Bachelor's Thesis

Optimierung des Tau-Identifikationsalgorithmus und Unterdrückung des Untergrundes von Elektronen am ATLAS-Experiment

Optimization of the Electron Veto for the identification of hadronically decaying Tau leptons with the ATLAS detector

prepared by

Alejandra Renee Pillado Gonzalez

from Sonora, Mexico

at the II. Physikalischen Institut

Abstract

Current research of the ATLAS physics program involves the study of final states with hadronically decaying tau leptons, which are important in studies such as Higgs physics and beyond Standard Model (BSM) physics. Electrons contribute in a significant way to the tau background observed at the detector, therefore, the electron veto plays an important role in the reconstruction and identification process of the tau lepton. Monte Carlo generated samples from 13 TeV events were implemented to train a boosted decision tree (BDT) to discriminate between background coming from Z$\to ee$ and signal from Z$\to \tau\tau$ decays. Background rejections of 85% for signal efficiencies of 90% were achieved for the prominent η barrel region, following different values for the end-cap and forward end-cap regions. In this thesis, the BDT method, as well as the different criteria considered to optimize the electron veto are implemented and described.

Abstract

Aktuelle Forschung im ATLAS-Physikprogramm umfasst die Untersuchung der finalen Zustände von Tau-Leptonen, die wichtig für die Forschung über der Physik des Higgs-bosons und jenseits von der Physik des Standard-Modells ist. Elektronen tragen signifikant zum Untergrund des Tau-Leptons bei, der im Detektor beobachtet wird. Deswegen spielt das Elektron-Veto eine Role im Prozess der Rekonstruktion und Identifikation der Tau-Leptonen. 13 TeV Monte-Carlo erzeugte Proben wurden in einem diskriminieren-den Algorithmus benutzt, in dem Z$\to ee$ und Z$\to \tau\tau$ die entsprechende Ereignisse des Hintergrunds bzw. Signal sind. Die Untergrundunterdrückung wurde für eine 90% Signaleffizienz einen Prozentsatz von 85% in der η "Barrel" Region erreicht, zugleich wurden unterschiedliche Werte für die andere Regionen erlangt. In dieser Arbeit wurden sowohl die BDT-Methode als auch die verschiedene Kriterien für die Elektron-Veto Optimierung eingeführt und beschrieben.

Contents

Nomenclature

Latin Letters

Variable	Meaning	Units
c	Velocity of light	$\frac{m}{s}$
E, E_T	Energy/Transverse energy	GeV
$\not{E}_T$	Missing transverse energy	GeV
m, m_T	Mass/Transverse mass	$\frac{GeV}{c^2}$
P, P_T	Momentum/Transverse momentum	GeV

Greek Letters

Variable	Meaning	Units
η	Pseudorapidity	–
γ	Photon	–
λ_f	Yukawa coupling	–
μ	Muon	–
$\pi^{\pm}$	Pion	–
ϕ	Azimuthal angle	; –
σ	Standard deviation	–
τ, τ_{had}, τ_{had}^{vis}	Tau lepton/Hadronically decaying tau/Visible hadronically decaying tau	–
θ	Polar angle	; –
ν	Neutrino	–

Indices

Index	Meaning
$e^{\pm}$	Electrons
g	Gluon
H	Higgs boson
l	Lepton
ΔR	Separation between particles in terms of η and ϕ
$\sqrt{s}$	Center of mass energy
$W^{\pm}$	W bosons
Z^0	Z boson

Abbreviations

Abbreviation	Meaning
Atlas	A Toroidal LHC ApparatuS
Alice	A Large Ion Collider Experiment
BDT	Boosted Decision Tree
BEH	Brout-Englert-Higgs
BSM	Beyond Standard Model
Cern	European Organization for Nuclear Research
Cms	Compact Muon Solenoid
EM	Electromagnetic
Fermilab	Fermi National Accelerator Laboratory
KS	Kolmogorov-Smirmov
Lhc	Large Hadron Collider
Lhcb	Large Hadron Collider beauty experiment
QCD	Quantum chromodynamics
QED	Quantum electrodynamics
SM	Standard Model
Slac	Stanford Linear Accelerator Center
TMVA	Toolkit for Multivariate Data Analysis
TRT	Transition Radiation Tracker
VEV	Vacuum Expectation Value

1. Introduction

The purpose of science is to understand how the Universe works. However, the phenomena in the Universe range from the smallest to the greatest scales ever imagined. The description of its smallest structures may help us in understanding the universe as a whole, and physicists have worked on possible theories for many centuries.

Today one of the most successful theories of the physics of elementary particles is the Standard Model (SM), which describes a zoo of particles and their interactions. The different particle accelerators around the world have put this theory to test with the discovery of many particles and gauge bosons predicted by the SM over the last 50 years. The SM, however, has its limitations, and this has opened the field to more theories and possible answers to questions which the SM cannot answer.

New energetic accelerators and precise detectors allow the possibility of observations of physics beyond the SM. The discovery of the Higgs boson in 2012 started a new chapter in physics and it is still extensively studied today. Particularly its decay to tau leptons can provide information not only about the nature of the new boson, but also about the mass generation mechanism for other particles. Similarly, the tau lepton is important in the research and measurements at the current accelerators, which makes it necessary to be able to efficiently identify it and discriminate it from backgrounds of other particles. In the following thesis, a discrimination method against electrons will be presented with data simulated of 13 TeV. A Boosted Decision Tree (BDT) will be implemented for this optimization to achieve a good separation between taus and electrons.

2. Theory

2.1. Standard Model of Particle Physics

The Standard Model (SM) describes the interactions of the fundamental components of matter and is a theory that describes the strong and unified electroweak interactions. Figure 2.1 shows the particle content of the SM and how the particles are categorized. The
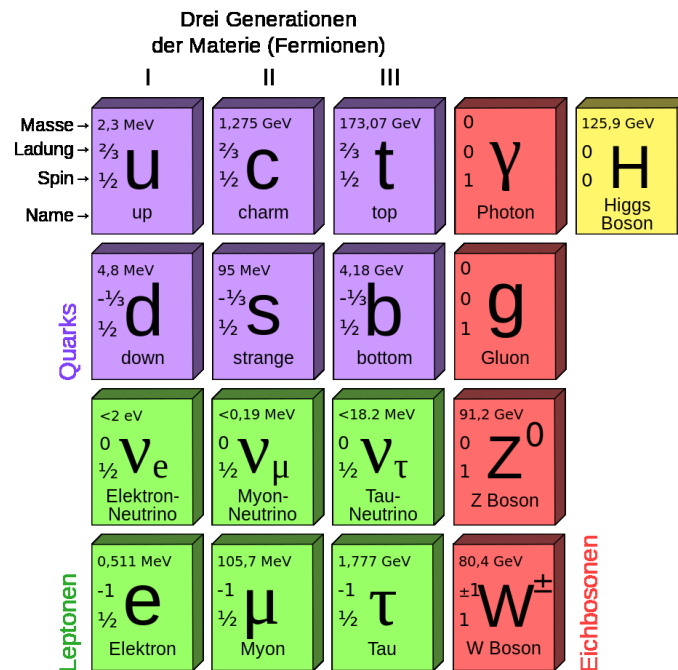

Figure 2.1.: The Standard Model's components of matter and its gauge bosons.

fermions are the fundamental particles of this theory which can be divided into quarks and leptons. The first ones are colour charged and are the components of greater particles known as hadrons, contrary to the leptons. As shown in Fig. 2.1, they are divided in three generations, being each generation more massive than the previous one. Particles composed of two quarks are called mesons and the ones composed of three are known as baryons. Together mesons and baryons are named hadrons, which can be observed at the detectors. Protons and neutrons are examples of hadrons.

2. Theory

The gauge bosons are the mediators of the interactions between these particles. The particles' interactions with the Higgs field [1] result in masses for those particles. A direct physical consequence of this field is the Higgs boson. This topic will be covered in a more detailed manner in Section 1.1.2. The properties of each boson are summarized in Table 2.1. The theory of the electromagnetic force is Quantum Electrodynamics (QED) and the theory for the strong interaction is Quantum Chromodynamics (QCD).

Boson	Charge	Mass	Interaction
Photon γ	Neutral	Massless	QED
Gluon g	Colour charged	Massless	QCD
Z^0	Neutral	91.1876±0.002 GeV [2]	Weak
$W^{\pm}$	+1/-1	80.375±0.023 GeV [3]	Weak
Higgs H	Neutral	125.09±0.24 GeV [4]	Higgs Mechanism

Table 2.1.: Basic properties of the gauge bosons.

The SM is a gauge theory, in which the electromagnetic and weak interactions are unified, thus describing two fundamental interactions: Strong and Electroweak. These are based on the gauge symmetry group of the SM:

$$\underbrace{SU(3)_C}_{Strong} \otimes \underbrace{SU(2)_L \otimes U(1)_Y}_{Electroweak}$$

The strong interaction is based on the non-Abelian symmetry gauge group SU(3), with N_c =3 being the number of colour charges (red, blue, green). There are eight gluons which are the mediators of this interaction. The gluon is coloured and can couple with itself unlike the photon (which only couples with electrically charged particles).

The electroweak model was first introduced by Glashow, Salam and Weinberg [5], which unifies the QED and weak forces and implies a direct relationship between its couplings. The spontaneous breaking of this symmetry for electroweak neutral currents is the one that gives the $W^{\pm}$ (responsible for the charged currents) and Z^0 bosons its masses while leaving the photon massless (the last two being involved in the neutral currents). This symmetry group consists of left handed fermions placed in weak isospin doublets and right

handed fermions in weak isospin singlets as shown below [5].

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L, \begin{pmatrix} u \\ d' \end{pmatrix}_L \cdots$$

$$e_R^-, \ \mu_R^-, \ \tau_R^-, \ u_R \ \cdots$$

2.2. The BEH Mechanism and the Higgs Boson

Even though many predictions of the SM are confirmed by different measurements, it is still not complete; it is known that the $W^\pm$ and Z^0 bosons have mass and this implied the symmetry of the SM was no longer respected. To solve this problem, the Brout-Englert-Higgs Mechanism was introduced, in which particles acquire their mass through interactions with the Higgs field. As shown in Figure 2.2, the Higgs model consists of
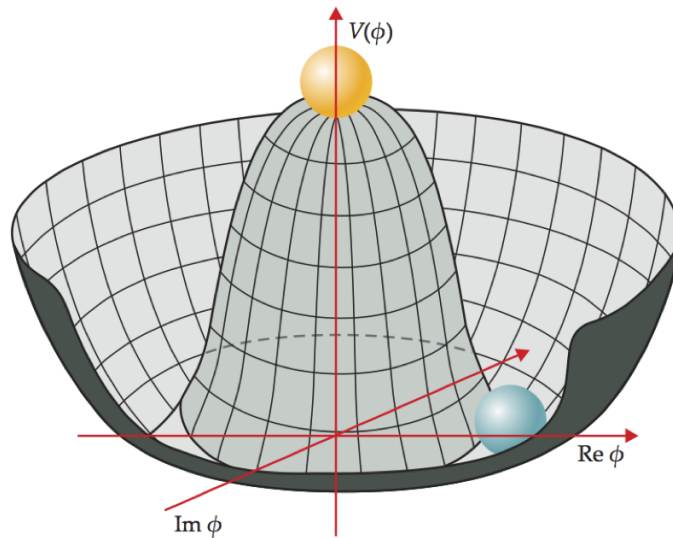

Figure 2.2.: The Higgs Potential.

a doublet of complex scalar fields which give mass to bosons and provides the necessary tools for the fermion masses. The potential is defined by:

$$V(\phi) = \mu^2(\phi^\dagger\phi) + \lambda(\phi^+\phi)^2$$

with two unknowns: μ and λ, being free parameters not predicted by the SM. ϕ is the complex scalar field doublet with four degrees of freedom:

$$\phi = \begin{pmatrix} \phi^\dagger \\ \phi^0 \end{pmatrix} = \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}$$

which through the electroweak symmetry breaking generates the mass of fermions (not predicted in this mechanism), of the vector bosons and leaves a massless photon. This scalar field is expected to have a non-zero vacuum expectation value (Eq. 2.1) as seen in Fig. 2.2 which is added to the Lagrangian with covariant derivatives:

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix} \tag{2.1}$$

where v is the VEV, related to the other parameters by $v^2 = \frac{-\mu^2}{\lambda}$, and $h(x)$ the physical Higgs field. Defining θ_W such that $\tan\theta_W = \frac{g'}{g_W}$, where g' and g_W are the coupling constants of U(1) and SU(2) gauge interactions respectively, the mass of each of the vector bosons can be expressed:

$$m_A = 0, \quad m_Z = \frac{1}{2}v\sqrt{g_W^2 + g'^2} = \frac{m_W}{cos\theta_W}, \quad m_W = \frac{1}{2}vg_W$$

In the above equations, m_A represents the mass of a massless neutral boson (photon), m_Z a massive neutral boson (Z^0) and m_W charged massive bosons ($W^\pm$). In the case of the fermion masses, the Yukawa Coupling plays a role in the process. The Yukawa Coupling term emerges from the Lagrangian density and Equation 2.1, making it invariant under SU(2)$_L \otimes$U(1)$_Y$ [6]:

$$\lambda_f = \sqrt{2}\frac{m_f}{v}$$

where λ_f is the Yukawa Coupling of the fermions to the Higgs field, and m_f is the mass of the fermion (obtained experimentally). This implies that the Higgs coupling with fermions is directly proportional to their mass, thus being greater for heavy fermions like the bottom-quark or the τ-lepton.

2.2.1. Discovery of the Higgs boson

The Higgs Boson was discovered in 2012 by the scientists working with the ATLAS and CMS detectors [7],[8]. At the LHC, where the discovery was made, data of 2011 and 2012

from ATLAS and CMS proportioned the significance needed to confirm the new boson's existence as shown in Figure 2.3. The measured mass at the time of its discovery was $m_H = 126.06 \pm 0.4(\text{stat}) \pm 0.4(\text{syst})$ GeV [7],[8]. The analysis used an integrated luminosity of $\int \mathcal{L}\, dt \sim$4.6-4.8 fb^{-1} at $\sqrt{s}$=7 TeV in 2011 and 5.8-5.9 fb^{-1} at $\sqrt{s}$=8 TeV in 2012 [7]. The Higgs boson has various decay modes, the five most sensitive at the LHC for
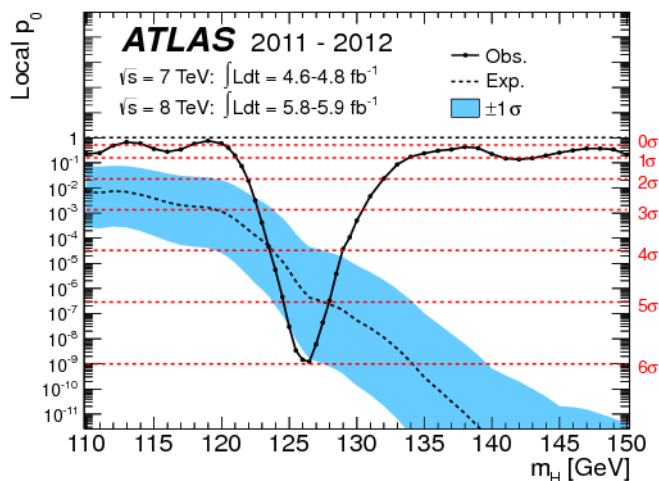

Figure 2.3.: Observed local p_0 as a function of the Higgs mass, corresponding each horizontal dashed line to significances from 1 to 6σ [7].

its discovery and precision measurements are H$\to \gamma\gamma$, H$\to ZZ \to \ell^+\ell^-\ell^+\ell^-$, H$\to \tau\tau$, H$\to W^+W^- \to \ell^+\nu\ell^-\bar{\nu}$ and H$\to \bar{b}b$ [9]. The combined data of all channels would help reach the discovery of the Higgs boson, being the H$\to \gamma\gamma$ channel the most significant. Table 2.2 presents the SM predictions of the branching ratios for these decay modes.

Decay Channel	Branching ratio [%]
H$\to \bar{b}b$	57.5$\pm$1.9
H$\to W^+W^-$	21.6$\pm$0.9
H$\to \tau\tau$	6.30$\pm$0.36
H$\to ZZ$	2.67$\pm$0.11
H$\to \gamma\gamma$	0.228$\pm$0.011

Table 2.2.: Branching ratios of the main decay modes for a Higgs boson mass of 125.09 GeV [10].

The main Higgs production modes studied in the collider were Gluon Fusion and Vector Boson Fusion (VBF). Their corresponding Feynman diagrams are shown in Fig. 2.4.
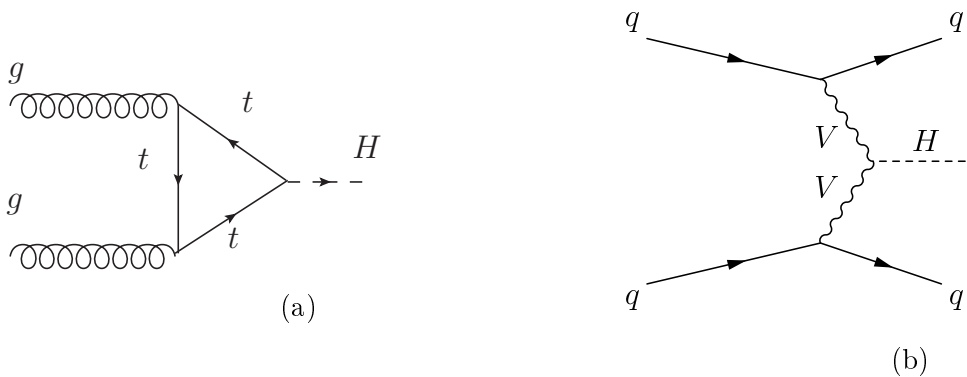
Figure 2.4.: (a)Gluon Fusion. (b)Vector Boson Fusion.

2.2.2. H→ $\tau\tau$ Channel

As mentioned in the previous section, the Higgs boson has a variety of decay modes, where H→ $\tau\tau$ is among the most significant ones. The existence of this decay was confirmed due to the combined measurements of both ATLAS and CMS experiments with a significance of 5.5σ [9]. Combined measurements of ATLAS and CMS yield a signal strength of μ=1.11$^{+0.24}_{-0.22}$ [9].

The tau lepton is able to couple to the Higgs boson because of its relatively large mass, which makes it possible to directly study the decay rates and coupling of the Higgs with fermions through the Yukawa interaction. This gives us the possibility to confirm the mass generation mechanism for fermions.

2.3. The Tau Lepton

The tau lepton is a third generation fermion which was discovered in 1975 with the SPEAR Collider at the SLAC National Accelerator Facility in the United States [11]. With a mass of 1776.82±0.16 MeV/c^2 [12], it is the heaviest of all leptons, giving it the ability to couple to particles such as the Higgs boson and being the only lepton that can decay into hadrons. This particle can also decay into fermions of the first two generations as shown in Figure 2.5. Because of its short lifetime, only the decay products of the tau lepton are observed in the detectors. The tau lepton currently plays an important role at the LHC, where tau final states provide clean measurements for the discovery of new physics, as well as the study of the Higgs boson's properties as mentioned in the previous section. The hadronically decaying tau lepton, denoted as τ_{had}, is the focus of this thesis. This decay comprises 65% of its branching ratio and can be 1-prong or 3-prong representing
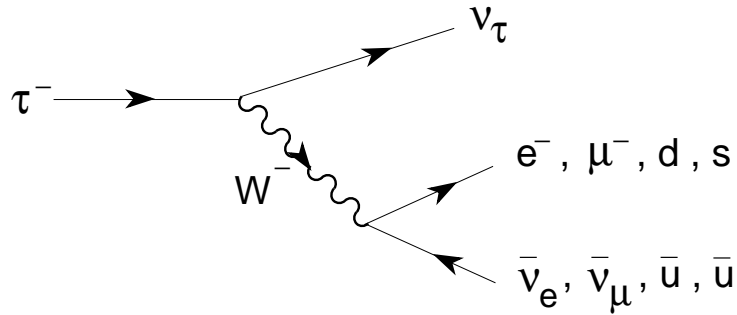
Figure 2.5.: Feynman diagram of a tau decay showing leptonic and hadronic modes [13].

one and three charged tracks respectively. The decay modes can be seen in Figure 2.6. The corresponding branching ratios for 1-prong and 3-prong hadronic decay modes are: $\tau \rightarrow \pi^- \nu_\tau$ with Γ=10.83±0.06% and $\tau \rightarrow \pi^- \pi^+ \pi^- \nu_\tau$ with Γ=8.99±0.06% [12]. The hadronic decay mode of the tau makes it possible to study hadronic weak currents and low energy aspects of QCD [13].
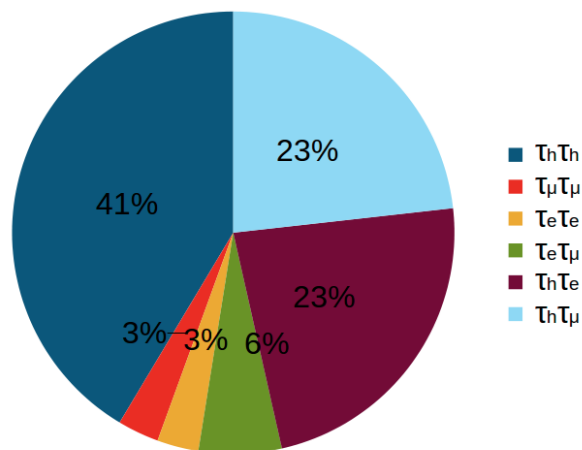

Figure 2.6.: Chart with the different decay modes of the tau lepton. τ_h denotes the hadronic decay, τ_e and τ_μ correspond to its leptonic decay to electrons or muons respectively.

3. Experimental Setup

3.1. The Large Hadron Collider

Located at Geneva, Switzerland in the facilities of the European Organization for Nuclear Research (CERN), the Large Hadron Collider is a big collaboration of scientists around the world with the goal of explaining the fundamental questions in particle physics. It consists of four experiments: ATLAS, CMS, ALICE and LHCb, each with specific functions and located at different interaction points. The proton-proton collider is 27 km long and has a design center of mass energy of 14 TeV.

The LHC Run-1 was a project spanning from 2009 to 2012 with the goal of finding the Higgs boson. The ATLAS and the CMS detectors collected data with a recorded luminosity of 5 fb^{-1} with $\sqrt{s}$ =7 TeV and 20 fb^{-1} with $\sqrt{s}$ =8 TeV, respectively [14]. Run-1 not only accomplished this objective, but gave rise to sensitive searches for BSM physics, and provided a great quantity of data of electroweak and QCD processes. In June 2015 began Run-2 and reached energies of 13 TeV, which is expected to grow to the full design energy of 14 TeV.

3.2. The ATLAS Experiment

ATLAS stands for "A Toroidal LHC ApparatuS" and has been designed to fulfill the requirements to explore phenomena at the TeV energy scale and with it, the search for the Standard Model Higgs boson. In the following subsections, the coordinate system used at ATLAS will be presented as well as a brief description of its components.

3.2.1. Coordinate System at ATLAS

At ATLAS there is a special nomenclature for the coordinate system consisting of a cylindrical symmetry. The interaction point, where the two beams collide, is considered the origin of this coordinate system. The beam travels in the z-direction while traversing the

x-y-plane, θ is the polar angle and ϕ the azimuthal one, representing the angle from the beam axis and the angle around the beam, respectively [15]. The kinematics of a particle are described by its four vector (P_x, P_y, P_z, E) where P_x and P_y are invariant for boosts in the z-direction. In this context, it is necessary to introduce invariant quantities. The pseudorapidity η is a geometrical variable defined by $\eta = -\ln(\tan\frac{\theta}{2})$, in which a value of $\eta = 0$ would mean a particle going perpendicular to the z-direction. The separation of two particles is described by $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$.

3.2.2. The Detector

The ATLAS detector is composed of various subsystems as shown in Figure 3.1. Table 3.1 summarizes the performance of each component of the detector.
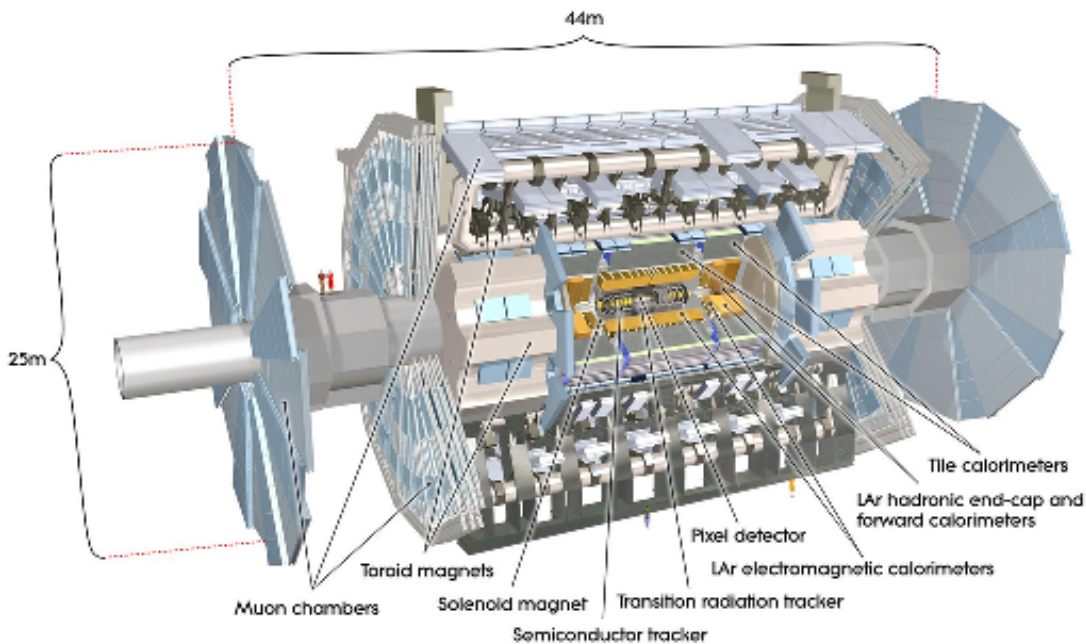

Figure 3.1.: View of the structure of the ATLAS detector [16].

The inner detector includes silicon pixel and microstrip layers, and a transition radiation tracker (TRT) which enhances the electron identification. It is immersed in a 2 T solenoid magnet field and is the closest of the detector layers to the interaction point, which makes it important for good measurements of a particle's momentum covering $p_T > 0.5$ GeV and up to $|\eta| = 2.5$.

The calorimetry of ATLAS is constructed with a liquid argon technology and is com-

Detector component	Resolution
Tracking	$\sigma_{PT}/p_T = 0.05\%p_T \oplus 1\%$
EM calorimetry	$\sigma_{PT}/E = 10\%/\sqrt{E} \oplus 0.7\%$
Hadronic calorimetry:	
barrel and end-cap	$\sigma_{PT}/E = 50\%/\sqrt{E} \oplus 3\%$
forward end-cap	$\sigma_{PT}/E = 100\%/\sqrt{E} \oplus 10\%$
Muon spectrometer	$\sigma_{PT}/p_T = 10\%$ at p_T=1 TeV

Table 3.1.: Performance summary of each ATLAS detector component with their corresponding goal resolutions with p_T[GeV] and E[GeV] [15].

posed of: an electromagnetic barrel and endcap (which are housed in a cryostat); and a hadronic endcap, barrel and forward calorimeters. The electromagnetic calorimeter is a sampling calorimeter, composed of layers of lead (absorber) and liquid argon (active material) (LAr) in an accordion structure, covering a pseudorapidity of $|\eta|$ <3.2. This calorimeter measures the energy of particles created in electromagnetic showers that take place in the absorber and whose signal is created in the active material. The hadronic calorimetry is placed outside the electromagnetic part and covers a range of $|\eta|$ <1.7. It has a scintillating-tiles technology which works together with the LAr at the endcaps, the plastic scintillators detect charged particles through the conversion of ionization energy to light signal in the optical range. In this calorimeter, steel is used as the absorber and the scintillating tiles as the active material. With the forward calorimeters, the pseudorapidity coverage is extended to $|\eta|$=4.9.

The muon spectrometer is located outside the calorimetry and is the outermost part of the ATLAS detector. It is covered by a system of air-core toroidal magnets with a magnetic field of 0.5 T. To measure the muon's momentum with high resolution, there are three layers of tracking and trigger chambers. The last one is important for providing P_T thresholds, identification and measuring of the muons' coordinates [17],[18].

ATLAS has also a trigger a system that filters selected events from the large number of data obtained. It is composed of different levels: L1, L2, and the event filter. Each of the levels choose events based on the detector's data and make a refinement on the selection made by the previous level plus other decision criteria. L1 receives the input from the calorimeter and looks for high transverse momentum objects: muons, electrons, photons, jets, and hadronically decaying tau leptons. Based on each event selected, Regions-of-Interest (RoI) are selected and cover the events of interest in η and ϕ coordinates. This level can make a decision in less than 2.5 μs out a rate of 75 kHz. The information of L1

is then passed to the L2 and then finally to the event filter, where the trigger rate gets reduced and an offline analysis follows [15].

4. Identification of Hadronically Decaying Tau Leptons at Atlas

In chapter 2, the role of the tau lepton in current particle physics research, as well as the need for good performing reconstruction and identification of hadronically decaying τ leptons for these final states was presented. At the Atlas experiment, algorithms are used and constantly updated to achieve an optimal separation between tau signal and background. With the new higher energies and luminosity at the Lhc, pile-up corrections are needed in order to make the tau reconstruction and identification less sensitive to them. Hence the necessity of their optimization for the Run-2 running conditions.

Since the tau lepton yields signatures in the detector that can be faked by jets and electrons, these two are its principal background sources. Information in the different parts of the detector about the tau tracks and energy deposited in calorimeters is used to differentiate between signal and background. The discriminating variables are the ones that contain this information from the detector and are used in discriminating algorithms. The separation of signal and background is usually done with methods such as a projective likelihood (LLH) and boosted decision trees (BDT). The first method will not be considered in this thesis.

4.1. Tau reconstruction and identification

The reconstruction of hadronic tau leptons is done with an algorithm that is "seeded" from jets reconstructed using the anti-k_T algorithm with a distance parameter of $R = 0.4$. The tau reconstruction associates each reconstructed jet (called a *seed*) with a τ_{had}^{vis} candidate, where the seed must satisfy $p_T >$10 GeV and $|\eta|<$2.5 [19]. The topoclusters of the jet seed are the three-dimensional clusters of calorimeter cells. The Local Hadronic Calibration (LC scale) is applied to the topocluster energy to improve the calibration of the energy deposition from hadrons [20]. From the tracking and calorimeter information obtained, kinematic variables are calculated. The reconstruction algorithm is run on all

jets within a η range of $|\eta| < 2.5$ and takes the energy of the τ_{had}^{vis} candidate from the sum over the uncalibrated (EM scale) energies of the cells within $\Delta R <$ 0.4 of the seed jet axis [21],[22].

The variables calculated in the reconstruction are then used for the identification process. Since they provide information about the detector signatures of the particles involved, they are implemented in multivariate discriminants to reject tau lepton backgrounds such as jets and electrons.

4.2. Electron Veto

Hadronically decaying 1-prong tau leptons have a signature which is well mimicked by electrons, representing a huge contribution of the background after suppressing the jets. However, there are certain properties which are helpful when trying to discriminate between them. An important discriminating characteristic is the shape of both showers. A shower produced by a tau lepton tends to be longer and wider than an electron shower. Transition radiation in the TRT is also a characteristic property of the electrons. For heavy particles such as the tau lepton passing through the TRT, transition radiation is much less prevalent. These and other properties are used for the discrimination of electrons from tau leptons.

During the reconstruction process of τ_{had}^{vis} there is little rejection against the background from electrons. Therefore the information provided from the discriminating variables calculated during the reconstruction is needed. The following are the descriptions of these variables which directly follows the description in [22]:

Leading track momentum fraction (f_{track}):

$$f_{track} = \frac{P_T^{track}}{\sum_{j\epsilon(all)}^{\Delta R_j < 0.4} E_{T,j}^{EM}},$$

where P_T^{track} corresponds to the leading P_T core track of the tau candidate and j runs over all cells in $\Delta R < 0.4$. The calibration used for the cells is at the EM scale.

Core energy fraction (f_{core}): Fraction of transverse energy in the core region $\Delta R < 0.1$

of the τ_{had}^{vis} candidate:

$$f_{core} = \frac{\sum_{i\epsilon(all)}^{\Delta R_i < 0.1} E_{T,i}^{EM}}{\sum_{j\epsilon(all)}^{\Delta R_j < 0.4} E_{T,j}^{EM}},$$

where i runs over all cells related to the tau candidate within $\Delta R < 0.1$ of the intermediate axis and j runs over all cells in $\Delta R < 0.4$. Cell calibration is at the EM scale.

Electromagnetic fraction ($\mathbf{f}_{EM}$): Fraction of transverse energy of the τ_{had}^{vis} candidate deposited in the EM calorimeter:

$$f_{EM} = \frac{\sum_{i\epsilon(EM0-2)}^{\Delta R_i < 0.4} E_{T,i}^{EM}}{\sum_{j\epsilon(all)}^{\Delta R_j < 0.4} E_{T,j}^{EM}},$$

where i runs over the cells in the first three layers of the EM calorimeter and j runs over all layers of the calorimeter.

TRT HT fraction ($\mathbf{f}_{HT}$): Ratio of high-threshold hits to low-threshold hits in the Transition Radiation Tracker (TRT) for the leading P_T core track.

Ring Isolation ($\mathbf{f}_{iso}$):

$$f_{iso} = \frac{\sum_{i\epsilon(all)}^{0.1 < \Delta R_i < 0.2} E_{T,i}^{EM}}{\sum_{j\epsilon(all)}^{\Delta R_j < 0.4} E_{T,j}^{EM}}.$$

where i runs over all cells in the associated topocluster of the tau candidate in an annular region within $0.1 < \Delta R < 0.2$ around the seed. The index j runs over cells in a cone of $\Delta R_j < 0.4$. Energy calibration is at the EM scale.

Presampler strip energy fraction ($\mathbf{f}_{PS}$):

$$f_{PS} = \frac{\sum_{l=0}^{N_{clus}} E_l^{PS}}{\sum_{l=0}^{N_{clus}} E_l}.$$

where l runs over the calorimeter clusters associated to the tau candidate, while E_l^{PS} is the cluster energy deposited in the Presampler layer of the calorimeter. E_l denotes the total energy of the calorimeter cluster. Calibration of the cluster energy was made at the LC scale.

Secondary energy deposits in the strip compartment ($\mathrm{E}_{T,max}^{strip}$): Sum of the energy in the strip layer over three cells in ϕ and local maxima are searched for cell sums in η centered around the impact point of the leading track associated to the τ_{had}^{vis} candidate. The energy associated with the leading track is excluded and the variable is only calculated for a region of $\eta \leq 1.7$. For this thesis, the variable used is divided by the P_T of the leading track.

Hadronic Leakage ($\mathrm{E}_{T,reco}^{Had/EM}$): Ratio of E_T reconstructed in the first compartment of the hadronic calorimeter in a region of $\Delta\eta \times \Delta\phi = 0.2 \times 0.2$ to the E_T reconstructed in the EM calorimeter [23].

Absolute distance in η ($\Delta\eta^{track}$): Absolute distance in η between the tau candidate and its leading track.

Absolute distance in ϕ ($\Delta\phi^{track}$): Absolute distance in ϕ between the tau candidate and its leading track.

4.2.1. Electron veto variables at 8 TeV

In Run-1 with energies of 8 TeV, discrimination against electrons was done with the variables described previously without $\mathrm{E}_{T,reco}^{Had/EM}$. The optimization done at the time made use of Monte Carlo generated events of Z$\rightarrow \tau\tau$ for signal and a Z$\rightarrow ee$ sample for background. The cuts applied to both signal and background candidates were:

- $P_T > 20$ GeV,

- signal candidate must be matched to a true hadronically decaying tau,

- background candidate must be matched to a true electron,

- BDT discriminant was trained in different regions of $|\eta|$: barrel $|\eta| < 1.37$, crack $1.37 < |\eta| < 1.52$, end-cap $1.52 < |\eta| < 2.0$ and forward end-cap $2.0 < |\eta| < 2.3$.

The results obtained in this optimization from 2012 can be seen in [22].

4.2.2. Electron veto variables at 13 TeV

With higher energies, luminosities and changes to some components of the detector, the Run-2 electron veto must be optimized and the identification variables updated. However,

not all the variables were available for the Run-2 discrimination against electrons, hence different results from the electron veto at 8 TeV were obtained. The missing variable was the TRT HT fraction (f_{HT}), which due to technical reasons was not available in the 13 TeV samples.

The cuts applied for Run-1 listed in section 4.2.1. were also implemented for Run-2, including $P_T >$20 GeV for the tau candidate. However, due to different behaviour of the variables in each region of the detector, the discriminant was trained only in the barrel $|\eta| < 1.37$, end-cap $1.52 < |\eta| < 2.0$ and forward end-cap $|\eta| > 2.0$ regions. The crack (or transition) region $1.37 < |\eta| < 1.52$ is no longer used for the tau reconstruction in Run-2 due to bad energy reconstruction in this part of the detector.

Samples of Monte Carlo simulated events of Z$\rightarrow \tau\tau$ for signal and Z$\rightarrow ee$ for background were used, where Pythia was implemented for the Parton Model Shower [24], Powheg-Box for the Matrix Element [25] and GEANT4 for the detector simulation [26]. A condensed description of these samples is presented in Table 4.1. To simulate the pileup present with the higher luminosities, a number of minimum bias interactions extracted from a Poisson distribution were added event-by-event. The average number of interaction per bunch crossing of the Poisson distribution was set as μ, which yielded a mean value of 25 covering the expected values of the Run-2 data.

| Sample	No. of Events	Parton Shower Model	Matrix Element	Detector
Z$\rightarrow \tau\tau$	111,777	Pythia 8.186	Powheg-Box v2	GEANT4
Z$\rightarrow ee$	5,893,703	Pythia 8.186	Powheg-Box v2	GEANT4

Table 4.1.: Summary about the details of the samples used for the discrimination against electrons.

Figures 4.1 and 4.2 show the distributions in the barrel region of all input discriminating variables normalized to unit area of the samples used. The distributions for the other two regions are presented in the Appendix. These plots show the matched tau candidate and electron of each of the variables. In the figures, one can see how well each variable separates between signal and background. The variables f_{EM} and $E_{T,max}^{strip}$ as shown in Figures 4.1d and 4.2e, provide a large separation power, whereas the separation between τ_{had} and electrons has gotten worse for f_{PS}. In addition, f_P^{EM} shows a subtle miscalibration in the background distribution, which is expected to peak at 1. In general, most of the variables do not yield a great separation power.
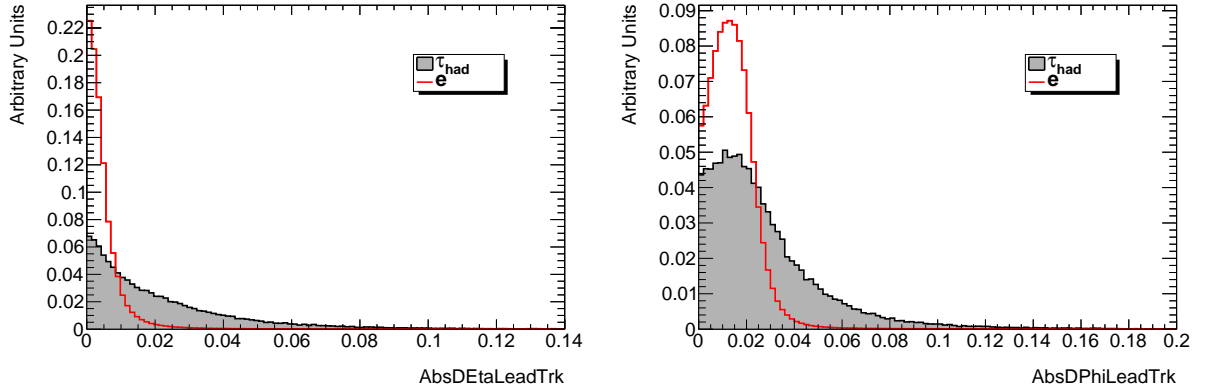
(a) Absolute distance in η between the tau candidate and its leading track ($\Delta\eta^{track}$).

(b) Absolute distance in ϕ between the tau candidate and its leading track ($\Delta\phi^{track}$).
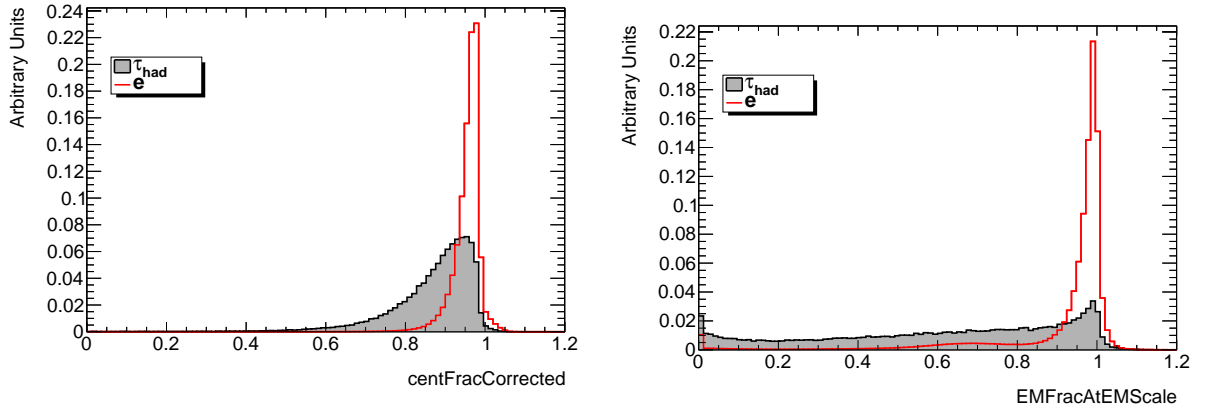
(c) Core energy fraction corrected for pile-up (f_{core}^{corr}).

(d) Electromagnetic fraction (f_{EM}).

Figure 4.1.: Distributions of some of the identification variables for 1-prong in the η barrel region for Run-2.
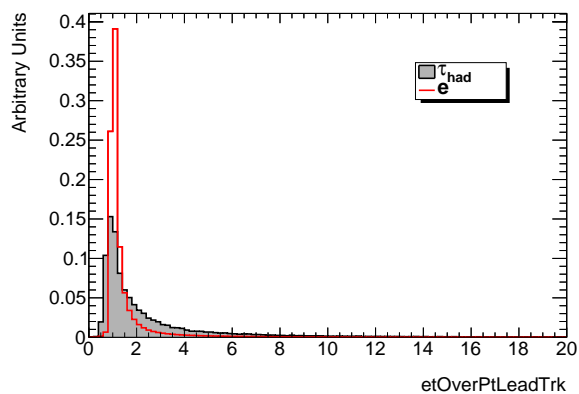
(a) Electromagnetic energy over momentum of track system (f_P^{EM}).
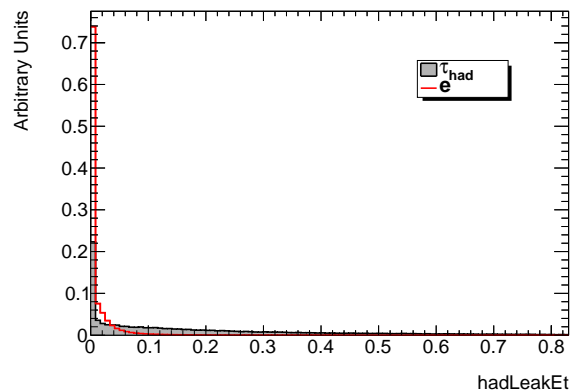
(b) Hadronic Leakage ($E_{T,reco}^{Had/EM}$).
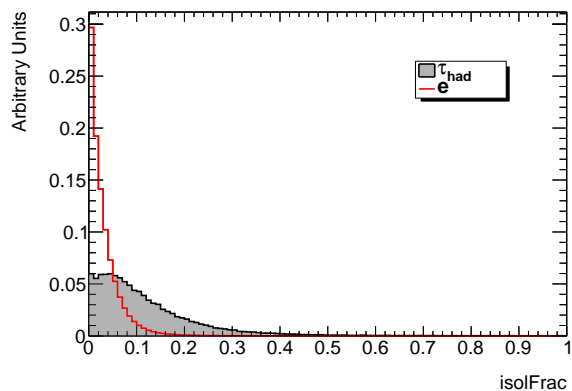
(c) Ring Isolation (f_{iso}).
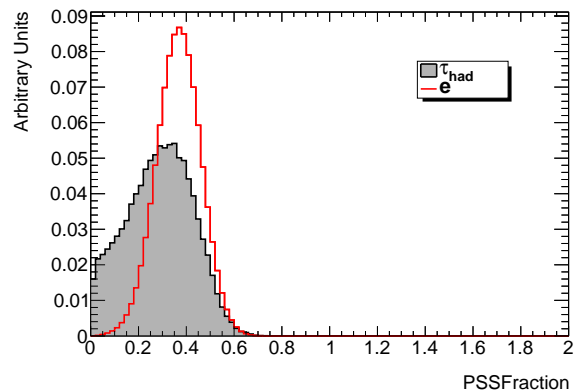
(d) Presampler strip energy fraction (f_{PS}).
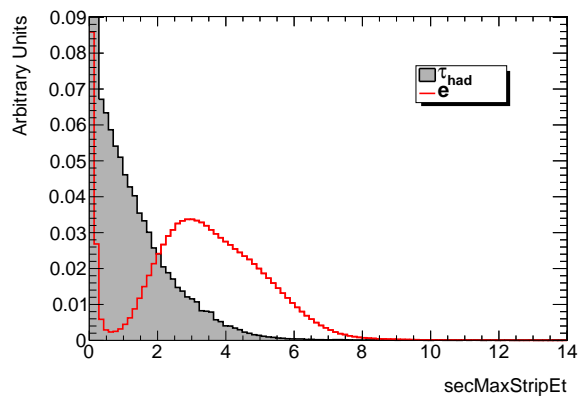
(e) Secondary energy deposits in the strip compartment ($E_{T,max}^{strip}$).

Figure 4.2.: Distributions of some of the identification variables for 1-prong in the η barrel region for Run-2.

5. Suppression of Electron Candidates in Tau Identification at 13 TeV

The rejection of electrons from tau leptons at 13 TeV is done with a new electron veto algorithm, which is trained using a multivariate discriminator. In Run-1, a BDT was used for the same purpose with the criteria described in section 4.2.1. For Run-2, a BDT is also trained with Z$\rightarrow \tau\tau$ and Z$\rightarrow ee$ MC samples simulated for events at 13 TeV. This algorithm is implemented in a multivariate analysis package known as TMVA [27].

5.1. Boosted Decision Trees

The BDT is a multivariate classification algorithm (or regressor) and is widely used for the discrimination against electrons from tau leptons. The method consists of a series of cut based decisions, and with each decision (which may be called a node), a certain variable is cut on based on how well it separates signal from background, in other words, increasing purity. A tree (as shown in Fig.5.1) is composed out of a set of nodes that dictate how a set of cuts, which separate signal from background data, can be applied to increase purity $p = \frac{S}{S+B}$.

The *boosting* of a decision tree consists of extending the process to a certain amount of trees called a forest. The forest is the collection of all decisions made through the training and combining them into a single regressor. The tree gives a weight to events that were misclassified, and for each tree there is a reweighting process of the events depending on how often they were used in the node cuts, which yields the different trees in the forest. The final regressor gives a weighted average of each of the trees based on their misclassification rate.

The Toolkit for Multivariate Data Analysis (TMVA) provides the methods necessary to analyze the information provided by identification variables and classify the events in a sample as signal or background, which is the objective of this thesis. TMVA is integrated to the analysis framework ROOT, the versions used for this thesis were TMVA 4.2.1 and
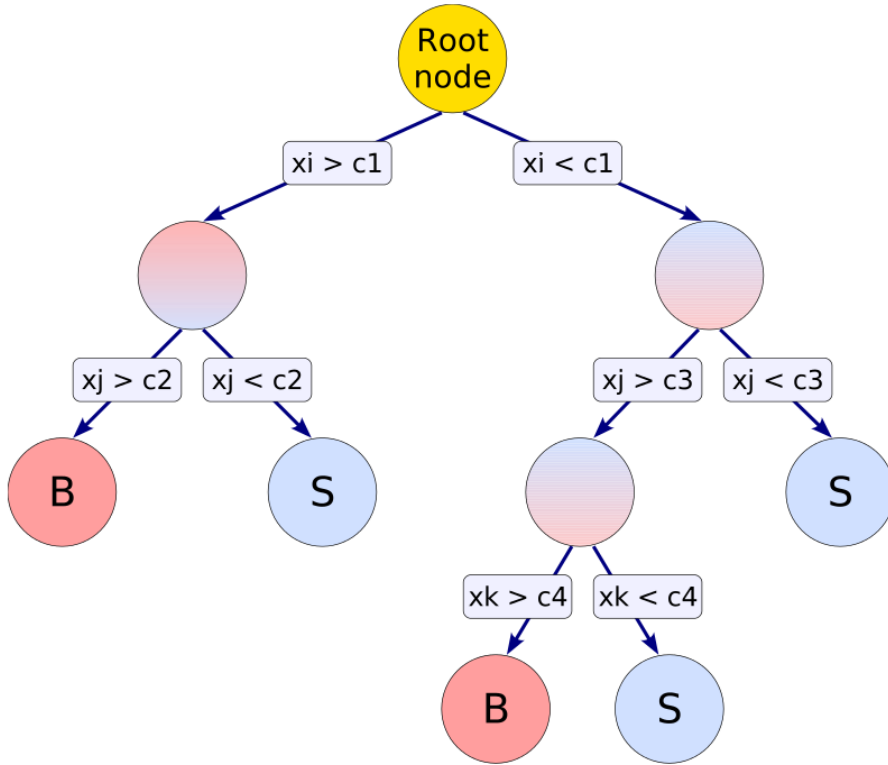
Figure 5.1.: View of a tree structure of a BDT, where each node represents a decision on a variable cut to select the events more background or signal like.

ROOT 6.04/14.

5.2. Performance Measures

In the algorithm it is possible to change a variety of parameters to achieve the best discrimination performance. During the tuning of the BDT, the following aspects were considered to measure its performance:

- the overtraining in signal and background events measured with the Kolmogorov-Smirmov (KS) test,

- the area under the ROC curve, and

- the background rejection ($r_{bkgd} = 1 - \epsilon_{bkgd}$) at 90% signal efficiency ϵ_{sig}.

The BDT works with a training and test sample to check for overtraining, which occurs when an excessive amount of decisions are made to a relatively small data sample, thus the TMVA compares the performance for both training and test samples. In this case

half of the events were the test and the other half was the training sample. This guarantees a statistically independent evaluation of the BDT [27]. One measurement of the overtraining is the Kolmogorov-Smirnov test, which gives a value ranging between 0 and 1. An indication of consistent distributions coming from the same parent distribution is given by calculated KS test values closer to 1 [28].

For the overtraining check of the BDTs, the KS test of the ROOT framework was implemented. The calculated value (denoted as simply KS_{sig} for signal and KS_{bkgd} for background) was obtained comparing the shapes of the histograms of the test and training samples. The minimum threshold of 10% for the KS test value is chosen to ensure that the BDT does not suffer too much from the effects of overtraining.

The correlations between variables is also an important aspect when working with statistical data. The BDT is not highly dependent on the correlations but this can still harm the performance if not modelled correctly.

The receiver operating characteristic curve (ROC) plots the background rejection $r_{bkgd} = 1 - \epsilon_{bkg}$ (being ϵ_{bkg} the background efficiency) and signal efficiency ϵ_{sig} obtained for different cuts on the BDT output, summarizing the overall separation power of the BDT. The ideal ROC curve would include the point ϵ_{sig}=100% and r_{bkgd}=100%, with a total integral under the curve being equal to unity, meaning that full separation of signal and background is possible. High performance is thus characterized by large area under de ROC curve.

During the training of the BDT, three parameters were tuned: the number of trees in the forest "NTrees", minimum node size "MinNodeSize" which is the minimum percentage of events in a node, and maximum depth of the decision tree "MaxDepth" as seen in Table 5.1. For the trainings in the barrel $|\eta| < 1.37$ and end-cap $1.52 < |\eta| < 2.0$ regions, all variables were kept before optimization. The distributions of the variables $\text{E}_{T,reco}^{Had/EM}$, $\text{E}_{T,max}^{strip}$ and f_{EM} were not available in the forward end-cap $|\eta| > 2.0$ region, and therefore, they were not used in the training in that particular η range (see Appendix B).

5.3. Results

The identification variables obtained from the samples were introduced into the BDT, which yielded output plots of the distributions of test and training samples for back-

Parameters	Definition	Default values
NTrees	Number of trees in the forest	850
MinNodeSize	Minimum node size	2.5%
MaxDepth	Maximum depth of the decision tree	3

Table 5.1.: List of the parameters tuned in the trainings with their corresponding definitions and default values.

ground and signal, the ROC curve, the correlation matrices of the input variables, as well individual distributions and other TMVA functions which were not significant for this thesis.

The first BDT training was carried out in the barrel region with the default values of the paramaters: NTrees=850, MinNodeSize=2.5%, and MaxDepth=3. In Figure 5.2, the plots obtained from the TMVA are shown. The BDT output shows a correlation matrix for signal and other for background, the distribution of training and test samples to check for overtraining and a background rejection versus signal efficiency plot (ROC curve). The results obtained presented good performance for the set of parameters, with an area under the ROC curve of 0.95 and a background rejection of r_{bkgd}=0.84±0.01 at ϵ_{sig}=90%. In addition, the KS test values for the methods described in section 5.2 were: KS_{sig}=0.96 and KS_{bkgd}=0.62, indicating very low overtraining in the process. The overall performance of the trainings with default parameters in all η regions can be seen in Table 5.2.

Region	NT	MNS	MD	KS_{sig}	KS_{bkgd}	$r_{bkgd}(\epsilon_{sig} = 90\%)$	ROC
barrel $|\eta| < 1.37$	850	2.5	3	0.96	0.62	0.84±0.01	0.95
end-cap $1.52 < |\eta| < 2.0$	850	2.5	3	0.68	0.99	0.76±0.01	0.93
forward end-cap $|\eta| > 2.0$	850	2.5	3	0.30	0.30	0.68±0.01	0.92

Table 5.2.: Performance summary of the data calculated from the first BDT trainings. The columns' letters stand as follows: NT=NTrees, MNS=MinNodeSize, MD=MaxDepth, KS_{sig} and KS_{bkgd} are the Kolmogorov Smirmov tests for signal and background respectively, ROC=area under the ROC curve, and $r_{bkgd}(\epsilon_{sig} = 90\%)$=background rejection at ϵ_{sig} =90%.

In the first training however, two variables were confirmed to be highly correlated. Figures 5.2a and 5.2b show the resulting correlation matrices for signal and background, respectively. The correlation plot of f_{core} and f_{iso} presented in Fig. 5.2e confirms how high they are correlated. Both variables were equally ranked in the first BDT with 8% of

variable importance and presented similar correlations with the other variables, hence two BDTs where trained: one without f_{core} and other without f_{iso}. Both trainings performed similarly, making no difference in choosing a variable over the other. It was decided to take f_{core} out of the trainings.

Because of the low ranking for the variable $E_{T,reco}^{Had/EM}$ (6%), some BDT trainings were done also without it, however, the performance did not show improvement and the rest of the trainings were done with all variables listed in Table 5.3.

Variable	$|\eta| < 1.37$	$1.52 < |\eta| < 2.0$	$|\eta| > 2.0$
f_{EM}	✓	✓	
f_{iso}	✓	✓	✓
f_P^{EM}	✓	✓	✓
f_{PS}	✓	✓	✓
$E_{T,max}^{strip}$	✓	✓	
$E_{T,reco}^{Had/EM}$	✓	✓	
$\Delta\eta^{track}$	✓	✓	✓
$\Delta\phi^{track}$	✓	✓	✓

Table 5.3.: List of recommended variables to apply for each η region in the BDT trainings.

After running an automation script for the different parameters and η sections, the settings for the BDT that yielded the best performance were selected, where a good performance is defined by high values of both ROC curve integral and background rejection. All correspond to a signal efficiency of 90%, which have the background rejections with least uncertainty and greater integral under the curve. Table 5.4 summarizes the recommended parameters for each of the η regions.

The results of the training with the best parameters in the barrel region can be seen in Figure 5.3, which compared to the other two, had most of the events and hence the least overtraining, with values of KS_{sig}=0.66 and KS_{bkgd}=0.84, as presented in Fig. 5.3c where test and training samples are superimposed. The ROC Curve shown in Fig. 5.3d presents an integral under the curve of 0.95 and a background rejection of $r_{bkgd} = 0.85 \pm 0.01$ for a signal efficiency of ϵ_{sig}=90%.

In Figure 5.4 the plots for the BDT training in the end-cap region are shown. The

Region	NT	MNS	MD	KS_{sig}	KS_{bkgd}	$r_{bkgd}(\epsilon_{sig} = 90\%)$	ROC
barrel $|\eta| < 1.37$	850	2.0	4	0.66	0.84	0.85±0.01	0.95
end-cap $1.52 < |\eta| < 2.0$	850	2.0	3	0.70	0.93	0.77±0.01	0.94
forward end-cap $|\eta| > 2.0$	900	2.5	2	0.54	0.74	0.69±0.01	0.92

Table 5.4.: Summary of the data calculated from the BDT trainings performed with the recommended parameters. The columns' letters stand as follows: NT=NTrees, MNS=MinNodeSize, MD=MaxDepth, KS_{sig} and KS_{bkgd} are the Kolmogorov Smirnov tests for signal and background respectively, ROC=area under the ROC curve, and $r_{bkgd}(\epsilon_{sig} = 90\%)$=background rejection at ϵ_{sig} =90%.

best performance parameters are similar to the ones in the barrel region except for the maximum depth of the tree. This difference is caused by the less number of events that occur in this η region compared to the barrel section, therefore, the tree needs less nodes to improve the decision process thus needing a smaller number for the MaxDepth parameter. The values obtained for the overtaining check were the following: KS_{sig}=0.70 and KS_{bkgd}=0.93. The area under the ROC curve shown in Fig. 5.4d is 0.94 and the background rejection calculated for ϵ_{sig}=90% is $r_{bkgd} = 0.77 \pm 0.01$.

The optimal parameters for the forward end-cap region show a relative difference to the other η regions. This BDT had much less events, thus making it more prone to overtraining due to the lack of statistics compared to the other two. The number of discriminating variables implemented has also an important effect in the performance and therefore, different values for the parameters are needed. In comparison with the trainings in the other two regions, the overtraining in this BDT was higher but still good enough to claim compatibility. The values calculated for the KS test were KS_{sig}=0.54 and KS_{bkgd}=0.74. Figure 5.5 presents the results of this BDT, where the integral under the ROC curve is 0.92 and $r_{bkgd} = 0.69 \pm 0.01$ for ϵ_{sig}=90%.
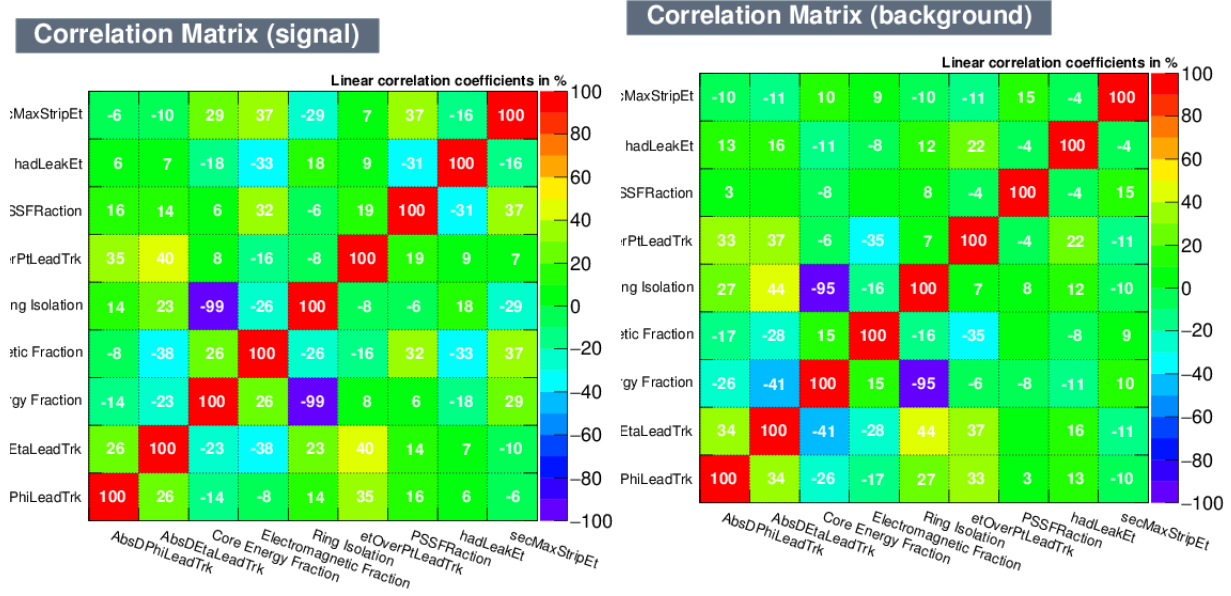
(a) Correlation Matrix for the signal events.

(b) Correlation Matrix for the background events.
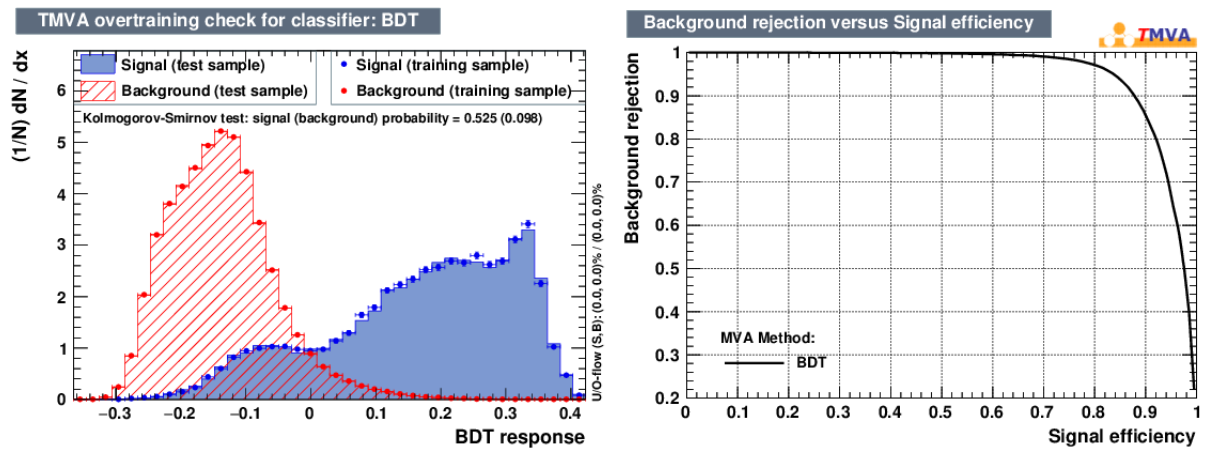
(c) Comparison of the test and sample distributions from signal and background events.

(d) ROC curve of the BDT performance. Background rejection= $1 - \epsilon_{bkgd}$.
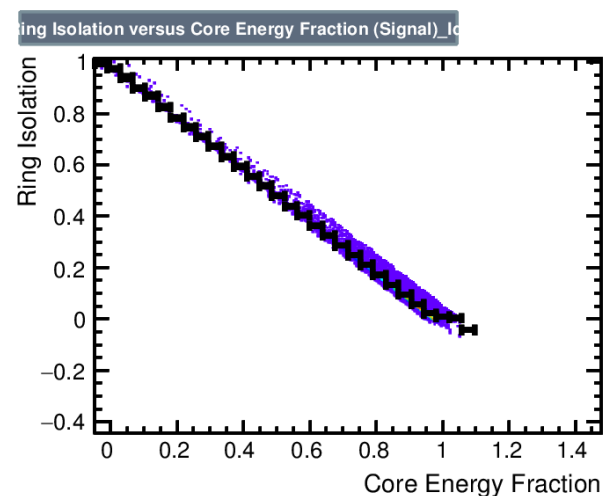
(e) Correlation plot of f$_{core}$ and f$_{iso}$.

Figure 5.2.: Resulting plots of the first BDT training with all variables and default parameters: NTrees=850, MinNodeSize=2.5%, and MaxDepth=3.

(a) Correlation Matrix for the signal events.

(b) Correlation Matrix for the background events.

(c) Comparison of the test and sample distributions from signal and background events.

(d) ROC curve of the BDT performance. Background rejection= $1 - \epsilon_{bkgd}$.
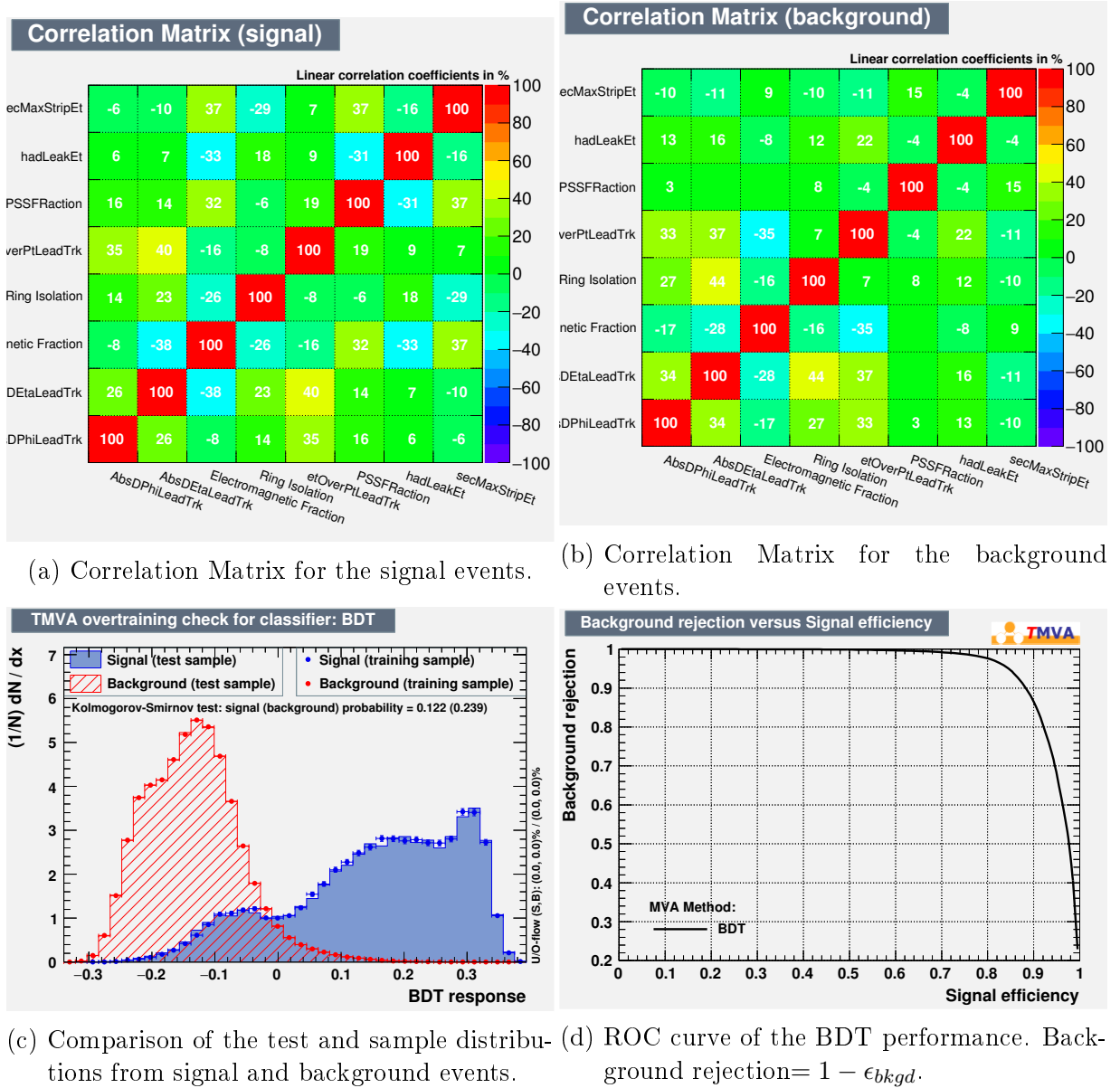
Figure 5.3.: Resulting plots of the BDT training in the η barrel region for the parameters: NTrees=850, MinNodeSize=2.0 and MaxDepth=4.

(a) Correlation Matrix for the signal events.

(b) Correlation Matrix for the background events.

(c) Comparison of the test and sample distributions from signal and background events.

(d) ROC curve of the BDT performance. Background rejection$= 1 - \epsilon_{bkgd}$.
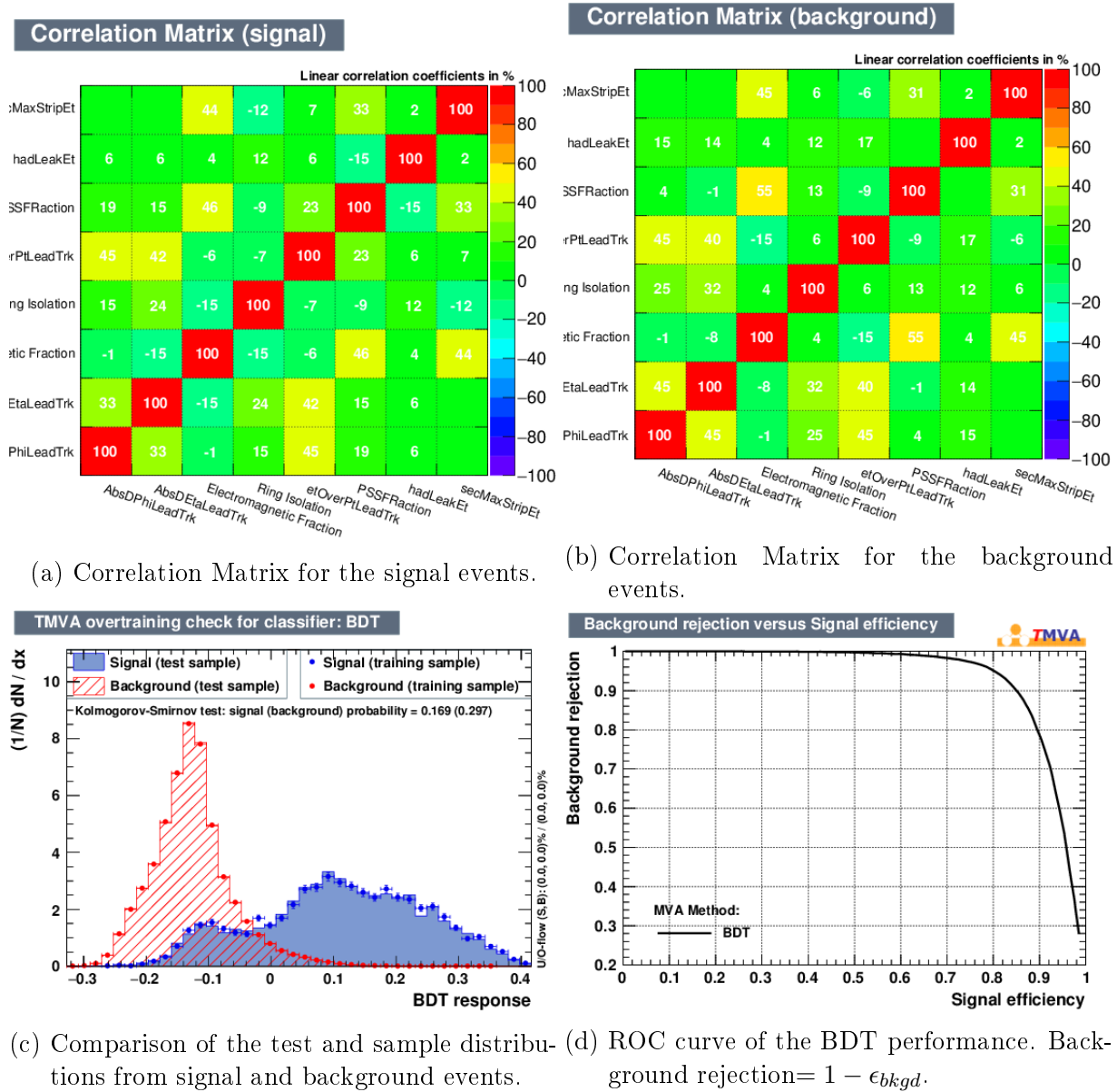
Figure 5.4.: Resulting plots of the BDT training in the η end-cap region for the parameters: NTrees=850, MinNodeSize=2.0 and MaxDepth=3.

(a) Correlation Matrix for the signal events.

(b) Correlation Matrix for the background events.

(c) Comparison of the test and sample distributions from signal and background events.

(d) ROC curve of the BDT performance. Background rejection= $1 - \epsilon_{bkgd}$.
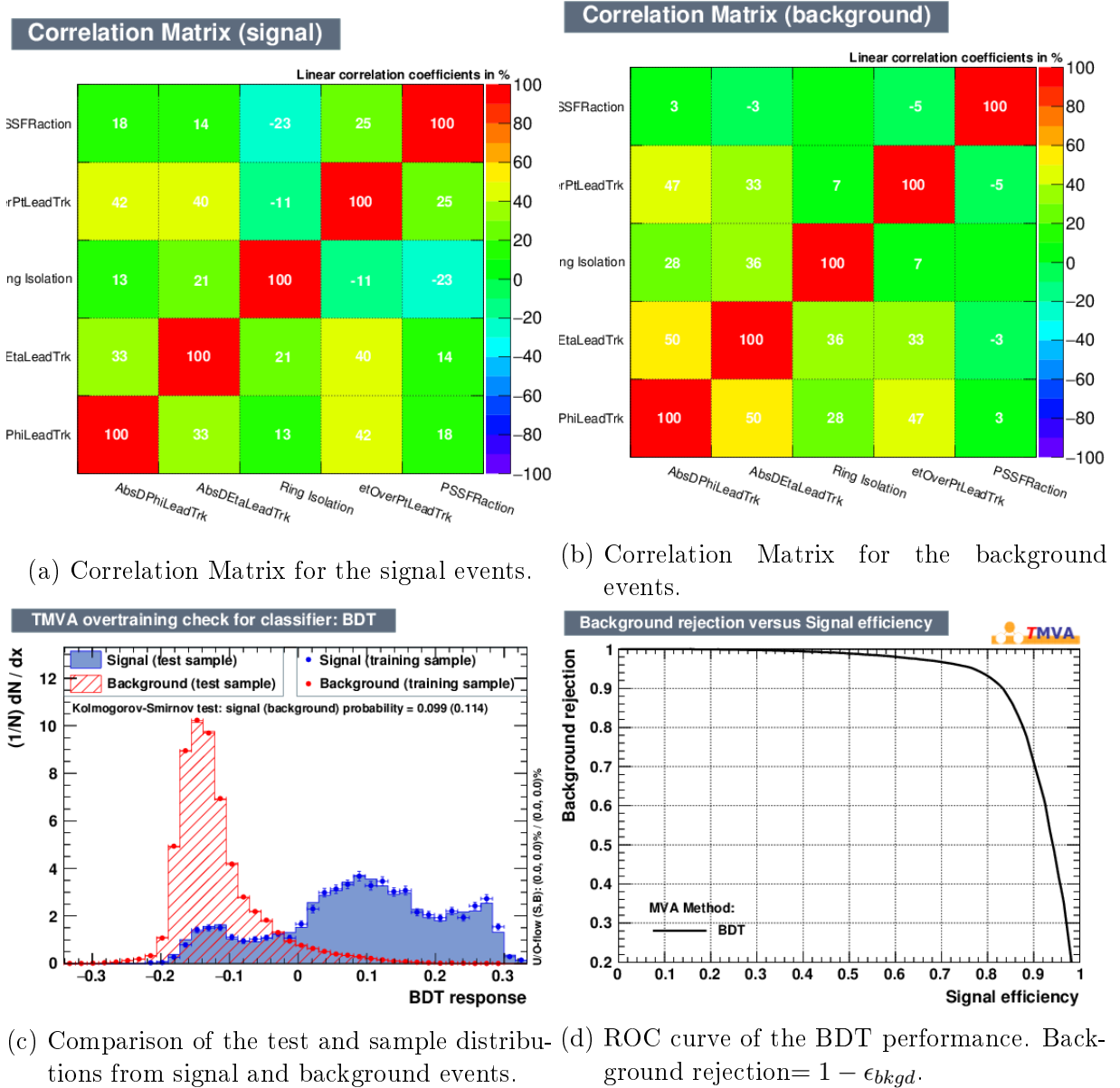
Figure 5.5.: Resulting plots of the BDT training in the η forward end-cap region for the parameters: NTrees=900, MinNodeSize=2.5 and MaxDepth=2.

6. Conclusion

In this thesis, an electron discriminant for 1-prong hadronically decaying tau leptons was optimized using a BDT method based on studies in Run-1. The BDT algorithm acts as a discriminant between performing a series of cuts on variables with high separation power and yielding a series of multiple decisions. The training was implemented on MC generated samples of Z→ $\tau\tau$ as signal and Z→ ee as background, and was divided in three pseudorapidity regions of the ATLAS detector.

The information about the detector signatures of both τ_{had} and electron candidates was introduced through the discriminating variables. One of the variables used in the electron veto for 8 TeV was not implemented for Run-2, and from the ones that were used, some were also taken out due to high correlations with other variables or suffered from a lack of information in the forward end-cap η region. This made the new optimization for 13 TeV important for the future studies and measurements in tau physics. The performance of the produced trainings was measured with calculated values of the overtraining test, the area under the ROC curve and the background rejection at 90% signal efficiency. It was also shown that the performance of two variables that presented high correlations was very similar, hence the separation power of future electron vetos should not be significantly affected by the choice of one variable over the other.

The results yielded a good background rejection for a signal efficiency of 90% with parameters relatively close to the TMVA default values. The possibility of automation of the BDT helped the understanding of the performance of this algorithm with a variety of different parameters and thus, it was easier to tune the BDT to values closer to the optimal.

6.1. Outlook

In addition to the above mentioned, it is suggested to implement the missing variable f_{HT} to newer electron vetos, since this variable contains valuable information that help

improve the performance of the BDT training. The discrimination against electrons for the identification of τ_{had} will benefit the current and future measurements in the 13 TeV LHC. In section 2.2.2., the importance of the H$\to \tau\tau$ was described. This decay channel provides information in the coupling of the Higgs with fermions, and the effective identification of the tau lepton plays an important role in the measurement of the Yukawa coupling and other BSM physics studies such as the search for supersymmetric heavy Higgs decaying leptonically. It is expected that an improvement in the identification of the tau leptons will provide the tools necessary to find new particles and discover new physics along the way.

A. Identification Variable Distributions
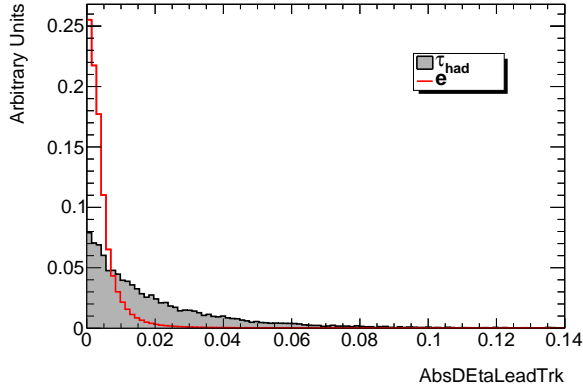
(a) Absolute distance in η between the tau candidate and its leading track.
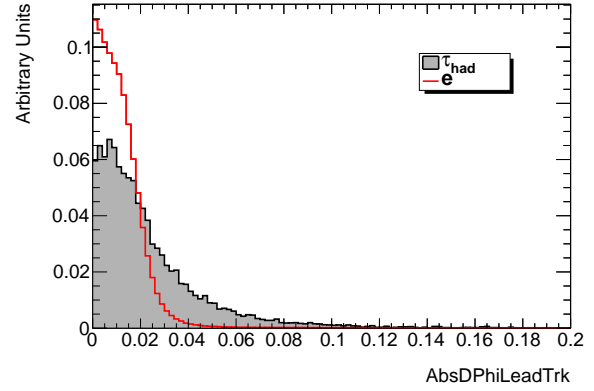
(b) Absolute distance in ϕ between the tau candidate and its leading track
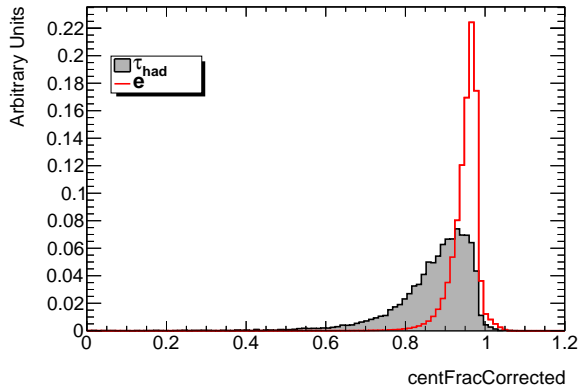
(c) Core energy fraction corrected for pile-up (f_{core}^{corr}).
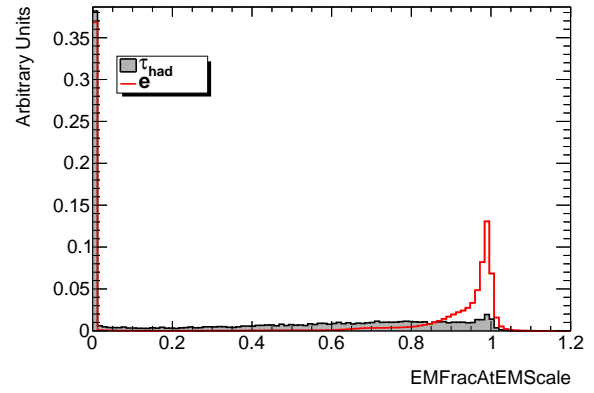
(d) Electromagnetic fraction (f_{EM}).

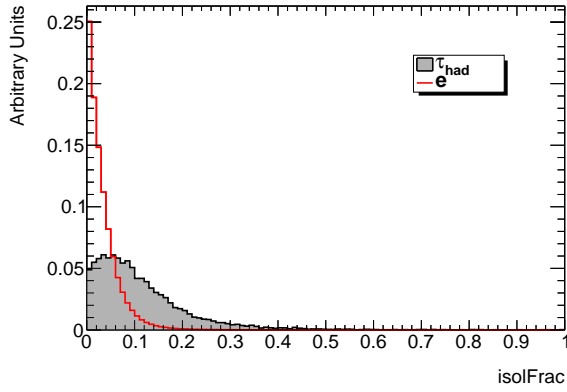Figure A.1.: Distributions of some of the identification variables for 1-prong in the η endcap region for Run-2.
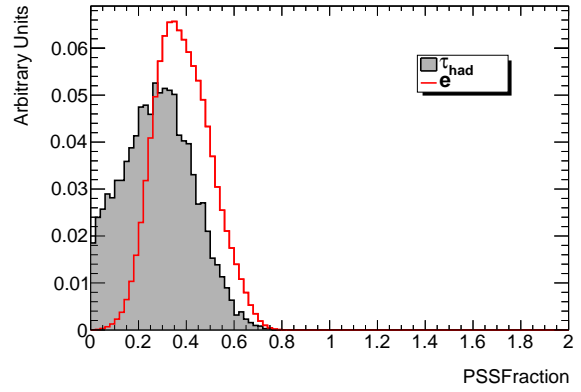
A. Identification Variable Distributions
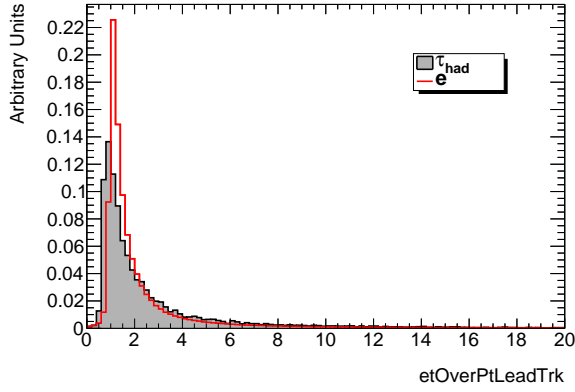
(a) Electromagnetic energy over momentum of track system (f_P^{EM}).
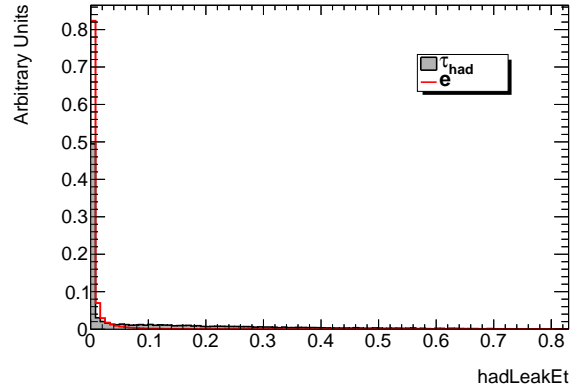
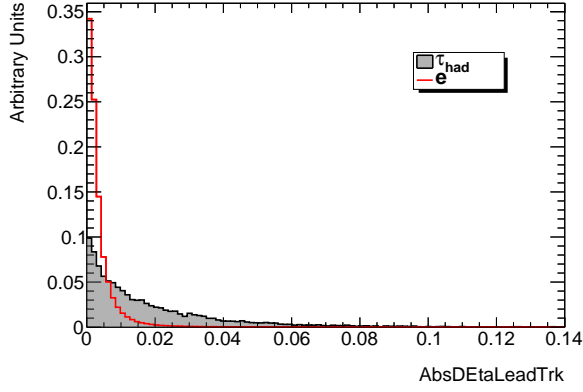(b) Hadronic Leakage ($E_{T,reco}^{Had/EM}$).

(c) Ring Isolation (f_{iso}).

(d) Presampler strip energy fraction (f_{PS}).
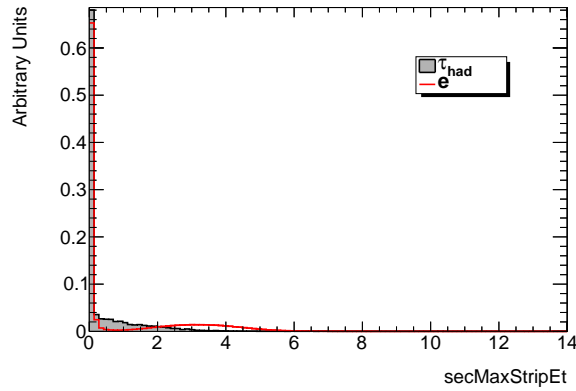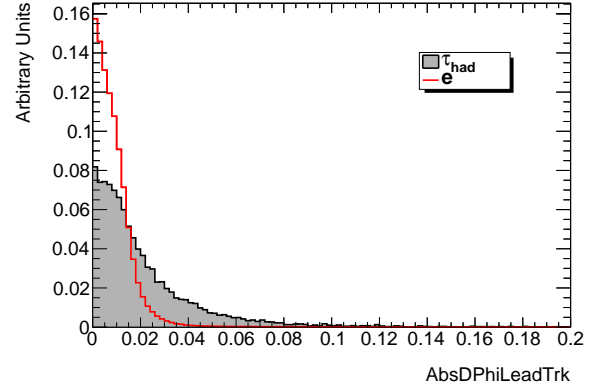
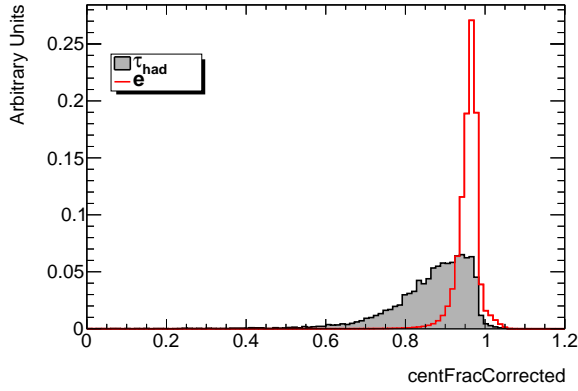(e) Secondary energy deposits in the strip compartment ($E_{T,max}^{strip}$).

Figure A.2.: Distributions of some of the identification variables for 1-prong in the η end-cap region for Run-2.
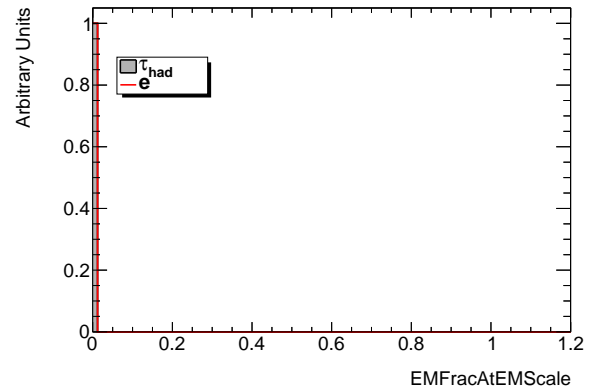
(a) Absolute distance in η between the tau candidate and its leading track.

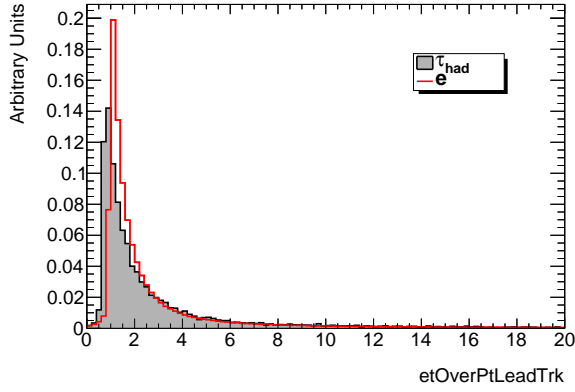(b) Absolute distance in ϕ between the tau candidate and its leading track

(c) Core energy fraction corrected for pile-up (f_{core}^{corr}).
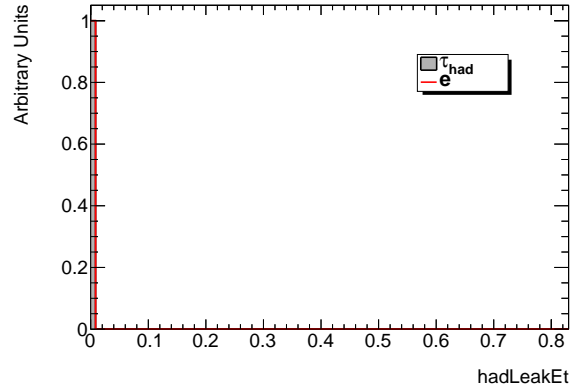
(d) Electromagnetic fraction (f_{EM}).

Figure A.3.: Distributions of some of the identification variables for 1-prong in the η forward end-cap region for Run-2.
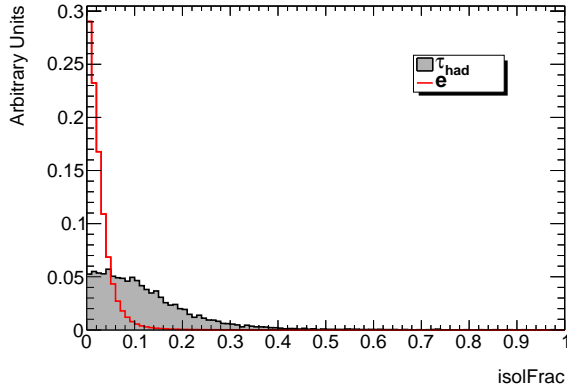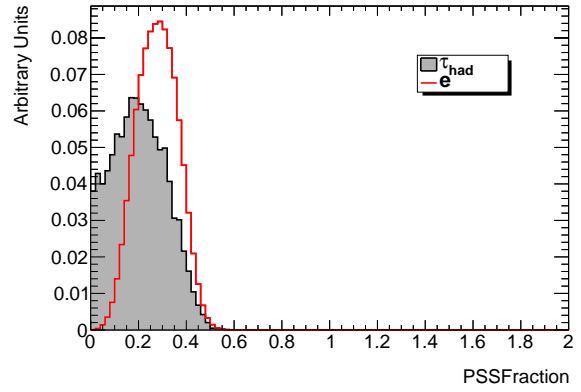
A. Identification Variable Distributions

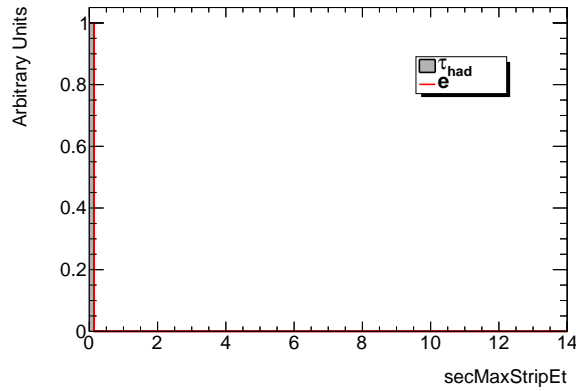(a) Electromagnetic energy over momentum of track system (f_P^{EM}).

(b) Hadronic Leakage ($E_{T,reco}^{Had/EM}$).

(c) Ring Isolation (f_{iso}).

(d) Presampler strip energy fraction (f_{PS}).

(e) Secondary energy deposits in the strip compartment ($E_{T,max}^{strip}$).

Figure A.4.: Distributions of some of the identification variables for 1-prong in the η forward end-cap region for Run-2.

Bibliography

[1] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. **145**, 1156 (1966)

[2] K. Nakamura, *et al.*, (Particle Data Group), JP G **37**, 075021 (2010) and 2011 partial update for the 2012 edition (URL: http://pdg.lbl.gov)

[3] V. M. Abazov, et al. (D0), *Measurement of the W Boson Mass with the D0 Detector*, Phys. Rev. Lett. **108**, 151804 (2012)

[4] G. Aad, et al. (ATLAS, CMS), *Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments*, Phys. Rev. Lett. **114**, 191803 (2015)

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

[6] M. Thomson, *Modern Particle Physics*, chapter 14–17, CUP (2013)

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

[8] S. Chatrchyan, et al. (CMS), *A New Boson with a Mass of 125 GeV Observed with the CMS Experiment at the Large Hadron Collider*, Science **338**, 1569 (2012)

[9] G. Aad, et al. (ATLAS, CMS), *Measurements of the Higgs boson production and decay rates and constraints on its couplings from a combined ATLAS and CMS analysis of the LHC pp collision data at $\sqrt{s} = 7$ and 8 TeV*, (2016)

[10] J. R. Andersen, et al. (LHC Higgs Cross Section Working Group), *Handbook of LHC Higgs Cross Sections: 3. Higgs Properties*, 10.5170/CERN-2013-004 (2013), `1307.1347`

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

Bibliography

[12] K. A. Olive, et al. (Particle Data Group), *Review of Particle Physics*, Chin. Phys. **C38**, 090001 (2014)

[13] A. Pich, *Precision Tau Physics*, Prog. Part. Nucl. Phys. **75**, 41 (2014)

[14] B. Mele, *Approaching a new Energy Frontier at the LHC*, Frascati Phys. Ser. **61**, 4 (2016)

[15] The ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation **3(08)**, S08003 (2008)

[16] G. Aad, et al. (ATLAS Collaboration), *Studies of the performance of the ATLAS detector using cosmic-ray muons*, Eur. Phys. J. **C71**, 1593 (2011)

[17] P. Campana, et al., *Physics Goals and Experimental Challenges of the Proton-Proton High-Luminosity Operation of the LHC*, `1603.09549`

[18] `https://twiki.cern.ch/twiki/pub/Atlas/AtlasTechnicalPaper/Main_jinst_d4.pdf`, as of 04.04.2016

[19] G. Aad, et al. (ATLAS Collaboration), *Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at $\sqrt{s}$=8 TeV*, Eur. Phys. J. **C75(7)**, 303 (2015)

[20] T. Barillari, et al., *Local Hadronic Calibration*, Technical Report ATL-LARG-PUB-2009-001-2. ATL-COM-LARG-2008-006. ATL-LARG-PUB-2009-001, CERN, Geneva (2008)

[21] The ATLAS Collaboration, *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons*, ATLAS-CONF-2011-077 (2011)

[22] The ATLAS Collaboration, *Performance of the Reconstruction and Identification of Hadronic Tau Decays in ATLAS with 2011 Data*, ATLAS-CONF-2012-142 (2012)

[23] `https://twiki.cern.ch/twiki/bin/view/Main/ElectronIDIsEM`, as of 30.06.2016

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

[25] S. Alioli, et al., *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06**, 043 (2010)

[26] S. Agostinelli, J. Allison, *Geant4a simulation toolkit*, NIM A **506(3)**, 250 (2003)

[27] `http://tmva.sourceforge.net/docu/TMVAUsersGuide.pdf`, as of 30.06.2016

[28] R. Barlow, *Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences*, volume A of *The Manchester Physics Series*, John Wiley Sons, LTD (1989)

Acknowledgments

First of all, I would like to thank Prof. Dr. Stan Lai and Prof. Dr. Arnulf Quadt for giving me the amazing opportunity to do my bachelor thesis in particle physics. It was a wonderful experience and learned a lot from it. Prof. Dr. Stan Lai was always very supportive in every moment and gave the best advice during both the analysis and the writing phases. I'm very grateful that I was able to do my bachelor thesis in his research group, it was a very interesting and incredible experience.

I want to also give a big thanks to Dr. Michel Janus, from whom I learned so much from particle physics research (and a lot more from programming too) during the bachelor thesis process. He always had a reason or solution for any problem that we encountered and was always very helpful.

I want to thank Nils too, we were lucky to be able to work together and support each other during these 14 weeks, that really helped me a lot and working with you just made it more fun!

And in general, I would like to thank all of my family, friends and teachers in Mexico, because even if we have been far away, they have always supported me and played an important role in where I am today.

For my mother:
Finalmente, quisiera agradecer a mi mamá, quien me ha apoyado incondicionalmente en todas las locas decisiones y oportunidades que he tomado desde que decidí estudiar física, además de haber sido un gran ejemplo para mi desde niña. Y aunque no sea fácil para ella que este lejos, ella entiende que es mi camino para cumplir mis metas y sueños, y esto es solo el comienzo.

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

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

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

<div align="center">

Göttingen, den December 9, 2016

(Alejandra Renee Pillado Gonzalez)

</div>