ which is used throughout the entire swROD project. The custom functionality has been dubbed the CSMswROD. This implementation is called via the ROBFragmentBuilder, instead of the built-in FullModeBuilder or GBTModeBuilder, right after the data is received.

The main implementation is the CSMBuilder shown in Figure 3.7, which is inherited from the built-in FullModeBuilder. The main task of the CSMBuilder is to receive L1ID information from the TTC and to compare this information with the L1ID retrieved from the data. It is also possible to operate it without the L1ID comparison and in this case the data is just simply forwarded to the ROBFragmentConsumer. For every instance of the `CSMBuilder` a `CSMWorker` is created. The task of the `CSMWorker` is to receive data, extract the L1ID and send the data fragment to the `CSMBuilder` for trigger information comparison. Before this is possible, the `CSMWorker` initializes the `CSMProcessor`, which is responsible for data rearrangement into ROD fragment.

The `CSMProcessor` receives data, strips off the separator and the filler words, and sorts the incoming data into different buffers based on the TDCs. Within these buffers, the BOT word and the trailer EOT words are also stored separately. The `CSMProcessor` builds the event by sorting the TDC data so all TDC event fragments are starting with the BOT word and terminated by the EOT word. When all the TDC inputs delivered an EOT word the `CSMProcessor` packs the TDC event fragments together. This event building procedure shown in Figure 3.6 is same as in the case of MDT Read Out Driver (MROD) and the most important feature of the `CSMProcessor`.

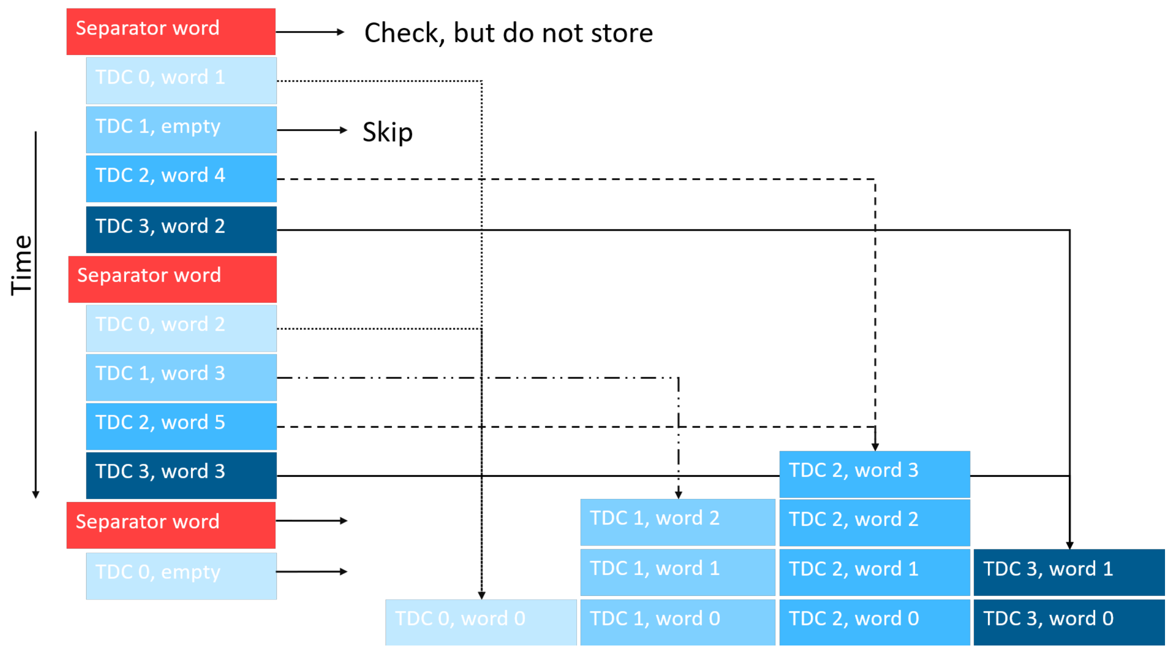


Figure 3.6: The representation of the `CSMProcessor` event building functionality, which was previously performed by the MROD. The left column indicates the incoming data format, which is stored into separated buffers based on the TDC input.

After the event is built, the `CSMProcessor` constructs the header and trailer words corresponding to the CSM and MROD data packets and inserts them to their predefined places within the data fragment. To build the header and trailer words defined in Section 3.1.1, additional information is needed about the MS chambers. These parameters needed to be incorporated to the general `swRODSegment` config file, which contains also other configurations needed for the swROD operation. The `CSMswROD` specific settings are the following:

- `MRODSerialNumber`: Contains the MROD serial number which is switched by the FELIX card. This information is needed for the BOL header word.
- `CSMInputNumber`: Contains the CSM number corresponding to the input link. This information also necessary for constructing the BOL header word.
- `HPTDC`: A simple flag, which determines the type of TDC used in this chamber. This flag is stored in the BOL header word as well.

- `ZeroSuppression` and `TrailerSuppression`: Flags which are indicates if zero suppression or trailer supression is turned on. These flags are also present in the BOL header word.
- `TDCChannelMask`: Indicated which TDCs are present in the given input link. This information is also presented in the TLP header word.

With the parameters above, it is possible to build every header and trailer word needed to construct the full ROD fragment, except the MROD header, the BOB word. This word contains the full L1ID, which is not possible to determine purely from the data. To be able to construct the BOB word, an additional implementation is needed, the `CSMCustomProcessor`. This plugin is called via the `ROBFragmentConsumer`, retrieves the full trigger information from the TTC input and places it to the BOB word. The plugin is also responsible for configuration and addition of the status words in the end of the data fragment.

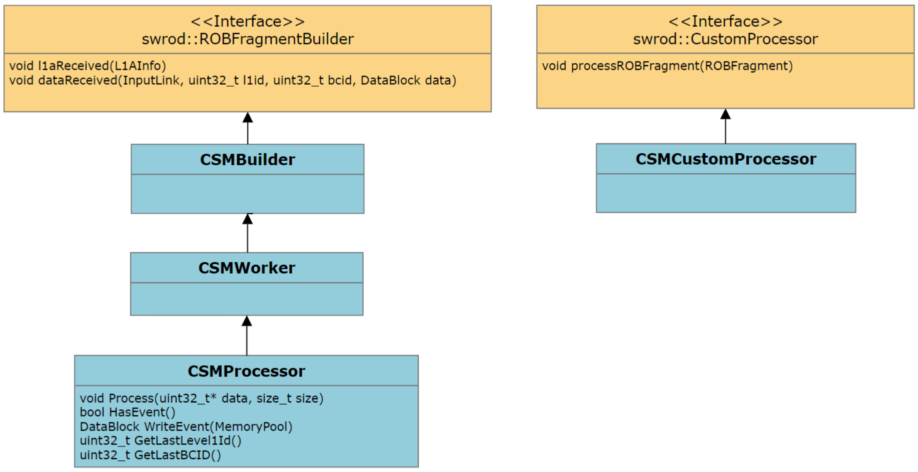


Figure 3.7: `CSMswROD` interface implementations. The `CSMBuilder` is part of the `ROBFragmentBuilder` interface, while the `CSMCustomProcessor` is part of the `CustomProcessor` interface.

In addition to the fragment building, the `CSMProcessor` also validates the data in each TDC fragment, as well performing a synchronization procedure, to make sure that the event fragments with the same L1ID are accepted from the TDC links and are sent out together. At this step the `CSMProcessor` can perform channel suppression in case oneof the TDC channels falls out of synchronization. In this case if multiple channels have data, but one or more is lagging behind, the given channel is turned off, and the TLP word is updated accordingly. The channel is checked on every L1ID rollover, and in case it is supplying data again, it is turned back on. Also an important part of the `CSMswROD` is the error reporting system, which includes all the error check and exeption conditions already defined by MROD. These errors are defined within the Error Reporting System (ERS) of the swROD and are reported alongside with the general swROD warnings or errors. Besides the already defined error conditions, a new feature is also added to the `CSMProcessor`, which was not part of the original MROD system. It is possible that the `CSMProcessor` creates a fake BOT or and EOT word on a given TDC channel when they are missing, which ensures the continuous data processing and building. Both the BOT and EOT word contains information which can be retrieved from the data itself, so it is possible to create them without any intervention. In case this is happens, a flag is set in the corresponding EOT word, to indicate that the EOT or the BOT were faked during data processing. When the complete event is build, it is sent to the `CSMBuilder` and then to the `ROBFragmentConsumer` interface for further processing.

3.4 Testing and integration

FELIX-MROD readout chain has been tested through different levels of integration. During the development phase, the FLX-712 card has been loaded with recorded MDT data and used as data generator. Next, the FELIX-MROD has been deployed in a test setup at CERN that included two BIS-MDT chambers. As a last step, the FELIX-MROD has been installed in the ATLAS counting room and connected to MDT chambers installed on the detector.

3.4.1 Software test with pre-recorded data

The first test of FELIX-MROD involved only `CSMswROD` and was performed without a FELIX card. A file containing pre-recorded MDT data was provided to `felixcore` (predecessor of `felix-star`) and events were sent to `CSMswROD` over the network.

This approach allowed to simulate up to 10 CSMs in a realistic scenario and evaluate the feasibility of implementing the fragment building algorithm in software. `CSMswROD` was found to be able to process incoming messages within 8 μs deploying one thread per CSM on an Intel Xeon Gold 5115 CPU. The processing time distribution is shown in 3.8: input messages containing no hits are processes in shorter time (first peak) compared to messages with hits (second peak).

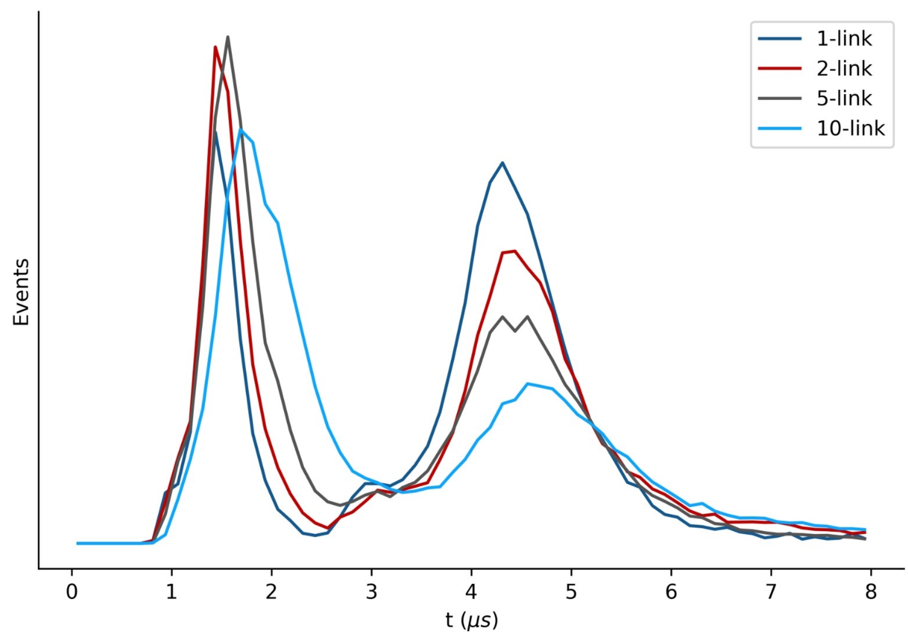


Figure 3.8: Processing time of FELIX messages. With more links more data flows through the software, resulting an interference between the data streams. This means a slightly worse per link performance, which is visible in the width of the peaks.

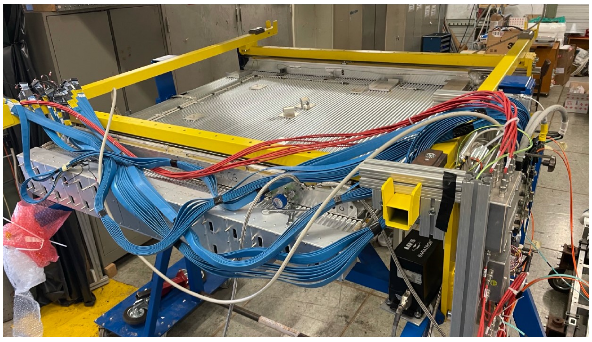
3.4.2 Test with MDT chambers

The first integration test with MDT chambers has been performed in the BB5 facility at CERN in 2021.

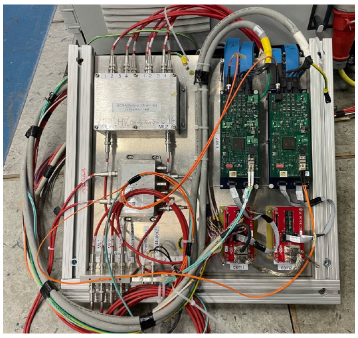
The test setup contain a BIS-MDT chamber, same as the barrel part of the ATLAS detector. For the duration of the test the MDT tubes were filled with gas and connected to a high voltage power supply. Being the Monitored Drift Tube (MDT) active and the BB5 facility on surface, the source of hits were cosmic rays and noise. The chambers were sending data to two CSMs, with a total of 18+13 TDCs present.

The system is connected to the TTCVi system, which simulates the L1 accept signal up to a rate of 100 kHz. Hits were not used to trigger data acquisition, trigger signals were produced independently. In response to L1 accept signals, CSMs produce data in the same format disregarding from the number of hits. Since FELIX-MROD is a data acquisition system, and does not perform any event reconstruction, the presence of hits is irrelevant.

The FELIX firmware and software were running on a PC equipped with an FLX-712 FELIX card. This setup was particularly suitable for testing FELIX-MROD, and it is shown in Figure 3.9.



(a)



(b)

Figure 3.9: Pictures of the test MDT chambers (a) and CSM modules (b) in the BB5 facility at CERN.

No firmware issues were encountered and link alignment was immediate. FELIX-MROD operated reading out two CSMs at a L1 accept rate of 143 kHz. The excess over 100 kHz was due to the imprecise calibration of TTCVi. `CSMswROD` has been run in two modes: data-driven, in which L1IDs are read exclusively from data messages, and TTC-driven in which L1IDs are obtained from the TTC link and searched for in data.

In both modes, stable operation has been recorded as shown in the IGUI screen capture of Figure 3.10. Data fragments produced by `CSMswROD` were dumped to a local disk for analysis, limiting the duration of data-taking periods. No data corruption has been observed in the recorded data. The performance of `CSMswROD` is shown in 3.11: the memory use reaches a plateau at 1 GB while 3.5 cores are used to process data from 2 CSM. However not all CPU resource are used for data processing in `CSMswROD`; a part is used for data reception by the underlying network libraries.

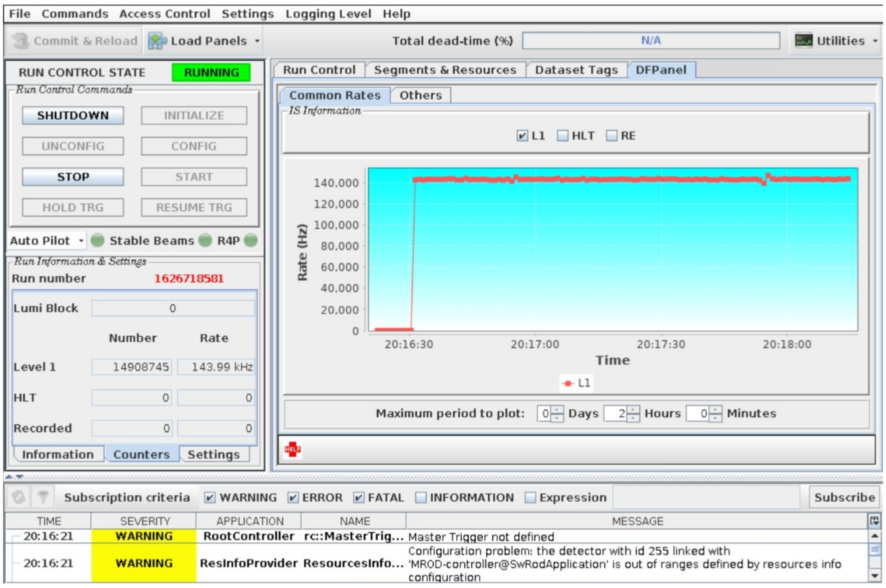
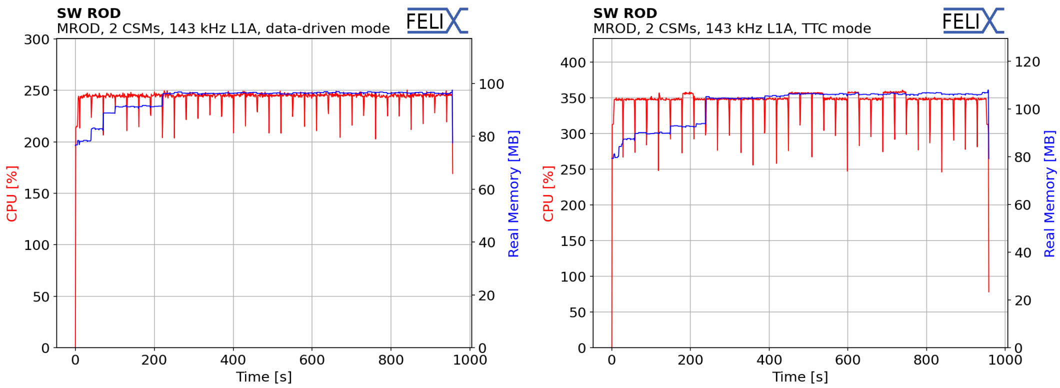


Figure 3.10: Fragment building rate of CSMswROD as displayed in the IGUI. Fragments are build in TTC-driven mode reading two CSMs with 18 and 13 TDCs at L1 accept rate of 143 kHz.



(a) (b)

Figure 3.11: CPU and memory use of CSMswROD while reading two CSMs with 18 and 13 TDCs at L1 accept signal rate of 143 kHz in both data-driven (a) and TTC-driven (b) mode.

3.4.3 Test with ATLAS

One FELIX-MROD PC has been installed in the ATLAS counting room (USA15) in 2021. The 6 inputs of the MROD installed in Y.24-19.A1, crate 7, slot 8 were provided to FELIX-MROD by passive optical splitting. The six inputs correspond to the MDTs in the barrel, side C, tower`0x620004`. The optical power measured by the FLX-712 was sufficient for 5 out of 6 CSMs.

Aiming to include FELIX-MROD in the combined ATLAS partition, `CSMswROD` had been integrated in the Muon segment. However the misconfiguration of the HLT request handler and TDC masks did not allow to run successfully during the 1 hour slot primary assigned for the project. Nevertheless data samples were recorded with low-level FELIX tools and `CSMswROD` during the pilot beam run of October 28-29. Data recorded by FELIX low-level was found to be consistent. The analysis of the 12137 fragments produced by `CSMswROD` revealed that valid fragments were built until the first Event Count Reset (ECR) signal, then empty fragments were produced until a roll-over of L1ID in data messages. The reason for this behaviour is that `CSMswROD` extended the 12 bits of data L1ID adding a local roll-over counter. This implementation keeps the L1ID retrieved from data synchronised with the 24-bit L1ID from the TTC link. However, this mechanism breaks when an ECR happens. An ECR, in fact, resets all the 24-bit of the L1ID from TTC but does not reset the local roll-over counter maintained by `CSMswROD` in the data. `CSMswROD` has been reworked to be able to support ECRs but tests with ATLAS were no longer available.

3.5 Conclusion

The test results, summarized in the previous section, demonstrate that the modified FELIX readout chain will be able to satisfy all requirements that come with the increased data rates in Run 3. `CSMswROD` performs the data processing entirely on software level, which will not only be able to cope with the recommended performance, but also will ensure the system compatibility with commercial hardware. It will also open up the possibility to implement further updates and patches without switching any hardware part of the system. In addition, the system already implements new data recovery processes which was not possible with the previous MROD based read out system. Further on, additional tests are also being preformed to exploit all the possibilities and shortcomings of the system.

The FELIX readout chain is already being utilized as the readout system of the newly installed New Small Wheels. However, the FELIX-MROD project is the only ongoing effort, where the FELIX system is combined with a previously installed readout chain. This gives the project a leading role in the future FELIX readout setup for other sub-detectors and also provides an example implementation for them to follow. This is not only important for the Run 3 operation, but also for the HL-LHC upgrade.

In addition, the project proves that any previous system can be adapted using FELIX cards not just in the scope of high energy physics experiments. The FELIX system can be implemented in any data acquisition systems where high-speed data processing is required. For example, it could be used in high-frequency trading or in production chains, where the production is monitored and automated using different sensors.

4 Object reconstruction

The conversion of the ATLAS detector signals to physics objects is a crucial step in all physics analyses. The signals recorded by the sub-detectors are analysed by event reconstruction algorithms to find the individual particles and determine their kinematics. ATLAS is able to identify electrons, photons, muons, jets, hadronically decaying b- and c-jets, which are then used to select the events with certain properties. As shown in Figure 4.1, different types of particles interact in particular ways with the various detector components.

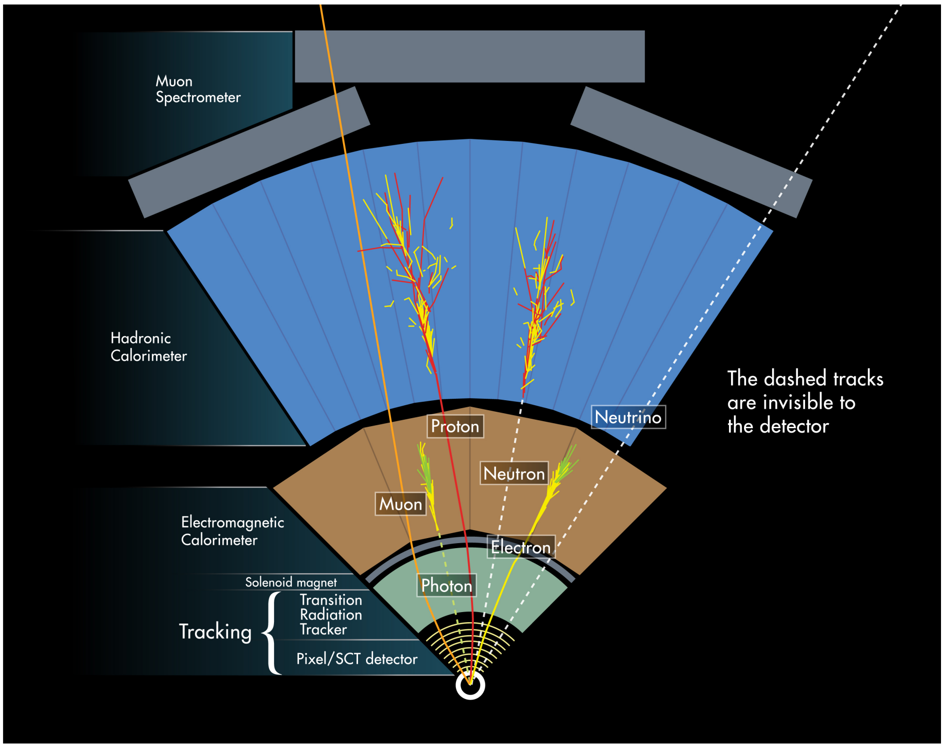


Figure 4.1: Transverse view of an ATLAS detector segment with a representation of the interaction of different types of particles in the different sub-detectors [83].

4.1 Tracks and vertices

A track corresponds to a trajectory of the charged particle within the ID. Tracks are described through a reference point and five track parameters. These are illustrated in Figure 4.2 and correspond to:

- The polar and azimuthal angles θ and ϕ.

- The transverse and longitudinal impact parameter d_0 and z_0, defined as the distance of the closest point of the track in the transverse and longitudinal plane to the reference point. The reference point used is the average position of the pp interactions.
- The ratio q/p of the reconstructed track where q is the charge and p is the momentum.

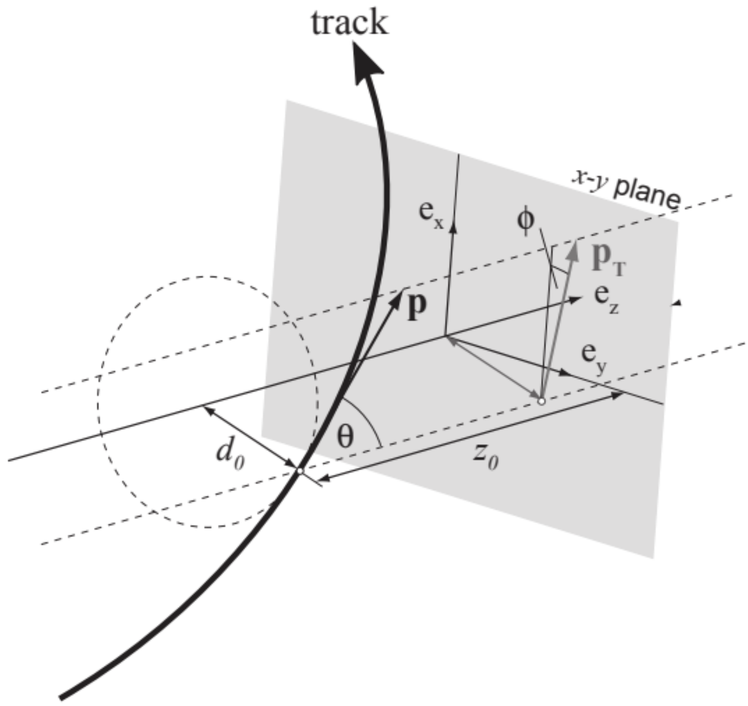


Figure 4.2: Illustration of the track parameters with respect to the perigee. Taken from [84].

In ATLAS, tracking starts using an inside-out track finding strategy, where tracks are formed in the innermost pixel layer at first followed by an outside-in tracking method, where the track formation built starting from the TRT [85]. The approach starts from the track reconstruction in the pixel and SCT detectors. The information is then propagated to the TRT using a combinatorial Kalman filter [86]. Each track candidate is labeled with a score, depending on the presence of missing hits and the weight of the measurement in each sub-detector depending on the precision. A preference is given to high p_T tracks in which improve the elimination of the low p_T tracks formed from incorrect combination of hits. The reconstructed track candidates are then divided in two different categories. The `Loose` category collects all tracks with $p_T >$ 500 MeV and $|\eta| <$ 2.5 with at least seven hits in the silicon detectors, not more than one shared module (pixel and SCT), not more than two missing hits in the silicon detectors and not more than one missing hit in the pixel detectors. To the `Tight Primary` belongs tracks passing the Loose requirements and having at least nine hits in the silicon detectors if $|\eta| \leq$ 1.65 or at least eleven hits if $|\eta| >$ 1.65 at least one hit in the two innermost pixel layers and no missing hit in the pixel detectors. The track reconstruction efficiency for `Loose` (`Tight Primary`) tracks, as calculated from minimum bias simulated events[1], varies from a maximum value of 91% (86%) in the region $|\eta| \leq 0.1$ to a minimum of 73% (63%) in the region $2.3 \leq |\eta| \leq 2.5$ [87].

Once tracks are created, primary vertices are reconstructed in order to identify the hard-scatter interaction, i.e. the interaction having the largest collision energy. Primary vertices are defined as the points in space where proton-proton interactions have occurred. Their efficient and precise reconstruction is an essential element of data analysis, as it is of direct relevance to the reconstruction of hard-scatter interactions and the full kinematic properties of the event. Among the reconstructed vertices, the primary vertex is defined as the vertex with the highest sum of squared transverse momenta of associated tracks. The remaining reconstructed vertices comprise vertices from pile-up and secondary vertices, which are fundamental for b-jet identification.

To be able to reconstruct vertices an iterative vertex finding algorithm [88] is developed, which uses the reconstructed tracks as input. The vertex seed is defined by finding the peak of the z_0 distribution. This seed, together with the tracks around it are added to the adaptive vertex fitting (AVF) algorithm [89], which uses a Kalman Filter based approach to fit the vertex. Tracks which are incompatible with the vertex are used as new seeds for the next iteration. The iterative vertex finding algorithm finishes when all tracks are associated with vertices or when no additional vertices can be found. The vertex reconstruction efficiency as a function of the number of tracks is shown in Figure 4.3, calculated from minimum-bias data and MC events. Only vertices with at least two tracks are considered. It is evident that when more tracks can be associated with the vertex, the higher is the efficiency and almost all vertices are found when they have more than four associated tracks [90].

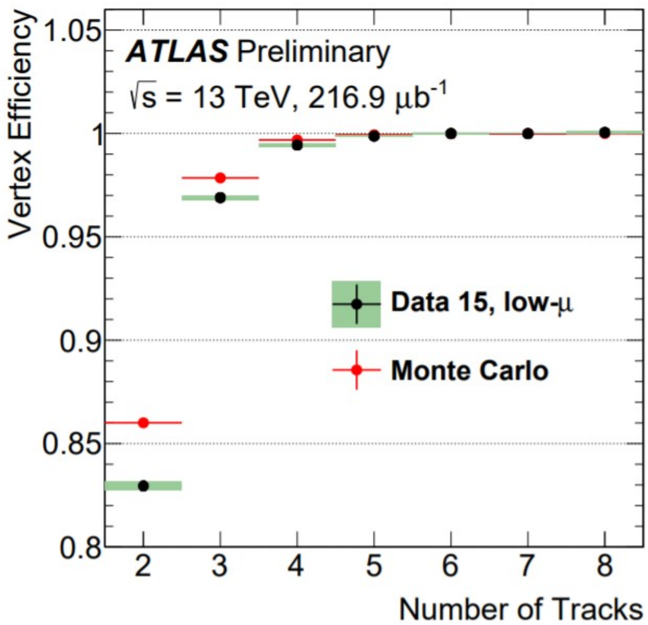


Figure 4.3: Efficiency of vertex reconstruction versus the number of tracks calculated form min-bias data and MC simulation [90].

[1]A minimum bias sample includes all of the inelastic collisions; single and double diffractive and non-diffractive interactions which are selected by a loose trigger with as little bias as possible.

4.2 Electrons and Photons

As it can be seen on Figure 4.1, electrons and photons initiate particle showers in the electromagnetic calorimeter, and therefore, their reconstruction is based on using information from the ECAL and the ID.

The reconstruction algorithm in the ECAL for photons and electrons is based on a so-called sliding window algorithm [91], which determines local energy deposits using dynamic, variable-size clusters of calorimeter cells, called superclusters [92]. The formation of the supercluster starts with clusters selected from energy deposits measured in topologically connected electromagnetic and hadronic calorimeter cells called topo-clusters. Every topo-cluster have to pass a selection criteria to be used as seed cluster candidate and to form the basis of supercluster. Topo-clusters near the seeds, which can emerge from bremsstrahlung radiation or topo-cluster splitting are reconstructed as satellite clusters. The final superclusters are defined when satellite clusters are added to the seed candidates. After the electron and photon superclusters are built, an initial energy calibration and position correction is applied to them.

Due to its charge, only the electron leaves a track in the ID. Therefore, for the electron reconstruction, tracks have to be matched to electron superclusters [93]. The clusters associated with electron candidates must satisfy a set of selection criteria, requiring their longitudinal and transverse profiles to be compatible with electromagnetic showers induced by electrons. Three identification levels, labelled `Loose`, `Medium` and `Tight`, have been defined to identify electrons based on the average electron efficiencies of 93%, 88% and 80% [94]. They provide an increasing level of background rejection at the cost of some efficiency loss. The selections are tightened by using more variables and stricter cuts.

The reconstruction of photons proceeds in parallel to the electrons [93]. Photons can interact with the detector subsystems in different ways, depending if the particle has undergone a conversion to an electron-positron pair, which, in ATLAS, occurs around 50% of the time. In case of a photon conversion before the calorimeter, two electron tracks originating from a displaced conversion vertex are expected to be detected in the ID. Tracks that are matched to the calorimeter cluster are used as input to a conversion vertex finding algorithm, to establish if a photon conversion has taken place. In case of more than one conversion vertex being found by the algorithm, a preference is given to the ones reconstructed from oppositely-charged track-pairs. Therefore, clusters to which neither a conversion vertex candidate nor any track has been matched are considered unconverted photon candidates, but clusters that are matched to a conversion vertex candidate are considered converted photon candidates. The reconstruction of electrons and photons is summarized in Figure 4.4.

The energy calibration of the photons is required to correct the energy response of electrons and photons originating from the energy of a cluster of cells in the electromagnetic calorimeter. The energy resolution of the photon (or electron) is optimized using a multivariate regression algorithm based on the shower properties in the electromagnetic calorimeter. The photon energy scale is corrected using $Z \to ee$ decays, since these corrections are similar for both electrons and photons, while including systematic uncertainties related to pile-up and material effects. In addition, a data-driven validation of the photon energy scale corrections is performed using

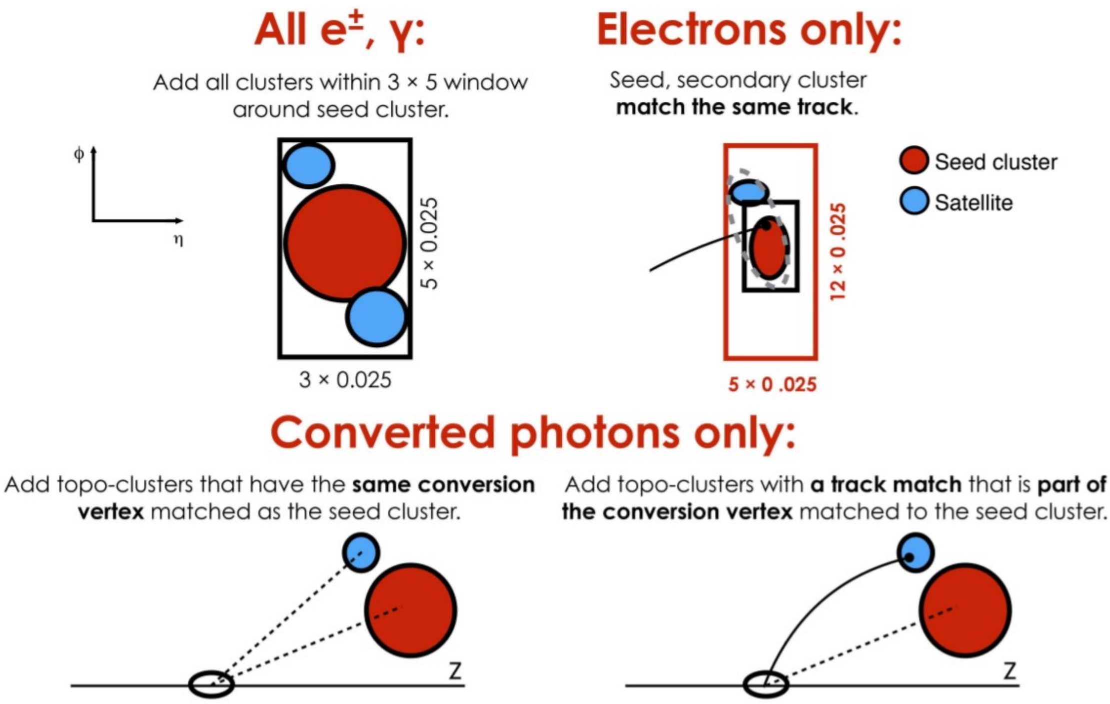


Figure 4.4: Illustration of the superclustering algorithm for electrons and photons. Seed clusters are drawn with red, satellite clusters with blue [93].

radiative decays of the Z boson [95].

To discriminate photons from their possible background, a photon identification algorithm is in place [96]. This helps to reject hadronic jet activity and select prompt photons with a given efficiency. The photon ID is constructed using cut-based selection on the shower shape variables. There are two main working points of identification, the choice of which depends on the purposes of the analysis: `Loose` and `Tight`. The `Loose` identification criteria is designed to provide a very loose level of background rejection along with a high signal efficiency. It is based on shower shapes in the second layer of the electromagnetic calorimeter and on the energy deposited in the hadronic calorimeter. The `Tight` identification criteria provides a high level of background rejection with a high level of identification efficiency. The tight selections add information from the finely segmented strip layer of the calorimeter, and are separately calculated for unconverted and converted photons, to account for the usually broader lateral shower profile of the latter. The photon ID efficiency values ranging from 50–60% at E_T = 10 GeV, to 95–99% (unconverted) and 88–96% (converted) for photons with E_T above 250 GeV.

To distinguish isolated photons from the activity coming either from energy deposits in the calorimeters or from the tracks of nearby charged particles, a photon isolation algorithm is in place [94]. Calorimeter isolation (E_T^{cone}) is defined as a sum of the transverse energy of topological clusters, while the track isolation variable, (p_T^{cone}), is defined as the sum of the transverse momentum of selected tracks within a cone around the photon cluster. Those tracks which are matched to the converted photon are excluded. Three photon isolation working point is defined. The `FixedCutLoose` photon isolation working point is defined primarily for the purpose of the diphoton channel of the Higgs boson search. An isolation cone of $\Delta R = 0.2$

for both calorimeter and track isolation has been introduced to reduce pile-up. The `FixedCutTight` and `FixedCutTightCaloOnly` photon isolation working points were optimised for high E_T photons. For the `FixedCutTight` working point, a selection is applied both on calorimetric isolation with cone size $\Delta R = 0.4$, and on track isolation with cone size of $\Delta R = 0.2$. For the `FixedCutTightCaloOnly` a selection is applied only on the calorimetric isolation. The precise definition of the photon isolation points can be found in Table 4.1.

Working point	Calorimeter selection	Track selection
`FixedCutLoose`	$E_T^{cone20} < 0.022 \times p_T + 2.45\ GeV$	$p_T^{cone20}/p_T < 0.05$
`FixedCutTight`	$E_T^{cone40} < 0.065 \times p_T$	$p_T^{cone20}/p_T < 0.05$
`FixedCutTightCaloOnly`	$E_T^{cone40} < 0.022 \times p_T + 2.45\ GeV$	-

Table 4.1: Definition of the photon isolation working points.

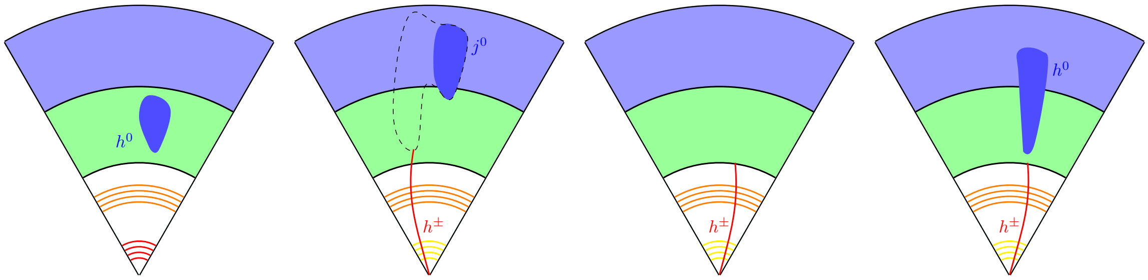
4.3 Jets

Jets are the experimental manifestation of quarks and gluons. Due to colour confinement, when those quarks and gluons are released in a collision, they hadronise with other particles into colourless states, and form a jet. These particles interact with the calorimeters (both the ECAL and HCAL) and with the ID.

To reconstruct a jet candidate the Particle Flow (pFlow) approach is used where calorimeter deposits are combined with the tracking information provided by the ID [97]. In ATLAS, the pFlow algorithm is based on matching a selection of `Tight` tracks to topoclusters. Considering that at very high p_T the momentum resolution of the tracks is lower than the calorimeter measurements, tracks having $p_T >$ 40 GeV are excluded together with tracks matched to medium quality electrons and muons.

The first step of the tracker-calorimeter combination procedure consists of matching between tracks and topocluster in order to identify calorimeter energy deposits caused by charged particles. Later these energy deposits are removed from the calorimeter in order to better identify neutral signals. The final objects provided by this algorithm represents ideal particles in the detector, generally known as Particle Flow Object (PFO). The Neutral PFOs composed of topoclusters not associated to any track or calorimeter signals surviving the subtraction procedure (see Figure 4.5a and 4.5b), while the Charged PFOs composed of isolated tracks or tracks and topoclusters matched by the pFlow algorithm (Figure 4.5c and 4.5d). The inputs to the jet finding algorithm are then represented by the full set of neutral PFOs, and the charged PFOs matched to hard-scatter vertex. This step is allowing to discard calorimeter energy deposits associated to pile-up interactions before building jets. The Particle Flow jets are finally formed by running the $anti - k_t$ algorithm on the input constituents.

The $anti - k_t$ algorithm [98] is used as the main jet reconstruction algorithm in ATLAS, and it iteratively merges pairs of jet candidates until the distance between nearby objects is greater than a predefined value. The distance can be defined, for two candidates i and j, as



(a) Neutral PFO, isolated topocluster. **(b)** Neutral PFO, subtraction remnant. **(c)** Charged PFO, isolated track. **(d)** Charged PFO, track and topocluster

Figure 4.5: Charged and neutral Particle Flow Object (PFO) indicated with the letter h in the ID (white), ECAL (green) and hadronic calorimeter (HCAL) (blue).

$$d_{ij} = min(p_{Ti}^{-2}, p_{Tj}^{-2})\frac{\Delta R_{ij}^2}{R^2}, \tag{4.1}$$

where ΔR is the distance in the $r-\phi$ plane, and R is an arbitrary parameter typically set to $R = 0.4$. In the $anti-k_t$ algorithm, the distance d_{ij} between a soft and a hard particle is dominated by p_T of the hard particle. Therefore the d_{ij} between soft particles will be much larger causing soft particles to cluster with hard ones instead of among each other. The output of the algorithm represents jet candidates within the detector. These candidates are then calibrated through a multiple step procedure to account for several effects, such as out-of-cone radiation or energy deposits below noise threshold. As last step in the jet reconstruction chain, jets are also tagged according to their substructure.

The pFlow approach presents various advantages against other methods, when the jet reconstruction happens purely from calorimetric information. For low momentum particles the tracker momentum resolution is better than the energy resolution of the HCAL, but the calorimeter provides a better measurement at high p_T. This feature is due that at high p_T particle tracks have a very small bending angle, making its momentum estimation difficult. However, the calorimeter energy measurement in this regime is effectively better resulting from the high number of particles in the calorimeter shower. At low p_T , the large stochastic uncertainty and large noise in the calorimeter makes the tracker transverse momentum estimation more suited for a precise p_T estimation. Therefore, it is natural to use the tracker information for low p_T jets while the calorimeter in the high p_T regime. Considering this, improvements can be seen expected in terms of energy and angular resolution for low p_T jets. Additionally, since charged pile-up constituents are already removed before building jets, the pile-up pFlow jets are partially suppressed prior to jet creation.

4.4 Taus

The τ lepton decays either leptonic ($\mathscr{B} \approx$ 35.2%) or hadronic ($\mathscr{B} \approx$ 64.8%) as it is presented in Table 4.2. Both type of decay produces neutrinos which are detected as part of the missing transverse momentum of the event. The remaining products are the visible decay products. In case of a leptonic τ decay this is an electron or muon, which can be reconstructed like any other isolated electrons or muons. However the visible decay products of a hadronic τ decay ($\tau_{\text{had-vis}}$) requires more specialised algorithms.

		τ decay mode	Branching fraction (%)
Leptonic		$\tau^{\pm} \to e^{\pm} + \bar{\nu}_e + \nu_\tau$	17.82
		$\tau^{\pm} \to \mu^{\pm} + \bar{\nu}_\mu + \nu_\tau$	17.39
Hadronic	One-prong	$\tau^{\pm} \to \pi^{\pm} + (\geq 0\pi^0) + \nu_\tau$	50.05
		$\tau^{\pm} \to \pi^{\pm} + \nu_\tau$	10.82
		$\tau^{\pm} \to \rho^{\pm}(\to \pi^{\pm} + \pi^0) + \nu_\tau$	25.49
		$\tau^{\pm} \to \pi^{\pm} + 2\pi^0 + \nu_\tau$	10.81
		$\tau^{\pm} \to \pi^{\pm} + 3\pi^0 + \nu_\tau$	1.34
Hadronic	Three-prong	$\tau^{\pm} \to \pi^{\pm} + \pi^{\mp} + \pi^{\pm} + (\geq 0\pi^0) + \nu_\tau$	14.55
		$\tau^{\pm} \to \pi^{\pm} + \pi^{\mp} + \pi^{\pm} + \nu_\tau$	8.99
		$\tau^{\pm} \to \pi^{\pm} + \pi^{\mp} + \pi^{\pm} + \pi^0 + \nu_\tau$	2.70

Table 4.2: Most common τ decay modes. Hadronic decays can be one-prong or three-prong depending on the number od associated charged products.

As hadrons, $\tau_{\text{had-vis}}$ can almost always be reconstructed as a hadronic jet. Therefore, the starting point of the reconstruction is the jets reconstructed by the method mentioned the previous section. All reconstructed hadronic jets are considered as possible $\tau_{\text{had-vis}}$ candidates.

Tracks detected in the inner detector are related to the $\tau_{\text{had-vis}}$ candidates if they have matching directions. The direction of a $\tau_{\text{had-vis}}$ candidate is defined by the barycentre of the clusters in the jet. Tracks that are in the "core region" $\Delta R < 0.2$ around the $\tau_{\text{had-vis}}$ direction are the associated tracks. Since a $\tau_{\text{had-vis}}$ consists of one or three charged particles over 99.9% of the time, it can be expected that jets originated from $\tau_{\text{had-vis}}$ have exactly one or three associated tracks. Therefore $\tau_{\text{had-vis}}$ candidates are classified based on the number of tracks into one-prong or three-prong.

It is also possible to reconstruct individual τ decay products instead of treating all of them as a collective object [99]. In this case, the momentum of $\tau_{\text{had-vis}}$ is calculated as the sum of individually reconstructed charged hadrons ($h^{\pm}$, predominantly pions $\pi\pm$) and neutral pions (π^0). The major challenge of this approach is to disentangle the energy deposits of the $h^{\pm}$'s and π^0s. To achieve that, again the pFlow approach is used. In Tau Particle Flow, $h^{\pm}$'s are first identified using the associated tracks. For $\tau_{\text{had-vis}}$ that has an energy around or below 100 GeV, the $h^{\pm}$'s can be correctly identified by the tracks about 98% of the time. Since π^0 showers very rarely extend beyond the ECAL, the energy deposits in the HCAL that matches the direction of the $h^{\pm}$ track can be fully matched to the $h^{\pm}$. By subtracting the energy deposits in the HCAL from the energy calculated using the track, one can then estimate the amount of $h^{\pm}$ energy deposits in the ECAL. Then, π^0 candidates are reconstructed using clusters in the ECAL. If the π^0 clusters are overlapping with a reconstructed $h^{\pm}$ shower, the estimated energy deposits from the $h^{\pm}$ are subtracted from the clusters. This systematically disentangles the energy deposits of the $h^{\pm}$'s and π^0 and avoids double counting. After that, a Boosted Decision Tree (BDT) classifier is used to identify real π^0's from the π^0 candidates by exploiting the relatively regular shape and size of π^0 showers. At last, after the $h^{\pm}$'s and π^0's are reconstructed, the full

$\tau_{\text{had-vis}}$ can then be reconstructed by treating the $h^{\pm}$'s and π^{0}'s as constituents and summing their momenta. The reconstruction efficiencies are shown in 4.6.

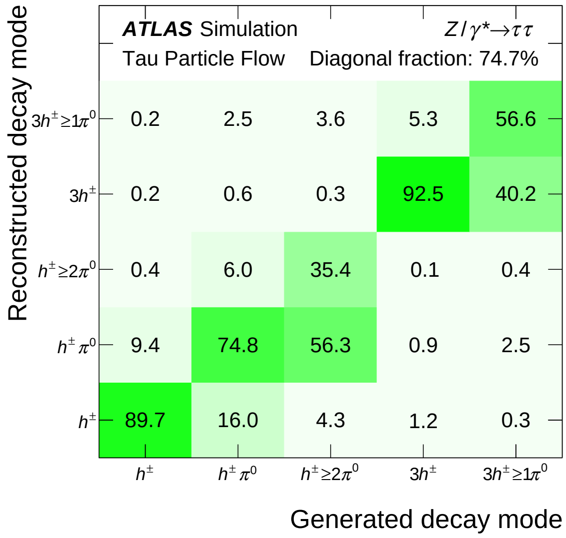


Figure 4.6: Decay mode classification efficiency matrix showing the probability for a given generated mode to be reconstructed as a particular mode by the Tau Particle Flow after final decay mode classification in simulated $Z \to \tau\tau$ events [99].

By taking into account the information from ID, the substructure reconstruction is able to determine the direction of $\tau_{\text{had-vis}}$ more accurate. For low-p_T $\tau_{\text{had-vis}}$ ($p_T <$ 100 GeV), the momentum resolution is also high. Even more importantly, it allows one to determine the actual decay mode of $\tau_{\text{had-vis}}$. The BDT algorithm is capable of classifying a $\tau_{\text{had-vis}}$ into one of the five modes: $\mathtt{1p0n}$, $\mathtt{1p1n}$, $\mathtt{1pXn}$, $\mathtt{3p0n}$ and $\mathtt{3pXn}$, where the number before p represents the number of $h^{\pm}$'s and the number before n represents the number of π^{0}'s, with X denoting $>$ 1 in $\mathtt{1pXn}$ and $\leq$ 1 in $\mathtt{3pXn}$.

Neither taus reconstructed from purely calorimeter information, nor taus reconstructed from the substructure information are optimal for the vast range of physics analyses performed in ATLAS. In this regard, an advanced energy calibration that combines the merits of both of the methods has been developed, which uses information only from the calorimeter. As the first step, the sum of energy of all topoclusters within ΔR <0.2 of the $\tau_{\text{had-vis}}$ candidate is calibrated using the local cell signal weighting method [100]. Furthermore, two additional steps are performed which calibrate the tau energy to the correct energy. First, the energy contribution originating from pile-up is subtracted. A response correction is then applied to account for the particles, whom decay products not reaching the calorimeter, not depositing enough energy to create topoclusters, or not detected within ΔR =0.2 of the reconstructed $\tau_{\text{had-vis}}$ candidate. Additionally, further corrections are are implemented to improve the energy resolution using $Z \to \tau\tau$ MC samples. The final step of the calibration relies on BDT regression algorithm [101]. The BDT is trained to minimise the mean squared error between its output and the training target, and the output of the BDT is a correction factor to the p_T value of the tau which is depending on the MC generated event information information.

Working point	true $\tau_{\text{had-vis}}$ efficiency	fake $\tau_{\text{had-vis}}$ rejection
Tight	60%	70
Medium	75%	35
Loose	85%	21
Very Loose	95%	9.9

Table 4.3: Efficiency for true and fake 1-prong $\tau_{\text{had-vis}}$ measured in simulated event $\gamma^* \to \tau\tau$ and dijet events.

4.4.1 Tau RNN algorithm

At this stage $\tau_{\text{had-vis}}$ candidates can corresponds to actual τ_{had} or be faked by jets initiated by quarks and gluons. The fake $\tau_{\text{had-vis}}$ are suppressed using an identification algorithm based on a Recurrent Neural Network (RNN) to fully exploit the discriminating power in low-level input variables of individual tracks and clusters [102]. Isolated $\tau_{\text{had-vis}}$ typically have one or three associated tracks in the core region, their showers are on average more collimated than jets, especially gluon-initiated ones. $\tau_{\text{had-vis}}$ with π^0's also have distinctive energy deposit patterns in the ECAL. Furthermore, given the relatively long lifetime of τ leptons, the associated tracks of real $\tau_{\text{had-vis}}$ usually form a displaced secondary vertex and have a large impact parameter with respect to the primary vertex.

The RNN input variables can be divided into two categories: low and high-level variables. Low-level input variables to the RNN ID include the impact parameters, angular distance of the track to the $\tau_{\text{had-vis}}$ axis and the number of inner detector hits of the individual tracks. Low-level variables also includes the cluster transverse energy, the angular distance of the cluster to the $\tau_{\text{had-vis}}$ axis and the moments that quantify the longitudinal and radial shapes of the individual clusters.

High-level variables are, for example, the central energy fraction (f_{cent}), the invariant mass of the track system (m_{track}) and the transverse flight path significance (S_T^{flight}). f_{cent} quantifies how collimated the shower is by calculating the fraction of energy deposited in the region $\Delta R < 0.1$ to that in the entire core region $\Delta R < 0.2$. m_{track} is defined as the sum of the four-momenta of all the tracks, assuming a pion mass for each track. S_T^{flight} is the displacement of the secondary vertex in the transverse plane with respect to the primary vertex divided by its estimated uncertainty. It also includes the calibrated and uncalibrated transverse momentum of the original seed jet, the maximum ΔR between the associated core track and the $\tau_{\text{had-vis}}$ direction, as well as the inverse momentum fraction of the leading track ($f^{-1}1_{\text{leadtrack}}$). This variable describes the deposited transverse energy sum divided by the transverse momentum of the highest-p_T core track. The full list of variables can be found in [102].

As the output of the RNN algorithm four working points are defined: `Tight`, `Medium`, `Loose` and `VeryLoose`. The working points are defined with increasing signal selection efficiencies as it is reported in Table 4.3.

Aside from jets, electrons are also a meaningful background source to the $\tau_{\text{had-vis}}$ identification. The RNN ID, developed specifically for discriminating against jets, and it does not provide acceptable discriminating power against electrons. For this reason, a specific e-veto BDT algorithm [103] specialised in telling electrons and $\tau_{\text{had-vis}}$ candidates apart is developed.

Electrons have features similar to those of 1-prong $\tau_{\text{had-vis}}$, notably for those with one neutral pion. Nonetheless, some differences in the detector response are very effective in differentiating electrons and $\tau_{\text{had-vis}}$. Since electron showers are purely electromagnetic, only a very small amount of the electron energy could leak to the HCAL. On the other hand, the $\tau_{\text{had-vis}}$ often deposit significant amount of energy in the HCAL due to the $h^{\pm}$'s. Another effective discrimination is indicated by the response of the TRT. Electrons are always ultrarelativistic due to their small rest mass, therefore they leave more hits in the TRT than the heavier $h^{\pm}$'s from $\tau_{\text{had-vis}}$. The e-veto BDT is trained to exploit these differences. Three working points are defined with increasing efficiencies, `Tight`, `Medium` and `Loose`, as it is shown on Figure 4.7.

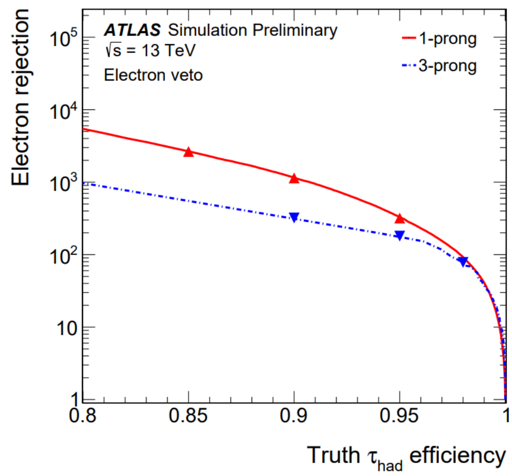


Figure 4.7: Electron rejection for misidentified $\tau_{\text{had-vis}}$ candidates as a function of the efficiency. The two lines refer to 1-prong and 3-prong tau candidates and the markers correspond to the three working points (Loose, Medium, Tight) [100].

4.5 Muons

The muon reconstruction and identification starts with the information obtained from the ID, the MS and to a lesser extent, from the calorimeter. The reconstruction starts independently in each sub-detector and then combined. The reconstruction in the ID is implemented in the same way as for all of the charged particles. The reconstruction in the MS begins with a search algorithm which looks for segments in each MDT and trigger chamber. Tracks in the MS are required to have at least three hits in two MDT chamber layers [104].

Based on the sub-detector information, five types of muons can be distinguished:

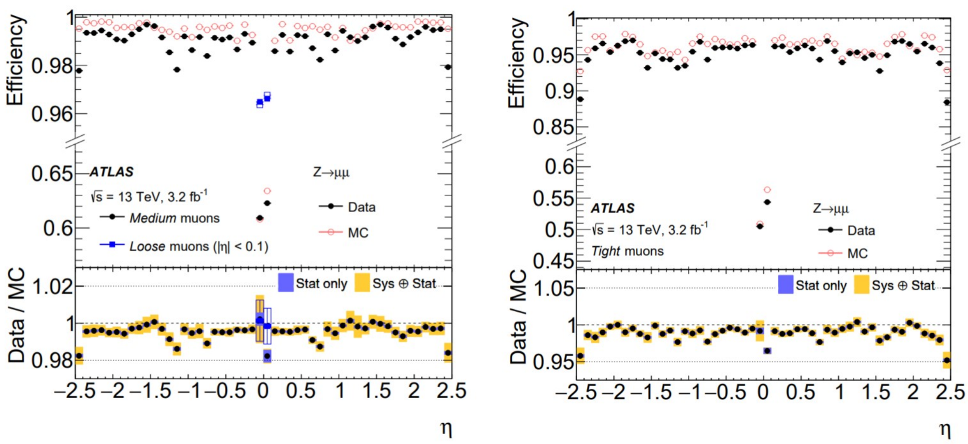
1. Combined (CB): the measurements from the ID and MS are combined using a global fit. Most of the muons are reconstructed with an outside-in algorithm, where the tracks are extrapolated inwards from the MS to the ID.
2. Inside-out combined (IO): using an inside-out algorithm, the hits from the ID are extrapolated to the MS, and loosely matched with at least three MS hits.
3. Extrapolated (ME): muons in this category are reconstructed using the MS

tracks only. These tracks need to be well reconstructed in the MS with segments present in at least two layers.

4. Segment-tagged (ST): a track in the ID associated with at least one local MS track segment. Generally, due to the low p_T, these muons only hit one layer of the MS chambers.
5. Calorimeter-tagged (CT): the track from the ID is combined with an energy deposit in the calorimeter, if the latter is compatible with the signature of a minimum-ionizing particle.

For muons three levels of identification criteria are defined aimed to improve either the purity or the acceptance of signal-like muons. The `Medium` selection is the most commonly used as it grants good acceptance with relatively small fake-rate and uncertainties. `Loose` identification maximises the reconstruction efficiency and it provides good-quality muon tracks, while `Tight` identification is designed to provide maximal purity of muons.

Figure 4.8 shows the muon reconstruction efficiency as a function of η for $p_T > 10$ GeV, as measured using the tag-and-probe method with $Z \to \mu\mu$ events. The `Loose` and `Medium` muons have a uniform muon reconstruction efficiency of about 99% over most detector regions, while `Tight` muons have approximately 95% reconstruction efficiency. The reconstruction efficiency of `Loose` muons is affected by acceptance losses mainly in the region at $\eta \sim 0$, where the MS is only partially equipped with muon chambers.



(a) **(b)**

Figure 4.8: Muon reconstruction efficiency as a function of η measured in $Z \to \mu\mu$ events for muons with p_T >10 GeV shown for the a) `Medium` and b) `Tight` muon selection [105].

Muon isolation algorithms are developed to separate prompt muons from fake muons originating from semi-leptonic decays, which are often enclosed in jets or from muons arising from light mesons. The track-based isolation variable, $p_T^{varcone30}$, equals the scalar sum of the transverse momenta of the tracks with p_T >1 GeV in a cone of size of $\Delta R =$ min$(10\ GeV/p_T^{\mu}, 0.3)$ around the muon of transverse momentum p_T^{μ}, excluding the muon track itself. The calorimeter-based isolation variable, $E_T^{topocone20}$, is defined as the sum of the transverse energy of topological clusters in a cone of size ΔR = 0.2 around the muon. In Run 2 there are seven isolation selection working points defined, which use the combination of track and calorimeter isolation [104].

4.6 Missing Transverse Momentum

The missing transverse momentum, E_T^{miss}, is an important observable for ATLAS as it allows to measure the amount of transverse momentum carried out by invisible particles produced in pp collisions. The transverse momentum in the event is a conserved property, meaning that the final state must count a total transverse momentum component compatible with zero. A significant deviation from a null value can indicate the presence of non-interacting particles, such as neutrinos. The reconstruction of E_T^{miss} is difficult because it require all detector subsystems and a complete and unambiguous representation of the hard interaction. This representation is complicated by the limitations introduced by the detector acceptance and by the impact of pile-up.

The missing transverse momentum, E_T^{miss} is defined as the negative of the vectorial momentum sum of all reconstructed objects in the detector, and it is is characterized by two contributions [106]. First, the hard-event signals, including fully reconstructed and calibrated particles and jets. Second, soft-event signals consisting of reconstructed charged-particle tracks associated with the hard-scatter vertex but not with hard objects. In order to avoid double counting of the same energy deposit, only objects from mutually exclusive detector signals are added in the E_T^{miss} calculation, in a particular order. A priority is given to electrons, followed by photons and other leptons, while jets are rejected if they overlap with accepted higher-priority particles. The lowest priority is given to the tracks belonging to the soft term.

It is important to mention that the calculation of E_T^{miss} is also affected by undetected particles and also detector miscalibrations, resolution errors. Figure 4.9 shows the E_T^{miss} distribution for data and MC. The data and the MC agree within 20% for most part of the distribution, except for high E_T^{miss} regions, where the larger differences are bigger than the total uncertainties [106].

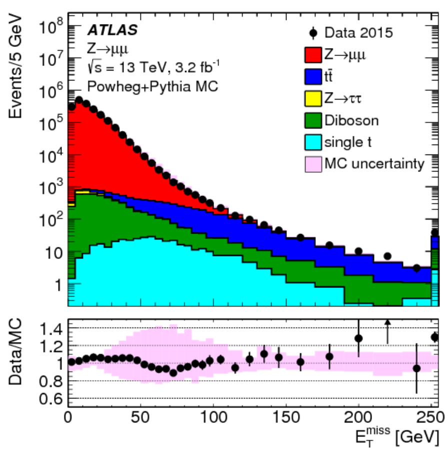


Figure 4.9: Distribution of E_T^{miss} for an inclusive $Z \to \mu\mu$ data sample compared to MC simulations including relevant backgrounds. The shaded areas display the total uncertainty for MC simulations, including the overall statistical uncertainty combined with systematic uncertainties from the p_T scale and resolution [106].

5 Search for $W \to \rho\gamma$ decay

As detailed in the first chapter, the exclusive W boson hadronic decays provide a new validation for QCD factorization approach and are sensitive to the W boson and photon coupling. However, none of these hadronic decays have been observed yet. Previous bounds exist for some of the decays, but not for the $W \to \rho\gamma$ channel. The search described in this chapter is the first ever attempt to perform a measurement of the $W \to \rho\gamma$ decay branching fraction.

The search for $W \to \rho\gamma$ decay uses pp collision data collected by the ATLAS experiment between 2015 and 2018. The dataset contains collision events at the centre-of-mass energy of 13 TeV corresponding to an integrated luminosity of 140 fb^{-1}. The event selection exploits photon and hadronically decaying tau lepton reconstruction algorithms to fully reconstruct the $W \to \rho\gamma$ decay. The W boson candidate is reconstructed selecting a γ and a $\tau_{\text{had-vis}}$, back-to-back in the transverse plane. The search is experimentally challenging due to the large multijet background and as an additional background, the $Z \to e^+e^-$ process is estimated with simulated event samples. The limit on the $W \to \rho\gamma$ branching fraction is estimated using a binned maximum likelihood fit on the W boson invariant mass distribution.

5.1 Signal MC simulation

The $pp \to W \to \rho\gamma$ MC samples have been generated at NLO precision in QCD with `Powheg` using the `CT10` PDF set [107]. `Powheg` is a MC simulation framework, which is able to carry out calculations at NLO level [108]. It also outputs Les Houches Event (LHE) files, which can be processed with other generators to simulate the parton shower and hadronisation. `Powheg` deals with a list of predefined decays and processes. Therefore, to generate $W \to \rho\gamma$ signal samples, first the $W \to \mu\nu$ process is simulated. In the next step, the LHE records are edited: muon and neutrino entries are removed replaced by a meson and a photon. The meson and photon momenta are first computed in the W boson rest frame according to Equation 5.1 and then boosted to the laboratory frame:

$$\begin{aligned} E_\gamma &= (m_W^2 + m_M^2)/2m_W, \\ E_M &= (m_W^2 - m_M^2)/2m_W, \\ |\vec{p}| &= \frac{1}{2}\frac{\sqrt{(m_W - m_M)^2(m_W + m_M)^2}}{m_W}. \end{aligned} \tag{5.1}$$

In Equation 5.1 m_M represents the meson mass and m_W the W boson invariant mass as generated by `Powheg`. This process makes possible to the W boson to decay isotropically into a photon and a ρ meson in the matrix element. In the next step, the LHE events are passed to `Pythia8` [109], where the ρ meson is recognised as such and is decayed to a charged and neutral π meson. `Pythia8` uses different approximations to perform high-multiplicity perturbative QCD calculations, and phenomenological attempts to address non-perturbative physics. This means that each model presents several free parameters which must be optimised to produce a reasonable description of measured observables. This optimisation process is known as tuning, and the resulting parameter sets are referred to as MC generator tunes. The parton shower, hadronisation and underlying event is modelled using the `CTEQ6L1` PDF set [110] and configured according to the `AZNLO` tune [111]. In all the generated events the detector response is simulated with `GEANT4` [64].

5.1.1 Polarisation

The generated $W \to \mu\nu$ process decays isotropically, therefore an additional reweighing is needed to be able to describe the correct angular distribution of the signal process. In this case the ρ meson further decays into two pions, which has a total angular momentum of $J = 0$. The cascade process of $W \to \rho(\to \pi\pi) + \gamma$ has four degrees of freedom represented by the angles of θ and ϕ describing the direction of ρ in the W rest frame and θ' and ϕ' the π emission angles in the ρ rest frame. For the exact representation the cloned cascade frame is used (Figure 5.1) which is defined using the W boson Collins-Soper frame [112]. This means the θ and ϕ angles are defined according to x, y, z in the Collins-Soper frame of the W boson (the z axis is aligned with the bisector of angle between one proton beam and the opposite of the other beam in W boson rest frame), and the θ' and ϕ' angles are defined according to x', y', z', which are exact geometrical clones of x, y, z, used now in the ρ rest frame.

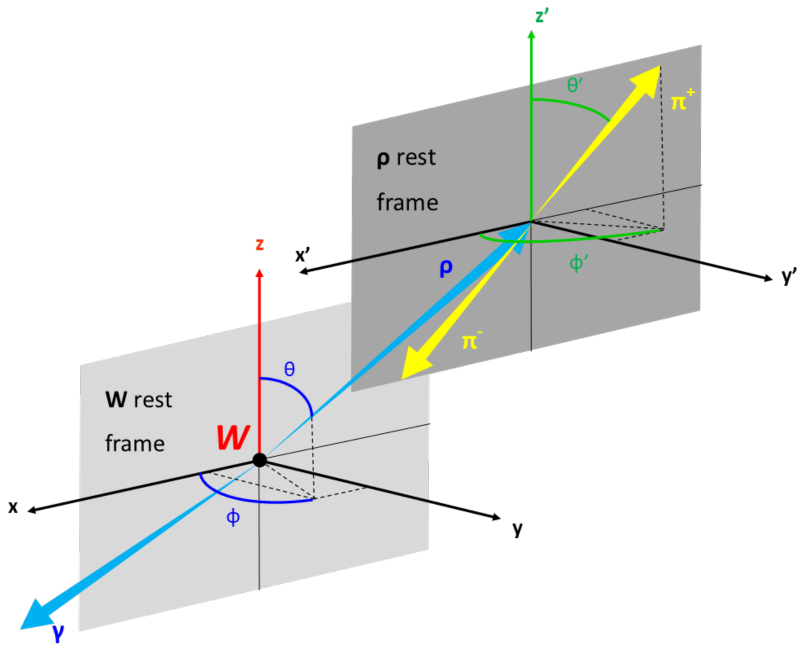


Figure 5.1: The definition of the cloned cascade (CC) polarization frame used in the description of the cascade decay of $W \to \rho(\to \pi\pi) + \gamma$ [113].

The angular distribution for this decay is calculated in Ref. [113] and can be written as:

$$
\begin{aligned}
& W_{\mathrm{CC}}^{\pi\pi}(\cos\theta, \phi, \cos\theta', \phi') \propto \\
& \quad 2 - A_0 \cos^2\theta' + 2\,(A_0 - 1)\cos^2\theta \\
& + (4 - 3A_0)\cos^2\theta\cos^2\theta' \\
& + A_0 \sin^2\theta \sin^2\theta' \cos 2(\phi - \phi') \\
& + \left(1 - \frac{A_0}{2}\right) \sin 2\theta \sin 2\theta' \cos(\phi - \phi') \\
& + \frac{A_2}{2}[\sin 2\theta \sin 2\theta' \cos(\phi + \phi') \\
& \quad + 2\sin^2\theta\sin^2\theta'(\cos 2\phi + \cos 2\phi') - 2\sin^2\theta\cos 2\phi] \\
& + \frac{A_5}{2}[\sin 2\theta \sin 2\theta' \sin(\phi + \phi') \\
& \quad + 2\sin^2\theta\sin^2\theta'(\sin 2\phi + \sin 2\phi') - 2\sin^2\theta\sin 2\phi] \\
& + A_1\{2\sin 2\theta' \cos\phi' \\
& \quad + \sin^2\theta \sin 2\theta'[\cos(2\phi - \phi') - \cos\phi'] \\
& \quad + \sin^2\theta' \sin 2\theta[\cos(2\phi' - \phi) - \cos\phi]\} \\
& + A_6\{2\sin 2\theta' \sin\phi' \\
& \quad + \sin^2\theta \sin 2\theta'[\sin(2\phi - \phi') - \sin\phi'] \\
& \quad + \sin^2\theta' \sin 2\theta[\sin(2\phi' - \phi) - \sin\phi]\}\,,
\end{aligned}
\tag{5.2}
$$

where A_i represent the coefficients of the angular distribution of the $Z \to \ell\ell$ decay. These coefficients can be determined using Z boson polarisation measurements [114] or in first approximation the parameters can be determined as a function of p_T/M and rapidity. As it can be seen in Figure 5.2 the coefficients are rapidly varying as a function of p_T. In the case of the $W \to \rho\gamma$ decay, the W meson p_T is significantly smaller than the boson mass, the polarization is approximately purely transverse in the CS frame, where all coefficients A_i vanish. In this case, the angular distributions of the two decay chain reduce to:

$$
W_{\mathrm{CC}}^{\pi\pi} \sim 1 - \cos^2\theta + 2\cos^2\theta\cos^2\theta' + \frac{1}{2}\sin 2\theta \sin 2\theta' \cos(\phi - \phi'). \tag{5.3}
$$

Generation-level kinematic distributions are shown, with and without reweighing applied, in Figure 5.3.

5.2 Data samples

The analysis is performed on the full Run 2 pp dataset, corresponding to 140 fb^{-1}. All events must belong to a luminosity block that is part of the recommended GRL, so passed the ”good for physics” criteria.

5.3 Object selection

For the analysis, it is necessary to reconstruct a photon and a ρ meson. Photons are required to have $p_T > 20$ GeV, pass the `Tight` identification and the `FixedCutTight` isolation criteria, as described in Section 4.2. To reconstruct the $\rho^{\pm}$ meson a $\tau_{\text{had-vis}}$

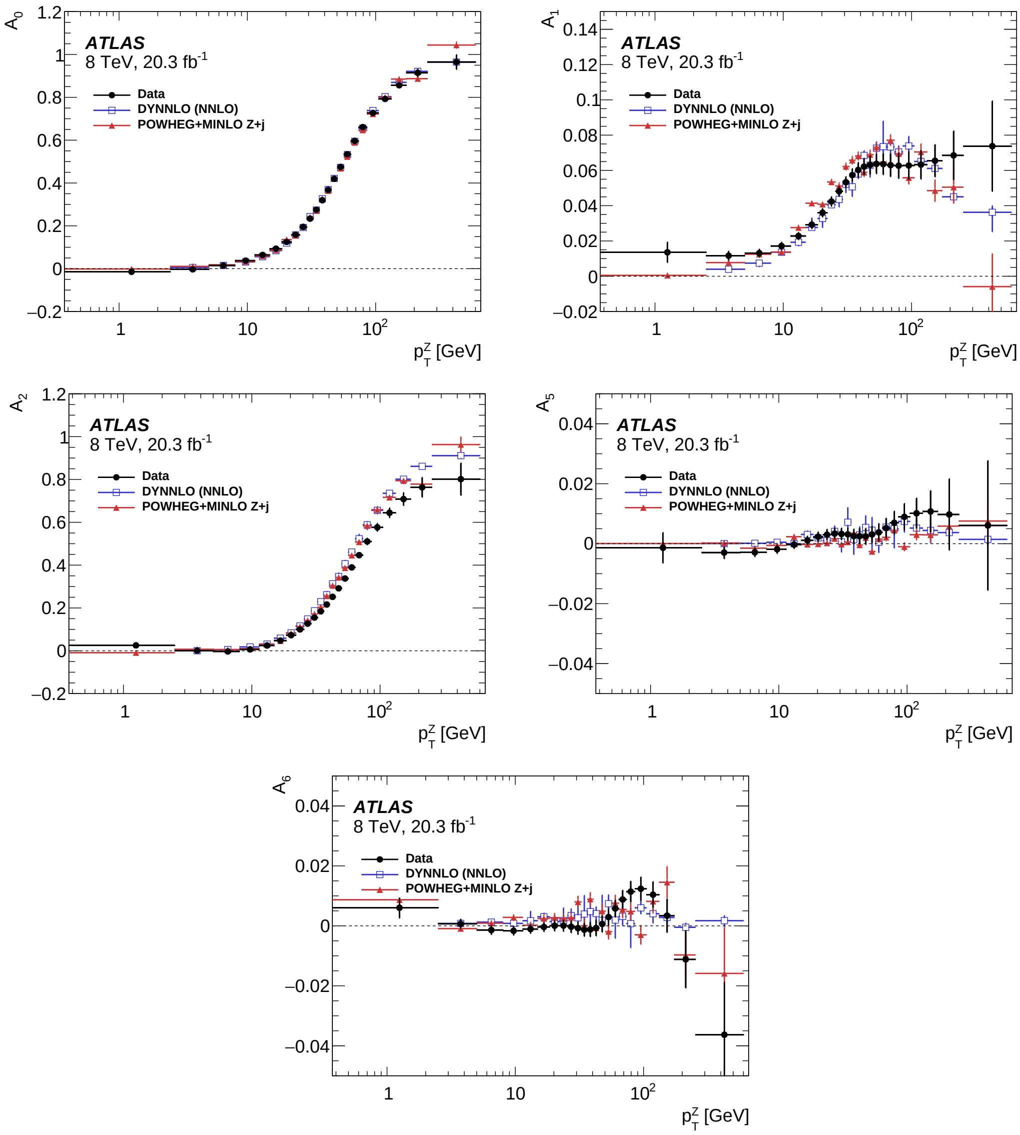


Figure 5.2: Distributions of the angular coefficients A_i as a function of p_T. The results from the measurements are compared to MC predictions [114].

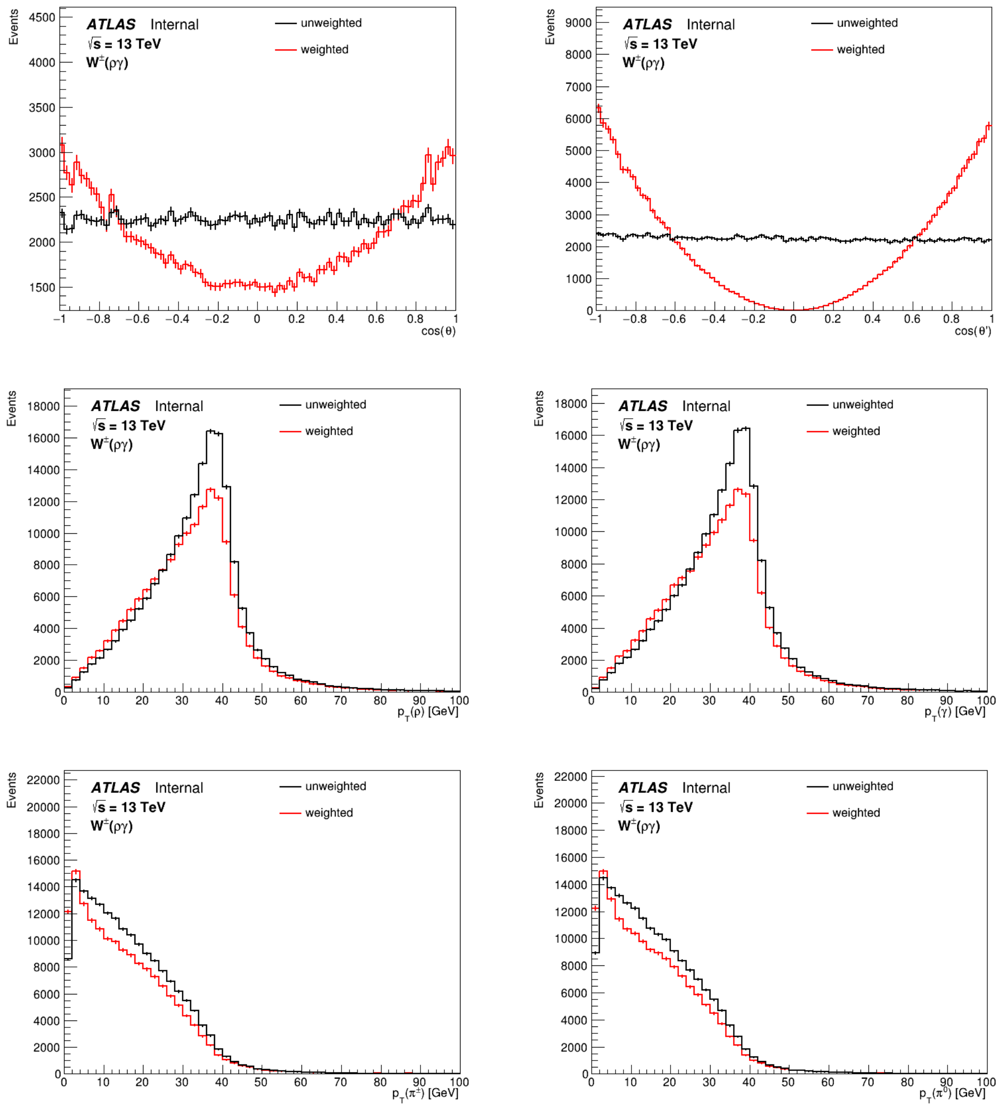


Figure 5.3: MC generated $\cos\theta$ and $\cos\theta'$ distribution and p_T distributions of the ρ, γ, $\pi^{\pm}$, π^0, where $\pi^{\pm}$ and π^0 are the decay products of the ρ meson.

object with $p_T > 20$ GeV and exactly one charged and one neutral PFO is used. The vast majority of $\tau \to \pi^{\pm}\pi^{0}\nu_{\tau}$ decays proceed in fact via an intermediate ρ meson: $\mathscr{B}(\tau \to \rho^{\pm}(\pi^{\pm}\pi^{0})\nu_{\tau}) = 25.5\%$ while $\mathscr{B}(\tau \to \pi^{\pm}\pi^{0}\nu_{\tau}, \text{non-}\rho) = 3.0 \times 10^{-3}$. The $\rho^{\pm}$ decays $\pi^{\pm}\pi^{0}$ with a branching fraction $\sim 100\%$ [5]. It follows that a isolated prompt $\rho^{\pm}$ and a $\tau_{\text{had-vis}}$ with exactly one charged and one neutral PFO are indistinguishable, except for a displacement of the $\pi^{\pm}$ track due to the decay length of the τ lepton. Having a single track, a secondary vertex cannot be reconstructed, therefore the $\pi^{\pm}/\tau_{\text{had-vis}}$ difference can manifest only in the longitudinal and transverse impact parameters. Bearing in mind this potential difference, reconstructing the $\rho^{\pm}$ as a $\tau_{\text{had-vis}}$ has several advantages:

- $\pi^{\pm}$ and $\pi^{0}(\to \gamma\gamma)$ are reconstructed using techniques refined over years of ATLAS operations that could hardly be improved re-implementing them at the level of the analysis,
- the decay mode classification provides a criterion to disentangle the $W \to \pi\gamma$ and $W \to \rho\gamma$ decay channels,
- complex multivariate identification algorithms allows to suppress the electron and the jet background. In comparison with a jet, the $\rho^{\pm}$ meson in $W \to \rho\gamma$ is much more similar to a $\tau_{\text{had-vis}}$ than a jet because of its isolation.

To discard data events which are not of interest in the analysis, a custom data derivation have been developed. The $\rho\gamma$ analysis relies on the `STDM14` derivation, which is characterized by the diphoton trigger and a custom event filter for events having at least one reconstructed photon and τ_{had}.

In addition all events are required to have a primary vertex. This requirements is satisfied by over 99.99% of events in the AOD sample of $W \to \rho\gamma$. The track of the reconstructed τ_{had} is not required to be associated to the primary vertex.

5.3.1 Overlap removal

The overlap removal procedure allows to resolve ambiguities between reconstructed objects using photons, hadronic taus, jets, electrons and muons. The objects provided to the overlap removal are subject to acceptance requirements listed in Table 5.1, which, for the τ and photon are the same described above. The overlap removal algorithm performs the following steps in order:

1. Remove `Loose` electron within $\Delta R < 0.2$ of a tau. This is the only step differing from the standard ATLAS which procedure does the opposite i.e. prefers electrons over hadronic taus.
2. Remove taus within $\Delta R < 0.2$ of a `Loose` muon.
3. Remove any calorimeter-tagged muon sharing a track with an electron within $\Delta R < 0.2$ and then remove any electrons sharing a track with remaining muons.
4. Remove photons within $\Delta R < 0.4$ of electrons and muons.
5. Reject jets within $\Delta R < 0.2$ of an electron and then reject electrons within $\Delta R < 0.4$ of remaining non-pileup jets (i.e. jets passing the Jet Vertex Tagger selection).
6. Reject jets if they are within $\Delta R < 0.2$ or ghost-matched [115] to a muon, and the number of tracks associated to the jet is less than three. Then reject muons within $\Delta R < 0.4$ of remaining non-pileup jets.
7. Remove jets within $\Delta R < 0.2$ of a tau.

8. Remove jets within $\Delta R < 0.4$ of a photon.

Photon	$p_T > 20$ GeV, $\|\eta\| < 2.37$
Hadronic τ	$p_T > 20$ GeV, $\|\eta\| < 2.5$
Jet	$p_T > 20$ GeV, $\|\eta\| < 2.5$
Electron	$p_T > 20$ GeV, $\|\eta\| < 2.47$, $\|d_0\|/\sigma_{d_0} < 5$ and $\|z_0 \sin(\theta)\| < 0.5$ mm
Muon	$p_T > 20$ GeV, $\|\eta\| < 2.5$

Table 5.1: Acceptance requirements applied in overlap removal procedure.

5.4 Event selection

The set of dedicated triggers developed to select $W \to \pi\gamma$ require that the track associated to the meson satisfies the $0.4 < E_T/p_T < 0.85$ requirement [1]. Such a criteria is optimal for targeting $W \to \pi\gamma$ but not $W \to \rho\gamma$ events, as it is evident from the distribution of E_T/p_T displayed in Figure 5.4. The distributions are obtained using simulated events containing one reconstructed τ lepton and one photon, prior the trigger selection.

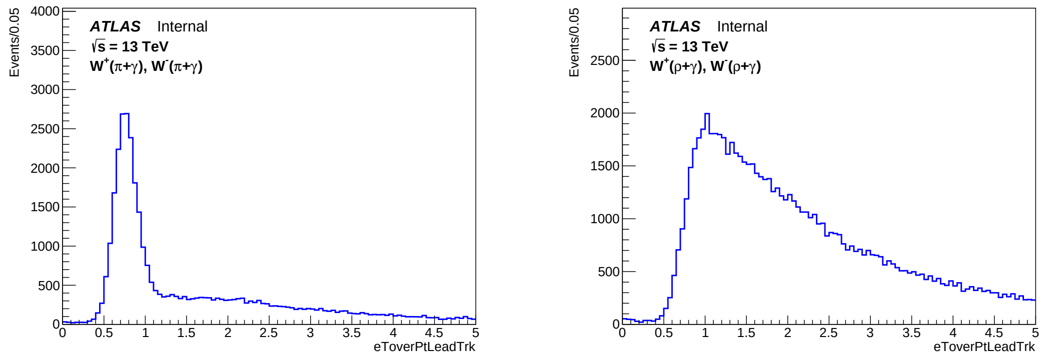


Figure 5.4: E_T/p_T distribution of the track associated to the π (left) and the ρ (right) in the $\pi\gamma$ and $\rho\gamma$ MC samples, respectively.

Because the lack of dedicated triggers, a survey of the triggers active during Run 2 has been carried out. In the absence of a trigger requiring a hadronic τ lepton and a photon, single electron and diphoton trigger were considered. Single τ triggers were not studied due to their high p_T thresholds. Table 5.3 shows the absolute and relative efficiencies of each trigger. Absolute efficiency is defined as the ratio of the number of selected events and the number of generated events, while relative efficiency is the ratio of the current selection and previously applied selection efficiency. The γ and ρ selection is summarised in Table 5.2. The diphoton triggers `HLT_g35_loose_g25_loose` and `HLT_g35_medium_g25_medium_L12EM20VH` are found to provide the highest combined absolute efficiency of 2.5% and relative efficiency of 10.2%. The possibility of combining the diphoton triggers with the dedicated $\pi\gamma$ triggers has been explored but discarded. The marginal efficiency gain ($< 1\%$) would not be beneficial because of the additional complication in the event selection.

[1] Electrons have E_T/p_T value close to 1, while muons peak at low values.

Objects	
photon	$\tau_{\textbf{had}}$ (ρ candidate)
$p_T > 20$ GeV	$p_T > 20$ GeV
Tight ID	1 track
Tight isolation	$h^{\pm}\pi^0$ decay mode [99]
Selections	
γ **selection:** ≥ 1 photon	
ρ **selection:** ≥ 1 photon & $\geq 1\tau_{\text{had}}$	

Table 5.2: Event selection applied on the $W \to \rho\gamma$ MC sample used to measure the trigger efficiency.

During event selection, all events must pass the trigger listed in Table 5.4. The triggers used are the lowest unprescaled diphoton triggers available in each period. `HLT_g35_loose_g25_loose` was active in 2015 and 2016. It requires the presence of two photons passing the `Loose` identification requirement with $p_T > 35$ GeV and $p_T > 25$ GeV respectively. The second trigger was active in 2017 and 2018. The `HLT_g35_medium_g25_medium_L12EM20VH` trigger increases the identification requirement to medium and is seeded by a L1 trigger requiring two energy deposits with $E_T > 20$ GeV (η-variable threshold) and a veto on hadronic energy deposits.

Analysing the trigger matched objects in the signal MC sample, it can be concluded that the diphoton trigger is activated by the selected photon and the tau. This means that the trigger is used on a not envisaged way and that for this purpose the trigger performance and uncertainties are not provided. Therefore, to estimate the performance of the photon trigger on the hadronically tau leptons, an analysis using the photon+muon trigger `HLT_g25_medium_mu24` together with the $Z \to \tau^+\tau^-$ MC sample has been carried out. This trigger require the presence of a photon passing the medium identification requirement with $p_T > 25$ GeV and a muon with $p_T > 24$ GeV respectively. In this trigger the corresponding photon leg is the same as the second photon leg in the `HLT_g35_medium_g25_medium_L12EM20VH` trigger. In the case of the $Z \to \tau^+\tau^-$ sample one of the tau is triggered by the muon leg, while the other tau is triggered by the photon leg.

A set of selection requirements have been applied to the data and MC, to suppress the presence of other background processes. These requirements are presented in Table 5.5, and the variables, before the selection requirement is applied, are presented on Figure 5.5. On top of the selection requirements, background samples also have been processed to further increase the agreement between the MC samples and data.

The result of the MC and data comparison is shown in Figure 5.6, where only the relevant background processes are presented. Although dijet, $Z \to e^+e^-$ and single photon samples were also processed, marginal amount of the events survived the selection, so they are excluded from the plot. As it can be seen, the agreement is satisfactory under the Z reconstructed mass peak where our signal region lies, meaning that no correction is needed on the tau objects triggered by a photon trigger.

2016				
Sample	$W^- \rightarrow \rho\gamma$		$W^+ \rightarrow \rho\gamma$	
Selection	Abs. ε	Rel. ε	Abs. ε	Rel. ε
γ selection	38.0%	38.0%	32.8%	32.8%
ρ selection	23.1%	60.8%	19.5 %	59.6%
HLT_e26_lhtight_nod0_ivarloose	0.15%	0.64 %	0.15%	0.75 %
HLT_g35_loose_g25_loose	3.23%	14.0 %	2.85%	14.6 %
HLT_2g20_tight	2.17%	9.39 %	1.88%	9.64 %
HLT_2g22_tight	2.04%	8.85 %	1.79%	9.19 %
Dedicated $\pi\gamma$ triggers	0.46%	2.01 %	0.39%	1.98%
All previous triggers combined	4.47%	19.4 %	3.87%	19.8%
2017				
Sample	$W^- \rightarrow \rho\gamma$		$W^+ \rightarrow \rho\gamma$	
Selection	Abs. ε	Rel. ε	Abs. ε	Rel. ε
γ selection	34.1%	34.1%	29.8%	29.8%
ρ selection	21.5%	63.0%	18.2 %	61.1%
HLT_e26_lhtight_nod0_ivarloose	0.12%	0.57 %	0.17%	0.93 %
HLT_g35_medium_g25_medium_L12EM20VH	1.95%	9.01 %	1.60%	8.77 %
HLT_2g20_tight_icalovloose_L12EM15VHI	1.62%	7.57 %	1.24%	6.80 %
HLT_2g22_tight_L12EM15VHI	1.82%	8.47 %	1.43%	7.84 %
HLT_2g25_tight_L12EM20VH	1.80%	8.41 %	1.45%	7.97 %
Dedicated $\pi\gamma$ triggers	0.52%	2.42 %	0.46%	2.53%
All previous triggers combined	3.16%	14.7 %	2.62%	14.4%
2018				
Sample	$W^- \rightarrow \rho\gamma$		$W^+ \rightarrow \rho\gamma$	
Selection	Abs. ε	Rel. ε	Abs. ε	Rel. ε
γ selection	34.5%	34.5%	30.1%	30.1%
ρ selection	21.8%	63.0%	18.6 %	61.9%
HLT_e26_lhtight_nod0_ivarloose	0.18%	0.81 %	0.12%	0.64 %
HLT_g35_medium_g25_medium_L12EM20VH	1.95%	8.97 %	1.57%	8.41 %
HLT_2g20_tight_icalovloose_L12EM15VHI	1.45%	6.66 %	1.17%	6.30 %
HLT_2g22_tight_L12EM15VHI	1.66%	7.61 %	1.33%	7.13 %
HLT_2g25_tight_L12EM20VH	1.67%	7.69 %	1.38%	7.38 %
Dedicated $\pi\gamma$ triggers	0.58%	2.67 %	0.44%	2.34%
All previous triggers combined	3.20%	14.7 %	2.56%	13.72%

Table 5.3: Absolute and relative trigger efficiency measured on the $W \rightarrow \rho\gamma$ MC sample applying the selections defined in 5.2.

2015-16	HLT_g35_loose_g25_loose
2017-18	HLT_g35_medium_g25_medium_L12EM20VH

Table 5.4: Triggers used in the $W \rightarrow \rho\gamma$ analysis.

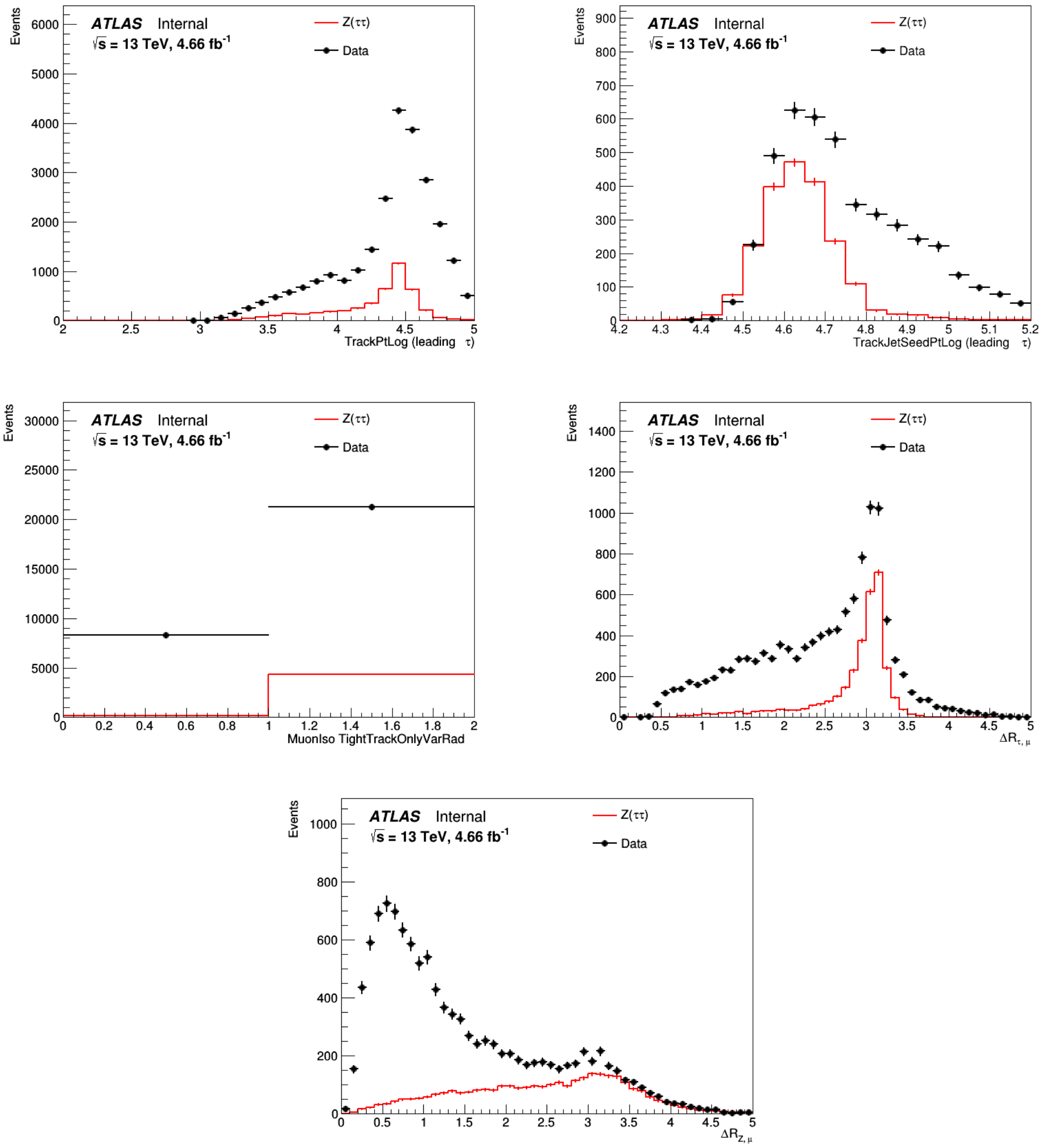


Figure 5.5: Distribution of the variables used to suppress background processes. The data correspond to a subset of the Run 2 dataset.

Tau requirements
$p_T > 26$ GeV, $\|\eta\| < 2.5$ + crack veto ($1.37 < \|\eta\| < 1.52$)
Tight ID, Tight isolation,
TauTrackPtLog> 4.0, TauTrackJetSeedPtLog< 4.8
Muon requirements
$p_T > 24$ GeV, $\|\eta\| < 2.7$
Medium working point,
`TightTrackOnly_VarRad` isolation
Global requirements
At least one tau and muon.
Photon, electron and jet veto.
$\Delta R(\tau, \mu) > 2.8$
$\Delta R(Z, \mu) > 2.0$

Table 5.5: Selection applied on data and on MC samples used to calculate the trigger performance. The MC samples include the $Z \to \tau\tau$ as the signal and the relevant background processes as well.

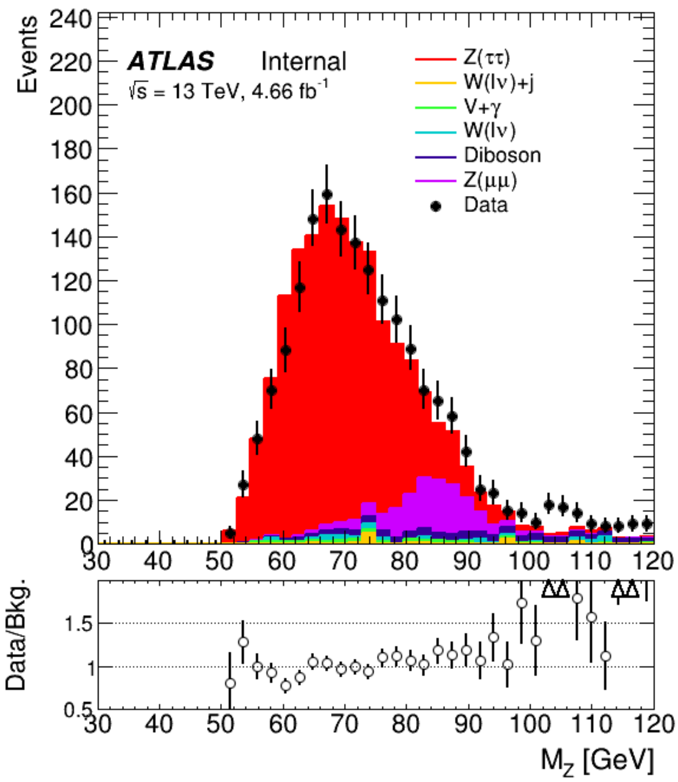


Figure 5.6: Reconstructed Z invariant mass. The data correspond to a subset of the Run 2 dataset.

5.5 Background modelling

The main source of background are the inclusive photon plus jet and dijet events. Here jets are reconstructed as a tau and possibly as a photon. Neither the shape nor the normalisation of these background processes can be precisely modelled with the MC approach due to the complex combination of contributing mechanisms. To include these processes a non-parametric data-driven approach were used, which makes it possible to model this background and to calculate its contribution.

The data-driven approach require different kinematic and isolation variable distributions of a large sample of loose $\rho\gamma$ data candidates. This region corresponding to the generation region (GR) and it covers a phase space similar to the one of the signal process but the selection is loose enough to make the signal contamination negligible. In the next step, these distributions and the correlations between them are used to generate an ensemble of toy $\rho\gamma$ candidates. These toy $\rho\gamma$ candidates are then required to pass the same selection criteria as the data. These selected toy candidates are used to form sample which models the inclusive background. The phase space described by these selected toy candidates is further on referred to as the Signal Region or SR. Three validation regions VR1, VR2 and VR3, are also defined to validate the background modelling procedure.

5.5.1 Region definition

All events considered in the analysis must pass the trigger selection, presented in Section 5.4, and the selection criteria listed in Table 5.6, which defines the GR. All the quantities used have been introduced in Section 5.3, except for the specific set of requirements applied to the τ objects to suppress the $Z \to ee$ background. The critical effect of the selection requirements and their optimisation is described in Section 5.5.3. In short, the $Z \to e^+e^-$ background affects the signal sensitivity and, as expected, it is not well modelled by the data-driven background estimation due to its resonant natures, so it has to be suppressed already in the GR.

The SR is defined by applying three additional selection requirements on the GR: $p_T(\tau) > 32$ GeV, $\Delta R_{\text{max}} < 0.067$, $\log(|d_0(\tau)|) < -1$, where ΔR_{max} is the maximum ΔR between the track associated with the $\tau_{\text{had-vis}}$ candidate and the $\tau_{\text{had-vis}}$ direction and d_0 represents the transverse impact parameter of the τ track. The application of one cut at a time defines three validation regions, as summarised in Table 5.7. The $p_T(\tau)$, ΔR_{max} and the absolute value of the τ track transverse impact parameter on its logarithm have been chosen as they are the most discriminant variables. Both the ΔR_{max} and $\log(|d_0(\tau)|)$ variable is an input variable of the RNN, as it is described in Section 4.4.1. Many other variables have been scrutinised, including track-associated quantities, but none displayed any useful separation. The effect of other RNN variables on $W \to \rho\gamma$ signal and background is studied in Section 5.5.2.

The specific value of the selection criteria has been determined by a three-dimensional optimisation with `blackbox` multidimensional optimisation [116] using the figure of merit $S/\sqrt{B}$ in the SR where S denotes the number of signal events using the signal MC sample and B is the number of estimated background events. The maximum number of iterations is set to 200, as visible in Figure 5.7. The final region definition is summarised in Table 5.7. The figure of merit has been chosen because of its linearity with respect to re-scaling of the signal normalisation.

To check the contamination of $W \to \pi\gamma$ and $W \to K\gamma$ final states, the selection described above was also applied on the $W \to \pi\gamma$ and $W \to K\gamma$ MC samples. All the events were discarded already during the GR selection, which means that the hadronic τ and photon final state only applicable for the $W \to \rho\gamma$ process.

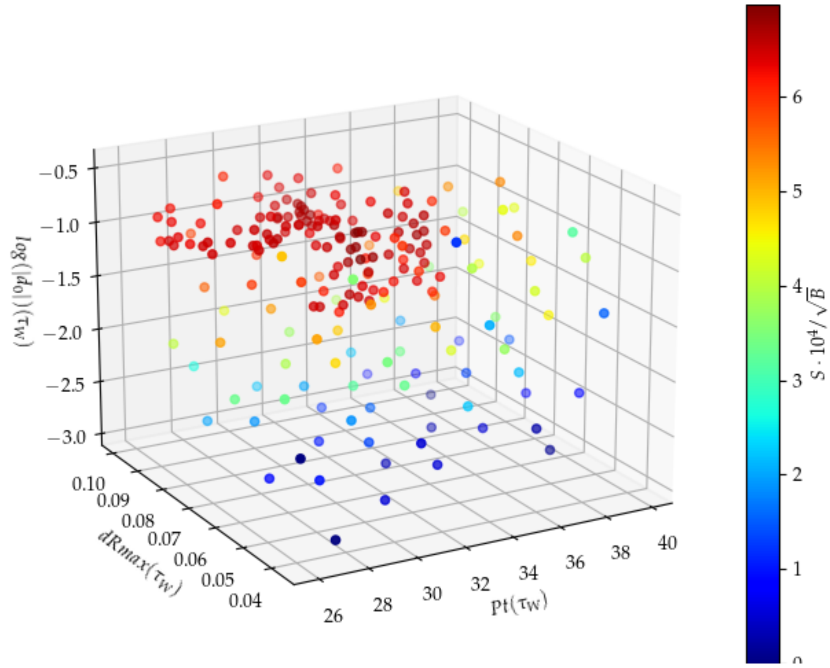


Figure 5.7: Optimisation of the signal region (SR) selection requirements with a total of 200 iterations. The optimisation were performed on a subset of the Run 2 dataset.

Almost all events passing the GR have one photon and one τ, making the reconstruction of the W candidate unambiguous. If more than one τ, γ pair can be build the one with the $\Delta\Phi$ closer to π is chosen. The true nature and origin of the candidate W decay products, τ_W and γ_W, has been checked on simulated signal events. For all signal events in the GR all selected τ and γ are associated to γ and ρ originating from the W boson decay. Figure 5.8 shows the photon and τ multiplicity as well as the correct W matching according to the event generation records.

5.5.2 Tau RNN input variables

As it is already mentioned in Section 5.5.1, many variables have been studied to find the ones with the most discriminating power between signal and background. The Tau RNN score showed enough discrimination, but the use of a high-level variable in the background model is not recommended. To have some insight on the impact of the RNN on the $W \to \rho\gamma$ analysis, the RNN input variables have been plotted in the GR and in the GR with the RNN `Tight` requirement applied. The RNN has been designed and trained with the aim to distinguish hadronic taus from jets originating from other kind of particles. A complex multivariate algorithm such as the RNN could be sensitive to the small differences between a prompt $\tau^{\pm}$ meson and a $\rho^{\pm}$. The full list RNN input variables are listed and explained in Section 4.4.1 and can be grouped in three categories: track variables, cluster variables and high-level variables. In the following some the most relevant track and high-level variables are shown in

Photon requirements
$p_T > 20$ GeV, $\lvert\eta\rvert < 2.37$ + crack veto ($1.37 < \lvert\eta\rvert < 1.52$)
Tight ID, Tight isolation,
τ **requirements**
$h^{\pm}\pi^{0}$ decay mode
$p_T > 20$ GeV, $\lvert\eta\rvert < 2.5$ + crack veto ($1.37 < \lvert\eta\rvert < 1.52$)
Medium RNN score
$Z \to ee$ veto requirements:
Tight TauEleBDTScore
etOverPtLeadTrack > 2.4
$\Delta R_{\text{max}} > 0.036$
eProbabilityHT (associated to the tau track) < 0.9
Global requirements
At least one photon with $p_T > 36$ GeV (trigger threshold $+1$ GeV)
At least one τ with $p_T > 26$ GeV (trigger threshold $+1$ GeV)
At least a τ, γ pair with $\Delta\Phi(\tau_{\text{had}}, \gamma) > 2$

Table 5.6: Selection defining the Generation Region.

VR1	GR requirements + $p_T(\tau) > 32$ GeV
VR2	GR requirements + $\Delta R_{\text{max}} < 0.067$
VR3	GR requirements + $\log(\lvert d_0(\tau)\rvert) < -1$
SR	All the requirements listed above (VR1 + VR2 + VR3): $p_T(\tau) > 32$ GeV, $\Delta R_{\text{max}} < 0.067$, $\log(\lvert d_0(\tau)\rvert) < -1$

Table 5.7: Selection applied on the GR to define the signal and validation regions.

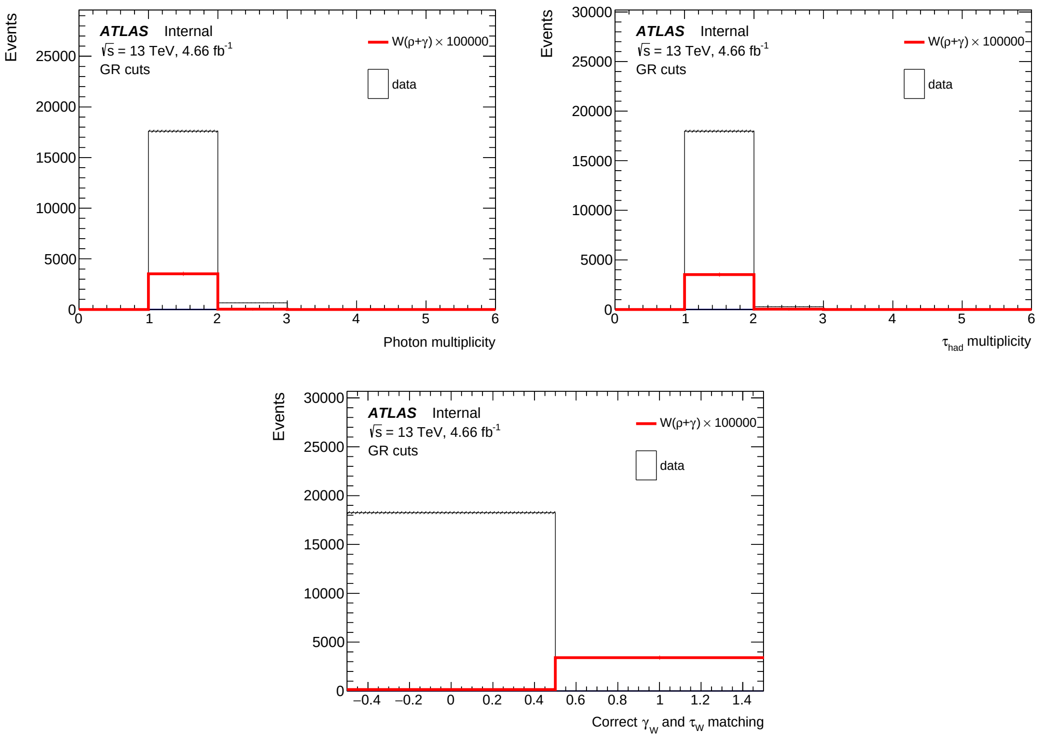


Figure 5.8: Multiplicity of photons and τ's in the GR; fraction of correct W boson-matched objects. The data correspond to a subset of the Run 2 dataset.

Figure 5.9 and Figure 5.10 respectively. Cluster variables are computed for each single cluster (there are up to 10 clusters in each event) and are not expected to differ between a prompt $\rho^{\pm}$ and a $\rho^{\pm}$ as shown in Figure 5.11. When the RNN is applied data distributions become more similar to the signal ones, while signal distributions are not particularly affected. In fact, the only variables exhibiting some discriminating power are the ΔR_{max} (introduced already in Section 5.5.3 and shown in Figure 5.10) and the track impact transverse impact parameter (Figure 5.9).

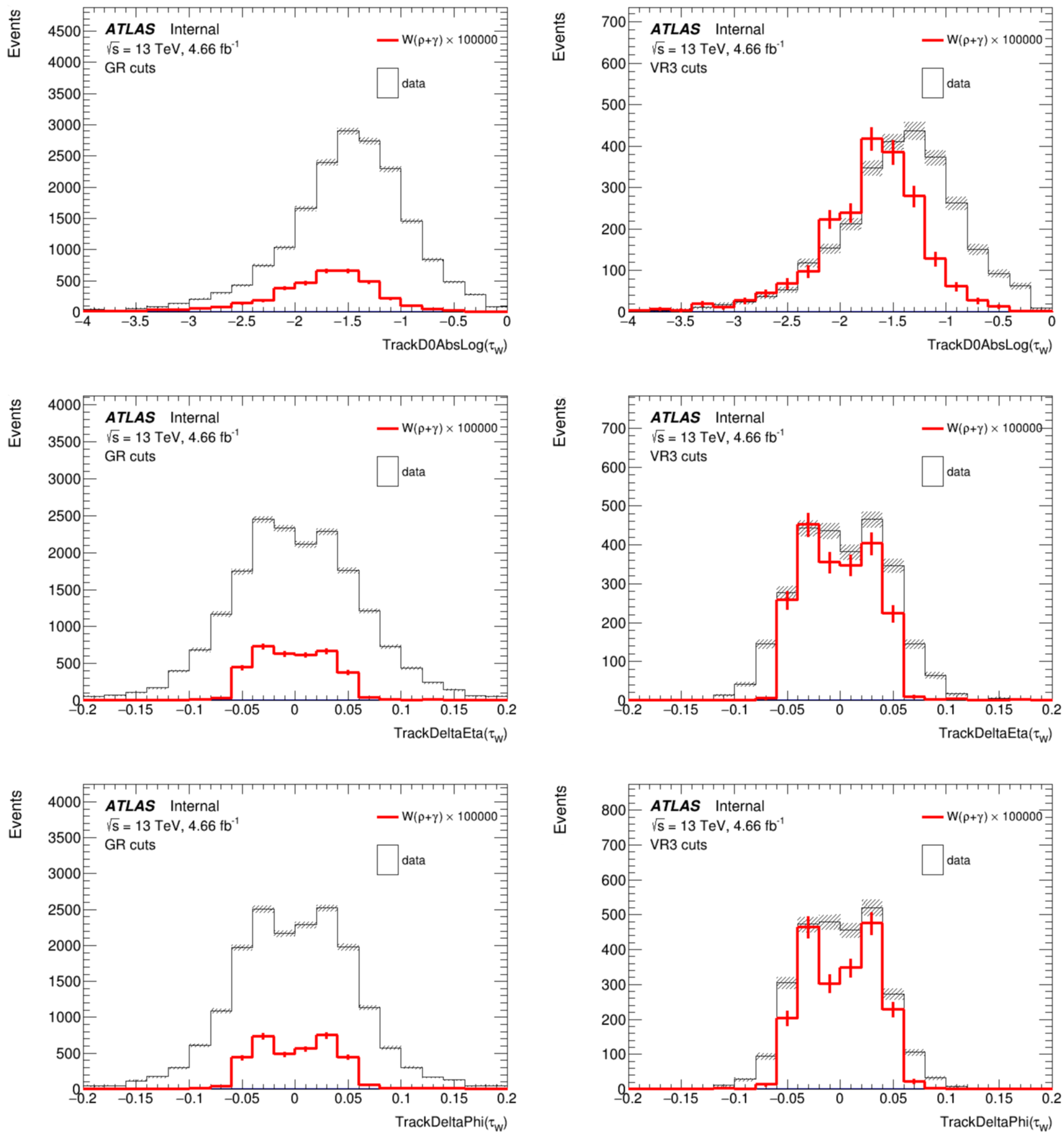


Figure 5.9: Track variables used as input by the RNN. The left column displays the GR, the right column the GR when the `Tight` RNN is also applied.

5.5.3 Suppression of the $Z \to e^+e^-$ background

It is observed that events selected with the diphoton trigger and that require presence of a reconstructed hadronic tau and a photon is dominated by the $Z \to e^+e^-$

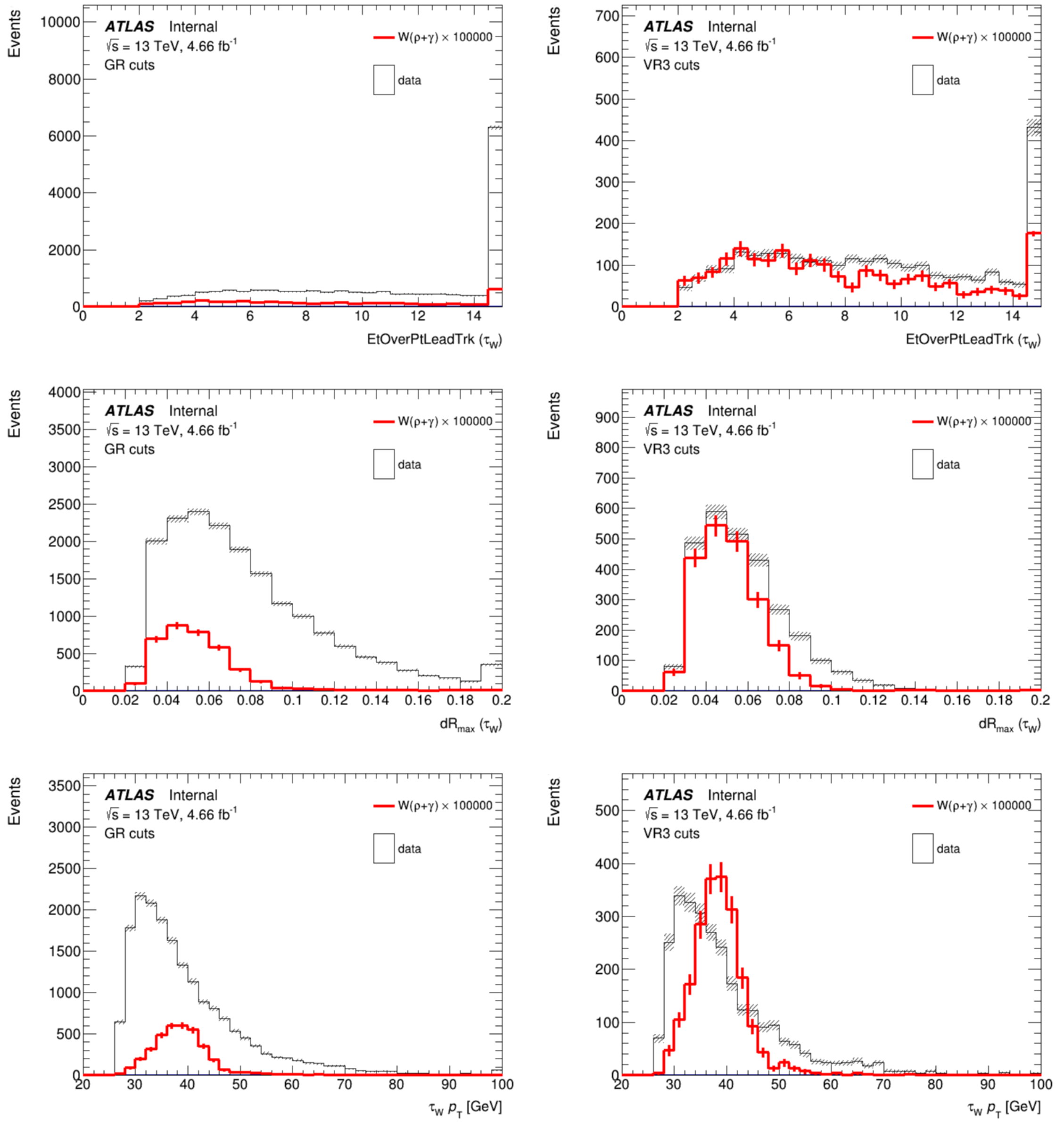


Figure 5.10: High level variables used as input by the RNN. The left column displays the GR, the right column the GR when the `Tight` RNN is also applied.

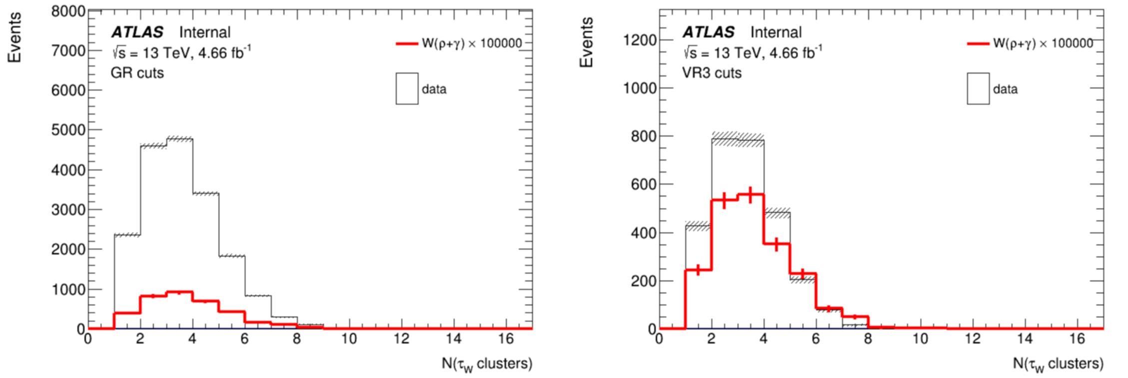


Figure 5.11: Cluster multiplicity used as input by the RNN. The left column displays the GR, the right column the GR when the `Tight` RNN is also applied.

background. Figure 5.12 displays, for illustrative purposes, the reconstructed $\rho\gamma$ invariant mass for events with `Tight` photon (both in identification and isolation) and a `Medium` RNN τ. The peak at 90 GeV observed in data and predicted by $Z \to e^+e^-$ MC sample highlight the sizeable contribution of the $Z \to e^+e^-$ background.

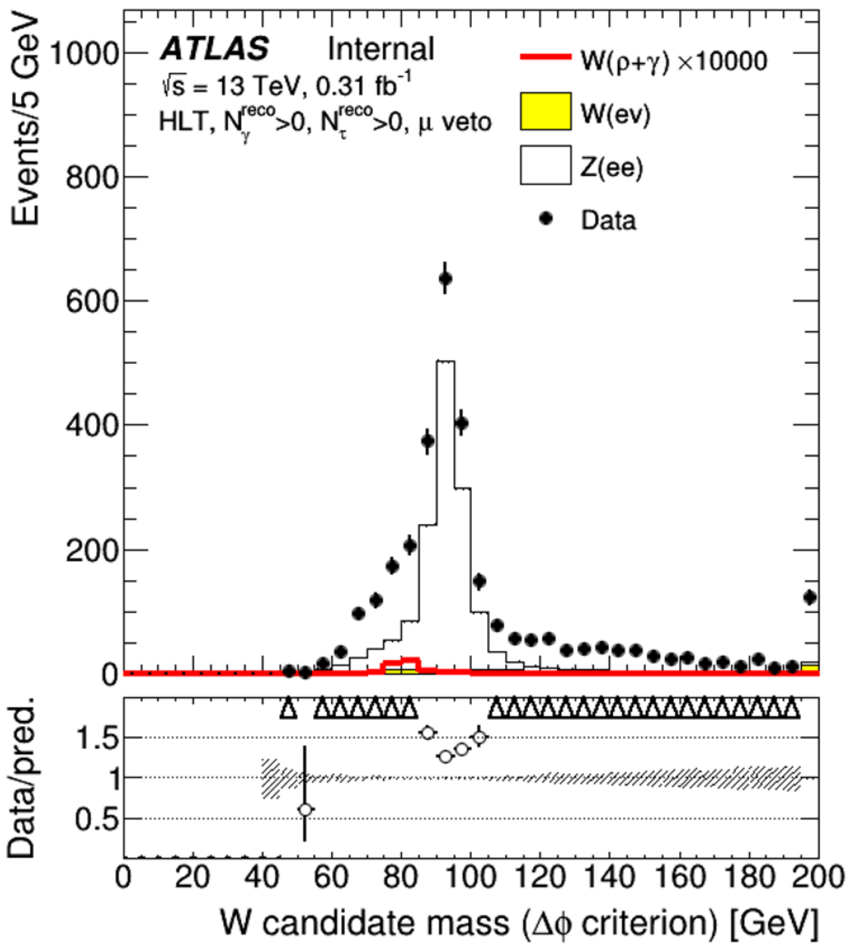


Figure 5.12: Reconstructed $\rho\gamma$ invariant mass prior to $Z \to e^+e^-$ background suppression.

In order to suppress the $Z \to e^+e^-$ background the following discriminant variables have been identified:

- `eProbabilityHT`: quality likelihood of TRT hits, including straw position, track-to-wire distance, and high-threshold status. The variable refers to the track retrieved from the $\tau_{\text{had-vis}}$ object.
- `etOverPtLeadTrk`: the transverse energy sum deposited in the core region of the TopoCluster of the $\tau_{\text{had-vis}}$ candidate, which is calibrated at the electromagnetic energy scale and divided by the transverse momentum of the highest p_T charged particle in the core region. This variable is also an input of the RNN discriminator.
- ΔR_{max}. This variable is also an input of the RNN discriminator.
- `eBDT` electron-veto discriminator. The optimised cut value is kept fixed at 0.25, corresponding to the `Tight` `eBDT` working point.

The three-dimensional optimisation has been performed again with the `blackbox`

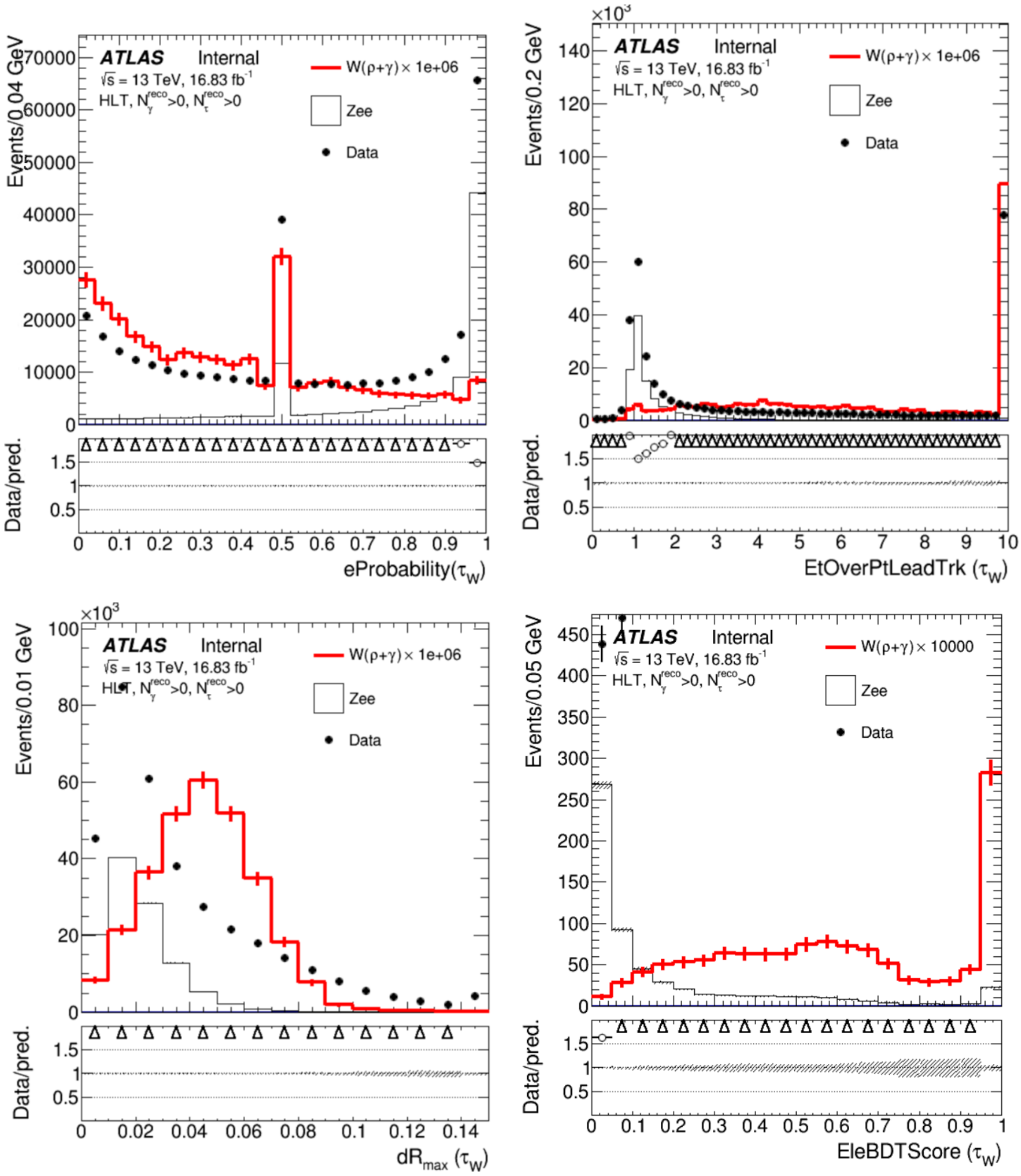


Figure 5.13: Discriminant variables against the $Z \to e^+e^-$ background.

multidimensional optimisation using the figure of merit $S/\sqrt{B}$ in the GR. The maximum number of iterations is set to 200, as visible in Figure 5.14. The distributions of these variables in the GR, defined as in Table 5.6, illustrated in Figure 5.13.

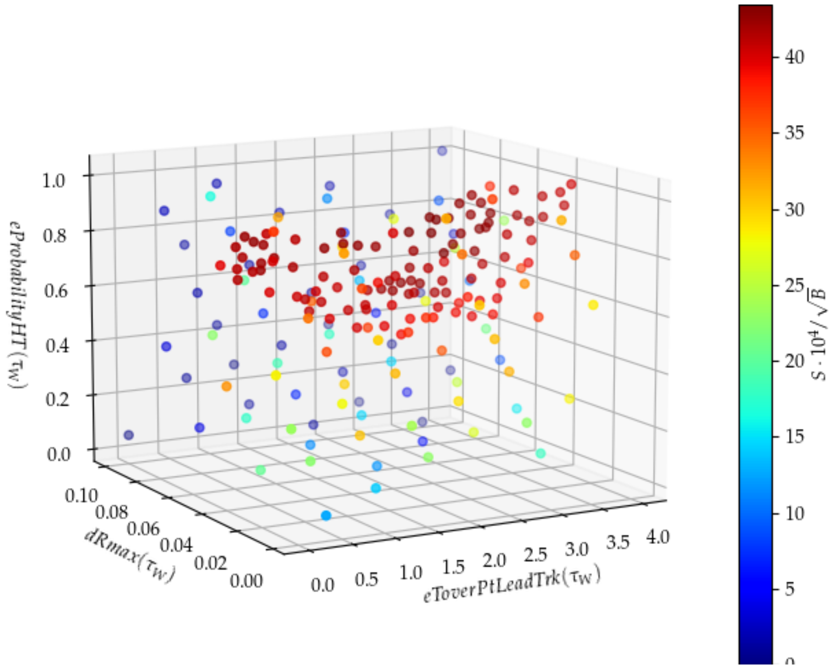


Figure 5.14: Optimisation of the $Z \to e^+e^-$ veto requirements with a total of 200 iterations. The optimisation were performed on a subset of the Run 2 dataset.

5.5.4 Model setup

Multijet and photon+jet background is estimated with the data-driven technique where pseudo-data events are generated in the GR using probability density functions determined using observed data. The generated events are subject to the selection requirements defining the VRs and are expected to model the background shapes in the SR. The pseudo-data events described only the variables necessary to model the W invariant mass in the SR and are generated according to the following procedure:

1. $p_T(\tau)$ is samples from the data distribution,
2. $p_T(\gamma)$ is drawn from the 2D distribution of $p_T(\gamma)$,$p_T(\tau)$, having $p_T(\tau)$ fixed at the value obtained in the previous step,
3. ΔR_{max} is drawn from the 3D distribution describing ΔR_{max}, $p_T(\tau)$ and $p_T(\gamma)$,
4. $\log(|d_0(\tau_W)|)$ is generated depending on $p_T(\tau)$ and ΔR_{max} variables previously sampled,
5. $\eta(\tau)$ is generated depending on the previously chosen $\log(|d_0(\tau_W)|)$ and ΔR_{max},
6. $\Delta\eta(\tau, \gamma)$ is generated depending on $\eta(\tau)$,
7. $\Delta\phi(\tau, \gamma)$ is generated depending on both $p_T(\tau)$ and $p_T(\gamma)$,
8. $\phi(\tau)$ is generated from the data distribution, independently from the other variables.

The sequence, shown in Figure 5.15, has been determined to capture the relevant correlations in the variables' multi-dimensional space. A useful check is the comparison of linear correlation coefficients in data and in the generated event sample. Figure 5.16 shows that the linear correlations are correctly reproduced. An additional check were performed on the $W \to q\bar{q}$ MC sample as well, to make sure this background does not form a peak, meaning that its contribution is well modelled. None of background events passed the selection, meaning that this contribution is not present in the hadronic τ and photon final state. Therefore, the GR is dominated by multijet and photon+jet background and the generated events are expected to model it in the VRs and SRs. The resulting background model is presented in the next section.

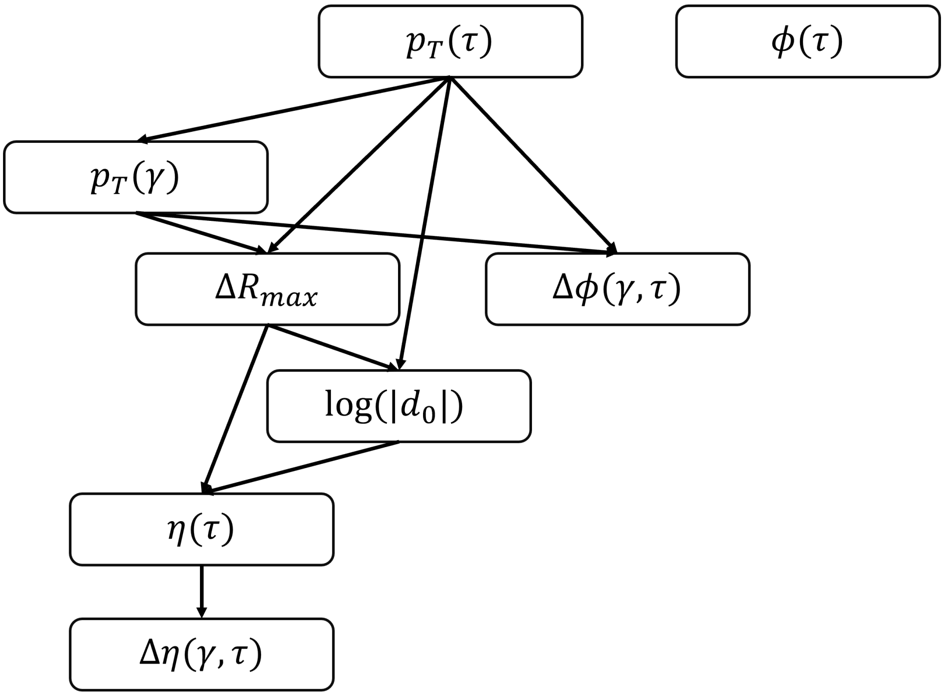


Figure 5.15: Schematic view of the background modelling algorithm. The top row of the variables have been generated independently, while everything below depends on the variable above as it is indicated with the arrows.

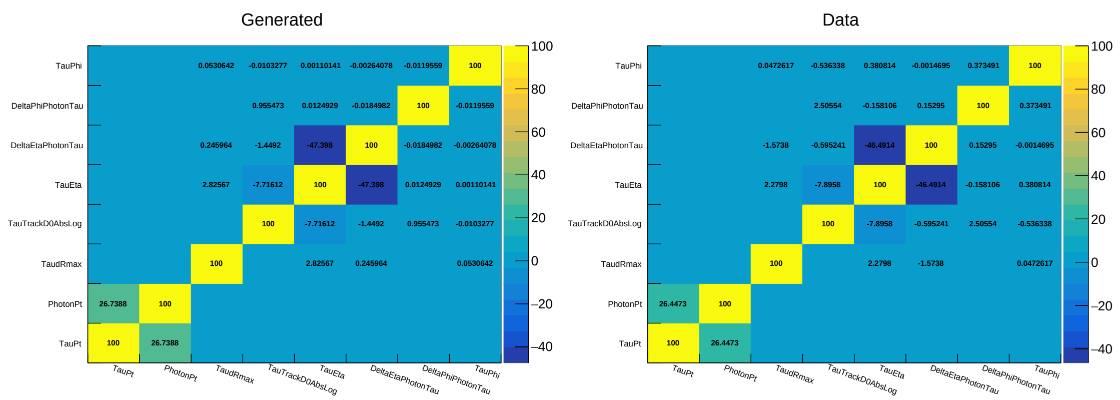


Figure 5.16: Linear correlation coefficients measured in the GR using the generated event sample on the left and observed data on the right.

5.5.5 Validation of the background estimation

In this section some relevant kinematic variables are presented which, are modelled by the background model. Figure 5.17 is dedicated to the GR, Figure 5.18, Figure 5.19 and Figure 5.20 to the three validation regions, and Figure 5.21 to the SR. The event count for all regions is reported in Table 5.8. The data-driven background model is normalised in the GR on the sidebands of the W invariant mass distribution. However, in the final fit the normalisation will be a floating parameter. Considering that no systematic uncertainties are included, all variables appear well modelled.

Inclusive selection				
Region	$W \to \rho\gamma \times 10^4$	$Z \to e^+e^-$	Background	Data
GR	10109 ± 377	1136 ± 20	$171735 \pm$	171850
VR1	9688 ± 368	1059 ± 19	146606 ± 146	146681
VR2	8089 ± 339	676 ± 15	97728 ± 205	97451
VR3	9022 ± 356	429 ± 12	116416 ± 194	116246
SR	7212 ± 320	193 ± 8	67513 ± 202	67210
$m_W \in$[70,90] GeV				
Region	$W \to \rho\gamma \times 10^4$	$Z \to e^+e^-$	Background	$(S \times 10^4)/\sqrt{B}$
GR	9999	1066	111761	30
VR1	9577	999	91034	32
VR2	8013	633	58746	33
VR3	8922	386	75401	32
SR	7136	172	39263	36

Table 5.8: Signal, background and data event counts. Only statistical uncertainties are included.For the Background binomial uncertainty is estimated moving from GR to any other region (therefore it is not defined for GR). The signal is scaled by a factor 10^4. Event counts are also shown for a selection the reconstructed W boson invariant mass.

5.6 Systematic uncertainties

The final results are affected by systematic uncertainties, which are arising form the experimental uncertainties on data, and from the $m_{\rho\gamma}$ distribution modelling.

5.6.1 Simulated event samples

Systematic uncertainties in the signal predictions arise from the measurement of the integrated luminosity, trigger efficiency and from the photon and tau detection and reconstruction efficiency. The uncertainties considered are discussed below and summarised in Table 5.9.

Factorisation and renormalisation scale variation

Since the nominal MC sample did not include LHE weights, for each renormalisation and factorisation variation one additional MC sample has been generated with the

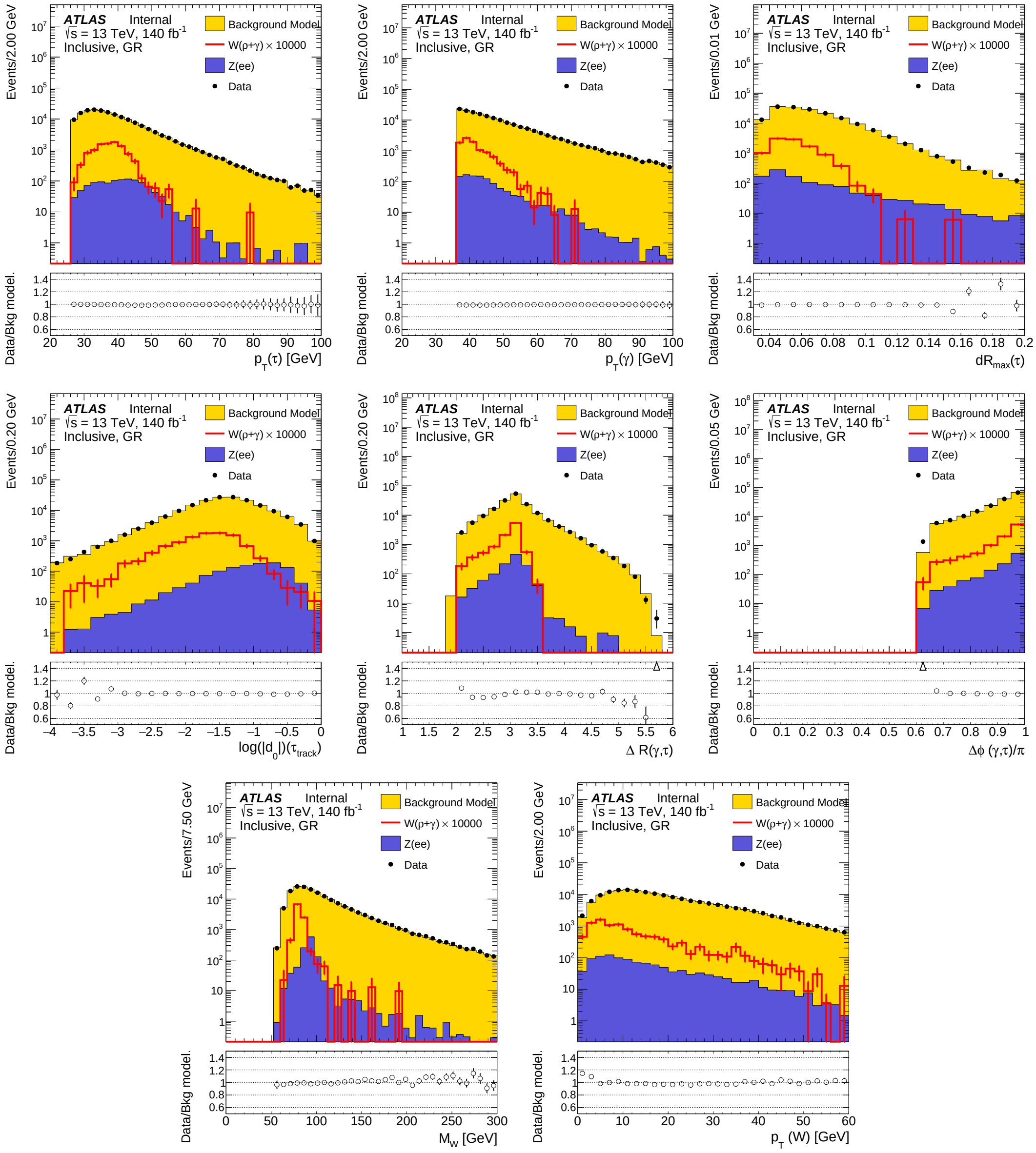


Figure 5.17: Kinematic distributions in the GR region. The statistical uncertainties are negligible for both the data and the data-driven background model, which consists of 60 million pseudo-data events.

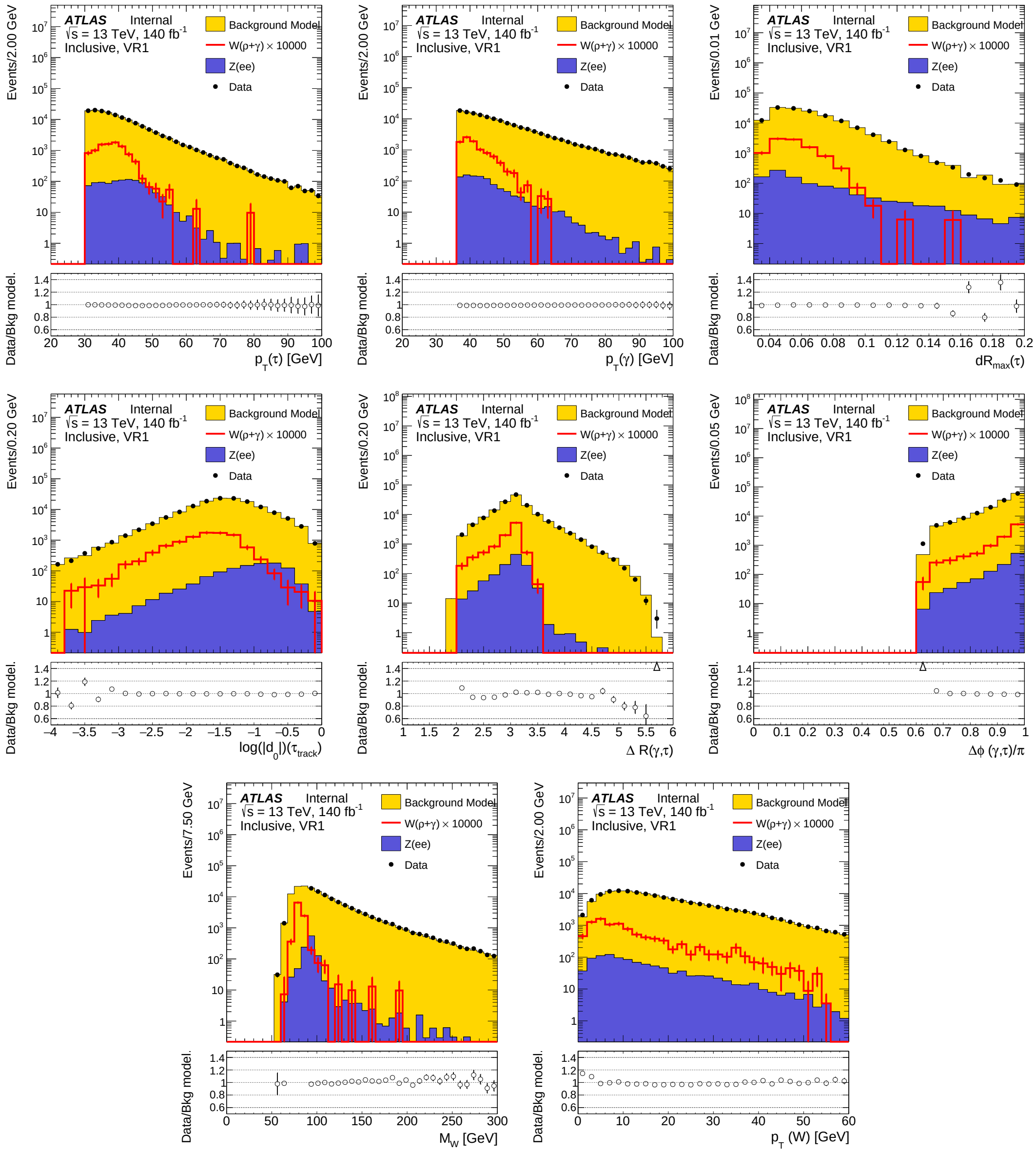


Figure 5.18: Kinematic distributions in the VR1 region. The statistical uncertainties are negligible for both the data and the data-driven background model, which consists of 60 million pseudo-data events.

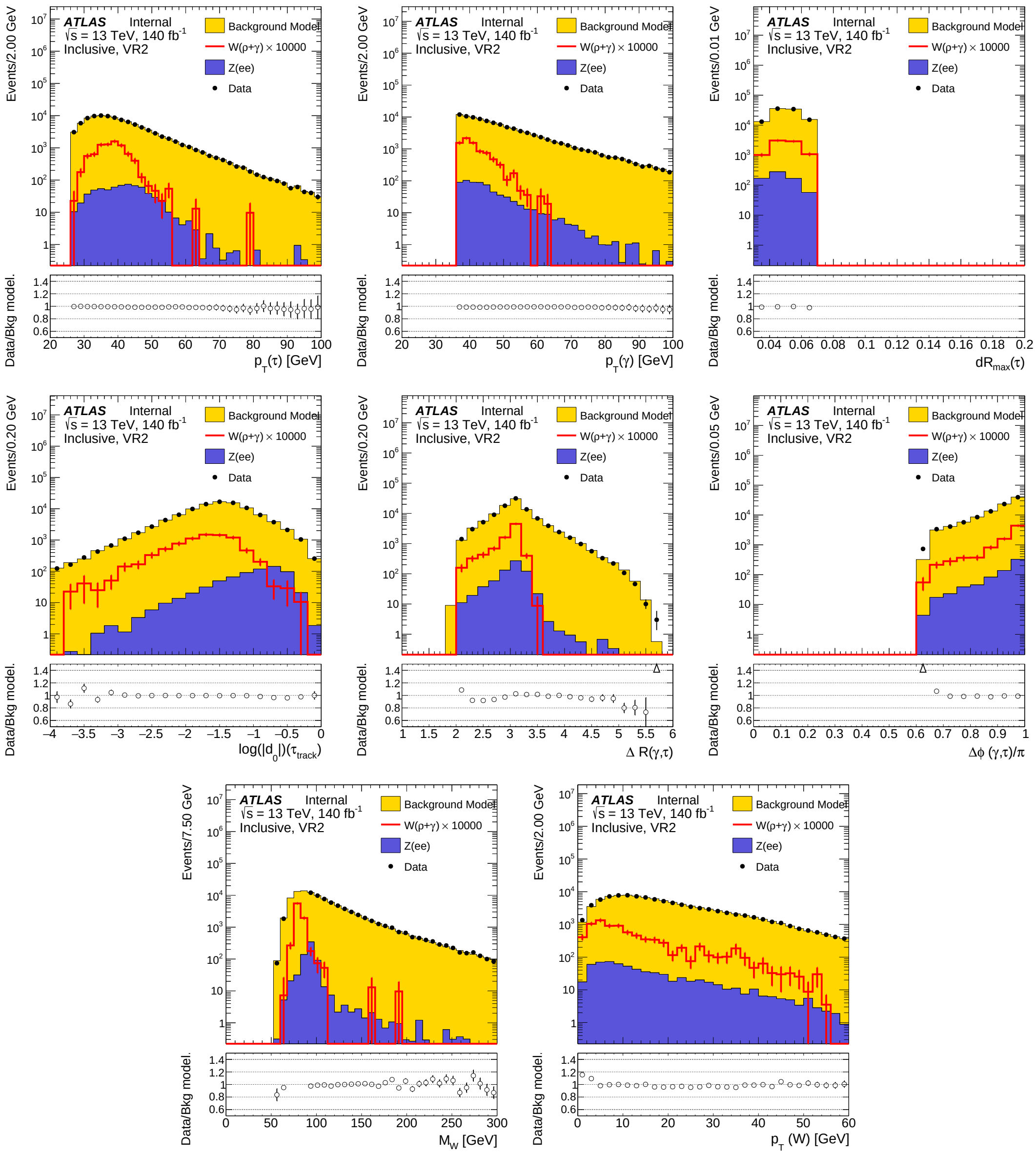


Figure 5.19: Kinematic distributions in the VR2 region.

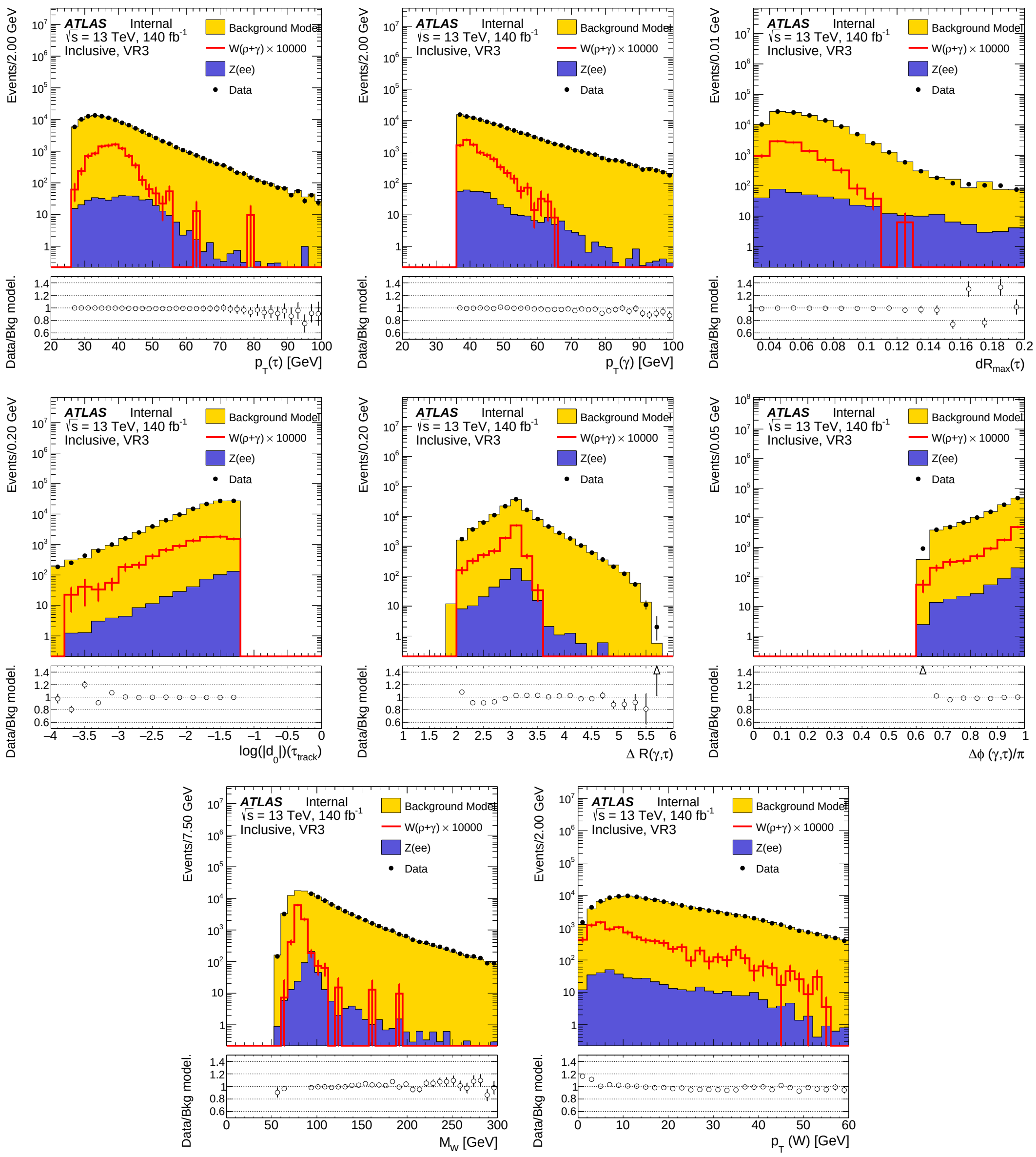


Figure 5.20: Kinematic distributions in the VR3 region.

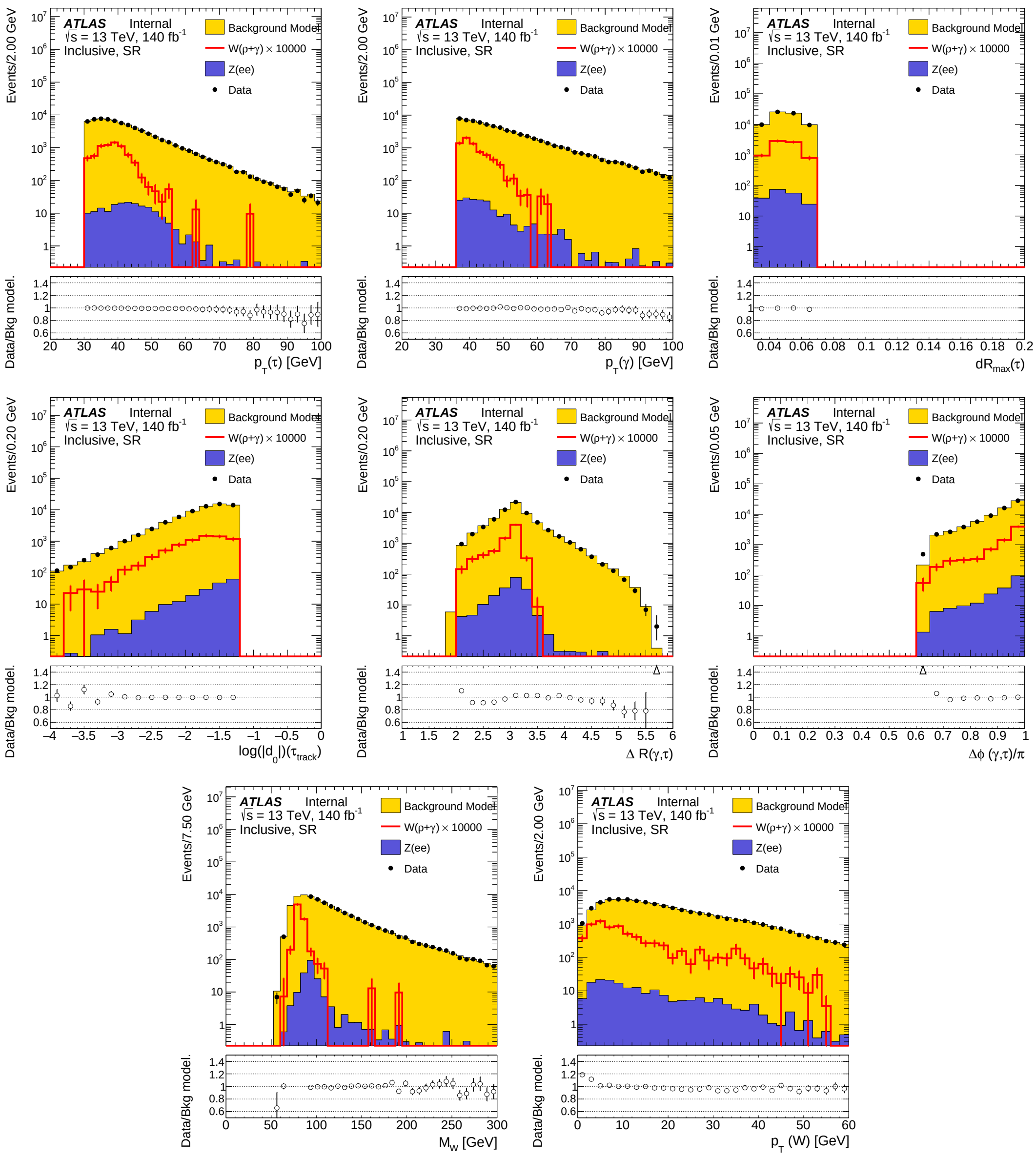


Figure 5.21: Kinematic distributions in the SR region.

same number of events as in the nominal sample. All samples are normalised to the nominal cross-section to account only for acceptance variations. The difference between each variation and the nominal sample has been estimated at the generation level applying all the kinematic requirements of the SR. The resulting uncertainty, shown in 5.22, corresponds to the envelope of all variations. Given that the variation is flat in the populated range ($m_W \in$ [77, 83] GeV) only the normalisation component (6.5%) is included in the fit.

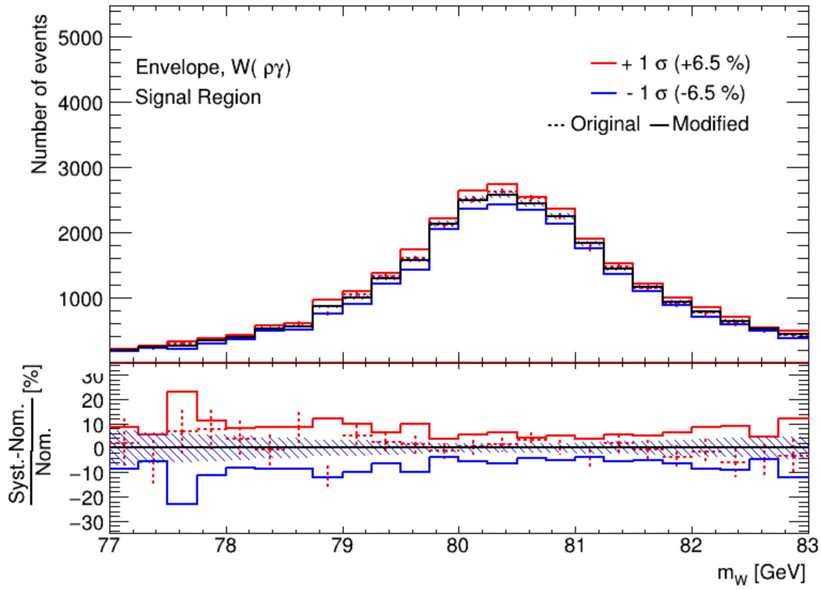


Figure 5.22: Envelope of factorisation and scale variations evaluated in SR at generation level.

Resolution and Energy Scale

Energy mis-calibration effects arise from different calorimeter response between simulated and real data. The calibration procedures are affected by systematic uncertainty. A simplified uncertainty model, which combines all the uncertainties for scale and resolution in two systematic variables, is used for this analysis. The uncertainties for each variation are summed in quadrature resulting a total uncertainty due the electron and photon energy scale is 3.0%, due the energy resolution is 4.9% and due the additional photon energy scale is 1.7% [95].

Photon Identification and Isolation

The normalisation uncertainty due the photon reconstruction and identification uncertainty is 1.1%, while the photon isolation uncertainty is estimated to be 1.6% [92].

Tau Efficiency

The systematic uncertainty associated with the $\tau_{\text{had-vis}}$ reconstruction is 1.2%, systematic uncertainties related to the RNN algorithm are 0.3%, 0.3% and 0.8%, and the uncertainties related to the Electron Veto algorithm are 1.2%, 5.7% and 1.3%

[103]. Uncertainties regarding to the tau energy calibration are also taken into account. The full list of tau uncertainties, with their definition can be found in Table 5.9.

Trigger Efficiency

Trigger scale factors provided by the CP group are not applicable in this analysis as a single photon is selected. To estimate an uncertainty a dedicated study has been performed, which were already described in Section 5.4. The study finds that differences between data and MC are contained in a 10% band along the invariant mass distribution m$(\mu, \tau_{\text{had-vis}})$. Therefore a 10% normalization uncertainty has been applied to $W \to \rho\gamma$ to account for efficiency variations.

Luminosity

The uncertainty on the combined 2015-2018 integrated luminosity is 0.83%, and is applied to signal and the $Z \to ee$ background estimated from MC predictions [34].

Pileup

The uncertainty on the pile-up weight is 5.5%, and is applied to signal [117].

Uncertainty source	Normalisation variation (%)
Cross Section	3.4
Luminosity	0.83
Pileup	5.5
Photon Identification	1.1
Photon Isolation	1.6
Electron and photon scale	3.0
Electron and photon resolution	4.9
Photon energy scale	1.7
Tau reconstruction	1.2
1 prong tau RNN Identification (30< p_T <40 GeV)	0.3
1 prong tau RNN Identification (p_T >40 GeV)	0.3
RNN Identification systematic uncertainty	0.8
Electron veto	1.2
eBT Identification statistical uncertainty	5.7
eBDT Identification systematic uncertainty	1.3
Tau MC simulation uncertainty	6.6
Tau energy scale uncertainty due to detector description	1.9
Tau energy scale uncertainty due to detector response	5.6
Tau energy scale uncertainty due fitting	6.5
Tau calibration model uncertainty	4.3
Trigger	10
Scale variations	6.5

Table 5.9: Systematic uncertainties of $W \to \rho\gamma$ in SR.

5.6.2 Uncertainties on the data driven background

The systematic uncertainties arising from the background modelling method are estimated by generating alternative background shapes. These alternative shapes provide flexibility to the inclusive background model, so it can freely adjust to the observed background.

During the generation of the first alternative model the same method was used to generate the nominal model, but the $p_T(\tau)$ distributions are artificially shifted by ± 3 GeV. In the case of the $\Delta\phi(\rho, \gamma)$ distortion the variation is implemented by scaling each bin by $(1 + \Delta\phi/\pi)^{10}$ and $1 + 10(1 - \Delta\phi/\pi)$ as up and down variation. These systematics lead to the lateral movement of the $m_{\rho\gamma}$ peak. To allow an overall tilt of the $m_{\rho\gamma}$ distribution, the model is re-weighted with a linear function. The parameters for this tilt are defined as:

$$\begin{aligned} \text{Up :} y &= -0.0026 \times mass + 1.34 \\ \text{Down :} y &= 0.0026 \times mass + 0.635. \end{aligned} \tag{5.4}$$

5.7 Statistical methods

In the $W \to \rho\gamma$ analysis, the hypotheses that are tested are the background-only hypothesis, when no signal events are present, against the signal-plus-background hypotheses, which is correlated with the presence of the $W \to \rho\gamma$ production.

To test these hypotheses, a binned maximum likelihood fit is performed on the $m_{\rho\gamma}$ invariant mass. The likelihood function is constructed as the product of the Poisson probability terms of the signal and the background:

$$\mathscr{L} = \prod_i^N \text{Pois}(n_i|\mu s_i + b_i) \prod_k G(\theta). \tag{5.5}$$

The first term of the likelihood indicated the product of the Poisson probability terms of all events, n_i is the number of data events, s_i is the expected amount of signal and b_i is the expected amount of background events for the given bin i. The second term introduces the nuisance parameter (NP) indicated with θ. In the context of the analysis, NPs are the systematic uncertainties introduced in Section 5.6, and most of them described with a Gaussian probability density functions.

The parameter of interest is the signal strength, μ, and it is defined as the ratio of the expected signal events and the expected events predicted by the SM:

$$\mu = \frac{N_{\text{signal}}}{N_{\text{SM}}}. \tag{5.6}$$

If $\mu = 1$, the test agrees with the signal-plus-background hypothesis, while the background-only hypothesis is represented by $\mu = 0$. Since $N_{signal} = \epsilon L_{\text{int}} \cdot \sigma_{pp \to W} \cdot \mathscr{B}(W \to \rho\gamma)$, where ϵ represents the selection efficiency, the signal strength can be also written as:

$$\mu = \frac{\mathscr{B}(W \to \rho\gamma)_{\text{measured}}}{\mathscr{B}(W \to \rho\gamma)_{\text{SM predicted}}}. \tag{5.7}$$

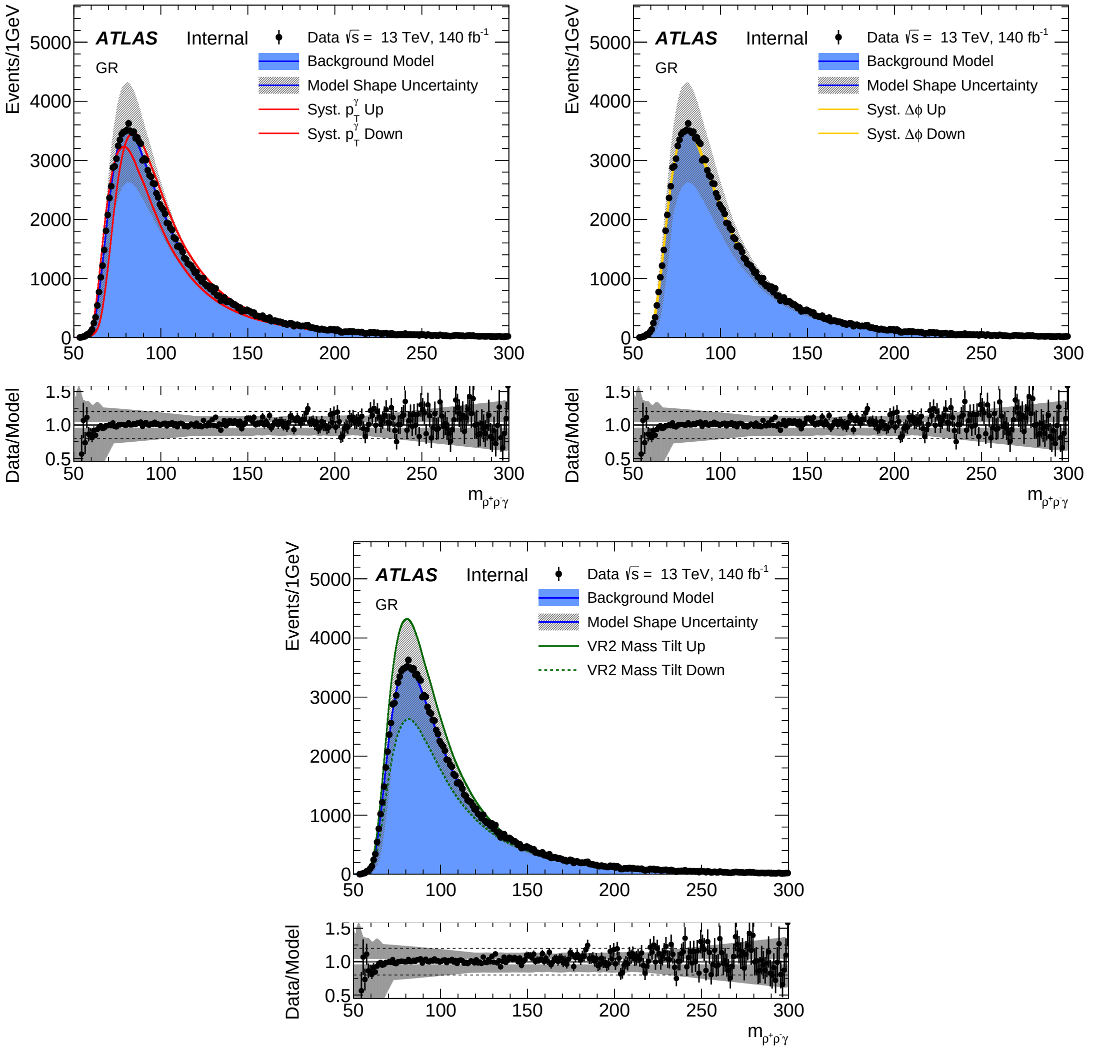


Figure 5.23: Effect of the systematic variations on $m_{\rho\gamma}$. The grey band indicates the maximum deviation from the nominal model.

To calculate the best estimated values of μ and θ, one need to maximise the likelihood function, or equivalently minimise the negative log-likelihood. With the estimated best-fit values of the θ, indicated with $\hat{\theta}$, and μ, indicated with $\hat{\mu}$, we can define the test statistics as:

$$q_\mu = \begin{cases} -2\ln\frac{\mathscr{L}(\mu,\hat{\hat{\theta}})}{\mathscr{L}(\hat{\mu},\hat{\theta})} & \text{if } \hat{\mu} \geq 0 \\ 0 & \text{if } \hat{\mu} < 0 \end{cases}, \tag{5.8}$$

where $\hat{\hat{\theta}}$ represents the best-fit value of the θ parameters for a given μ. The ratio of $\mathscr{L}(\mu,\hat{\hat{\theta}})/\mathscr{L}(\hat{\mu},\hat{\theta})$ defines the profile likelihood ratio [118].

The signal-plus-background hypothesis can only be accepted, if the background-only hypothesis can be rejected with a certain probability. If $\hat{\mu} = 0$ or negative, the q_0 value also zero, meaning a true null hypothesis. Otherwise, one needs to evaluate the level of agreement with the hypothesis, which is done by calculating the p_0-value:

$$p_0 = \int_{q_0^{obs}}^{\infty} f(q_0|\mu = 0)dq_0, \tag{5.9}$$

where $f(q_0|\mu = 0)$ represents is the test statistics q_0 probability function under the null hypothesis. The p_0-value indicates the probability of observing a dataset as signal-like presuming that there is no signal. Therefore, small p_0-values mean the incorrectness of the null hypothesis. In particle physics it is common to define the results in term of the significance, that corresponds as a number of standard deviations, Z, for a given p_0-value. In this case the measured p_0 value equal to the one sided tail area of a Gaussian distribution with zero mean and 1 variance. The common practice is to claim a discovery at Z$\geq$5.

If it is not possible to reject the null hypotheses, with the profile likelihood approach one can reject the signal-plus-background hypothesis for different μ values. This will result a range of μ values excluded by the data with a certain confidence level (CL), and can be used to set a limit on the parameter of interest. The most common practice is to use CL = 95%, which is equal of Z=1.64.

To set a limit, the fist step is to calculate the μ value by solving the p_μ equation for $p_0 = 1 -$ CL. The p_μ is defined analogously to the p_0:

$$p_\mu = \int_{q_\mu^{obs}}^{\infty} f(q_\mu|\mu)dq_\mu, \tag{5.10}$$

where the test statistics q_μ defined as:

$$q_\mu = \begin{cases} -2\ln\frac{\mathscr{L}(\mu,\hat{\hat{\theta}})}{\mathscr{L}(\hat{\mu},\hat{\theta})} & \text{if } \hat{\mu} \leq \mu \\ 0 & \text{if } \hat{\mu} > \mu \end{cases}. \tag{5.11}$$

If the experiment has a very low sensitivity, the distributions of the test statistic q_μ for the background-only and background-plus-signal hypothesis are very close to each other. To be able to include models to which one is not sensitive enough, the modified frequentist formalism with the CL method [119] can be used, which is defined as:

$$CL_s = \frac{CL_{s+b}}{CL_b} = \frac{p_\mu}{1-p_0}. \tag{5.12}$$

In this case, if $CL_s <$ 0.05 the signal-plus-background hypothesis can be excluded at 95% confidence level.

5.8 Expected sensitivity

The signal extraction is performed using a binned profile likelihood ratio [120] fit of the invariant mass of the W boson candidate in the SR. The likelihood $L(\mu, \theta)$, function of the signal strength μ and the nuisance parameters θ, is built using continuous signal and background probability distribution functions. The likelihood ratio, used as test statistic, is defined in 5.13. At the denominator $\hat{\mu}$ and $\hat{\theta}$ are the unconditional maximum likelihood estimators of the signal strength and nuisance parameters. At the numerator $\hat{\hat{\theta}}(\mu)$ corresponds to the value of $\hat{\theta}$ that maximises the likelihood for a given value of μ.

$$\Lambda(\mu) = \frac{L\big(\mu, \hat{\hat{\theta}}(\mu)\big)}{L(\hat{\mu}, \hat{\theta})}. \tag{5.13}$$

Upper limits are computed using the asymptotic approach [120] and the CL_s modified frequentist method [119]. Shape systematic uncertainties are implemented using the interpolation technique described in Ref. [121].

During the fitting, the $W \to \rho\gamma$ signal shape is represented by a Voigt function multiplied by an efficiency function. A Voigt distribution, reported in Equation 5.14, is the convolution between a Breit-Wigner and a Gaussian:

$$P_{\text{Voigt}}(\mathsf{m}, \mathsf{m}_W, \Gamma, \sigma) = \int_{-\infty}^{+\infty} d\mathsf{m}' \left(\frac{1}{\pi} \frac{\Gamma/2}{(\mathsf{m}' - \mathsf{m}_W)^2 + \Gamma^2/4} \right) \left(\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\mathsf{m}-\mathsf{m}')^2}{2\sigma^2}} \right). \tag{5.14}$$

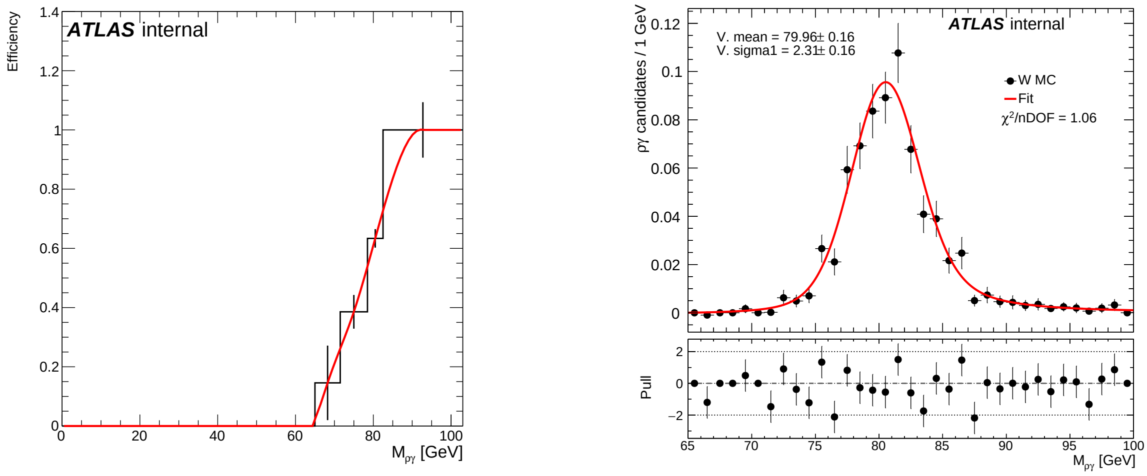
The width in the Breit-Wigner component is fixed and equal to the W boson width value (2.085 $\pm$ 0.042 GeV). The width of the Gaussian component, meant to model the experimental resolution is obtained from a fit on simulated events. The free parameters in the signal model fit are thus the centroid m and width σ of the Voigt distribution. The efficiency turn-on along $m_{\rho\gamma}$ is obtained fitting a cubic spline on the histogram obtained with simulated events at generation level. The efficiency function is displayed in Panel (a) of Figure 5.24. Panel (b) of Figure 5.24 shows the resulting function used to model the signal $m_{\rho\gamma}$ shape. The post-fit values of the Voigt function parameters are $m = 80.3 \pm 0.2$, $\sigma = 2.3 \pm 0.2$.

The multijet background template is obtained by smoothing the data-driven background model using a Gaussian kernel density estimator (KDE) [122]. The smoothing is performed in the adaptive mode for the core of the distribution (where there are more statistics), while a high smoothing parameter is imposed for the tail (where there are fewer events, and the adaptive mode was found to not provide a smooth enough distribution). This is done by obtaining two smoothing predictions - KDE_{core} and KDE_{tail}. KDE_{core} is obtained using the KDE adaptive mode, while a high smoothing parameter of 2 is employed for the tail of the distribution.

The $Z \to e^+e^-$ component estimated using MC prediction is also smoothed with the KDE algorithm. The same algorithm applied as in case of the multijet background, but with a high smoothing parameter of 1.5. The $Z \to e^+e^-$ prediction is given by:

$$(1 - \text{Erf}(m_{\text{track},\gamma} - 110)) \times \text{KDE}_{\text{core}} + ((\text{Erf}(m_{\text{track},\gamma} - 110) + 1) \times \text{KDE}_{\text{tail}}. \tag{5.15}$$

The resulting shape is shown in Figure 5.25.



(a) **(b)**

Figure 5.24: (a) Efficiency function determination for the $W \to \rho\gamma$ process. (b) Function to model the $W \to \rho\gamma$ $m_{\rho\gamma}$ shape.

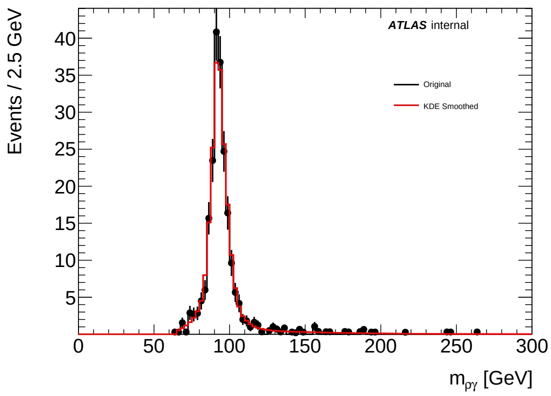


Figure 5.25: $Z \to e^+e^-$ track+photon invariant mass SR distribution, and KDE smoothing result, used to model $Z \to e^+e^-$ component.

5.8.1 Expected sensitivity with expected background Asimov dataset

The systematic uncertainties described in 5.6 are introduced in the fit as nuisance parameters. The unconstrained parameters are the signal strength, data-driven background normalisation and tilt uncertainty. The number of events in SR is reported in Table 5.10 and the $m_{\rho\gamma}$ distribution is displayed in Figure 5.26. A maximum likelihood signal-plus-background fit on the Asimov dataset is run as a consistency check. The post-fit parameter correlations are shown in Figure 5.27. The expected upper limit is computed using an Asimov dataset [2] constructed under the background-only hypothesis. The resulting expected upper limits are presented in Table 5.11 with and without systematic uncertainties.

Process	Number of events
$W \to \rho\gamma$	0.72
Multijet background	43133
$Z \to ee$	179

Table 5.10: Pre-fit number of events corresponding to μ=1.

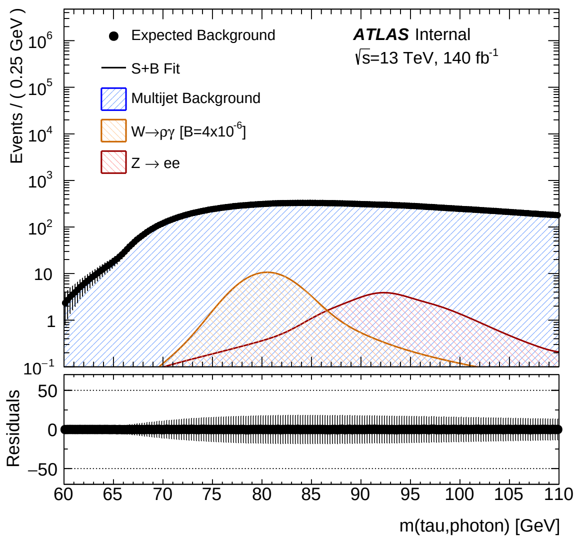


Figure 5.26: Binned maximum likelihood fit to the Asimov dataset constructed under the background-only hypothesis, in the $\rho\gamma$ final state.

5.8.2 Expected sensitivity with sideband fit Asimov dataset

A background-only fit to data (i.e. with $\mu(W \to \rho\gamma)$ fixed at 0) was performed on SR sidebands defined by excluding the candidate W boson mass range $[76.5, 84.5]$ GeV. Data and background distribution are shown in Figure 5.28. The obtained post-fit parameter values are listed in Table 5.12. An Asimov dataset was then created using

[2]Asimov datasets are built as binned datasets, in which the event count in each bin is set to the expected event yield for the chosen model parameters

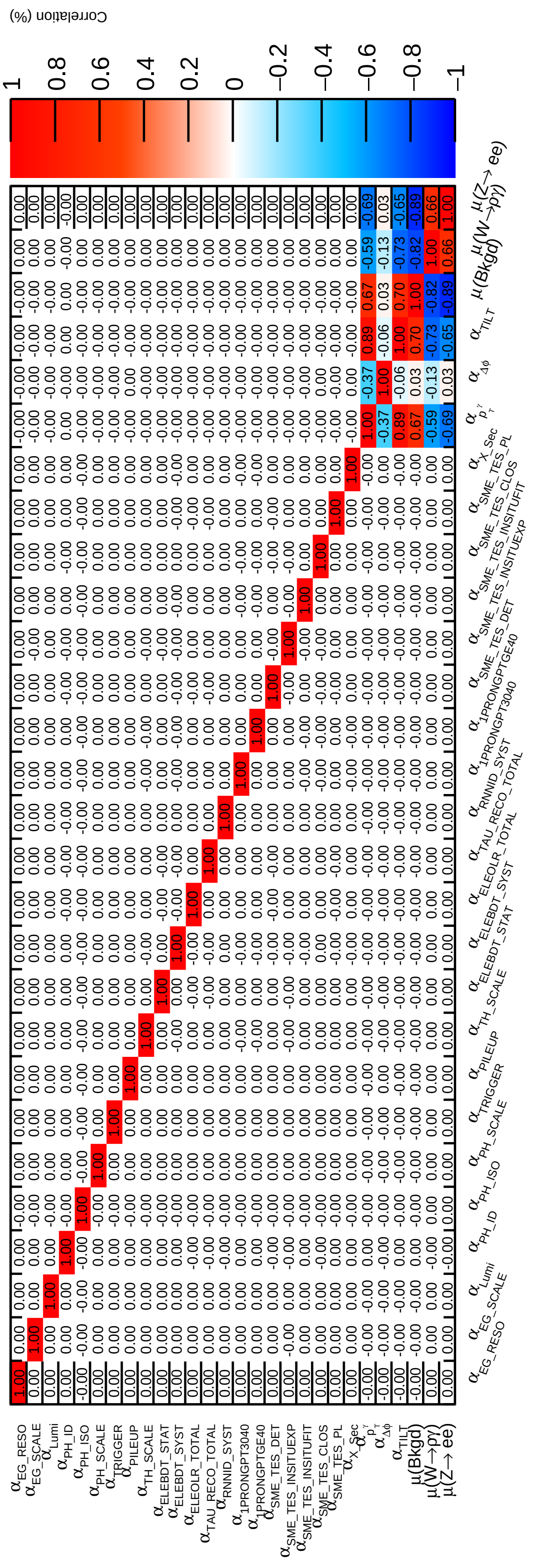


Figure 5.27: Correlation matrix between parameters in fit to Asimov dataset constructed from the expected background in the tau+photon final state.

	Expected ($\times 10^{-6}$)	$\pm 1\sigma$	$\pm 2\sigma$
No systematics	4.02	5.59/2.90	7.50/2.16
Shape	5.70	7.94/4.11	10.64/3.06
Norm	4.33	6.02/3.12	8.07/2.32
Shape + Norm	6.14	8.54/4.42	11.45/3.29

Table 5.11: Expected $W \to \rho\gamma$ branching fraction limit at 95% CL. The limits are estimated with and without normalisation (Table 5.9) and shape (Section 5.6.2) systematic uncertainties.

these values to what a signal+background fit was performed. New expected limits at 95% CL were derived and can be found in Table 5.13. The difference between the obtained upper limit and the one listed in Table 5.11 is negligible.

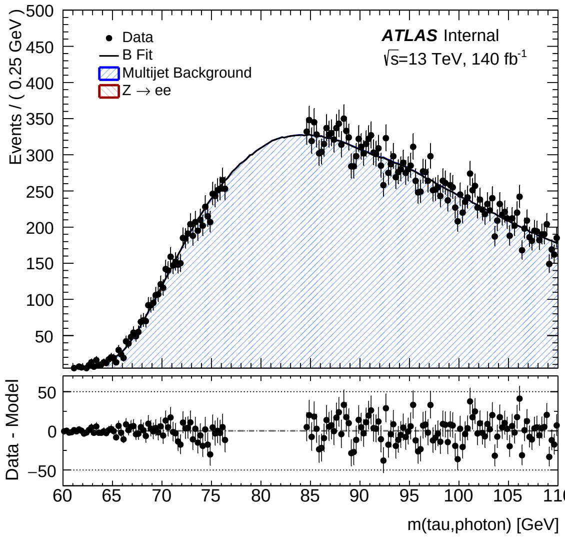


Figure 5.28: Background only fit to the tau+photon invariant mass distribution sideband data. The $Z \to e^+e^-$ and multijet are drawn separately in red and blue, respectively. The black full line shows the post-fit background model (multijet $+Z \to e^+e^-$). The signal components are fixed at a value of 0, and as such are not represented in the plot.

Parameter	Final Value	Uncertainty
$\alpha_{\text{background shape } p_T^\gamma}$	0.089	-0.156/+0.106
$\alpha_{\text{background shape } \Delta\phi}$	-0.142	-0.242/+0.291
$\alpha_{\text{background shape Tilt}}$	0.268	-0.490/+0.583
μ (Background)	0.998	-0.013/+0.013
μ ($Z \to ee$)	-0.197	-1.956/+1.963

Table 5.12: Values of the background normalisation parameters and shape systematics nuisance parameters after sideband fit.

	Expected ($\times 10^{-6}$)	$\pm 1\sigma$	$\pm 2\sigma$
Shape + Norm	6.38	8.88/4.60	11.90/3.42

Table 5.13: Expected upper limits at 95% CL using Asimov data built with the parameter values obtained in the sideband fit. Both normalisation (Table 5.9) and shape (Section 5.6.2) systematic uncertainties are included.

5.9 Results

The results of the fit to the full dataset in the tau+photon final state are shown in Figure 5.29, Table 5.14 lists the post-fit number of events for each contribution. The observed upper limit is equal

$$\mathscr{B}(W \to \rho\gamma) < 6.29 \times 10^{-6}. \tag{5.16}$$

Process	Number of events
$W \to \rho\gamma$	-9 $\pm$ 252
Multijet background	43007 $\pm$ 618
$Z \to ee$	-80 $\pm$ 385
Data	42918

Table 5.14: Post-fit number of events.

5.10 Combined fit

An other approach for measuring the $W \to \rho\gamma$ branching fraction is to reconstruct the ρ meson using a track. In this case only the $\pi^{\pm}$ meson originating from the ρ meson decay is reconstructed as a track, while the π^0 meson is not explicitly identified. This method was originally developed for the search of $W \to \pi\gamma$ decay, but it also has sensitivity in the $W \to \rho\gamma$ domain. Therefore, the search for the $W \to \rho\gamma$ is also performed in this case, and it is referred as the track-plus-photon final state.

Similarly to the tau-photon final state, all events must pass the trigger selection. Track-photon triggers are derived from τ-lepton triggers [123] and modified to select $W \to \pi\gamma$ events. These triggers were activated in 2016 and collected a dataset of 137 fb^{-1}. Tracks must satisfy $p_T >$ 30 GeV, and $|\eta| <$ 2.5 selection requirements and the "Tight Primary" criteria [87]. As it is already mentioned, tracks are considered as the meson candidates. Photons are required to have $p_T >$ 30 GeV or $p_T >$ 35 GeV, depending on the trigger. The highest p_T photon and meson constitute the candidate W boson if the difference in azimuthal angle $\Delta\phi(M, \gamma)$ is larger than $\pi/2$. If both the meson and the photon are reconstructed in the endcap regions ($|\eta(M)| > 1.5$ and $|\eta(\gamma)| > 1.37$) the photon and meson candidates are required to have $\eta(M) \times \eta(\gamma) \geq 0$, to suppress multijet background.

The $Z \to e^+e^-$ background is suppressed based on two criteria:

- the probability that the track is an electron based on TRT information exceeds 10%,

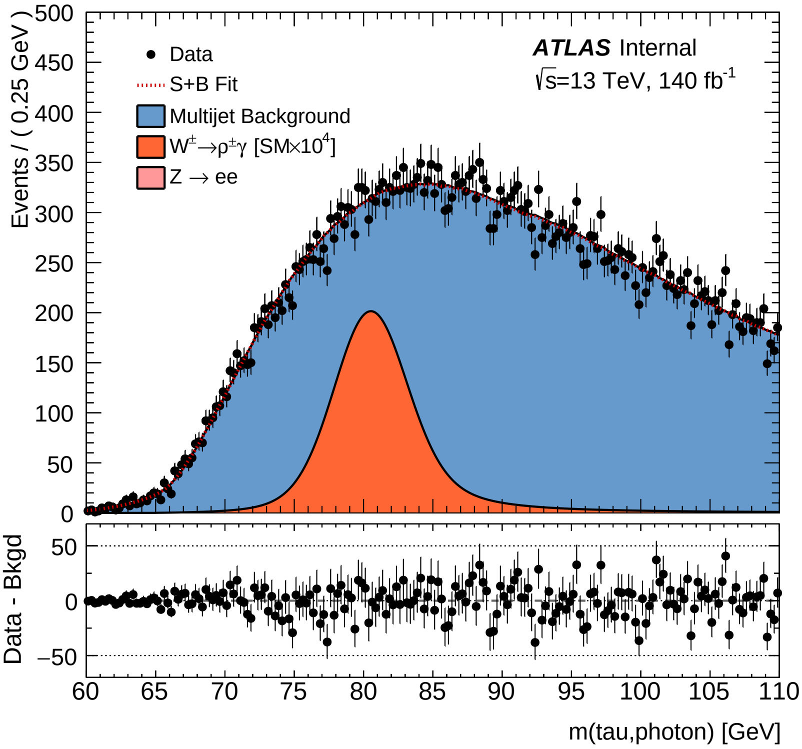


Figure 5.29: Fit to the full dataset after unblinding in the tau-plus-photon final state. The backgrounds are shown using the post-fit value. The post fit $Z \to e^+e^-$ is not visible since it has a negative post-fit value. The dotted red line shows the post-fit signal+background model. The $W \to \rho\gamma$ is set to the SM expected multiplied by a factor of 10^4 for visibility. The bottom panel shows the differences between the data and the fitted background.

- the hadronic leakage of the energy deposit matched to the meson track is at least 3%. The hadronic leakage is defined as the ratio between transverse energy deposited in the hadronic and electromagnetic calorimeters.

The same background modelling techniques have been used as in the tau-photon case. The requirements listed above and applied to the selection define the GR. The track-plus-photon SR is defined by adding three requirements to the GR selection criteria: $p_T(M) > 33$ GeV; photon `FixedCutTight` isolation; and the sum of transverse momenta of tracks within a cone of $\Delta R = 0.2$, excluding the meson candidate track, is required to be less than 14% of $p_T(M)$ (meson isolation). Both isolation requirements reduce the contribution of photon and meson candidates faked by jets. The track-plus-photon SR efficiency (including the trigger selection) is 0.5% for $W \to \rho\gamma$ decay. The results of the track-plus-photon fit are shown in Figure 5.30.

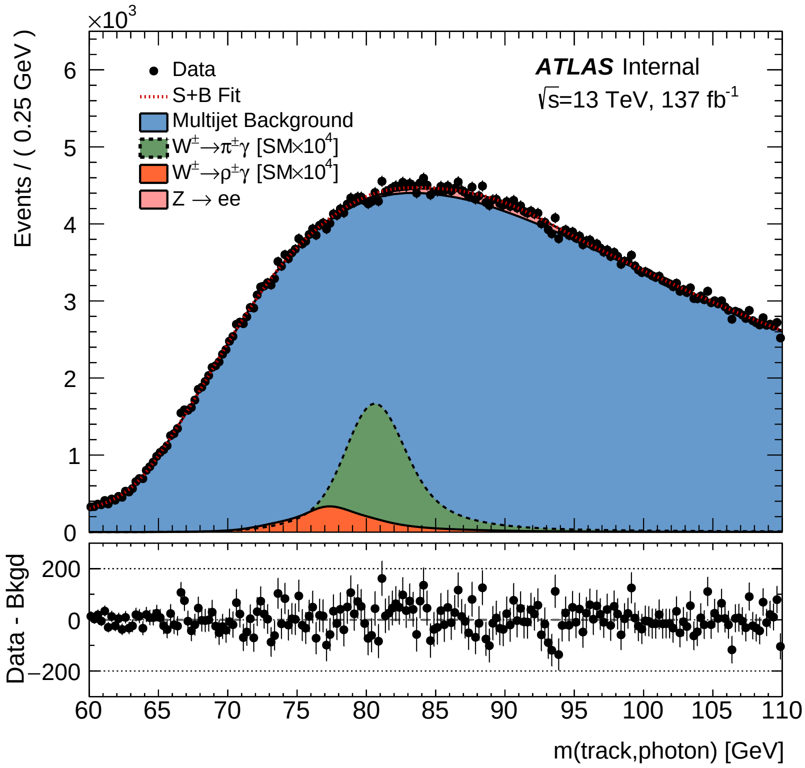


Figure 5.30: Fit to full dataset in the track-plus-photon final state. The $Z \to e^+e^-$ and multijet are normalised to their post-fit value and drawn separately in red and blue, respectively. The $Z \to e^+e^-$ contribution can be seen on top of the multijet background. The dotted red line shows the post-fit signal+background model. The $W \to \pi\gamma$ and $W \to \rho\gamma$ are set to the SM expected multiplied by a factor of 10^4 for visibility. The bottom panels show the differences between the data and the fitted background.

The track-plus-photon and tau-plus-photon signal regions are found to be orthogonal. This allows to perform a simultaneous fit of the track-plus-photon and tau-plus-photon SRs. Background uncertainties are considered uncorrelated as the trigger and event selection strategies probe different phase spaces, except for those associated with the W boson production cross section and the integrated luminosity. Upper limits on the branching fractions, computed fitting simultaneously both track-plus-photon and tau-plus-photon selections, are reported in Table 5.15. The expected limit on the track-plus-photon channel is approximately three times smaller than on the tau-plus-photon channel, therefore the combined limit is driven by the later.

		Expected limit ($\times 10^{-6}$)	Observed limit ($\times 10^{-6}$)
$W \to \rho\gamma$	track+photon	$17.57^{+6.88}_{-4.91}$	12.58
	tau + photon	$6.38^{+2.48}_{-1.78}$	6.29
	combined	$5.95^{+2.33}_{-1.66}$	5.17

Table 5.15: Observed upper limits, compared to the expected upper limits estimated using the post sideband fit Asimov.

5.11 Conclusion

The search for $W \to \rho\gamma$ decay is performed using ATLAS pp collision data collected during Run 2. This search provides, for the first time, a measurement of the $W \to \rho\gamma$ branching fraction with the limit of $\mathscr{B}(W \to \rho\gamma) < 5.17 \times 10^{-6}$ at 95% CLsm which is 592 times the predicted SM value. This result demonstrates the versatility of the ATLAS general purpose detector and the relevance of specialised event selection criteria in the online system. It also paves the way for a direct W boson mass measurement and stricter tests of theoretical prediction based on the QCD factorisation approach, both possible with larger datasets that will become available in the upcoming decades.

6 Identifying D_s decays at the LHC

As it is shown in the previous chapter, one can identify ρ or any other meson by observing its properties and comparing them to the properties of the particles originating from background processes and applying the corresponding selection requirements. However, this can be extremely difficult due the large number of variables and the correlations between them. Alternatively, with the use of Machine Learning (ML) techniques, it is possible to build an algorithm which is able to distinguish signal mesons from other mesons and most importantly from quark and gluon jet background.

Gauge bosons can also have radiative decays into mesons containing other quarks. Of special interest is the decay $W \to D_s\gamma$, because the combination $c\bar{s}$ is Cabibbo allowed, the decay has a relative large expected branching ratio. D mesons are pseudoscalar mesons and a subset of D mesons, called the strange D mesons (D_s) are composed of a combination of charm and strange quark, and they are defined as D_s^+ ($c\bar{s}$) or D_s^- ($s\bar{c}$). D mesons are the lightest mesons, including the mass of the D_s meson of 1968.35 MeV [5]. During their decay the charm (anti)quark must change into an (anti)quark of another type. Such transitions involve the change of the internal charm quantum number, therefore can only take place via the weak interaction. In the most common cases, the charm quark changes into a strange quark via an exchange of a W boson, therefore the D_s meson predominantly decays into kaons (K) and pions (π). D_s mesons cannot be directly detected due to their short mean lifetime of $5.04 \cdot 10^{-13}$ s. Hence, they are measured via the reconstruction of their decay products [5]. Unlike the π and ρ which have a unique signature in the detector, the D_s has several different decay modes into a number of charged and neutral particles. These decay product are typically identified as a jet of particles in the detector.

Jets are sprays of particles formed in large amounts when quarks and gluons undergo fragmentation, a process by which these particles 'dress' themselves by attracting additional quarks from the vacuum to form color-neutral hadrons. This results in a cone of various particles, observed in a detector as a jet. In contrast, mesons produced from the decay of a color singlet state-like W boson decaying into a D_s meson and a photon-result from the decay of a particle that is already color-neutral. This decay does not involve the complex process of fragmentation but rather a simpler transformation where the original particle's mass is directly converted into other particles. The distinguishing feature of color singlet jets and for instance a jet formed by a c quarks, which can also contain a D meson, is that the color singlet decay products are isolated from the rest of the event, where in a quark jet, the D meson would be surrounded by fragmentation tracks.

The main characteristic of the radiative decay is that the mesons are produced without accompanying fragmentation tracks, therefore they produce isolated jets.

These jet signatures are typically much simpler structures compared to the multi-particle jets initiated by quarks and gluons undergoing fragmentation and hadronization. The ML model is trained to pick up the differences in the underlying structure of the jet. This also means that with retraining, the algorithm offers an opportunity to identify other mesons originating from hadronic decays as well. This would improve future searches for these rare decays and could improve the measurement precision using data to be collected during the ongoing LHC Run 3.

This chapter describes the development of a machine learning algorithm to identify jets from a color singlet decay into a D meson against a background of quark and gluon jets and also other color singlet decays. It has been published in abbreviated form in [3].

6.1 Machine Learning

To distinguish D_s mesons from the background quarks and gluons a supervised ML algorithm has been used. In this technique the model is trained with labeled data which allows the model to make predictions about the unseen data based on the previous examples. An example of a model differentiating different kinds of particles is shown on Figure 6.1. First, the model is trained to understand the differences between these categories by providing the properties of these particles, like mass, electric and color charge, etc. After the training procedure, if a picture of an particle is presented to the model, the algorithm will check all the features and identify the object. One type of supervised learning is classification, where the categorical labels are discrete, unordered values. In this case the ML algorithms learns to distinguish between two possible cases: signal and background.

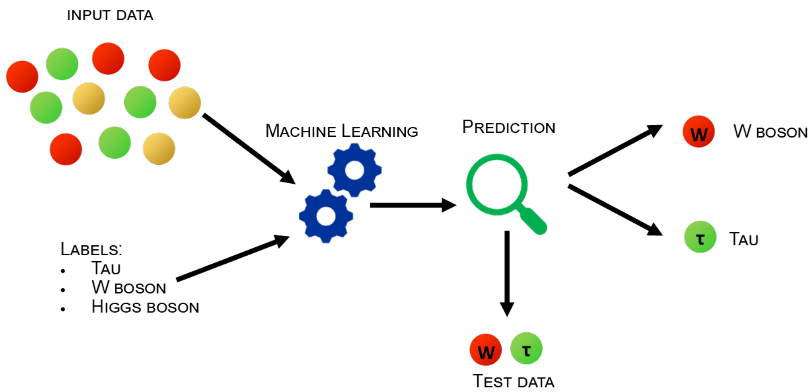


Figure 6.1: Workflow of an example of supervised learning. The input dataset is labeled with categorical values, here particle types. After the training the model is able to assign the unknown data point to the correct category.

6.1.1 Neural networks

Neural networks are a subset of ML techniques [124]. These networks are inspired by the structure of the human brain, and by the process of biological neurons interacting with each other.

The building blocks of the neural networks are called perceptrons or neurons. The perception makes its predictions based on calculating weighted sum over the entries of a vector $\vec{z}$ with addition of a constant value called bias b. An activation function $\phi(\vec{z})$ is also defined that takes this linear combination of the input values and the corresponding weight $\vec{w}$ resulting in an output z':

$$z' = \phi(\vec{w} \cdot \vec{z} + b). \tag{6.1}$$

With the combination of multiple perceptrons one can define a layer. In this case the perceptrons are producing a vectorial output $\vec{z'}$ by acting on the same input vector $\vec{z}$. In a layer instead of $\vec{w}$ a weight matrix ($\mathbf{W}$) is used and b becomes a bias vector $\vec{b}$:

$$\vec{z}' = \phi(\mathbf{W} \cdot \vec{z} + \vec{b}). \tag{6.2}$$

The activation function is needed to decide if a neuron needs to be activated or not. The activation function is needed to add non-linearity into the output of a neuron. Without the activation function, the sequential application of the layers on the input vector would produce the same results as the application of a single layer. There are multiple activation functions one can choose from and the most used ones are shown in Figure 6.2:

- Sigmoid is one of the most widely used non-linear activation function. It is commonly used for binary classification problems, since its output ranges between the 0 and 1:

$$f(x) = \frac{1}{1 + \exp(-x)}. \tag{6.3}$$

- The hyperbolic tangent function, $\tanh(x)$ is similar to the Sigmoid, but it is symmetric around the origin. This gives us an output value between -1 and 1:

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \tag{6.4}$$

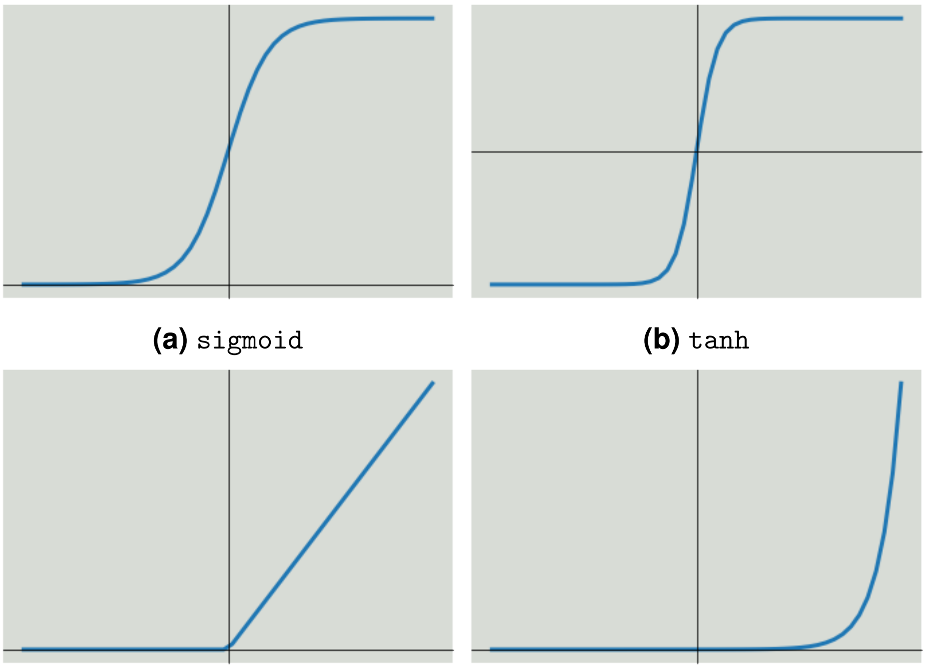
- The rectified linear unit (ReLU) [125] is another non-linear activation function. The advantage of this function is that it does not activate all the neurons at the same time, since the neurons are only activated if the output is bigger than 0. This makes `ReLU` a popular choice, since it is more computationally efficient than `sigmoid` or `tanh`:

$$f(x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} . \tag{6.5}$$

- The softmax is widely used for multi-class classification problems, since this function returns the probabilities of class memberships:

$$f(x)_i = \frac{e^{x_i}}{\Sigma_{j=1}^{n} e^{x_j}}. \tag{6.6}$$

A neural network which consists of sequentially applied layers is a Multilayer Perceptron (MLP) [124]. Figure 6.3 shows an example MLP made out of three layers: one input layer, two hidden layers, and one output layer. The input layer



(a) `sigmoid` **(b)** `tanh`

(c) `ReLU` **(d)** `softmax`

Figure 6.2: Visualisation of the most commonly used activation functions.

receives the input data, while the prediction or classification is performed by the output layer. In the case of MLP each layers linear combination is propagated to the next layer, meaning that each layer is providing their own results to the next one. This goes all the way through the hidden layers to the output layer. If a network has more than one hidden layer, it is defined as a Deep Neural Network (DNN). In DNNs all the neurons of the layers are connected to every neuron of its preceding layer, forming consecutive dense layers. Dense layers are connected on a general way, without any assumption about the features in the input data. This generality makes DNNs very expensive in terms of memory and computation.

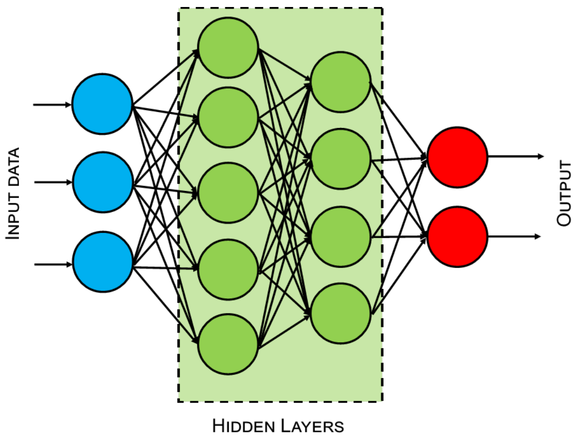


Figure 6.3: Graphical representation of a deep neural network. The network contains an input, and output and two hidden layers, consisting out of respectively 3, 5, 4 and 2 nodes (here circles).

A network which is using only a subset of the weights of a dense layer is called

Convolutional Neural Network (CNN). Only nearby inputs are connected together into a convolution, using the same set of weights for every neuron. To make use of the local connection patterns, the data needs to be spatial with local features, which are equally likely to exist in any part of the input data. CNNs are most commonly used on image data, where the features are local, and can occur anywhere within the image. The local connections and the shared weights makes CNN networks cheaper in terms of memory and computing power.

Convolutional layers are based on filters, which are small units that are applied across the data through a sliding window. The dimensions of the filter and the amount of the sliding are so called hyperparameters or tuning parameters of the network, which are defined prior the learning process. The output from multiplying the filter once with the input array is a single value. As the filter is applied multiple times the result is a two-dimensional array of output values called the feature map. These feature maps summarize the presence of the features in the input. In CNNs, a convolutional layer is usually followed by a pooling layer. During average pooling a filter on the feature map is applied, and selecting the average value within that filter, while during max pooling the maximum value is selected. The result of using a pooling layer is a summarized version of the features detected in the input. It is common practice that in case of classification each value in the final feature map is passed through one or more fully connected layer. This layer makes sure that the final decision is made based on the whole image, and not only the local features. An schematic representation of a CNN is shown in Figure 6.4.

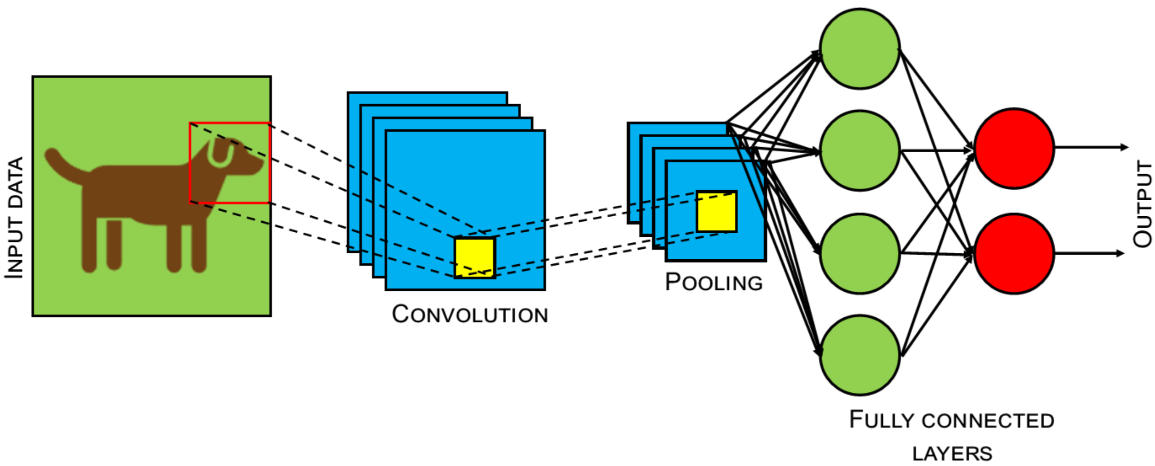


Figure 6.4: Graphical representation of a convolutional neural network. The network contains an input image, a convolutional layer and a pooling layer, followed by a number of fully connected layers.

Model training

Before creating and training a model, it is recommended to split the dataset into train and test data. The training dataset is used to train the machine learning model, while the test set is used to evaluate the model's performance. There are no rules what the size of these sets should be, but one needs to compromise. The larger training set results in a better trained model, while a smaller test set results in a less reliable performance test. On the other hand, a larger test set would improve the quality of the performance measure, but it would also mean that with less data the model could not be trained as well.

The goal of training any neural network is to make precise predictions $\hat{y}$ for data x which is not used during the training. The prediction should be as close as possible to the real y. During the training, by adjusting the weights and the biases in the network, one can minimize the loss, the difference between the actual value and the predicted value by the model. This technique is called backpropagation [126]. The degree of adjustment is regulated by the gradients of the loss-function with respect the model's parameters, defined as ∂loss$/\partial W$ and ∂loss$/\partial b$, where W represents the weight and b is the bias.

This technique can be repeated to determine the most optimal values by updating the parameters in each iteration. One of this iteration is called an epoch, while the learning rate controls the rate at which an algorithm updates the parameter estimates or learns the values of the parameters. This approach is called gradient descent [127], however this approach often raise a problem: the loss function is almost never fully convex leading to multiple local minima of the function and making the model's performance extremely sensitive to the initial parameters. This problem is usually solved by introducing the stochastic gradient descent algorithm [127] where a few samples are selected randomly instead of the whole data set for each iteration. The model is trained by updating the weights and biases in each samples, rather than in the entire training dataset. An extension of the stochastic gradient descent algorithm is the Adaptive moment estimation (Adam) [128]. Adam is an adaptive learning rate method, which means, it keeps the average of the previous parameter changes and uses this knowledge to update the parameter values and learning rates.

Model validation

Once model is trained, it can be used on data not used in the training process. However before making any predictions, it is necessary to know if the model is accurate, so that it makes the correct predictions. Estimating the performance on the training data is a biased measure, since it was trained with it. Instead, one should use the test set to find this performance, as this data were not used during the training procedure.

In our classification problem, most of the validation measures can be derived from the confusion matrix. In the binary classification problem the prediction can be either a positive case, signal, or a negative case, background. If the model predicted the signal event as signal, the classification for this event will be true positive (TP). Similarly, if the model correctly predicted a background event, it will fall into the true negative (TN) category. When the model categorizes the background event as a signal, the event will fall into the false positive category (FP), and similarly a false negative (FN) is an event which is a signal, but the model will predict as a background.

Using these quantities the following, most common performance measures can be constructed:

- Accuracy: The accuracy gives a comprehensive picture of the model's performance, since it is including all the predictions. However, it is only useful when all the classes are equally important, since it assigns the same weight for false positives and false negatives.

$$\text{accuracy} = \frac{TP + TN}{TP + TN + FP + FN}. \tag{6.7}$$

- Precision: The precision calculates the probability that a data point with a positive prediction is correctly categorized.

$$\text{precision} = \frac{TP}{TP + FP}. \tag{6.8}$$

- Sensitivity: the ratio of the true positive cases and the total actual positive cases.

$$\text{sensitivity} = \frac{TP}{TP + FN}. \tag{6.9}$$

All of there measurements can be used to determine a model's performance and to choose the best performing model. However, optimizing all parameters at the same time is rarely possible. To deal with this problem, a commonly used method in ML is the construction of the Receiver Operating Characteristic (ROC) curve. The ROC curve is a graph illustrating the model's performance at various classification thresholds, by plotting the true positive rate (sensitivity) against false positive rate. The false positive rate is defined as:

$$\text{FPR} = \frac{FP}{TN + FP}. \tag{6.10}$$

The Area Under the Curve (AuC) is the measure used as a summary of the ROC curve. An excellent model has the area under the ROC near to 1, meaning it has a good measure of separability. In the case of a poor model, the area equals 0, meaning it is reciprocating the results. When the area under the curve is 0.5, it means that the model has no separation capacity and classifies data randomly. In this case the ROC curve looks like a diagonal line. An example of the ROC curve can be seen in Figure 6.5.

Underfitting and overfitting

As it is described above, model's performance on a test set is a reliable indicator of its performance. However, it can also be used to determine if the model has picked up any patterns from the training data that do not exist in the test data, such as noise[1]. This can be accomplished by comparing the model's performance on the training and test data. If the model performs significantly better on the training set, it means that the model has "memorized" the training data instead of learning the underlying patterns and relationships. In other words, the model is overfitted.

Similarly, a model can also be underfitted. In this case the performance on the test and the training set are comparable, but both could be improved, The model is

[1] In the context of machine learning, noise refers to random fluctuations or irrelevant data in the training set that do not represent the underlying patterns or relationships that the model is trying to learn. Noise can be caused by various factors, such as measurement errors, outliers, or irrelevant features.

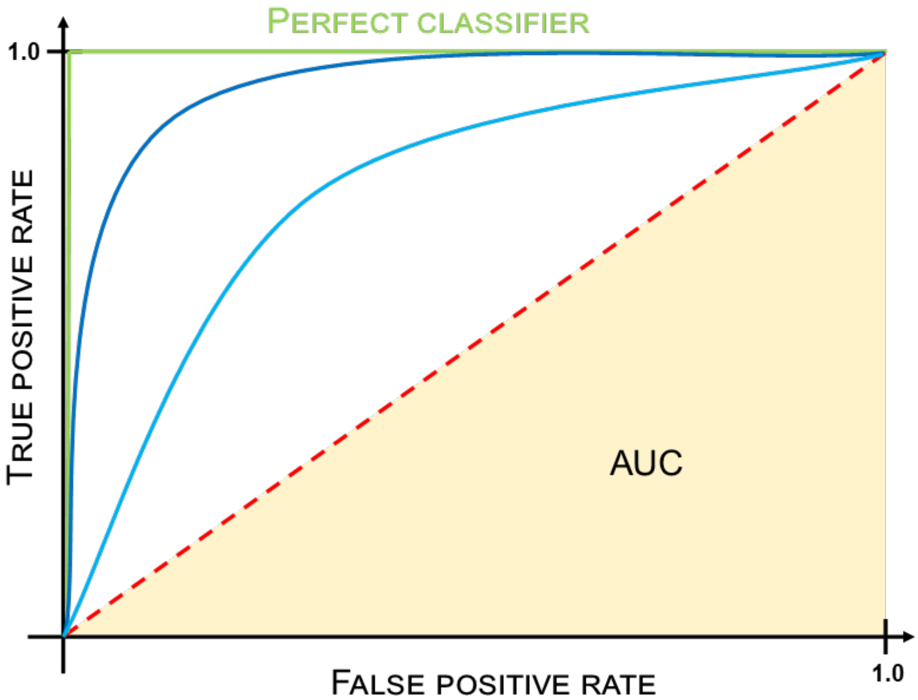


Figure 6.5: Examples of different ROC curves. The ROC curve of a perfect classifier is indicated with green, while the ROC curve of a classifier with random separation capacity is presented with red. ROC curves of the models, which are performing better than random but worse than the perfect classifier are indicated with blue. The AuC is also presented in the case of model with no separation capacity and it is coloured with yellow.

not complex enough to be able to describe the patterns in the training data, meaning that it will produce low performance on the testing data as well.

One way to prevent overfitting is by randomly dropping out nodes during the training. This is called dropout [129], and it causes the layer to appear and behave as if it were a layer with a different number of nodes and connection. Because of the dropout layer, only a few selected neurons are involved with the training, while the rest is ignored. After every iteration different sets of neurons are activated, and this is preventing some neurons from dominating the process, reducing the menace of overfitting. Figure 6.6 shows the neural network presented in Figure 6.3 when a dropout layer is used.

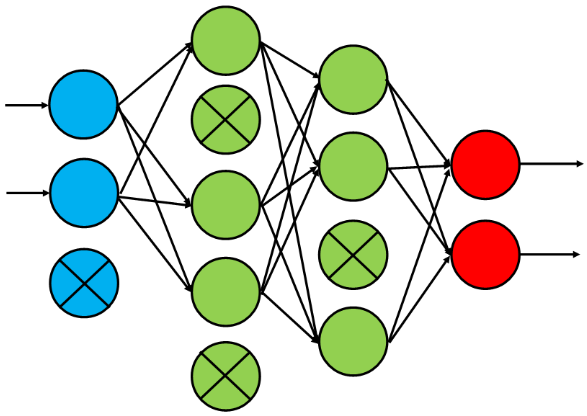


Figure 6.6: Neural network from Figure 6.3 with dropout layer applied. The dropped nodes are indicated with a cross.

6.1.2 General machine learning software and tools

The most popular programming language in data science is `Python` [130] therefore also the most used in ML. `Python` can be used with a large number of additional libraries, and it has a reactively simple syntax, which makes it easy to build ML algorithms. Although `Python` is not a high performance language for computation-intensive tasks, `Python` libraries `NumPy` [131] and `SciPy` [132] are implemented in `C` and perform well on vectorized operations of multidimensional arrays. They are widely used for data preparation and manipulation. `Python` also can be easily linked to `ROOT` [133], which is the most commonly used data analysis framework used in high energy physics.

The gist of the ML process happens in `Python` ML libraries like `TensorFlow` [134] or `PyTorch` [135]. `TensorFlow` is an open source library created by Google. It combines a variety of machine learning and deep learning models and algorithms into a single, flexible tool. It uses `Python` to create a user-friendly front-end API for developing applications, which is then executed in high-performance `C++`. `TensorFlow` applications can also be executed on GPUs, which allows high parallelisation yielding better performance.

While the interface of `TensorFlow` is rather complex, a built-in library have been developed to ease the development and evaluation of deep learning models. `Keras` [136] wraps the efficient numerical computation of `TensorFlow` but also makes it possible to define and train neural network models in just a few lines of code. The biggest advantage of Keras is that it enables fast experimentation with deep neural networks, while it is user-friendly, modular, and extensible.

6.1.3 Machine learning in high energy physics

Machine learning is widely used in every aspects of our life. It is not different with high energy physics either. ML is applied for a range of applications, starting from theoretical calculations to data analysis. Up to now the most frequently used machine learning algorithms in high energy physics are Boosted Decision Trees (BDTs) and Neural Networks (NN). There are different types of NN used: fully-connected (FCN), convolutional (CNN) and recurrent (RNN).

Particle physics has presented many opportunities for the application of machine learning, as various tasks require the classification of high-dimensional variable spaces. At the most basic level, machine learning tools can be used in hit reconstruction or track finding in individual detector systems. These tools are also capable of performing object identification by using data from different detector systems. Lastly, machine learning tools have been extensively utilized to categorize complete events as signal-like or background-like, both in the initial trigger decision and the final statistical analysis.

An overview of the ML applications in LHC physics can be found in Ref. [137] and [138].

6.2 Data generation and simulation

As it was already mentioned, multiple generators can be used to model physics processes. In the case of the D_s tagger, all the samples were generated at 13.6 TeV energy level, to match the Run 3 data taking period of the LHC, using two generators: Madgraph5 [139] and Pythia8 [109].

MadGraph5 [139] is a general purpose matrix-element based event generator. It can automatically generate matrix elements for any Lagrangian based process and it produces a dedicated output for Pythia8. For the simulation of the ATLAS detector response Delphes [140] were used. Delphes is a fast-simulation framework, which is able to simulate any general purpose detector response. The simulation includes tracking surrounded with a magnetic field, electromagnetic and hadronic calorimeters and muon identification systems, taking the granularity and resolution of the sub detector systems into account.The framework can be interfaced with any common event generator and outputs various physics objects, like isolated leptons, photons, collection of jets and missing transverse energy.

The Delphes package already includes the ATLAS detector specification by default. According to the default setup, every stable charged particle with a transverse momentum higher than 0.9 GeV, pseudorapidity of $|\eta| \leq 2.5$ and lying inside the detector volume covered by the tracker provides a track. The response of calorimeters to the incoming particles' energy deposits is dependent the segmentation and resolution of the calorimeters and the type of the particle. Delphes assumes that the ECAL and HCAL have identical segmentations and that the detector is symmetric in ϕ with respect to the $\eta = 0$ plane. By default, Delphes assumes that the ECAL and HCAL covers the pseudorapidity range of $\eta < 3$, and that muons and neutrinos will not interact with the calorimeters. In contrast, electrons and photons are assumed to deposit their energy in the electromagnetic parts of the calorimeters, while charged and neutral final-state hadrons are assumed to deposit their complete energy by interacting with the hadronic parts. In terms of geometrical sampling, the smallest unit of the calorimeters is a cell, which segments the (η, ϕ) plane for energy measurement purposes. The calorimeter response is parametrised through Gaussian smearing of the accumulated cell energy.

Photons are reconstructed solely from the ECAL information. Within `Delphes` photon conversion into electron positron pair is ignored and true photons and electrons without reconstructed tracks are both reconstructed as photons. Jets are reconstructed as pFlow jets [97]. In this case, `Delphes` uses both the tracking and calorimeter information to reconstruct the jet. The pFlow tracks contain the charged particles, while the pFlow towers consist of a mixture of neutral particles or charged particles without a corresponding reconstructed track. During the analysis the anti-k_t [98] jet clustering algorithm was used with a parameter of $\Delta R =$ 0.4 with a minimum threshold for transverse momentum of $p_T = 20$ GeV. `Delphes` includes the overlap removal by default meaning that it removes jets from the event if they were previously reconstructed as an isolated electron, muon or photon.

6.2.1 Signal process

Within the algorithm, D_s mesons originating from the radiative decay of $W \to D_s\gamma$ are considered as signal. The analysis is performed on the ATLAS detector with Run 3 parameters. The expected amount of signal events is approximately 2000 corresponding of $\sqrt{s}$=13.6 TeV centre of mass energy and 300 fb^{-1} pp collision data, using the $W \to D_s\gamma$ branching fraction from Table 1.1.

The signal samples were generated in two steps. The process $pp \to W$ was modelled via `MADGRAPHv5` using the default `NN23LO1` PDF set [12]. The output had been interfered with `Pythia`, which not only took care of the $W \to D_s\gamma$ decay, but also the hadronisation process. The $W^- \to D_s^-\gamma$ and the $W^+ \to D_s^+\gamma$ has been generated separately.

The decay products of the D_s are reconstructed as jets within the detector, therefore the signal sample consists of jets and photons. Jets are selected for further analysis if they satisfy the $p_T > 25$ GeV and $|\eta| < 2.1$ selection cuts. Jets are considered as a D_s meson if the angular distance to the generated D_s particle is $\Delta R < 0.2$. The full set of generated signal samples can be found in Table 6.1.

Signal validation

The validation of the signal samples have been performed on the generated particles. An event is considered valid, if it contains a generated D_s particle, with a generated photon sibling, and a generated W boson mother particle. The most important object variables were also validated. The p_T and energy distributions of the generated particles is presented in Figure 6.7, together with the ΔR, $\Delta\phi$ and $\Delta\eta$ between the D_s and the photon. As it is expected, the transverse momentum of the produced W boson is peaking at low values, resulting that the D_s and the γ is back-to-back.

A primitive reconstruction also have been carried out to study the detector effects. For this matter reconstructed jets and photons with $p_T > 25$ GeV have been selected. The recommended $|\eta|$ cut of 2.37 has been applied to photons, while jets

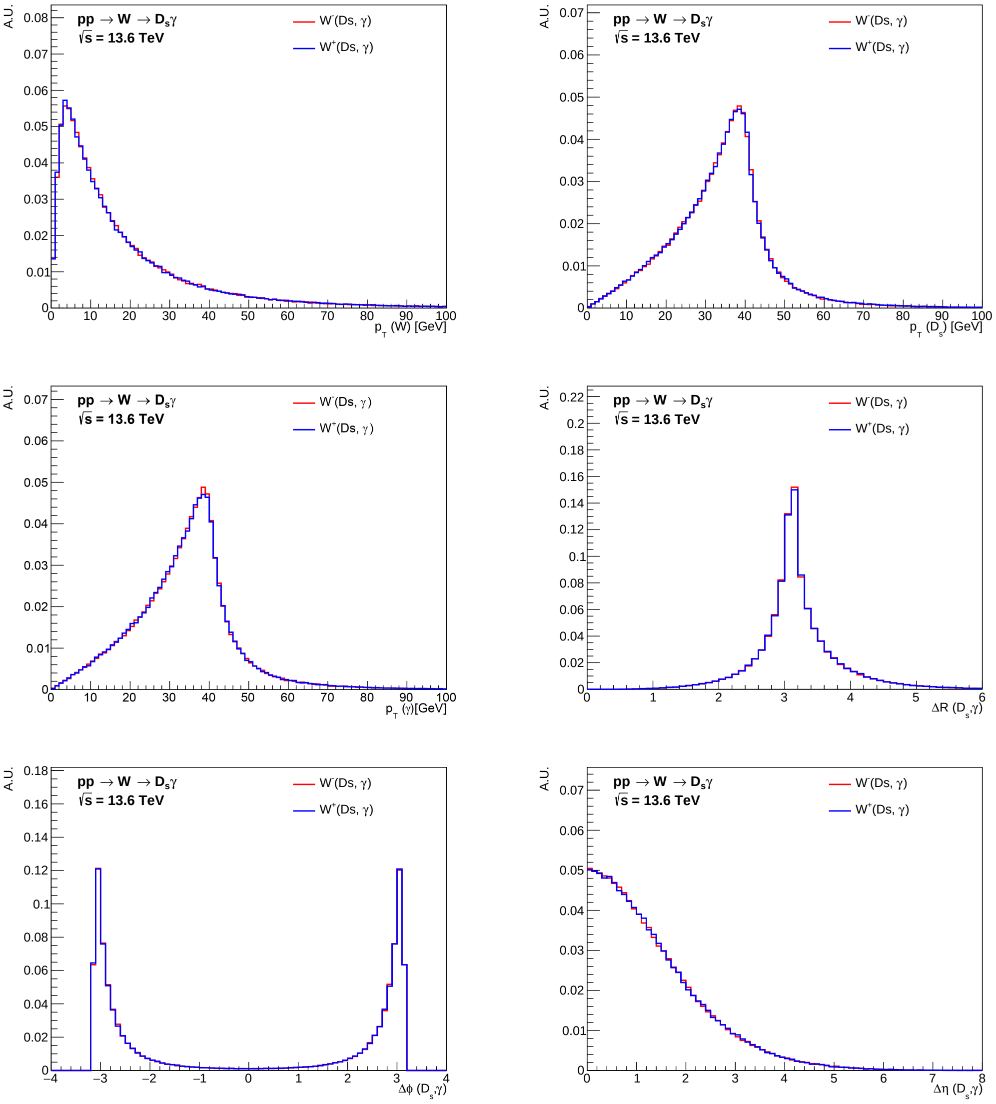


Figure 6.7: Kinematic distributions of the generated objects. Events are generated using `MadGraph5` and `Pythia8`. The plots are normalized to unity.

are selected if the satisfy $|\eta| < 2.1$. The W boson mass is reconstructed using a jet and a photon with the maximum $\Delta\pi$ value between them. The transverse momentum of the objects, together with the reconstructed W boson mass and ΔR, $\Delta\phi$ and $\Delta\eta$ is shown in Figure 6.8.

6.2.2 Background processes

The main background processes are $pp \to gg$ and $pp \to qq$ where g and q denotes the gluons and quarks respectively. The background samples were generated separately. In both cases the main process was simulated with `MadGraph5` again with the `NN23LO1` PDF set The output was interfered with Pythia to include the hadronisation process. Again, only jets with $p_T >$ 25 GeV and $|\eta| < 2.1$ selected for further analysis if they within $\Delta R < 0.2$ to the corresponding quark or gluon generated particle. Besides the quark and gluon backgrounds, $Z \to \Upsilon/(J/\psi)/\phi{+}\gamma$ events are also generated as background to ensure that the network is able to reject other color singlet states. Here again the same selection applied and events are selected if the jet is within $\Delta R < 0.2$ to the corresponding generated particle. The full set of generated background samples can be found in Table 6.1.

	Sample	Number of generated events	Cross-section (pb)
Signal	$pp \to W^+ \to D_s^+\gamma$	250 000	0.006954
	$pp \to W^+ \to D_s^+\gamma$	250 000	
Background	$pp \to gg$	3 000 000	5.03$\cdot 10^{11}$
	$pp \to qq$	3 000 000	2.14$\cdot 10^{10}$
	$pp \to Z \to \Upsilon\gamma$	50 000	0.003149
	$pp \to Z \to (J/\psi)\gamma$	50 000	0.004686
	$pp \to Z \to \phi\gamma$	50 000	0.000608

Table 6.1: List of generated signal and background processes with number of generated events and cross-section.

6.3 D_s identification using machine learning algorithm

The full set of $W \to D_s\gamma$ signal sample consists of 180k events, the qq background sample contains of 45k and the gg background sample contains 30k events. Besides qq and gg, 30k, 30k and 45k $Z \to \Upsilon/(J/\psi)/\phi{+}\gamma$ events are added to the background composition. This makes the full background sample with 160k events comparable to the signal. Before the training all the samples were divided into training and testing set, consisting of 70% and 30% of the full dataset respectively. To create the machine learning algorithm `TensorFlow` [134] and `Keras` [136] libraries were used. To determine the model performance the Receiver Operating Characteristic (ROC) curve and in particular the area under the ROC curve (AuC) has been determined. The network hyperparameters, such as the amount of layers and number of nodes in each layers were optimized with grid search to make sure that the best performing models are used to obtain the results.

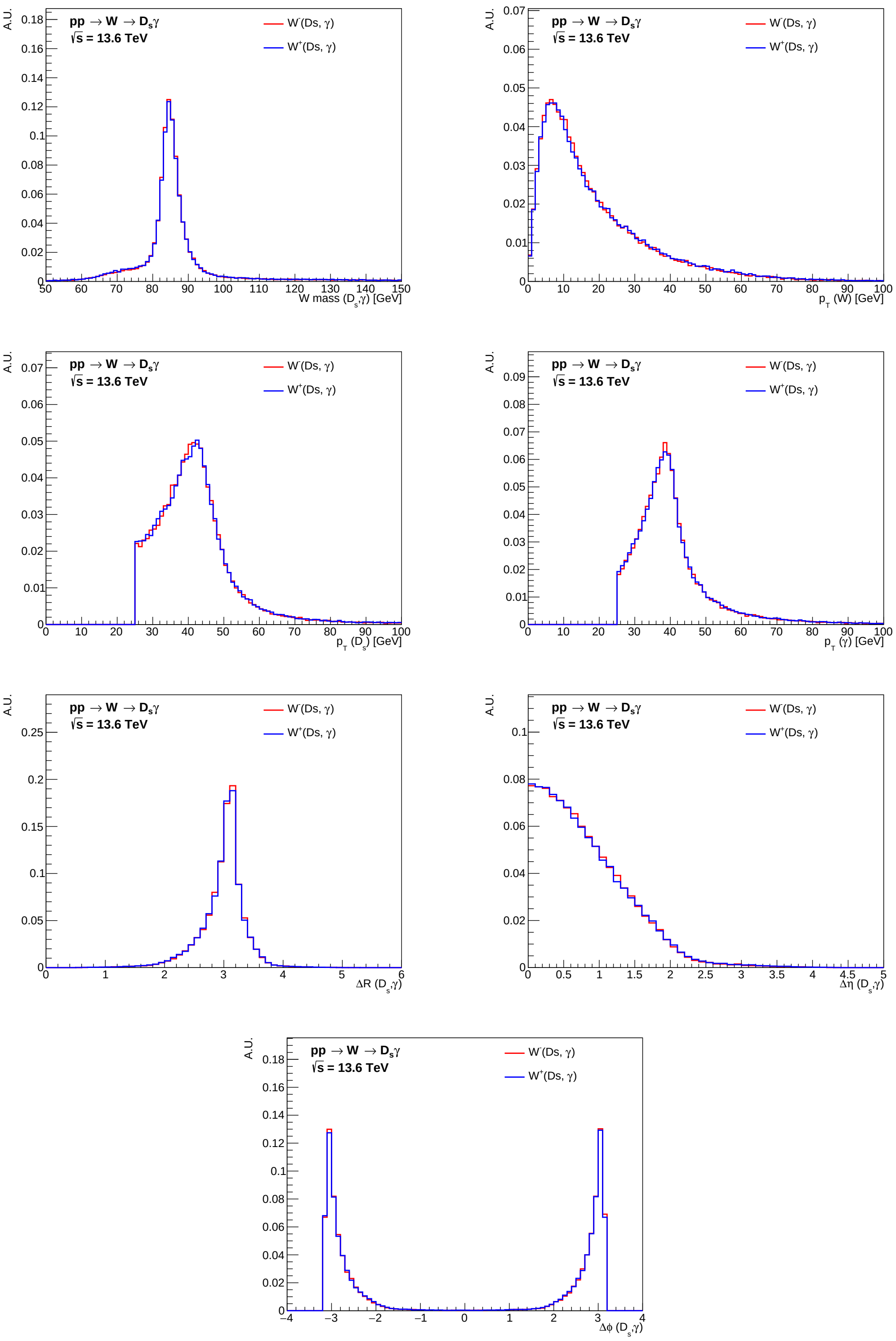


Figure 6.8: Kinematic distributions of the reconstructed objects after the detector simulation. Events are generated using `MadGraph5` and `Pythia8`, and the response of the ATLAS detector is simulated with `Delphes`. The plots are normalized to unity.

6.3.1 Deep Neural Network

In the previous chapter it is shown that using τ reconstruction variables can discriminate very well between the signal and background jets. A similar approach were used in the case of $W \rightarrow D_s\gamma$ studies as well. As it can be seen from the $\Delta\phi_{track}$ and $\Delta\eta_{track}$ variables presented on Figure 5.9, the background jets are more collimated than the jets originating from a $\tau_{had-vis}$ candidate. This is particularly true for gluon jets, since gluon-initiated jets have higher particle multiplicity and a softer fragmentation function, due to the large color factor. In the $D_s\gamma$ analysis, this is introduced through the variables of $\Delta\phi$ and $\Delta\eta$, which measures the width of the jet in the ϕ and η direction and through the R_{em} and R_{track} which measures the ΔR with respect to the jet axis in case of tracks and electromagnetic clusters. The choice of these variables is also supported by the fact, that within the $W \rightarrow \rho\gamma$ analysis the ΔR_{max} is one of the most discriminative variable (as it also can be seen on Figure 5.10), which is quite similar to the R_{track}. Another important variable used for τ reconstruction is the number of charged and neutral particle multiplicity, n_{ch} and n_0. The charged particle multiplicity is retrieved by simply counting the reconstructed tracks within the jet, while the number of neutrals is defined by the number of jet constituents which are not associated with a charge. Tau jets, and jets originating from D_s have lower multiplicity (n_{ch} and n_0) than quark and gluon jets. From the lower constituent multiplicity it can also be deducted, that signal jets have lower invariant mass. m_{tr} measures the invariant mass of of all charged tracks while m_j defines the invariant mass of all constituents in the jet. Jets emerging from D_s mesons are also less surrounded with hadronic activity caused by the fragmentation. p_{core} and f_{core} measuring the ratio of scalar sum p_T of the tracks in the jet cone and the jet p_T and the ratio of scalar sum E_T of the jet constituents in the jet cone and the jet total E_T respectively.

Since the usage of the Deep Neural Network (DNN) is inspired by the study presented in [141], the variables are further extended with the absolute values of the total charge and the jet-charge, the p_T weighted charge sum [142]:

$$q_j = \sqrt{\sum_i q_i |\vec{j} \cdot \vec{p}_i|^{1/2}} \sum_i |\vec{j} \cdot \vec{p}_i|^{1/2}, \tag{6.11}$$

where q_i is the i^{th} jet constituents charge, $\vec{p}_i$ is the momentum and $\vec{j}$ is the jet direction unit vector. The charge is expected to peak at zero for gluon jets, at one for signal jet, and have a higher average value for quark jets. in addition, with the b-jet identification some discriminating power against b-jets is also gained.

In addition a particular class of generalized angularities [143] are also added to the algorithm, which are efficient in distinguishing quark jets from gluon jets. The angularities are defined as:

$$\lambda_\beta^k = \sum_i z_i^k \theta_i^\beta, \tag{6.12}$$

where z_i is the momentum fraction of the i^{th} jet constituent, θ_i is the azimuth angle with respect to the jet axis while k and β are parameters. Furthermore, the E_{had}/E_{em} variable is also added to the list together with the N-Subjettiness variable τ_N [144].

The N-subjettiness is defined as:

$$\tau_N = \frac{1}{d_0} \sum_i p_{T,i} \min\left[\Delta R_{1,i}, \Delta R_{2,i}, ..., \Delta R_{1,N}\right], \tag{6.13}$$

where, i defines the i^{th} jet constituent, $p_{T,i}$ are their transverse momenta, and ΔR is the is the angular distance between a candidate and the constituent particle i. d_0 defines the normalization factor as $d_0 = \Sigma_i p_{T,i} R_0$ where R_0 is the pre defined jet radius by the jet clustering algorithm. τ_N shows to what degree the jet is composed of N subjets. For the signal jets the N-subjettiness expected to be close to zero, since all the radiation is aligned with the direction of the jet, meaning N (or fever) subjets. gg background jets have $\tau_N >> 0$, since large fraction of their energy distributed away from the jet direction, meaning they have at least N + 1 subjets.

All the variables used for the ML algorithm is listed in Table 6.2 and also shown in Figure 6.9.

Name	Description
$\Delta\eta$	width of the jet in η
$\Delta\phi$	width of the jet in ϕ
m_{tr}	invariant mass of all charged tracks in the jet
m_j	invariant mass of all constituents of the jet
n_{ch}	charged particle multiplicity
n_0	neutral particle multiplicity
$\lvert Q\rvert$	absolute value of the total charge
$\lvert q_j\rvert$	jet charge
b-tag	output of the b-tagging algorithm
R_{em}	Average ΔR with respect to the jet axis weighted by electromagnetic energy
R_{track}	p_T weighted average ΔR for tracks
f_{em}	fraction of EM energy over total neutral energy of the jet
p_{core1}	ratio of sum p_T in a cone of ΔR <0.1 and the jet p_T
p_{core2}	ratio of sum p_T in a cone of ΔR <0.2 and the jet p_T
f_{core1}	ratio of sum ET in a cone of $\Delta R <$ 0.1 and the jet total ET
f_{core2}	ratio of sum ET in a cone of $\Delta R <$ 0.2 and the jet total ET
f_{core3}	ratio of sum ET in a cone of $\Delta R <$ 0.3 and the jet total ET
$(p_T^D)^2$	λ_0^2
LHA	Les Houches Angularity; $\lambda_{0.5}^1$
Width	λ_1^1
Mass	λ_2^1
E_{had}/E_{em}	ratio of the hadronic versus electromagnetic energy deposited in the calorimeter
τ_0, τ_1, τ_2	N-Subjettiness

Table 6.2: DNN input parameters.

6.3.2 Convolutional Neural Network

An other approach for developing a D_s tagger is to use Convolutional Neural Network (CNN). In this case the input variables are the low level variables, like the energy deposit in the electromagnetic and the hadronic calorimeter and the track transverse momentum. These variables are plotted as 2D image and fed into a CNN. Using low

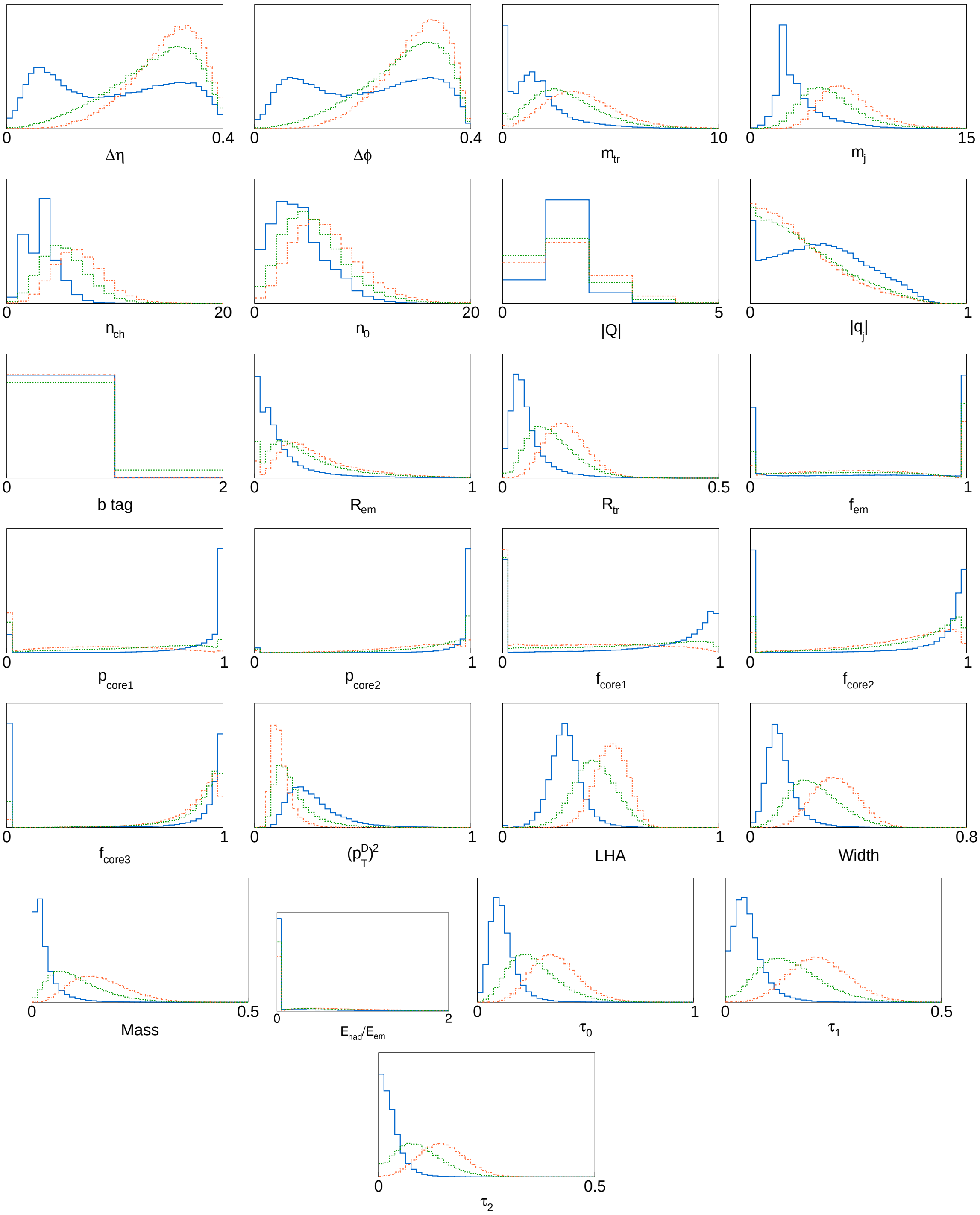


Figure 6.9: Distributions of the variables used for D_s identification, using DNN. The signal is presented with a solid blue line, while the gg and qq backgrounds are drawn with dashed red and dotted green lines respectively. The signal contains equal amount of $W^+ \to D_s\gamma$ and $W^- \to D_s\gamma$ events.

level variables make CNN favourable above DNN, because the later require carefully constructed high level variables, and thus does not guarantee that every variables are covered. This also means that no particle reconstruction is needed, one can feed the raw detector output into the algorithm, thus it is possible to implement the CNN algorithm in hardware level.

In the context of the $D_s\gamma$ analysis, these energy deposits and the track transverse momentum are converted into a 20×20 grid jet image. Since the jet reconstruction parameter is ΔR =0.4, and the segmentation of the ATLAS electromagnetic calorimeter is 0.02×0.02, the grid size of the jet image is equals of the smallest possible tower size in the η-ϕ plane. The variables introduced in three different channels as it is the case of a RGB picture, where the position of the hadronic deposit is indicated with blue, the electromagnetic deposit with green and the position of track is indicated with red. The intensity of colour notes the intensity of the deposit, while in the case of track, the transverse momentum. The images are scaled to the maximum values of each input. The schematic illustration of the jet image is shown in Figure 6.10.

6.3.3 Combined network

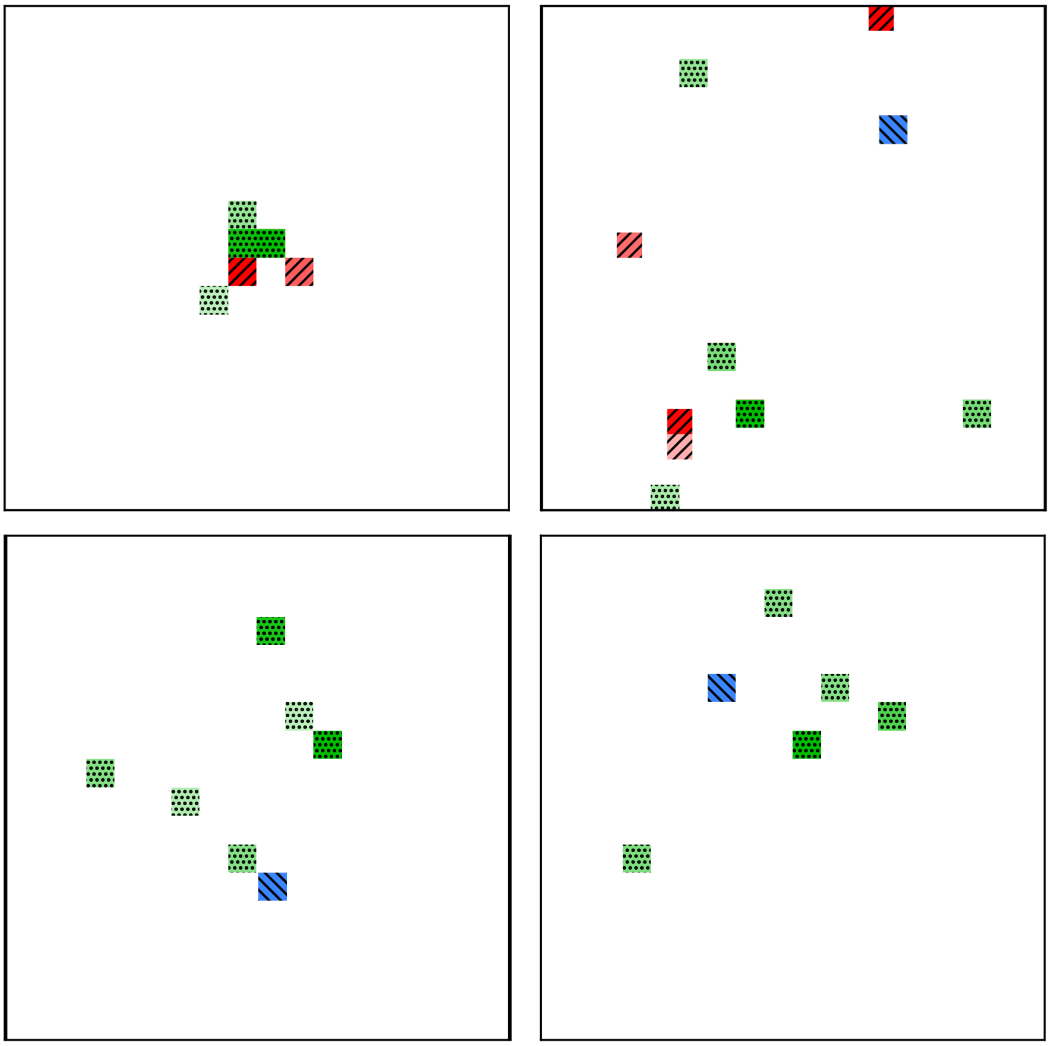
It is also possible to combine the DNN and the CNN approaches into a single network. In this case, the output of the DNN and the output of the CNN are the inputs of the next layer. The last layer of the model perform the classification, while the results are depends both on the output of the CNN and the DNN. This approach helps us to exploit the advantages of both network and increases the performance, since events that can only be classified with either CNN or DNN are all included in the final model. A schematic view of this combined network can be seen in Figure 6.11.

6.3.4 Model description and optimisation

To optimize the model's hyperparameters a separate dataset has been used, which is different from both the test and training data. For the approach using DNN, the optimization of the number of layers, the number of nodes in each layer, the activation function and the dropouts layers were performed. The results of the optimisation procedure is presented in Figure 6.12, where the best parameters are shown for the input layer, the first and second hidden layer and the necessity of a dropout layer.

During the optimisation it was determined that the best performing models have 35 nodes in the input layer and these nodes are activated with the `tanh` function. For the second layer the optimal number of nodes are considered to be 20. Although, during the optimisation the `softmax` activation function showed a slightly better performance on the second layer, it has been concluded that with the combination of the first layer parameters, the model benefits the most again from the `tanh` activation function on both the second and third layer. The optimal number of nodes in the third layer is 12. In addition the last plot in Figure 6.12 shows, that the model does not benefit from the dropout layers. Two dropout layers were considered, one after the input layer and one after the first hidden layer. Those models performed the best, where both dropout layer were excluded during the training.

Based on the optimisation results, the final DNN model configuration is consist of one input layer and two hidden layer with 35, 20 and 12 nodes respectively. The



(a) Signal **(b)** Background

Figure 6.10: Jet image construction from low level variables. The hadronic deposit is noted with blue back slash pattern, the electromagnetic deposit with dotted green and the track transverse momentum with red with forward slash pattern composing an RGB input picture to the CNN algorithm. The intensity of colour notes the intensity of the deposit, while in the case of track the transverse momentum.

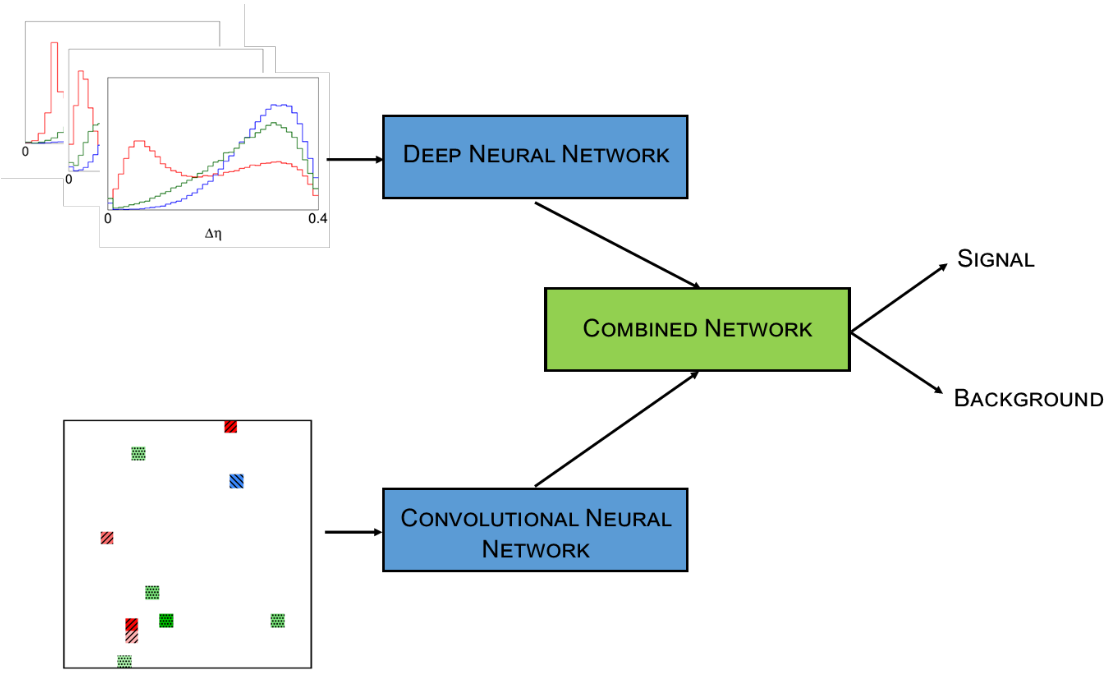


Figure 6.11: Combined network. The high level variables are input to a DNN, while the low level jet images are input to a CNN. The output of the DNN and CNN are input to an other ML model which carries out the classification.

activation function for the input layer and both hidden layer is `tanh`. No dropout layers were used. As it is common with the classification problems, the output layer is activated with the sigmoid function. The full set of hyperparameters is summarized in Table 6.3.

As mentioned in Section 6.1.1, CNNs are based on filters, what are applied throughout the full image as a sliding window. The dimension of the window and the amount of sliding are both hyperparameters one needs to optimize. In addition, CNNs are usually concluded with one or mode fully connected layer with a different set of parameters. These hyperparameters are added to the default set of parameters, as the number of nodes and layers and activation function type.

In the case fo the D_s tagger during the optimization the number of nodes of the dense and convolutional layer, the sliding window size and the activation functions, and the amount of max pooling layers were determined. As it can be seen from the Figure 6.13, the best performing models have [3$\times$3] windows sizes on both the first and second convolutional layer. For the third convolutional layer, the window size needs to be increased to [5$\times$5] The `tanh` activation function is determined to be the best choice for all the convolutional layers. As it can be seen from the plots, the number of nodes of the first layer should be around 30, and in the second layer and third layer it should be 8. The fully connected layer should have 10 nodes activated with the `ReLU` activation function, and also a max pooling layer should be added or after the first or after the second convolutional layer, but not both.

Taking the optimization results into account, the final model consists of 5 layers: 3 convolutional and 2 fully connected dense layer. The number of nodes in the convolutional layers are 30, 8 and 8 respectively. The window sizes are [3$\times$3] and [5$\times$5] in the last layer, while the activation function is `tanh` in all three cases. There is a max pooling layer added after the second convolutional layer. The number of

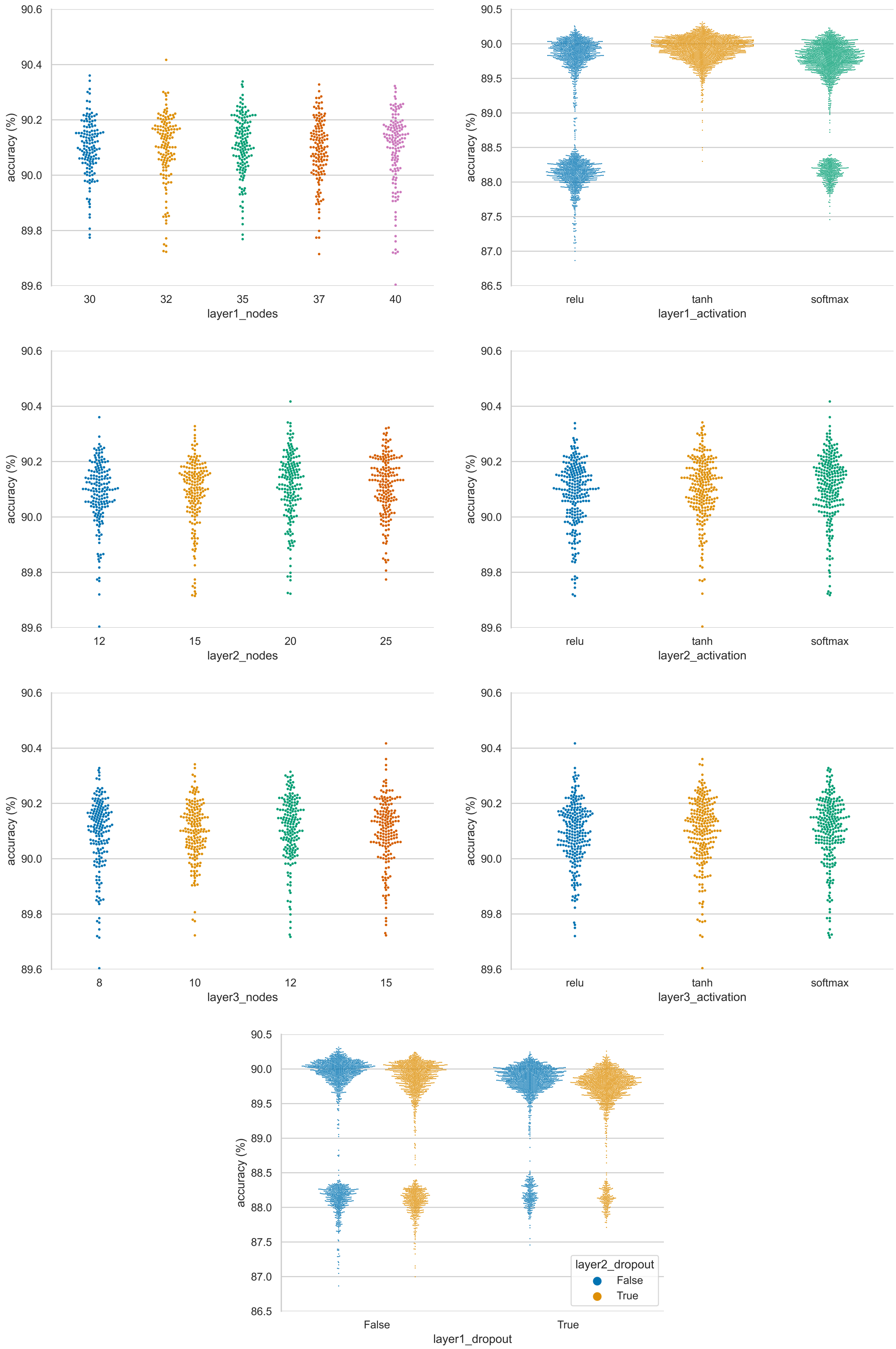


Figure 6.12: Optimisation of the network hyperparameters. The top row corresponds to the input layer, while the second and third row correspond to the first and second hidden layer respectively. The last plot shows the necessity of a dropout layer after the input or the first hidden layer.

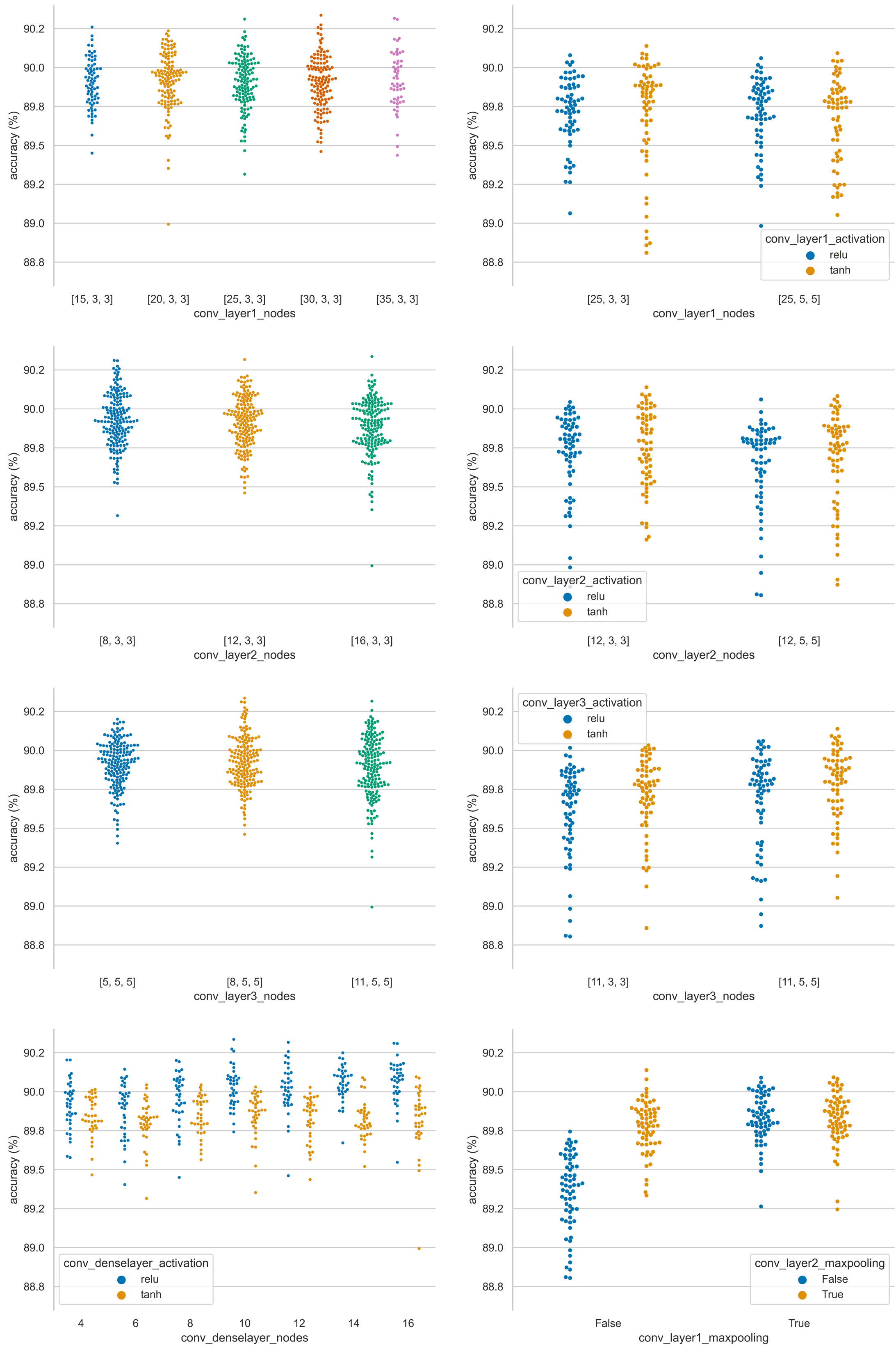


Figure 6.13: Optimisation of the CNN network hyperparameters. The top row corresponds to the input layer, while the second and third row correspond to the first and second CNN layer respectively. The plots in the last row show the parameters of the final dense layer and the necessity of a the max pooling layer.

nodes in the first dense layers is 10 with the `ReLU` activation function. The output layer is again a dense sigmoid layer. The parameters of the final CNN model are summarized in Table 6.3.

A similar optimisation approach has been used on the combined model as well. The best performing combined network has slightly different number of nodes within the DNN layers: 33, 20 and 14 respectively. An other significant change compared to the previously introduced models is the absence of the dense layers after the convolutional layers. Instead a combined dense layer is introduced with 8 nodes and `ReLU` activation function. The classification happens in the last `sigmoid` layer. The parameters of the combined model are summarized in the last column of Table 6.3.

Parameter	DNN	CNN	Combined
Dense layer nodes	35 - 20 - 12 - 1	–	33 - 20 - 14
Dense layer activation	`tanh` - `tanh` - `tanh` - `sigmoid`	–	`tanh` - `tanh` - `tanh`
Convolutional layer nodes	–	30 - 8 - 8	30 - 8 - 8
Window size	–	[3×3], [3×3], [5×5]	[3×3], [3×3], [5×5]
Convolutional layer activation	–	`tanh` - `tanh` - `tanh`	`tanh` - `tanh` - `tanh`
Max pooling	–	After the 1st convolutional layer	
Dense layers after convolution	–	10(`ReLU`) - 1(`sigmoid`)	–
Combined layer nodes	–	–	8 - 1
Combined layer activation	–	–	`ReLU` - `sigmoid`
Loss function	binary cross-entropy		
Optimizer	Adam		
Training epochs	40		
Batch size	1024		

Table 6.3: Hyperparameters of the different network types.

6.4 Results

The Receiver Operating Characteristic (ROC) curves of the different models are presented in Figure 6.14, while the output distributions of the models can be seen in Figure 6.15. Table 6.4 shows the Area Under the Curve (AuC) values of the different networks defined previously. As is expected, the combined model performs the best with 0.956, which corresponds to a signal efficiency of 47(15)% at a background rejection factor of 100(1000). Using DNN only one can reach a signal efficiency of 38(15)%, while using only CNN the efficiency is 35(9)% at 100(1000) times background rejection. As it can be seen, the performance is significantly better against a single background of gluon jets then against quark jets. This can be further improved if one uses only a gluon sample for training to an AuC of 0.991.

6.4.1 Network validation with signal-like samples

The tagging rate of the network for various samples used and not used during the training is presented in Table 6.5. Here a cut-off value of 0.75 is used. We find

Network type	Test sample	Training sample	AuC
DNN	D_s vs mixed	D_s vs mixed	0.939
CNN	D_s vs mixed	D_s vs mixed	0.938
Combined	D_s vs mixed	D_s vs mixed	0.956
	D_s vs gluon	D_s vs mixed	0.987
	D_s vs quark	D_s vs mixed	0.935
	D_s vs gluon	D_s vs gluon	0.991
	D_s vs quark	D_s vs quark	0.946

Table 6.4: Overview of the training results using the combined network. Mixed background test samples contain 50% quark and 50% gluon jets.

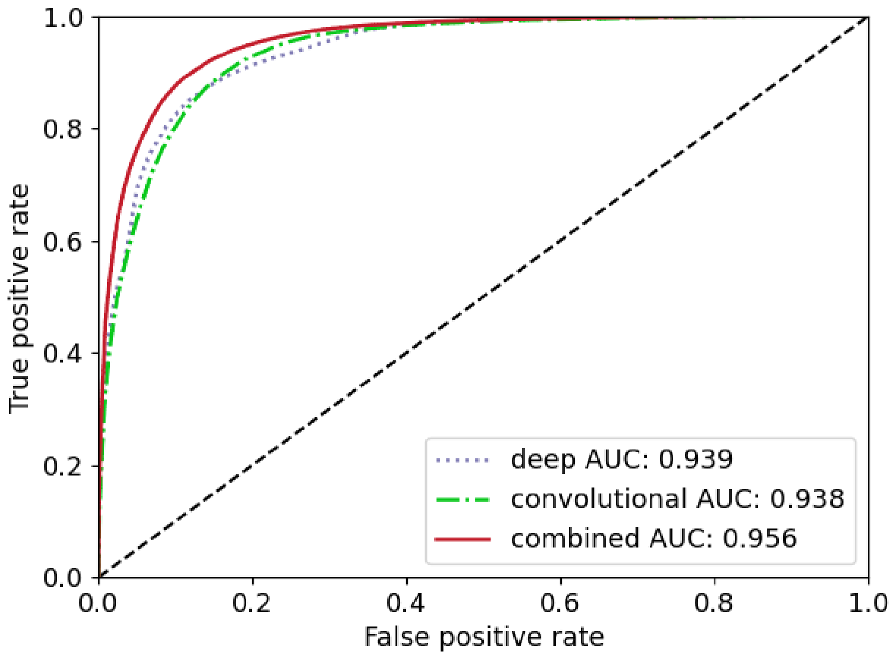
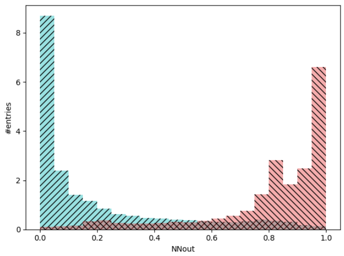
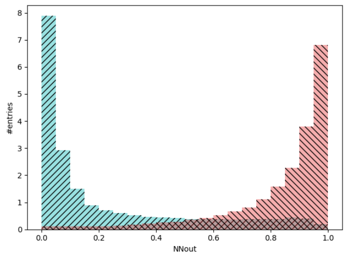


Figure 6.14: ROC curves for the different network types.

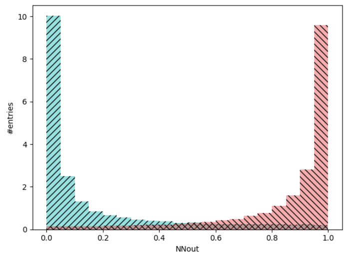
that for charm jets the results are not materially different from the generic quark-jet sample and this indicates that the absence of fragmentation tracks around the jets and a narrow jet with low multiplicity are more important than the exact D-meson decay topology. For hadronic τ decays, we find a high tagging rate, which is not surprising, given that τ leptons are also produced in a colour-singlet state and more than 5% of the D_s mesons decay to τs.



(a) DNN



(b) CNN



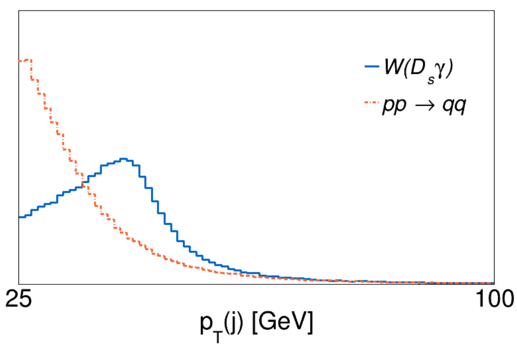
(c) Combined network

Figure 6.15: Output of the different networks for signal (red backward slash pattern) and background (blue forward slash pattern).

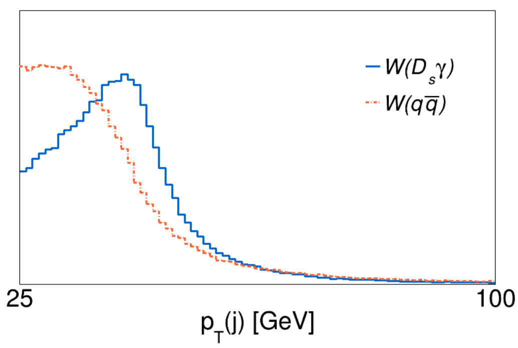
Sample	Tagging Rate	
$pp \to W \to D_s\gamma$	79%	
$pp \to qq$	9%	
$pp \to gg$	1%	
$pp \to Z \to \tau\tau$	62%	
$pp \to Z \to \Upsilon\gamma$	3%	
$pp \to Z \to (J/\psi)\gamma$	16%	
$pp \to Z \to \phi\gamma$	12%	
	Jet with a truth D_s	Jet without a truth D_s
$pp \to Z \to c\bar{c}$	9%	7%
$pp \to Z \to b\bar{b}$	1%	3%

Table 6.5: Jet tagging rate for different samples. For $c\bar{c}$ and $b\bar{b}$ samples, the tagging rate is separately evaluated for events, where the jet contains a truth D_s.

The p_T dependence of the DNN has also been tested, to make sure, that the network does not differentiate between the signal and background events purely due p_T difference. This has been tested by checking the network performance against $W \to q\bar{q}$ sample. In this case the quark jets are originating from the W boson have similar p_T distribution as the signal. Figure 6.16 shows the p_T distribution of the signal together with $pp \to qq$ and $W \to q\bar{q}$ sample.



(a)



(b)

Figure 6.16: p_T distribution of $W \to D_s\gamma$ against (a) $pp \to qq$ and (b) $W \to q\bar{q}$. The distributions are normalized to unity.

The performance of the combined model is 0.939 AuC. This can be compared against with the results of the combined network tested and trained only on quark sample presented in Table 6.4. The results show a 2% drop in the performance due to p_T dependence. This dependence is also visible on the feature importance plots presented on Figure 6.17, where the blue bars represent the weight of each feature (variable) within the network. Some of the p_T dependent variables, such as the τ_2 and τ_1 have high impact on the network performance.

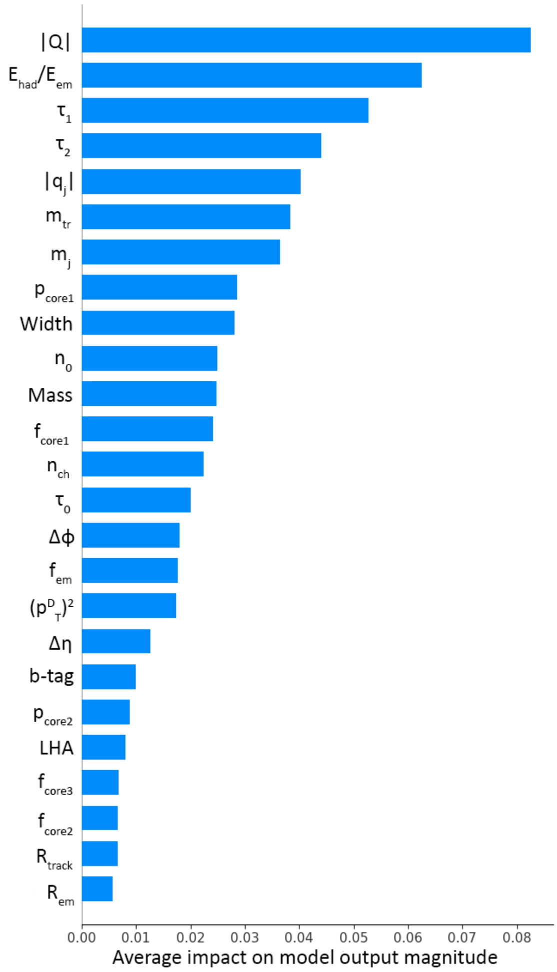


Figure 6.17: Feature importance plot of DNN. The blue bars represent the weight of each feature (variable) within the network.

6.4.2 Simulation uncertainties

To simulate data collected by ATLAS detector, `Pythia8` has been tuned based on Run 1 data, resulting in the A14 tunes, which is sensitive to the underlying event variables (evolution of transverse activity with leading track & calorimeter jets), the jet structure (track jet properties, jet masses & other substructure variables, jet shapes) and observables sensitive to additional jet emissions above the lowest-order process (dijet azimuthal decorrelation, $t\bar{t}$ gap fraction, the 3/2 jet ratio, and Z-boson p_T). To estimate the reliability of these approximations, systematic variations are included in `Pythia8`, which provide good coverage of the experimental and modelling uncertainties implicit in the tuning. Thus, the recommended variations of the

Pythia8 MC generator are defined to cover these processes: variation 1 is related to the underlying event activity, variation 2 is covering the jet shapes and substructure and the three variations 3 cover the effects of initial (ISR) and final state radiation (FSR) [145].

The stability of the network performance under these variations of the simulation parameters are investigated. The results were compared to the performance of the combined model, where both the training and the testing is performed on the mixed background sample. The results of the variance in the model performance is presented in Table 6.6, and as it can be seen, the network performance is very stable for the different tunes, the results are comparable with the previous study presented in Reference [141].

Parameter	+variation	-variation
Var1: UE activity	-0.008	0.003
Var2: jet shapes and substructure	-0.001	0.010
Var3a: ISR/FSR $t\bar{t}$ gap	-0.002	0.007
Var3b: ISR/FSR 3/2 jet ratio	-0.011	0.002
Var3c: ISR	-0.007	0.006

Table 6.6: Variations in the AuC for different Pythia8 tunes.

6.4.3 Consideration of pile-up

The effect of pile-up is also taken account during the analysis. In ATLAS, pile-up interactions are identified by means of vertex reconstruction. In the Delphes framework, the additional tracking and vertexing information is not available, meaning that the pile-up estimate is less reliable with respect real life conditions. In addition, pile-up mitigation techniques [146] are also not included in the Delphes simulation. In Delphes, pile-up interactions are extracted from a pre-generated QCD sample. These minimum bias interactions are randomly placed along the beam axis according to a predefined longitudinal spread. The actual number of pile-up interactions per bunch crossing is randomly extracted from a Poisson distribution.

The samples used for studying the pile-up effects, were simulated with pile-up of $\langle\mu\rangle = 40$ meaning on average 40 pile-up interaction, which was the expected amount for LHC Run 3 conditions at the time this study has been performed. The effect of pile-up on some of the DNN variables is shown on Figure 6.18. It can be seen that pile-up makes the signal more background-like, meaning that it is expected that the network will perform worse on the pile-up samples. The retrained network, without further optimisation shows a drop of 0.076 in the AuC, meaning that while pile-up has a significant effect, the model is still able to identify D_s mesons.

6.5 Estimated upper limit for $W \to D_s\gamma$

In this section prospects for the measurement of $\mathscr{B}(W \to D_s\gamma)$ using the method described previously is studied. For the purpose of this exercise it is assumed that low-pileup data corresponding to the integrated luminosity of 1 fb^{-1} is collected during LHC Run 3. Events are required to have one jet tagged as D_s and an isolated

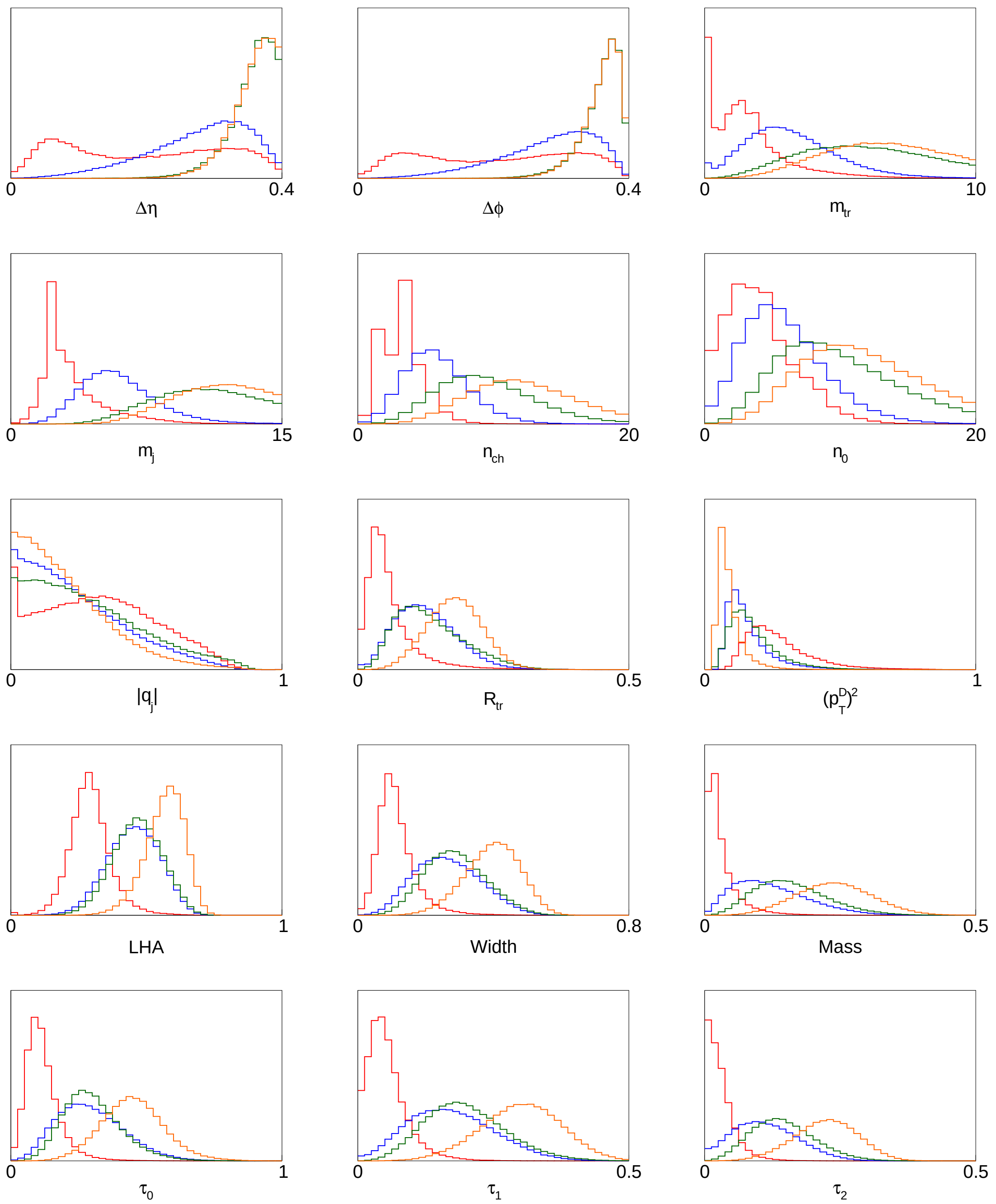


Figure 6.18: Effect of pile-up on the signal and background samples. Here the signal represents both W^+ and W^- samples, and background is both gg and qq. Red represents the signal without pile-up, blue is signal with pule-up, green and orange is background without and with pile-up.

photon with $p_T > 30$ GeV. Events with invariant mass of jet-photon system ± 10 GeV around W boson mass are selected. Triggering efficiency is assumed to be 100%. The optimised network cut-off of 0.75 provides the best sensitivity. Total signal efficiency for $W^+ \to D_s\gamma$ $(W^- \to D_s\gamma)$ is estimated to be 15.5% (18.7%) respectively.

In order to estimate background level, large MC samples of $pp \to gg$ and $pp \to qq$, as well as $pp \to q\gamma$, $Z \to ee$ and $Z \to \tau\tau$ are generated with `MADGRAPHv5` and `Pythia8`. The detector response is simulated via `Delphes` package using the ATLAS detector configuration files. Backgrounds are normalised according to their generated cross sections. The total level of background is estimated to be 930000 events corresponding to the integrated luminosity of 1 fb^{-1}. The background is dominated by QCD process while less than 1% the total background arises from Z boson events. Figure 6.19 show distribution of D_s tagged jet-plus-photon invariant mass for the backgrounds and $W \to D_s\gamma$ signal normalised to the integrated luminosity of 1 fb^{-1}. The signal histogram is overlaid and scaled by a factor of 10^4.

The CL_s method [119, 147] is used to calculate upper limit on the branching fraction of the $W \to D_s\gamma$ decay. Signal uncertainty is assumed to be 10% and has only marginal impact on the calculated limit. The uncertainty on the background level is assumed to be 0.5% as obtained in the ATLAS search for radiative Higgs boson decay [148]. The calculated CL_s exclusion as a function of branching fraction of $W \to D_s\gamma$ is shown in Figure 6.20.

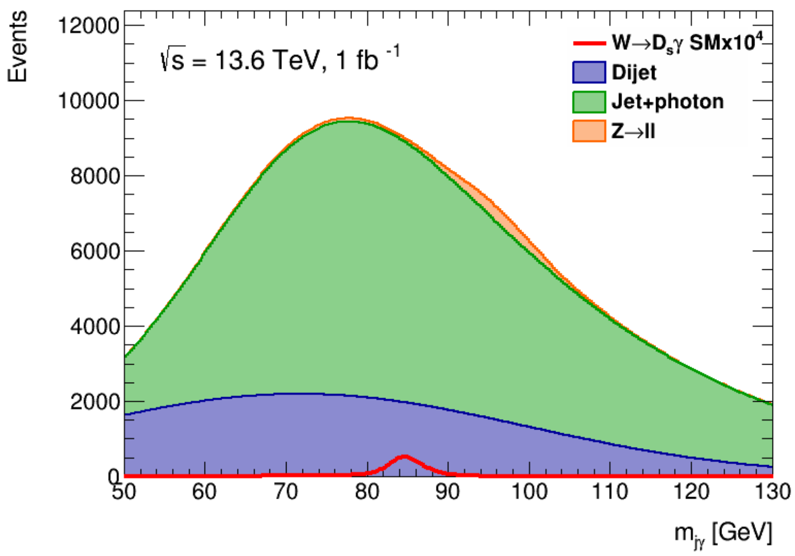


Figure 6.19: Distribution of the invariant mass of D_s tagged jet-plus-photon system.

The expected upper limit at the 95% confidence level is determined to be:

$$\mathscr{B}(W \to D_s\gamma) < (2.87 \pm 0.22) \times 10^{-4}, \tag{6.14}$$

which is by a factor of two compared to the observed upper limit from LHCb [24].

With the entire Run 3 dataset corresponding to about 300 fb^{-1}, assuming trigger efficiency of 40% and taking into account deterioration of the D_s tagger due to high pileup the expected upper limit improves to $\mathscr{B}(W \to D_s\gamma) < 1.6 \times 10^{-4}$. Development of dedicated trigger is needed to achieve corresponding precision.

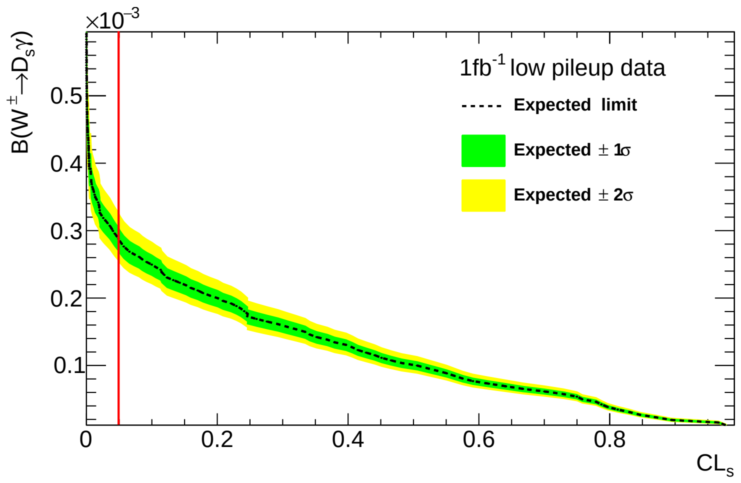


Figure 6.20: Expected upper limit on branching fraction of the $W \to D_s\gamma$ decay.

6.6 Conclusion

The algorithm to identify jets originating from D_s mesons in radiative W boson decays presented in this chapter shows a good efficiency of 47% for signal with a 100 times rejection of a background of quarks and gluons. Against a single background of gluon jets the algorithm works even better. The algorithm is stable under the variations of the simulation parameters and it also works in the presence of pile-up but with a 0.076 drop in the AuC. The algorithm opens up the possibility to further improve measurements and searches involving D_s mesons, especially in case of the rare decays that suffer from low statistics. The performance of the deep neural network and the convolutional neural network is very similar. The combined network performs slightly better than either. With low pileup dataset corresponding to the integrated luminosity of 1 fb^{-1} upper limit on branching fraction of $W \to D_s\gamma$ decay can be determined at the level of $\mathscr{B}(W \to D_s\gamma) < 2.9 \times 10^{-4}$.

List of abbreviations

ALICE A Large Ion Collider Experiment
AOD Analysis Object Data
ATLAS A Toroidal LHC ApparatuS
AuC Area Under the Curve

BCID Bunch Crossing Identification
BDT Boosted Decision Tree
BEH Brout-Englert-Higgs
BSM Beyond Standard Model

CDF Collider Detector at Fermilab
CERN European Organisation for Nuclear Research
CL confidence level
CMS Compact Muon Solenoid
CNN Convolutional Neural Network
CSC Cathode Strip Chambers
CSM Chamber Service Module
CTP Central Trigger Processor

DAQ Data Acquisition
DNN Deep Neural Network
DUNE Deep Underground Neutrino Experiment

ECAL electromagnetic calorimeter
ECR Event Count Reset
EF Event Filter
EW electroweak

FCal Forward Calorimeter
FELIX Front-End Link eXchange
FPGA Field-Programmable Gate Array
FSR Final State Radiation

GBT GigaBitTransceiver
GRL Good Runs List

HCAL hadronic calorimeter
HGTD High-Granularity Timing Detector
HL-LHC High Luminosity LHC
HLT High Level Trigger
HS hard scatter

IBL Insertable B-Layer
ID Inner Detector
ISR Initial State Radiation
ITk Inner Tracker

L1 Level-1
L1ID Level 1 Identification or Event number
L2 Level-2
LAr Liquid Argon
LCG LHC Computing Grid
LEP Large Electron–Positron Collider
LHC Large Hadron Collider
LHC-b Large Hadron Collider beauty
LHE Les Houches Event
LINAC2 Linear Accelerator

MC Monte Carlo
MDT Monitored Drift Tube
ML Machine Learning
MLE Maximum Likelihood Estimation
MLP Multilayer Perceptron
MROD MDT Read Out Driver
MS Muon Spectrometer

NLO Next-to-Leading Order
NP nuisance parameter
NSW New Small Wheels

PCIe Peripheral Component Interconnect Express
PDF Parton Distribution Function
pFlow Particle Flow
PFO Particle Flow Object
pp proton-proton
PS Parton Shower
PS Proton Synchotron
PSB Proton Syncrotron Booster

QCD Quantum Chromodynamics
QED Quantum Electrodynamics
QFT Quantum Field Theory

RF Radiofrequency
RNN Recurrent Neural Network
ROB Read Out Buffer
ROC Receiver Operating Characteristic
ROD Read Out Driver
RoI Regions-of-Interest
ROS Read Out System
RPC Resistive Plate Chambers

SCT Semi-Conductor Tracker

SM Standard Model
SPS Super Proton Synchotron
swROD software ROD

TDAQ Trigger and Data Acquisition
TDC Time to Digital Converter
TGC Thin Gap Chambers
TileCal Tile Calorimeter
TRT Transition Radiation Tracker
TTC Timing, Trigger and Control

Bibliography

[1] Evelin Bakos et al. *Search for the exclusive W boson hadronic decays $W^\pm \to \pi^\pm\gamma$, $W^\pm \to K^\pm\gamma$ and $W^\pm \to \rho^\pm\gamma$ with the ATLAS detector*. Tech. rep. Geneva: CERN, 2023. URL: https://cds.cern.ch/record/2867619.

[2] Georges Aad et al. "Search for the exclusive W boson hadronic decays $W^\pm \to \pi^\pm\gamma$, $W^\pm \to K^\pm\gamma$ and $W^\pm \to \rho^\pm\gamma$ with the ATLAS detector". In: (Sept. 2023). arXiv: 2309.15887 [hep-ex].

[3] E. Bakos, N. de Groot, and N. Vranjes. "Identifying D mesons from radiative W decays at the LHC". In: *Symmetry* 15.10 (2023). DOI: 10.3390/sym15101948. arXiv: 2207.13587 [hep-ph].

[4] Abdus Salam. "Gauge Unification of Fundamental Forces". In: *Rev. Mod. Phys.* 52 (1980). Ed. by A. Ali et al., pp. 525–538. DOI: 10.1103/RevModPhys.52.525; Chris Quigg. *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions: Second Edition*. USA: Princeton University Press, Sept. 2013. ISBN: 978-0-691-13548-9, 978-1-4008-4822-5; Peter W. Higgs. "Broken symmetries, massless particles and gauge fields". In: *Phys. Lett.* 12 (1964), pp. 132–133. DOI: 10.1016/0031-9163(64)91136-9; Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (1964). Ed. by J. C. Taylor, pp. 508–509. DOI: 10.1103/PhysRevLett.13.508; Peter W. Higgs. "Spontaneous Symmetry Breakdown without Massless Bosons". In: *Phys. Rev.* 145 (1966), pp. 1156–1163. DOI: 10.1103/PhysRev.145.1156.

[5] R. L. Workman et al. "Review of Particle Physics". In: *PTEP* 2022 (2022), p. 083C01. DOI: 10.1093/ptep/ptac097.

[6] Michael E Peskin and Daniel V Schroeder. *An introduction to quantum field theory*. Includes exercises. Boulder, CO: Westview, 1995. URL: https://cds.cern.ch/record/257493.

[7] Roel Aaij et al. "Test of lepton universality in beauty-quark decays". In: *Nature Phys.* 18.3 (2022), pp. 277–282. DOI: 10.1038/s41567-021-01478-8. arXiv: 2103.11769 [hep-ex].

[8] Peter W. Higgs. "Broken Symmetries and the Masses of Gauge Bosons". In: *Phys. Rev. Lett.* 13 (1964). Ed. by J. C. Taylor, pp. 508–509. DOI: 10.1103/PhysRevLett.13.508.

[9] A. B. Arbuzov. "Quantum Field Theory and the Electroweak Standard Model". In: (2017). Ed. by Martijn Mulders and Giulia Zanderighi, pp. 1–34. DOI: 10.23730/CYRSP-2017-004.1. arXiv: 1801.05670 [hep-ph].

[10] Michael H. Seymour and Marilyn Marx. “Monte Carlo Event Generators”. In: *69th Scottish Universities Summer School in Physics: LHC Physics*. Apr. 2013, pp. 287–319. DOI: 10.1007/978-3-319-05362-2_8. arXiv: 1304.6677 [hep-ph].

[11] S. D. Drell and Tung-Mow Yan. “Partons and their Applications at High-Energies”. In: *Annals Phys.* 66 (1971), p. 578. DOI: 10.1016/0003-4916(71)90071-6.

[12] Richard D. Ball et al. “Parton distributions from high-precision collider data”. In: *Eur. Phys. J. C* 77.10 (2017), p. 663. DOI: 10.1140/epjc/s10052-017-5199-5. arXiv: 1706.00428 [hep-ph].

[13] M. Banner et al. “Observation of Single Isolated Electrons of High Transverse Momentum in Events with Missing Transverse Energy at the CERN -p p Collider”. In: *Phys. Lett. B* 122 (1983), pp. 476–485. DOI: 10.1016/0370-2693(83)91605-2.

[14] M. Baak et al. “The global electroweak fit at NNLO and prospects for the LHC and ILC”. In: *Eur. Phys. J. C* 74 (2014), p. 3046. DOI: 10.1140/epjc/s10052-014-3046-5. arXiv: 1407.3792 [hep-ph].

[15] T. Aaltonen et al. “High-precision measurement of the W boson mass with the CDF II detector”. In: *Science* 376.6589 (2022), pp. 170–176. DOI: 10.1126/science.abk1781. URL: %5Curl%7Bhttps://www.science.org/doi/abs/10.1126/science.abk1781%7D.

[16] Georges Aad et al. “Measurement of the W-boson mass and width with the ATLAS detector using proton-proton collisions at $\sqrt{s}$ = 7 TeV”. In: (Mar. 2024). arXiv: 2403.15085 [hep-ex].

[17] Georges Aad et al. “Test of the universality of τ and μ lepton couplings in W-boson decays with the ATLAS detector”. In: *Nature Phys.* 17.7 (2021), pp. 813–818. DOI: 10.1038/s41567-021-01236-w. arXiv: 2007.14040 [hep-ex].

[18] Georges Aad et al. “Measurement of the $W \to \tau\nu_\tau$ cross section in pp collisions at $\sqrt{s}$ =7 TeV with the ATLAS experiment”. In: *Phys. Lett. B* 706 (2012), pp. 276–294. DOI: 10.1016/j.physletb.2011.11.057. arXiv: 1108.4101 [hep-ex].

[19] Yuval Grossman, Matthias König, and Matthias Neubert. “Exclusive Radiative Decays of W and Z Bosons in QCD Factorization”. In: *JHEP* 04 (2015), p. 101. DOI: 10.1007/JHEP04(2015)101. arXiv: 1501.06569 [hep-ph].

[20] Georges Aad et al. “Measurement of $W^\pm$ and Z-boson production cross sections in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector”. In: *Phys. Lett. B* 759 (2016), pp. 601–621. DOI: 10.1016/j.physletb.2016.06.023. arXiv: \textbf{1603.09222} [hep-ex].

[21] Michelangelo Mangano and Tom Melia. “Rare exclusive hadronic W decays in a $t\bar{t}$ environment”. In: *Eur. Phys. J. C* 75.6 (2015), p. 258. DOI: 10.1140/epjc/s10052-015-3482-x. arXiv: 1410.7475 [hep-ph].

[22] T. Aaltonen et al. “Search for the Rare Radiative Decay: $W \to \pi\gamma$ in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV”. In: *Phys. Rev. D* 85 (2012), p. 032001. DOI: 10.1103/PhysRevD.85.032001. arXiv: 1104.1585 [hep-ex].

[23] CMS Collaboration. "Search for W Boson Decays to Three Charged Pions". In: *Phys. Rev. Lett.* 122 (2019), p. 151802. DOI: 10.1103/PhysRevLett.122.151802. arXiv: 1901.11201 [hep-ex].

[24] "Search for the rare decays $W^+ \to D_s^+\gamma$ and $Z \to D^0\gamma$ at LHCb". In: (Dec. 2022). arXiv: arXiv:2212.07120 [hep-ex].

[25] F. Abe et al. "Search for the rare decay $W^\pm \to D_s^{+-}\gamma$ in $p\bar{p}$ collisions at $\sqrt{s} = 1.8$ TeV". In: *Phys. Rev. D* 58 (1998), p. 091101. DOI: 10.1103/PhysRevD.58.091101.

[26] Lyndon Evans and Philip Bryant. "LHC Machine". In: *Journal of Instrumentation* 3.08 (Aug. 2008), S08001–S08001. DOI: 10.1088/1748-0221/3/08/s08001. URL: https://doi.org/10.1088/1748-0221/3/08/s08001.

[27] G. Aad et al. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *JINST* 3 (2008), p. 08003. DOI: 10.1088/1748-0221/3/08/S08003.

[28] S. Chatrchyan et al. "The CMS Experiment at the CERN LHC". In: *JINST* 3 (2008), S08004. DOI: 10.1088/1748-0221/3/08/S08004.

[29] G. Aad et al. "Expected Performance of the ATLAS Experiment - Detector, Trigger and Physics". In: (Jan. 2009). arXiv: 0901.0512 [hep-ex].

[30] G. L. Bayatian et al. "CMS technical design report, volume II: Physics performance". In: *J. Phys. G* 34.6 (2007), pp. 995–1579. DOI: 10.1088/0954-3899/34/6/S01.

[31] K. Aamodt et al. "The ALICE experiment at the CERN LHC". In: *JINST* 3 (2008), S08002. DOI: 10.1088/1748-0221/3/08/S08002.

[32] A. Augusto Alves Jr. et al. "The LHCb Detector at the LHC". In: *JINST* 3 (2008), S08005. DOI: 10.1088/1748-0221/3/08/S08005.

[33] Esma Mobs. "The CERN accelerator complex - August 2018. Complexe des accélérateurs du CERN - Aout 2018". In: (Aug. 2018). General Photo. URL: https://cds.cern.ch/record/2636343.

[34] G. Aad et al. "Luminosity determination in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J. C* 83.10 (2023), p. 982. DOI: 10.1140/epjc/s10052-023-11747-w. arXiv: 2212.09379 [hep-ex].

[35] G. Avoni et al. "The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS". In: *JINST* 13.07 (2018), P07017. DOI: 10.1088/1748-0221/13/07/P07017.

[36] V. Cindro et al. "The ATLAS beam conditions monitor". In: *JINST* 3 (2008), P02004. DOI: 10.1088/1748-0221/3/02/P02004.

[37] *The ATLAS Collaboration Luminosity Public Results Run 2*. https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2. 2019.

[38] Georges Aad et al. "Improved luminosity determination in pp collisions at sqrt(s) = 7 TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J. C* 73.8 (2013), p. 2518. DOI: 10.1140/epjc/s10052-013-2518-3. arXiv: 1302.4393 [hep-ex].

[39] Morad Aaboud et al. "Luminosity determination in pp collisions at $\sqrt{s}$ = 8 TeV using the ATLAS detector at the LHC". In: *Eur. Phys. J. C* 76.12 (2016), p. 653. DOI: 10.1140/epjc/s10052-016-4466-1. arXiv: 1608.03953 [hep-ex].

[40] Burkhard Schmidt. "The High-Luminosity upgrade of the LHC: Physics and Technology Challenges for the Accelerator and the Experiments". In: *Journal of Physics: Conference Series* 706 (Apr. 2016), p. 022002. DOI: 10.1088/1742-6596/706/2/022002. URL: https://doi.org/10.1088/1742-6596/706/2/022002.

[41] G. Aad et al. "The ATLAS Inner Detector commissioning and calibration". In: *Eur. Phys. J. C* 70 (2010), pp. 787–821. DOI: 10.1140/epjc/s10052-010-1366-7. arXiv: 1004.5293 [physics.ins-det].

[42] *Track Reconstruction Performance of the ATLAS Inner Detector at* $\sqrt{s} = 13$ *TeV*. Tech. rep. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-018. Geneva: CERN, July 2015. URL: https://cds.cern.ch/record/2037683.

[43] M. Capeans et al. "ATLAS Insertable B-Layer Technical Design Report". In: (Sept. 2010).

[44] A. Ahmad et al. "The Silicon microstrip sensors of the ATLAS semiconductor tracker". In: *Nucl. Instrum. Meth. A* 578 (2007), pp. 98–118. DOI: 10.1016/j.nima.2007.04.157.

[45] T. Akesson et al. "Status of design and construction of the Transition Radiation Tracker (TRT) for the ATLAS experiment at the LHC". In: *Nucl. Instrum. Meth. A* 522 (2004). Ed. by C. Favuzzi et al., pp. 131–145. DOI: 10.1016/j.nima.2004.01.033.

[46] "ATLAS liquid argon calorimeter: Technical design report". In: (Dec. 1996).

[47] *ATLAS muon spectrometer: Technical Design Report*. Technical Design Report ATLAS. Geneva: CERN, 1997. URL: https://cds.cern.ch/record/331068.

[48] F. Bauer et al. "Construction and Test of MDT Chambers for the ATLAS Muon Spectrometer". In: *Nucl. Instrum. Meth. A* 461 (2001). Ed. by G. Batignani et al., pp. 17–20. DOI: 10.1016/S0168-9002(00)01156-6. arXiv: 1604.02000 [physics.ins-det].

[49] Theodoros Argyropoulos et al. "Cathode strip chambers in ATLAS: Installation, commissioning and in situ performance". In: *IEEE Trans. Nucl. Sci.* 56 (2009), pp. 1568–1574. DOI: 10.1109/TNS.2009.2020861.

[50] G. Aielli et al. "The RPC first level muon trigger in the barrel of the ATLAS experiment". In: *Nucl. Phys. B Proc. Suppl.* 158 (2006). Ed. by S. K. Park, S. P. Ratti, and R. Santonico, pp. 11–15. DOI: 10.1016/j.nuclphysbps.2006.07.031.

[51] S. Majewski et al. "A THIN MULTIWIRE CHAMBER OPERATING IN THE HIGH MULTIPLICATION MODE". In: *Nucl. Instrum. Meth.* 217 (1983), pp. 265–271. DOI: 10.1016/0167-5087(83)90146-1.

[52] Y Arai et al. “ATLAS Muon Drift Tube Electronics”. In: *JINST* 3 (2008). This is a backup for the ATLAS Detector paper., P09001. DOI: 10.1088/1748-0221/3/09/P09001. URL: https://cds.cern.ch/record/1100199.

[53] Paolo Branchini et al. “Global Time Fit for Tracking in an Array of Drift Cells: The Drift Tubes of the ATLAS Experiment”. In: *Nuclear Science, IEEE Transactions on* 55 (Mar. 2008), pp. 620–627. DOI: 10.1109/TNS.2007.914020.

[54] S Aefsky. *Alignment of the Muon Spectrometer in ATLAS*. Tech. rep. ATL-MUON-PROC-2011-003. Geneva: CERN, Sept. 2011. DOI: 10.1016/j.phpro.2012.02.355. URL: https://cds.cern.ch/record/1380912.

[55] Giulio Aielli. “Cavern background measurement with the ATLAS RPC system”. In: *PoS* RPC2012 (2012), p. 025. DOI: 10.22323/1.159.0025.

[56] A. Yamamoto et al. “The ATLAS central solenoid”. In: *Nucl. Instrum. Meth. A* 584 (2008), pp. 53–74. DOI: 10.1016/j.nima.2007.09.047.

[57] “ATLAS barrel toroid: Technical design report”. In: (Apr. 1997).

[58] “ATLAS endcap toroids: Technical design report”. In: (Apr. 1997).

[59] William Panduro Vazquez. “The ATLAS Data Acquisition system in LHC Run 2”. In: *J. Phys. Conf. Ser.* 898.3 (2017). Ed. by Richard Mount and Craig Tull, p. 032017. DOI: 10.1088/1742-6596/898/3/032017.

[60] Morad Aaboud et al. “Performance of the ATLAS Trigger System in 2015”. In: *Eur. Phys. J. C* 77.5 (2017), p. 317. DOI: 10.1140/epjc/s10052-017-4852-3. arXiv: 1611.09661 [hep-ex].

[61] W Buttinger. *The ATLAS Level-1 Trigger System*. Tech. rep. ATL-DAQ-PROC-2012-024. Geneva: CERN, June 2012. DOI: 10.1088/1742-6596/396/1/012010. URL: https://cds.cern.ch/record/1456546.

[62] “ATLAS high-level trigger, data acquisition and controls: Technical design report”. In: (July 2003).

[63] Georges Aad et al. “ATLAS data quality operations and performance for 2015–2018 data-taking”. In: *JINST* 15.04 (2020), P04003. DOI: 10.1088/1748-0221/15/04/P04003. arXiv: 1911.04632 [physics.ins-det].

[64] S. Agostinelli et al. “GEANT4–a simulation toolkit”. In: *Nucl. Instrum. Meth. A* 506 (2003), pp. 250–303. DOI: 10.1016/S0168-9002(03)01368-8.

[65] ATLAS Collaboration. *The ATLAS Collaboration Software and Firmware*. ATL-SOFT-PUB-2021-001. 2021. URL: https://cds.cern.ch/record/2767187.

[66] K. Bos et al. *LHC computing Grid: Technical Design Report. Version 1.06 (20 Jun 2005)*. Technical design report. LCG. Geneva: CERN, 2005. URL: http://cds.cern.ch/record/840543.

[67] Georges Aad, Abajyan, et al. *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System*. Tech. rep. Final version presented to December 2013 LHCC. Sept. 2013. URL: https://cds.cern.ch/record/1602235.

[68] Weihao Wu. “FELIX: the New Detector Interface for the ATLAS Experiment”. In: *IEEE Trans. Nucl. Sci.* 66 (June 2018). This manuscript for conference record of 21st IEEE Real Time conference only, 986–992. 7 p. DOI: 10.1109/TNS.2019.2913617. arXiv: 1806.10667. URL: https://cds.cern.ch/record/2627658.

[69] M (CERN) Aleksa et al. *ATLAS Liquid Argon Calorimeter Phase-I Upgrade: Technical Design Report*. Tech. rep. Final version presented to December 2013 LHCC. Sept. 2013. URL: https://cds.cern.ch/record/1602230.

[70] Bernd Stelzer. *The New Small Wheel Upgrade Project of the ATLAS Experiment*. Tech. rep. Geneva: CERN, Oct. 2014. DOI: 10.1016/j.nuclphysbps.2015.09.182. URL: https://cds.cern.ch/record/1958265.

[71] Lorenzo Massa. “The BIS78 Resistive Plate Chambers upgrade of the ATLAS Muon Spectrometer for the LHC Run-3”. In: *JINST* 15.10 (2020). Ed. by Barbara Liberti, Alessandro Paoloni, and Paolo Camarri, p. C10026. DOI: 10.1088/1748-0221/15/10/C10026. arXiv: 2004.12693 [physics.ins-det].

[72] *Technical Design Report for the Phase-II Upgrade of the ATLAS TDAQ System*. Tech. rep. Geneva: CERN, Sept. 2017. URL: https://cds.cern.ch/record/2285584.

[73] Y. Arai et al. “ATLAS muon drift tube electronics”. In: *JINST* 3 (2008), P09001. DOI: 10.1088/1748-0221/3/09/P09001.

[74] Soo Ryu and. “FELIX: The new detector readout system for the ATLAS experiment”. In: *Journal of Physics: Conference Series* 898 (Oct. 2017), p. 032057. DOI: 10.1088/1742-6596/898/3/032057. URL: https://doi.org/10.1088/1742-6596/898/3/032057.

[75] Y. Arai and J. Christiansen. “Requirements and Specifications of the TDC for the ATLAS Precision Muon Tracker”. In: (May 1997).

[76] J Chapman et al. “On-chamber readout system for the ATLAS MDT Muon Spectrometer”. In: *IEEE Trans. Nucl. Sci.* 51 (2004), pp. 2196–2200. DOI: 10.1109/TNS.2004.834701. URL: https://cds.cern.ch/record/818333.

[77] H. Boterenbrood et al. “The read-out driver for the ATLAS MDT muon precision chambers”. In: *IEEE Trans. Nucl. Sci.* 53 (2006). Ed. by J. P. Dufey, pp. 741–748. DOI: 10.1109/TNS.2006.874307.

[78] T. A. M. Wijnen. “The MROD data format and the tower partitioning of the MDT chambers.” In: (2003).

[79] C P Bee et al. *The raw event format in the ATLAS Trigger and DAQ*. Tech. rep. Revised version number 5 submitted on 2016-11-03 11:47. Geneva: CERN, Feb. 2016. URL: https://cds.cern.ch/record/683741.

[80] Andrea Borga et al. “FELIX based readout of the Single-Phase ProtoDUNE detector”. In: *EPJ Web Conf.* 214 (2019), 01013. 8 p. DOI: 10.1051/epjconf/201921401013. URL: https://cds.cern.ch/record/2702140.

[81] P. Moreira et al. “The GBT Project”. In: *Topical Workshop on Electronics for Particle Physics*. CERN, 2009. DOI: 10.5170/CERN-2009-006.342.

[82] Serguei Kolos, William P. Vazquez, and Gordon Crone. *New Software-Based Readout Driver for the ATLAS Experiment*. Tech. rep. Geneva: CERN, June 2021. DOI: 10.1109/TNS.2021.3083987. arXiv: 2010.14884. URL: https://cds.cern.ch/record/2742929.

[83] *Diagram of particle paths in the detector*. https://cds.cern.ch/images/CERN-EX-1301009-01. Accessed: 2022-10-07.

[84] *ATLAS Tracking Software Tutorial*. https://atlassoftwaredocs.web.cern.ch/trackingTutorial/idoverview/. Accessed: 2022-10-07.

[85] *Performance of the reconstruction of large impact parameter tracks in the ATLAS inner detector*. Tech. rep. Geneva: CERN, July 2017. URL: https://cds.cern.ch/record/2275635.

[86] R. Frühwirth. "Application of Kalman filtering to track and vertex fitting". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 262.2 (1987), pp. 444–450. ISSN: 0168-9002. DOI: https://doi.org/10.1016/0168-9002(87)90887-4. URL: https://www.sciencedirect.com/science/article/pii/0168900287908874.

[87] *Early Inner Detector Tracking Performance in the 2015 data at $\sqrt{s}$ = 13 TeV*. Tech. rep. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-051. Geneva: CERN, 2015. URL: https://cds.cern.ch/record/2110140.

[88] Federico Meloni. *Primary vertex reconstruction with the ATLAS detector*. Tech. rep. Geneva: CERN, Oct. 2016. DOI: 10.1088/1748-0221/11/12/C12060. URL: https://cds.cern.ch/record/2222390.

[89] R Frühwirth, Wolfgang Waltenberger, and Pascal Vanlaer. *Adaptive Vertex Fitting*. Tech. rep. Geneva: CERN, Mar. 2007. URL: https://cds.cern.ch/record/1027031.

[90] *Vertex Reconstruction Performance of the ATLAS Detector at $\sqrt{s} = 13\ TeV$*. Tech. rep. Geneva: CERN, July 2015. URL: https://cds.cern.ch/record/2037717.

[91] W Lampl et al. *Calorimeter Clustering Algorithms: Description and Performance*. Tech. rep. Geneva: CERN, Apr. 2008. URL: https://cds.cern.ch/record/1099735.

[92] Morad Aaboud et al. "Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton-proton collision data at $\sqrt{s}$ = 13 TeV". In: *Eur. Phys. J. C* 79.8 (2019), p. 639. DOI: 10.1140/epjc/s10052-019-7140-6. arXiv: 1902.04655 [physics.ins-det].

[93] *Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach*. Tech. rep. Geneva: CERN, Dec. 2017. URL: https://cds.cern.ch/record/2298955.

[94] Georges Aad et al. "Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton-proton collision data". In: *JINST* 14.12 (2019), P12006. DOI: 10.1088/1748-0221/14/12/P12006. arXiv: 1908.00005 [hep-ex].

[95] Morad Aaboud et al. "Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton-proton collision data". In: *JINST* 14.03 (2019), P03017. DOI: 10.1088/1748-0221/14/03/P03017. arXiv: 1812.03848 [hep-ex].

[96] Morad Aaboud et al. "Measurement of the photon identification efficiencies with the ATLAS detector using LHC Run 2 data collected in 2015 and 2016". In: *Eur. Phys. J. C* 79.3 (2019), p. 205. DOI: 10.1140/epjc/s10052-019-6650-6. arXiv: 1810.05087 [hep-ex].

[97] Morad Aaboud et al. "Jet reconstruction and performance using particle flow with the ATLAS Detector". In: *Eur. Phys. J. C* 77.7 (2017), p. 466. DOI: 10.1140/epjc/s10052-017-5031-2. arXiv: 1703.10485 [hep-ex].

[98] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. "The anti-k_t jet clustering algorithm". In: *JHEP* 04 (2008), p. 063. DOI: 10.1088/1126-6708/2008/04/063. arXiv: 0802.1189 [hep-ph].

[99] ATLAS Collaboration. "Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment". In: *Eur. Phys. J. C* 76 (2016), p. 295. DOI: 10.1140/epjc/s10052-016-4110-0. arXiv: 1512.05955 [hep-ex].

[100] *Reconstruction, Identification, and Calibration of hadronically decaying tau leptons with the ATLAS detector for the LHC Run 3 and reprocessed Run 2 data*. Tech. rep. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2022-044. Geneva: CERN, 2022. URL: https://cds.cern.ch/record/2827111.

[101] *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV*. Tech. rep. Geneva: CERN, May 2017. URL: https://cds.cern.ch/record/2261772.

[102] ATLAS Collaboration. *Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment*. ATL-PHYS-PUB-2019-033. 2019. URL: https://cds.cern.ch/record/2688062.

[103] *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*. Tech. rep. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2015-045. Geneva: CERN, 2015. URL: https://cds.cern.ch/record/2064383.

[104] Georges Aad et al. "Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at $\sqrt{s} = 13$ TeV". In: *Eur. Phys. J. C* 81.7 (2021), p. 578. DOI: 10.1140/epjc/s10052-021-09233-2. arXiv: 2012.00578 [hep-ex].

[105] Georges Aad et al. "Muon reconstruction performance of the ATLAS detector in proton–proton collision data at $\sqrt{s}$ =13 TeV". In: *Eur. Phys. J. C* 76.5 (2016), p. 292. DOI: 10.1140/epjc/s10052-016-4120-y. arXiv: 1603.05598 [hep-ex].

[106] Morad Aaboud et al. “Performance of missing transverse momentum reconstruction with the ATLAS detector using proton-proton collisions at $\sqrt{s}$ = 13 TeV”. In: *Eur. Phys. J. C* 78.11 (2018), p. 903. DOI: 10.1140/epjc/s10052-018-6288-9. arXiv: 1802.08168 [hep-ex].

[107] Hung-Liang Lai et al. “New parton distributions for collider physics”. In: *Phys. Rev. D* 82 (2010), p. 074024. DOI: 10.1103/PhysRevD.82.074024. arXiv: 1007.2241 [hep-ph].

[108] Carlo Oleari. “The POWHEG-BOX”. In: *Nucl. Phys. B Proc. Suppl.* 205-206 (2010). Ed. by Johannes Blümlein, Sven-Olaf Moch, and Tord Riemann, pp. 36–41. DOI: 10.1016/j.nuclphysbps.2010.08.016. arXiv: 1007.3893 [hep-ph].

[109] Torbjörn Sjöstrand et al. “An introduction to PYTHIA 8.2”. In: *Comput. Phys. Commun.* 191 (2015), pp. 159–177. DOI: 10.1016/j.cpc.2015.01.024. arXiv: 1410.3012 [hep-ph].

[110] Wu-Ki Tung. “New generation of parton distributions with uncertainties from global QCD analysis”. In: *Acta Phys. Polon. B* 33 (2002). Ed. by J. Tran Thanh Van, pp. 2933–2938. arXiv: hep-ph/0206114.

[111] Georges Aad et al. “Measurement of the Z/γ^* boson transverse momentum distribution in pp collisions at $\sqrt{s}$ = 7 TeV with the ATLAS detector”. In: *JHEP* 09 (2014), p. 145. DOI: 10.1007/JHEP09(2014)145. arXiv: 1406.3660 [hep-ex].

[112] John C. Collins and Davison E. Soper. “Angular Distribution of Dileptons in High-Energy Hadron Collisions”. In: *Phys. Rev. D* 16 (1977), p. 2219. DOI: 10.1103/PhysRevD.16.2219.

[113] Pietro Faccioli and Carlos Lourenço. *Particle Polarization in High Energy Physics*. Springer Cham, 2022. DOI: 10.1007/978-3-031-08876-6. URL: https://link.springer.com/book/10.1007/978-3-031-08876-6.

[114] Georges Aad et al. “Measurement of the angular coefficients in Z-boson events using electron and muon pairs from data taken at $\sqrt{s} = 8$ TeV with the ATLAS detector”. In: *JHEP* 08 (2016), p. 159. DOI: 10.1007/JHEP08(2016)159. arXiv: 1606.00689 [hep-ex].

[115] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez. “The Catchment Area of Jets”. In: *JHEP* 04 (2008), p. 005. DOI: 10.1088/1126-6708/2008/04/005. arXiv: 0802.1188 [hep-ph].

[116] Paul Knysh and Yannis Korkolis. “Blackbox: A procedure for parallel optimization of expensive black-box functions”. In: *CoRR* abs/1605.00998 (2016). arXiv: 1605.00998. URL: http://arxiv.org/abs/1605.00998.

[117] William Buttinger. *Using Event Weights to account for differences in Instantaneous Luminosity and Trigger Prescale in Monte Carlo and Data*. Tech. rep. Geneva: CERN, 2015. URL: https://cds.cern.ch/record/2014726.

[118] Glen Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *Eur. Phys. J. C* 71 (2011). [Erratum: Eur.Phys.J.C 73, 2501 (2013)], p. 1554. DOI: 10.1140/epjc/s10052-011-1554-0. arXiv: 1007.1727 [physics.data-an].

[119] Alexander L. Read. “Presentation of search results: The CL(s) technique”. In: *J. Phys. G* 28 (2002). Ed. by M. R. Whalley and L. Lyons, pp. 2693–2704. DOI: 10.1088/0954-3899/28/10/313.

[120] Glen Cowan et al. “Asymptotic formulae for likelihood-based tests of new physics”. In: *The European Physical Journal C* 71.2 (Feb. 2011). ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-011-1554-0. URL: http://dx.doi.org/10.1140/epjc/s10052-011-1554-0.

[121] Max Baak et al. “Interpolation between multi-dimensional histograms using a new non-linear moment morphing method”. In: *Nucl. Instrum. Meth. A* 771 (2015), pp. 39–48. DOI: 10.1016/j.nima.2014.10.033. arXiv: 1410.7388 [physics.data-an].

[122] David W. Scott and Simon J. Sheather. “Kernel density estimation with binned data”. In: *Communications in Statistics - Theory and Methods* 14.6 (1985), pp. 1353–1359. DOI: 10.1080/03610928508828980. eprint: https://www.tandfonline.com/doi/pdf/10.1080/03610928508828980. URL: https://www.tandfonline.com/doi/abs/10.1080/03610928508828980.

[123] ATLAS Collaboration. *The ATLAS Tau Trigger in Run 2*. ATLAS-CONF-2017-061. 2017. URL: https://cds.cern.ch/record/2274201.

[124] Charu C. Aggarwal. *Neural Networks and Deep Learning*. Cham: Springer, 2018, p. 497. ISBN: 978-3-319-94462-3. DOI: 10.1007/978-3-319-94463-0.

[125] Vinod Nair and Geoffrey Hinton. “Rectified Linear Units Improve Restricted Boltzmann Machines Vinod Nair”. In: vol. 27. June 2010, pp. 807–814.

[126] David E. Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams. “Learning representations by back-propagating errors”. In: *Nature* 323 (1986), pp. 533–536.

[127] Shai Shalev-Shwartz and Shai Ben-David. *Understanding Machine Learning: From Theory to Algorithms*. Cambridge University Press, 2014. DOI: 10.1017/CBO9781107298019.015.

[128] Diederik P. Kingma and Jimmy Ba. *Adam: A Method for Stochastic Optimization*. 2017. arXiv: 1412.6980 [cs.LG].

[129] Nitish Srivastava et al. “Dropout: A Simple Way to Prevent Neural Networks from Overfitting”. In: *Journal of Machine Learning Research* 15.56 (2014), pp. 1929–1958. URL: http://jmlr.org/papers/v15/srivastava14a.html.

[130] Guido Van Rossum and Fred L. Drake. *Python 3 Reference Manual*. Scotts Valley, CA: CreateSpace, 2009. ISBN: 1441412697.

[131] Charles R. Harris et al. “Array programming with NumPy”. In: *Nature* 585.7825 (Sept. 2020), pp. 357–362. DOI: 10.1038/s41586-020-2649-2. URL: https://doi.org/10.1038/s41586-020-2649-2.

[132] Eric Jones, Travis Oliphant, and Pearu Peterson. “SciPy: Open Source Scientific Tools for Python”. In: (Jan. 2001).

[133] Philippe Canal Rene Brun Fons Rademakers et al. *root-project/root: v6.18/02*. Version v6-18-02. June 2020. DOI: 10.5281/zenodo.3895860. URL: https://doi.org/10.5281/zenodo.3895860.

[134] Martin Abadi et al. "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems". In: *CoRR* abs/1603.04467 (2016). arXiv: 1603.04467. URL: http://arxiv.org/abs/1603.04467.

[135] Adam Paszke et al. "PyTorch: An Imperative Style, High-Performance Deep Learning Library". In: *CoRR* abs/1912.01703 (2019). arXiv: 1912.01703. URL: http://arxiv.org/abs/1912.01703.

[136] François Chollet et al. *Keras*. https://github.com/fchollet/keras. 2015.

[137] Dan Guest, Kyle Cranmer, and Daniel Whiteson. "Deep Learning and its Application to LHC Physics". In: *Ann. Rev. Nucl. Part. Sci.* 68 (2018), pp. 161–181. DOI: 10.1146/annurev-nucl-101917-021019. arXiv: 1806.11484 [hep-ex].

[138] "Machine learning at the energy and intensity frontiers of particle physics". In: *Nature* 560 (2018), pp. 41–48. DOI: 10.1038/s41586-018-0361-2. URL: http://cds.cern.ch/record/2647378.

[139] Johan Alwall et al. "MadGraph 5 : Going Beyond". In: *JHEP* 06 (2011), p. 128. DOI: 10.1007/JHEP06(2011)128. arXiv: 1106.0522 [hep-ph].

[140] J. de Favereau et al. "DELPHES 3, A modular framework for fast simulation of a generic collider experiment". In: *JHEP* 02 (2014), p. 057. DOI: 10.1007/JHEP02(2014)057. arXiv: 1307.6346 [hep-ex].

[141] Nicolo de Groot and Sergi Castells. "Identifying hadronic charmonium decays in hadron colliders". In: *SciPost Phys. Core* 2 (2 2020), p. 8. DOI: 10.21468/SciPostPhysCore.2.2.008. URL: https://scipost.org/10.21468/SciPostPhysCore.2.2.008.

[142] R. D. Field and R. P. Feynman. "A Parametrization of the Properties of Quark Jets". In: *Nucl. Phys. B* 136 (1978). Ed. by L. M. Brown, p. 1. DOI: 10.1016/0550-3213(78)90015-9.

[143] Andrew J. Larkoski, Jesse Thaler, and Wouter J. Waalewijn. "Gaining (Mutual) Information about Quark/Gluon Discrimination". In: *JHEP* 11 (2014), p. 129. DOI: 10.1007/JHEP11(2014)129. arXiv: 1408.3122 [hep-ph].

[144] Jesse Thaler and Ken Van Tilburg. "Identifying Boosted Objects with N-subjettiness". In: *JHEP* 03 (2011), p. 015. DOI: 10.1007/JHEP03(2011)015. arXiv: 1011.2268 [hep-ph].

[145] *ATLAS Pythia 8 tunes to 7 TeV data*. Tech. rep. All figures including auxiliary figures are available at https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PUBNOTES/ATL-PHYS-PUB-2014-021. Geneva: CERN, 2014. URL: https://cds.cern.ch/record/1966419.

[146] "Constituent-level pile-up mitigation techniques in ATLAS". In: (Aug. 2017).

[147] Thomas Junk. "Confidence level computation for combining searches with small statistics". In: *Nucl. Instrum. Meth. A* 434 (1999), pp. 435–443. DOI: 10.1016/S0168-9002(99)00498-2. arXiv: arXiv:hep-ex/9902006.

[148] M. Aaboud et al. "Search for exclusive Higgs and Z boson decays to $\phi\gamma$ and $\rho\gamma$ with the ATLAS detector". In: *JHEP* 07 (2018), p. 127. DOI: 10.1007/JHEP07(2018)127. arXiv: arXiv:1712.02758 [hep-ex].

[149] ATLAS Collaboration. *Athena*. Version 21.2.79. July 2020. DOI: 10.5281/zenodo.3933810. URL: https://doi.org/10.5281/zenodo.3933810.

Acknowledgement

I've tried my best to show in this thesis how much I've learned during my PhD journey, however it would not be complete without saying thank you for the people, who helped me through it.

I am forever grateful for my supervisors, Nenad Vranješ and Nicolo de Groot for supporting me in the last many years. Nicolo, you always had an open door for me and patience to answer all my questions during my first years in the Netherlands. Also, thank you for your great ideas and insights any time I've felt stuck with my research. Nenad, thank you for your support during my years in Serbia, for always checking on me, helping wherever you can, and making sure that I always progress. I also want to thank you for making it possible to pursue my carrier outside of academia in my final years of PhD. For both of you, thank you for the reading of these chapters, for the input and feedback and for helping me navigate through the endless paperwork of my joint PhD.

I would also like to thank Carlo for helping me through one of my toughest and busiest year at CERN. Without you I would have been lost and clueless. You thought me many things, not just about hardware and software, but also helped me shape my presenting and writing skills, which are of great use nowadays too. Thank you for all your patience and support but also for the outside-of-work conversations and memorable moments.

I am greatly honored to have Maja Burić, Duško Lataš, Marija Vranješ Milosavljević, Wim Beenakker, Antonio Pellegrino, Mengqing Wu and Gerco Onderwater forming the thesis committee and I would like to thank you for reviewing this manuscript. A particular thank you to Antonio for being my C3 member and making sure things were coming along smoothly throughout the PhD.

I would like to thank all my colleagues at Radboud, Nikhef, IPB and CERN for being part of my journey. I've received a great amount of help and made many unforgettable memories. I won't forget the PhD dinners, the coffee breaks and the fruitful discussions I've had with all of you.

A huge thank you for all my friends outside of my physics career. For the people once lived or connected to the Marialaan, thank you for accepting me as part of your bunch, for helping me learning about the Dutch culture and being patient with me while I was endlessly ranting after a hard day. For my friends back in Serbia, thank you for showing the other aspects of life and for always being there for a glass of pálinka. All my colleagues at Alliander, thank you for being interested in my PhD topic and supporting me through it in the last couple years.

Anya, Apa, szavakkal nem tudom kifejezni mennyire hálás vagyok, hogy támogattatok az elmúlt években. Hogy minden őrült ötletem ellenére mellettem áltatok, és végelen segítséget nyújtottatok. Köszönöm, hogy engedtétek, hogy kibontakozzam és olyan karriert választjak, amiben igazán megtaláltam önmagam. Nélkületek ez nem sikerült volna. Ervin, köszönöm, hogy példaképet mutattál, mint nagy testvér, és hogy bármikor számíthattam rád.

And last but not least, Thijs, thank you for being my biggest support not just in my PhD but in my life. I could not choose a better partner and I am looking forward to our next chapter together with our (now) family.

And, in case I've missed you, who made it this far (or reading only this page), thank you for making my PhD unforgettable.

Biography

Evelin was born on June 1, 1994, in Senta, Serbia. She graduated from the ”Svetozar Marković” gymnasium in Subotica, and in 2013 she enrolled in physics at the Faculty of Science in Novi Sad. She graduated in 2017 from the research major with an average grade of 9.47. In 2018, she completed her master's studies in nuclear physics with an average grade of 10.00, defending her master's thesis, done in cooperation with the ATLAS group at the Institute of Physics in Belgrade, entitled *Possibilities of the ATLAS experiment for the detection of the triple production of W bosons*. The results obtained during the preparation of the master thesis are an integral part of a monograph from *CERN Yellow Reports* CERN-2019-007 arXiv:1902.10229.

In 2018, she enrolled in doctoral studies at the Faculty of Physics of the University of Belgrade and Radboud University in Nijmegen, the Netherlands, as a joint PhD student. She passed the scheduled exams for doctoral studies at the Faculty of Physics with an average grade of 10.00. Since October 2018, she has been elected to the position of researcher intern at the Institute of Physics in Belgrade. The topic of her research was the *Search for the exclusive W boson hadronic decay of $W \to \rho\gamma$*. In 2019, Evelin was engaged in performing calculation exercises for Radboud University students in the subject *Standard Model and beyond*. Since 2019, he has been a qualified author of the ATLAS collaboration, having completed his qualification task on the *MROD-FELIX read out system integration*. She presented the preliminary results of this project at the ATLAS week at CERN in 2020 in the form of a poster. During her doctoral studies, Evelin attended two international schools: the *CERN Summer School* in Geneva during the summer of 2018 and the *BND School* in Spa, Belgium, in 2019.

Изјава о ауторству

Име и презиме аутора: **Евелин Бакош**

Број индекса: **2018/8003**

Изјављујем

да је докторска дисертација под насловом

Radiative W boson decay studies and the upgrade of the ATLAS muon spectrometer readout system
(Изучавање радијативних распада W бозона и унапређење система за очитавање мионског спектометра детектора АТЛАС)

- резултат сопственог истраживачког рада;
- да дисертација у целини ни у деловима није била предложена за стицање друге дипломе према студијским програмима других високошколских установа;
- да су резултати коректно наведени и
- да нисам кршио/ла ауторска права и користио/ла интелектуалну својину других лица.

Потпис аутора

У Београду, **1. 3. 2024.**

Bakoš

Изјава о истоветности штампане и електронске верзије докторског рада

Име и презиме аутора: **Евелин Бакош**

Број индекса: **2018/8003**

Студијски програм: **Физика - Физика високих енергија и нуклеарна физика**

Наслов рада: **Radiative W boson decay studies and the upgrade of the ATLAS muon spectrometer readout system (Изучавање радијативних распада W бозона и унапређење система за очитавање мионског спектометра детектора АТЛАС)**

Ментор: **др Ненад Врањеш**

Изјављујем да је штампана верзија мог докторског рада истоветна електронској верзији коју сам предао/ла ради похрањивања у **Дигиталном репозиторијуму Универзитета у Београду.**

Дозвољавам да се објаве моји лични подаци везани за добијање академског назива доктора наука, као што су име и презиме, година и место рођења и датум одбране рада.

Ови лични подаци могу се објавити на мрежним страницама дигиталне библиотеке, у електронском каталогу и у публикацијама Универзитета у Београду.

Потпис аутора

У Београду, **1. 3. 2024.**

Bakoš

Изјава о коришћењу

Овлашћујем Универзитетску библиотеку „Светозар Марковић“ да у Дигитални репозиторијум Универзитета у Београду унесе моју докторску дисертацију под насловом:

Radiative W boson decay studies and the upgrade of the ATLAS muon spectrometer readout system
(Изучавање радијативних распада W бозона и унапређење система за очитавање мионског спектометра детектора АТЛАС)

која је моје ауторско дело.

Дисертацију са свим прилозима предао/ла сам у електронском формату погодном за трајно архивирање.

Моју докторску дисертацију похрањену у Дигиталном репозиторијуму Универзитета у Београду и доступну у отвореном приступу могу да користе сви који поштују одредбе садржане у одабраном типу лиценце Креативне заједнице (Creative Commons) за коју сам се одлучио/ла.

1. Ауторство (CC BY)

2. Ауторство – некомерцијално (CC BY-NC)

3. Ауторство – некомерцијално – без прерада (CC BY-NC-ND)

4. Ауторство – некомерцијално – делити под истим условима (CC BY-NC-SA)

5. Ауторство – без прерада (CC BY-ND)

6. Ауторство – делити под истим условима (CC BY-SA)

(Молимо да заокружите само једну од шест понуђених лиценци.
Кратак опис лиценци је саставни део ове изјаве).

Потпис аутора

У Београду, **1. 3. 2024.**

1. **Ауторство.** Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце, чак и у комерцијалне сврхе. Ово је најслободнија од свих лиценци.

2. **Ауторство – некомерцијално**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела.

3. **Ауторство – некомерцијално – без прерада**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца не дозвољава комерцијалну употребу дела. У односу на све остале лиценце, овом лиценцом се ограничава највећи обим права коришћења дела.

4. **Ауторство – некомерцијално – делити под истим условима**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца не дозвољава комерцијалну употребу дела и прерада.

5. **Ауторство – без прерада**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, без промена, преобликовања или употребе дела у свом делу, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце. Ова лиценца дозвољава комерцијалну употребу дела.

6. **Ауторство – делити под истим условима**. Дозвољавате умножавање, дистрибуцију и јавно саопштавање дела, и прераде, ако се наведе име аутора на начин одређен од стране аутора или даваоца лиценце и ако се прерада дистрибуира под истом или сличном лиценцом. Ова лиценца дозвољава комерцијалну употребу дела и прерада. Слична је софтверским лиценцама, односно лиценцама отвореног кода.