

# Optimisation of the Track Classification of Hadronically Decaying Tau Leptons at the ATLAS Detector

**Diplomarbeit**

zur Erlangung des akademischen Grades

**Diplom Physiker**

**Richard Hartmann**

geboren am 01.11.1988 in Rodewisch

TECHNISCHE UNIVERSITÄT DRESDEN

Institut für Kern- und Teilchenphysik

Fakultät Physik

Bereich Mathematik und Naturwissenschaften

2017

**Eingereicht am:** 16. Oktober 2017

**1. Gutachter:** Prof. Dr. Arno Straessner

**2. Gutachter:** Dr. Frank Siegert

## Abstract

The ATLAS experiment at the Large Hadron Collider (LHC) provides a great opportunity to search for new phenomena in particle physics. Especially in cases, where lepton universality is not realized, searches involving hadronically decaying tau leptons are of great interest. In the hadronic tau lepton reconstruction process, a track selection is applied to associate reconstructed tracks from the detector to a designated tau candidate. Since the LHC provides proton-proton collisions with multiple interactions per bunch crossing, tau decays don't appear isolated, but in coincidence with many other events, making the correct tau track association, from the pool of available reconstructed tracks, a non-trivial task. Driven by the technical improvements over the last years, the amount of interactions, and hence the number of simultaneously reconstructed tracks even increased for the 2016 data taking period. To improve the tau lepton reconstruction and identification, this thesis investigates the tau track association by using multivariate classification techniques.

## Kurzfassung

Das ATLAS-Experiment am Large Hadron Collider (LHC) bietet eine großartige Möglichkeit für die Suche nach neuartigen Phänomenen in der Teilchenphysik. Besonders wenn Leptonuniversalität nicht vorausgesetzt werden kann, genießen Suchen mit hadronisch zerfallenden Tau-Leptonen eine hohe Aufmerksamkeit. Während der Rekonstruktion von Tau-Lepton-Zerfällen wird eine Spurauswahl angewandt um im Detektor rekonstruierte Spuren einem bestimmten Tau-Zerfall zuzuordnen. Da am LHC Proton-Proton-Kollisionen stattfinden, die mehrere Interaktionen pro Strahlkreuzung liefern, treten Zerfälle von Tau-Leptonen nicht isoliert auf, sondern gleichzeitig mit mehreren anderen Ereignissen. Das macht die richtige Zuordnung von Spuren, aus der verfügbaren Ansammlung von rekonstruierten Spuren, nicht einfach. Durch technische Verbesserungen in den letzten Jahren wurde, für die Datennahme von 2016, die Anzahl an Interaktionen und damit die Anzahl an gleichzeitig rekonstruierten Spuren sogar noch erhöht. Um die Rekonstruktion und Identifikation von Tau-Leptonen zu verbessern, wird in dieser Arbeit die Spurauswahl mit Hilfe einer multivariaten Analyseverfahren, der Boosted Decision Trees, untersucht.



# Contents

<b>1. Introduction</b>	<b>1</b>
<b>2. Theoretical Foundations</b>	<b>3</b>
2.1. The Standard Model of Particle Physics	3
2.2. Extensions of the SM and New Physics	6
<b>3. Experimental Apparatus</b>	<b>9</b>
3.1. The Large Hadron Collider	9
3.2. The ATLAS Detector	11
3.2.1. The ATLAS Coordinate System	11
3.2.2. The Inner Detector	13
3.2.3. The Calorimeter System	14
<b>4. Reconstruction and Identification of Tau Leptons</b>	<b>17</b>
4.1. Tau Lepton Related Physics	17
4.2. General Track Reconstruction	19
4.3. Reconstruction of Hadronic Tau Decays	20
4.3.1. Vertex Association	20
4.3.2. Track Selection	20
4.4. Identification of Hadronic Tau Decays	21
<b>5. Multivariate Tau Track Classification</b>	<b>23</b>
5.1. Boosted Decision Trees	25
5.2. BDT in Track Classification	27
5.3. Discriminating Variables	30
5.4. Tuning the BDT Algorithm	36
5.5. Results	37
5.6. Multi-Classification BDT	38
5.7. Performance Uncertainty	39
<b>6. Summary and Outlook</b>	<b>45</b>
<b>A. Data Sets</b>	<b>47</b>
<b>B. Discriminating Variables</b>	<b>49</b>

*Contents*

---

<b>C. BDT Algorithm Options</b>	<b>57</b>
<b>Bibliography</b>	<b>59</b>

# 1. Introduction

When on 4 July 2012, the two representatives of both the ATLAS and the CMS experiment announced the discovery of the Higgs boson[1, 2], a new milestone was achieved in the understanding of the universe. This discovery is so impressive, because the existence of that new particle was already predicted about 50 years before the physical proof. Now this event lines up in a row with discoveries, such as that of the first fundamental particle, the electron, in 1897 by J.J. Thomson[3], or the first observation of the neutron in 1931[4], and all the other discoveries of fundamental particles. In most cases, these great achievements are based on the work of both theorists and experimentalists.

The theory, that currently describes particle physics in the most comprehensive way, is the Standard Model (SM) of Particle Physics. It successfully explains all observed particles and most of the existing particle phenomena. The SM combines the quantum chromodynamics of strong interaction with the quantum theory of electroweak interaction. With the Higgs mechanism, an elegant way to explain the origin of particle masses is integrated. Although the SM's validity reaches over a wide range, there are still observations that are not covered by the SM. This opens the field for new or extending theories. Above all, the group of supersymmetric theories got a lot of attention, since Supersymmetry can describe the phenomena of the SM and also can deal with some of its deficiencies. However, that has the effect that new particles have to be introduced, which haven't been observed yet. The search for predicted particles is now task of experimentalists.

In the same way, as particle theories developed from classical mechanics to quantum field theories, also the experiments underwent a drastic change. From one-man-laboratories, they evolved to huge, international facilities, with detectors of impressive size. Some of the biggest installations are the two neutrino experiments Kamiokande[5] and IceCube[6], and the Large Hadron Collider (LHC)[7]. The latter is the largest particle accelerator and collider with a centre of mass energy of 13 TeV.

The ATLAS experiment is one of the experiments located at the LHC and runs the biggest multi-purpose detector ever build. For the huge amount of data collected at the detector, sophisticated methods of data acquisition and data analysis have been developed. Especially the grid computing, which allows many scientists all over world to access and process the data, is a major improvement.

One of the main purposes of the ATLAS experiment is, among others, the search for new particles. Since tau leptons appear in decay channels of many new particles, the searches that involve tau leptons play an important role. This thesis focuses on the hadronic decay of tau leptons by investigating sophisticated reconstruction techniques.

The success of every analysis depends heavily on a good reconstruction of particles from the collected detector information. Hadronic decays of tau leptons leave a distinct signature in the detector, which is exploited in the tau reconstruction and identification. In this process, a track selection is applied to assign reconstructed tracks measured in the detector to the tau decay. The method, currently used to accomplish this, is based on simple requirements on track properties. This thesis investigates the new approach to select and classify tracks through a multivariate analysis, namely Boosted Decision Trees, which is described in Chapter 5. Beforehand, Chapter 2 gives a brief overview on the SM for the understanding of the tau lepton physics. Chapter 3 describes the ATLAS experiment and Chapter 4 the procedure of the hadronic tau reconstruction and identification. Finally, the results are summarised in Chapter 6.



## 2. Theoretical Foundations

Particle physics is the science that studies the constituents of matter and their behaviour. With the emergence of quantum physics in the beginning of the 21st century, this field evolved to one of the leading branches in modern physics and probes now in scales far below what we experience in everyday life. The smallest pieces in our universe, the fundamental particles, are considered as objects with no spacial dimensions, which cannot be further divided. The fundamental forces are, descending in accordance to their range, the gravitational, the electromagnetic, the weak and the strong force. In a microscopic regime, the influence of gravitation is hardly noticeable and thus, although part of all grand unification approaches, plays a minor role in particle physics. The Standard Model (SM) combines the other forces in one theory, describing the fundamental particles and their interactions.

### 2.1. The Standard Model of Particle Physics

The SM consists of a set of particles that are divided into bosons and fermions. Bosons have integer spin and are the mediators of the interactions, while fermions have half-integer spin and are the building blocks of matter. The fermions are further divided into leptons and quarks, with both six particles in each group. Quarks come with a colour, the charge of strong interaction, while leptons are colourless and thus, cannot interact strongly. The intrinsic charge for weak interactions is the weak isospin and for electromagnetism the electric charge<sup>1</sup>. A summary of the particle content of the SM is shown in Table 2.1. Additionally, for every particle an anti-particle exists with oppositely signed charges but the same spin and mass.

A further classification of fermions is those in “generations”. The lightest and most stable particles make up the first generation, a pair of leptons (electron and

---

<sup>1</sup>Since the charge of the strong interaction is named after colours, the theory is called chromodynamics. When not extra specified, the property “charged” always refers to the electric charge.

## 2. Theoretical Foundations

electron neutrino) and a pair of quarks (up and down quark). All stable matter in the universe is build on particles of the first generation, because the heavier particles of the second and third generation quickly decay to a more stable level. Since there are no natural sources for heavier particles like the tau lepton, they are studied in collider experiments, like the ATLAS experiment (described in the next chapter), where the necessarily high energies are available to temporarily produce these unstable particle.

	particle	electric charge $Q$	weak isospin $I_3^W$	colour charge	mass [MeV]
quarks					
$u$	up quark	$\frac{2}{3}$	$\frac{1}{2}$	r, g, b	2.2
$d$	down quark	$-\frac{1}{3}$	$-\frac{1}{2}$	r, g, b	4.7
$c$	charm quark	$\frac{2}{3}$	$\frac{1}{2}$	r, g, b	$1.27 \times 10^3$
$s$	strange quark	$-\frac{1}{3}$	$-\frac{1}{2}$	r, g, b	96
$t$	top quark	$\frac{2}{3}$	$\frac{1}{2}$	r, g, b	$173.1 \times 10^3$
$b$	bottom quark	$-\frac{1}{3}$	$-\frac{1}{2}$	r, g, b	$4.18 \times 10^3$
leptons					
$e$	electron	-1	$-\frac{1}{2}$	none	0.511
$\mu$	muon	-1	$-\frac{1}{2}$	none	105.7
$\tau$	tau	-1	$-\frac{1}{2}$	none	$1.777 \times 10^3$
$\nu_e$	electron neutrino	0	$\frac{1}{2}$	none	$< 2 \times 10^{-6}$
$\nu_\mu$	muon neutrino	0	$\frac{1}{2}$	none	$< 2 \times 10^{-6}$
$\nu_\tau$	tau neutrino	0	$\frac{1}{2}$	none	$< 2 \times 10^{-6}$
bosons					
$g$	gluon	0	0	8 colours	0
$\gamma$	photon	0	0	none	$< 1 \times 10^{-24}$
$W$	W boson	$\pm 1$	$\pm 1$	none	$80.39 \times 10^3$
$Z$	Z boson	0	0	none	$91.19 \times 10^3$
$H$	Higgs boson	0	$-\frac{1}{2}$	none	$125.1 \times 10^3$

Table 2.1.: Overview of the particles of the SM with quantum numbers and masses[8].

The SM is a relativistic gauge quantum field theory[9, 10], build upon the electroweak  $SU(2)_L \times U(1)_Y$  group and the  $SU(3)_C$  group of quantum chromodynamics (QCD). The basis is a Lagrange density  $\mathcal{L}$ , which is invariant under local gauge transformations. The equations of motions or information about the interactions of particles can all be derived from it. The invariance under local gauge transformations leads to gauge fields, whose physical equivalent are the gauge bosons.

The QCD introduces 8 different gauge bosons, the gluons. They carry a combination of two colour charges, which means, that they can interact with themselves and with quarks. A special phenomenon, concerning colour-charged particles, is the *color confinement*. In principle, quarks or gluons cannot be isolated; they clump together and form hadrons. The two main types of hadrons are the mesons (consisting of a quark and antiquark pair) and the baryons (consisting of three quarks).

The electroweak theory unifies the quantum electrodynamics (QED) with the quantum flavour dynamics (QFD) under the gauge group  $SU(2)_L \times U(1)_Y$ . The corresponding generators are the three components of the weak isospin  $I^W$  for the  $SU(2)$  and the weak hypercharge  $Y$  for the  $U(1)$ . The gauge bosons are three  $W$  bosons and a  $B$  boson, respectively. After spontaneous symmetry breaking, the gauge group transforms to  $SU(2)_L \times U(1)_Y \rightarrow U(1)_{QED}$ , with the electrical charge,  $Q$ . The relation between the charges is  $Y = 2(Q - I_3^W)$ . Also the gauge fields mix under symmetry breaking and lead to the physical gauge bosons, the  $Z$  boson and the photon,  $\gamma$ :

$$\begin{pmatrix} \gamma \\ Z \end{pmatrix} = \begin{pmatrix} \cos \theta_W & \sin \theta_W \\ -\sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} B \\ W_3 \end{pmatrix}. \quad (2.1)$$

The mixing angle<sup>2</sup>,  $\theta_W$  is one of the free parameters of the SM and follows at leading order the equation:

$$\cos \theta_W = \frac{M_W}{M_Z}. \quad (2.2)$$

The physical  $W^\pm$  bosons are the result of a combination of the two bosons  $W_1$  and  $W_2$ :

$$W^\pm = \frac{1}{\sqrt{2}} (W_1 \mp iW_2). \quad (2.3)$$

Since the QFD has a chiral symmetry and only left-handed particles take part in weak interactions, the fermions are grouped in left-handed doublets and right-handed singlets, like

$$\begin{pmatrix} \nu_l \\ l \end{pmatrix}_L, \begin{pmatrix} U \\ D \end{pmatrix}_L, l_R, U_R \text{ and } D_R, \quad (2.4)$$

with the leptons  $l = e, \mu, \tau$  and up-type quarks  $U = u, c, t$  and down-type quarks  $D = d, s, b$ . There are no right-handed neutrinos in the SM.

---

<sup>2</sup>The index in  $\theta_W$  stands for Weinberg, a major contributor to the theory of electroweak interaction.

The invariance of the Lagrangian under gauge transformations does not allow simple mass terms, so in consequence all gauge bosons and fermions are massless. This problem is solved by extending the theory with a Higgs<sup>3</sup> sector[11], which includes an additional complex scalar doublet field, the Higgs field, and a Higgs potential. The potential is shaped in such a way, that the vacuum expectation value<sup>4</sup> (VEV) is non zero, which breaks the  $SU(2)_L \times U(1)_Y$  symmetry. One of the four degrees of freedom of the Higgs field becomes the massive Higgs boson, while the other three components are absorbed to give masses to the  $W$  and  $Z$  gauge bosons. The masses of all fermions are also a consequence of the symmetry breaking since the Higgs doublet is postulated to couple to the fermions through *Yukawa* interactions. Nevertheless, mass terms for neutrinos are omitted in the SM, because the origin of their mass is not entirely resolved yet<sup>5</sup>. The Higgs boson is the only scalar boson (spin 0), while the other gauge bosons are vector bosons (spin 1). Since its coupling is mass dependent, the Higgs boson also underlies self-coupling, but does not interact directly with the massless gluons or photons.

## 2.2. Extensions of the SM and New Physics

Although the Standard Model covers most phenomena of particle physics, there are still some unsolved problems and aesthetic flaws. One of the latter is the simple question of why there are exactly that many fermions in the SM; their number is not fixed by the theory. Also, the kind of arbitrary looking mass spectrum of the particles follows no satisfactory symmetry.

Indeed, the mass of the Higgs boson bares a more important issue, the *hierarchy problem*. Not only the huge difference between the gravitation and the  $10^{24}$ [15] times stronger weak force seems unphysical, but also does the Higgs boson mass compared to the Planck mass<sup>6</sup>. According to the expectations, large quantum contributions (higher loop corrections) to the square of the Higgs boson mass would inevitably

---

<sup>3</sup>Named after Peter Higgs, who developed the Higgs mechanism along with Robert Brout, François Englert, Gerald Guralnik, Carl Hagen and Tom Kibble[11, 12, 13].

<sup>4</sup>The VEV  $v \approx 246$  GeV is well measured through the muon decay and sets the scale of the electroweak symmetry breaking.

<sup>5</sup>Another theory for neutrino masses is the Majorana neutrino theory[14].

<sup>6</sup>The Planck scale is the scale, where the effective range of the SM ends and the grand unification takes place.

make the mass huge, unless there is an incredible *fine-tuning* cancellation between the corrections and the bare mass. This fine-tuning would require a precision up to 32 decimal places.

Another issue is the existence of *dark matter*[16], which could explain particular discrepancies in astronomical observations. Dark matter is not directly proven and cannot be described by the SM. It may consist of weakly interacting massive particles, whose main interaction would be gravitation. There are, in fact, the neutrinos in the SM, which fulfil these conditions. However, the upper bounds on neutrino masses are already too low to make them appropriate candidates for dark matter. In this case, new particles are the most promising solution.

A great step towards a Grand Unified Theory is the unification of gauge couplings. In the SM, the couplings are no constants, but functions of the energy transmission  $q$  of the interaction. In case of the strong coupling, which decreases with  $q$ , this leads to phenomena like the *color confinement* and the *asymptotic freedom* (where at high energies, and only there, quarks and gluons appear quasi-free). The couplings of the SM nearly meet at very high energies, leading the way to an exact unification and superordinated symmetry.

A group of theories, which deal with the mentioned problems of the SM, is called *Supersymmetry*[17]. SUSY proposes a new space-time symmetry, that relates fermions and bosons. Each particle of the SM is associated with a superpartner, whose spin is shifted by  $\pm\frac{1}{2}$ , making bosons to fermions and vice versa. Since the other quantum properties remain the same, in particular the particle mass, super-particles should have already been observed. Thus, if SUSY exists, it must be a spontaneously broken symmetry, so that the superpartners may differ in mass. Especially masses are considerable, which could not have been investigated with recent technologies.

One specific extension of the SM is the *Minimal Supersymmetric Standard Model* (MSSM), which assumes a minimum number of new particles and interactions. Among the superpartners also an additional Higgs doublet is introduced to avoid gauge anomalies. The combined Higgs fields, again, lead to the masses of the gauge bosons, but also to several new bosons. The bosons with actual observable mass eigenstates are the 5 following Higgs bosons: two CP-even<sup>7</sup> scalar neutral Higgs bosons  $h$  and  $H$ , a CP-odd pseudo-scalar neutral  $A$  boson and two charged Higgs bosons  $H^\pm$ . One

---

<sup>7</sup>CP refers to the charge conjugation (C) and parity (P) symmetry.

## 2. Theoretical Foundations

---

of the neutral scalar Higgs bosons is associated with the SM Higgs boson, most likely the lighter  $h$ . The other neutral Higgs bosons, but also some of the superpartners, serve as possible candidates for dark matter.

Another advantage of the MSSM is the cancellation of the loop corrections from partners and superpartners, which solves the fine-tuning problem, and also a unification of gauge couplings could be realised. But so far, neither SUSY particles were found, nor were new particles from other SM extending theories.

# 3. Experimental Apparatus

## 3.1. The Large Hadron Collider

The *European Organization for Nuclear Research*, short CERN, is with 22 member states the biggest international organisation in particle physics. The main facility near Geneva (Switzerland) hosts the world's largest and most complex scientific instruments for particle physics. One of them is the *Large Hadron Collider* (LHC), a powerful particle accelerator, which started its operation in September 2008. The LHC is placed in a 27 kilometre long ring tunnel, which was former used by the Large Electron-Positron Collider (LEP). In current experiments protons and/or lead ions are accelerated and brought to collision at four interaction points; at which each, a different particle detector is located – ATLAS[18], CMS[19], ALICE[20] and LHCb[21] (see Figure 3.1).

The ALICE detector is the only heavy-ion detector at the LHC and is designed to study strongly interacting matter at extreme energy densities. The LHCb experiment focuses on the  $b$  quarks to investigate differences between matter and antimatter. ATLAS and CMS are general-purpose detectors, which both study a wide range of high energy physics by using different technical solutions.

The LHC is the last element in a chain of accelerators, in which the particles are consecutively pushed to higher energies. For proton collisions, the procedure starts with a hydrogen gas, which is ionized in an electric field to gain protons. The protons then are accelerated in a first step by the Linac 2 accelerator to an energy of 50 MeV. Then, the particle beam is injected into the Proton Synchrotron Booster (PSB), which accelerates the protons to 1.4 GeV, followed by the Proton Synchrotron (PS) (25 GeV) and the Super Proton Synchrotron (SPS) (450 GeV). Finally, the protons are transferred to the two beam pipes of the LHC. The pipes guide the protons in opposite directions through the tunnel and only intersect at the four detectors, where

### 3. Experimental Apparatus

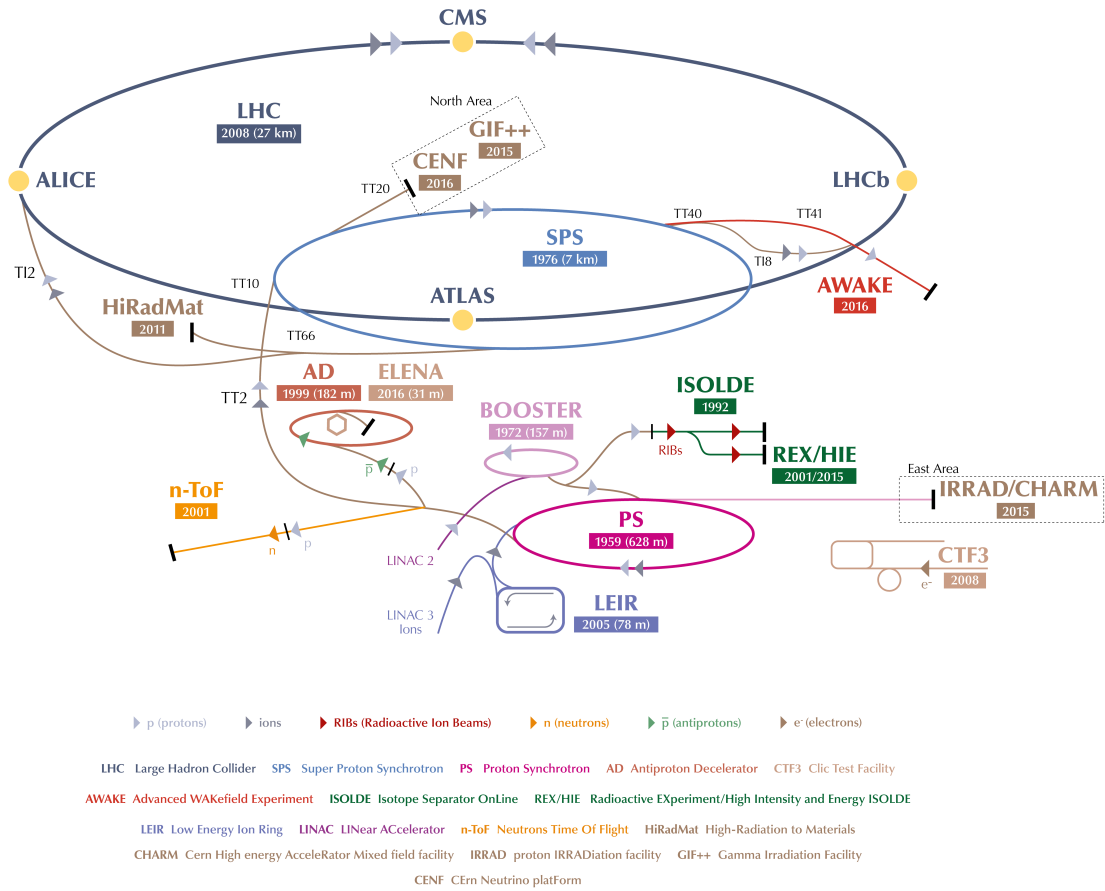


Figure 3.1.: Schematic view on the accelerators at CERN.[22]

the beams collide with a total energy of  $13 \text{ GeV}^1$ . With this high energy, the protons nearly move at the speed of light, which requires enormous technical efforts on the machinery to maintain stable beams. The tubes are kept at ultra-high vacuum and a strong magnetic field is needed to keep the beams in place. Superconducting magnets operate at the low temperature of  $1.9 \text{ K}$ , cooled down by liquid helium, to bend and focus the proton beams.

The power of the LHC can be valued with the *Luminosity*  $L$ , which is a measure of the number of potential collisions per second and per unit of area. Designed for a luminosity of  $10^{34} \text{ cm}^{-2} \text{ s}^{-1}$ , the LHC reached a record of  $1.58 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  in June 2017; made possible with the following specifications: 2556 proton bunches

<sup>1</sup>Effective for Run 2, the second data-taking period, that officially started in December 2015 and is still ongoing.



circulating in each direction of the accelerator, spaced 25 nanoseconds apart, with each one containing more than 100 billion protons[23]. This huge amount of collisions, respectively data, should give insight on even the rare physical processes.

## 3.2. The ATLAS Detector

The ATLAS detector is a general-purpose detector, build with the goal to search for new particles, as realized with the discovery of the Higgs boson in 2012[1], but also to confirm known physics with unmatched precision. Therefore, designed to detect as many particles as possibly, the detector covers almost all spatial directions around the collision point. With its length of 46 m, height and width of 25 m, the 7000-tonne ATLAS detector is one of the largest volume particle detectors ever constructed.

The detector consists of several components that are arranged cylindrical around the beam pipe, as shown in Figure 3.2. The sub-detectors are each divided in a concentrically layered barrel part and two vertically arranged end-caps at both sides of the barrel. From the centre on outwards, first comes the Inner Detector (ID), followed by the Calorimeter System. Both are described further in the following sections. The outer shell is the muon spectrometer, which occupies most of the ATLAS detector's total volume. Since muons are one of the few particles that can cross the calorimeter system without being stopped, the muon detector is an important component to also provide muon momentum measurements and tracking information. It has its own trigger system and magnetic field, maintained by three superconducting air-core magnets.

The interaction rate of 30 million collisions per second is too much to handle for the computing units of the ATLAS detector, hence, a Trigger System[24] is installed to reduce the data flow to a manageable amount. The hardware-based Level-1 trigger reduces the rate to 100 kHz, a limit given by the bandwidth of the read-out system. Then, the software-based high level trigger selects about 1000 interesting events per second, which are finally passed to the storage system.

### 3.2.1. The ATLAS Coordinate System

The ATLAS coordinate system sets its origin to centre of the beam collision and the  $z$ -axis along the beam pipe. The transverse plane is commonly spanned by the polar

### 3. Experimental Apparatus

---

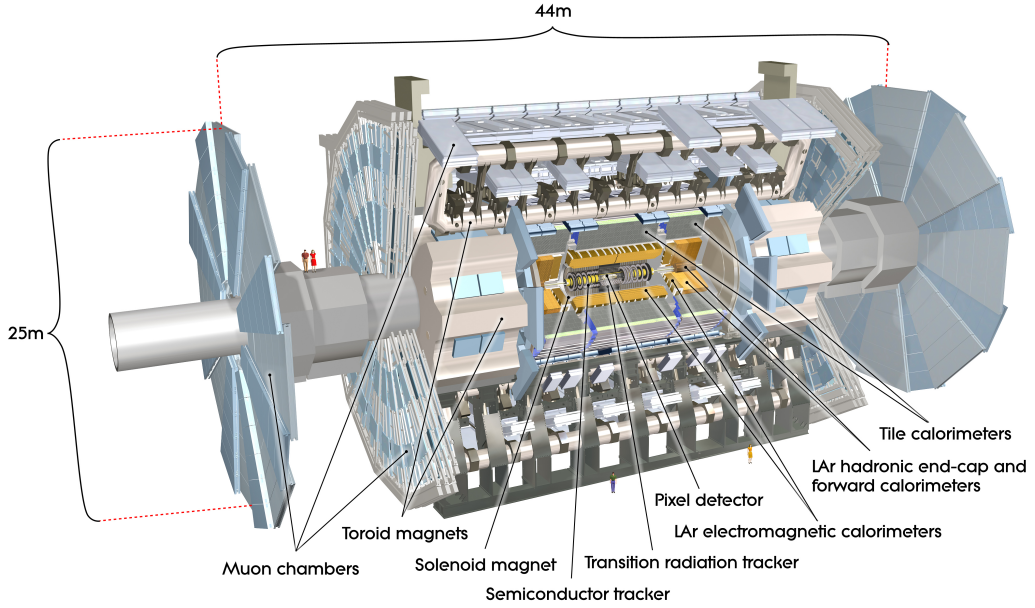


Figure 3.2.: Sketch of the ATLAS detector.[25]

coordinates  $(r, \varphi)$ , but also  $x$ -axis (showing to the ring centre) and  $y$ -axis (showing upwards) are introduced. The polar angle  $\theta$  is measured in the  $(z, r)$ -plane, but is mostly substituted by the pseudorapidity  $\eta$ :

$$\eta = -\ln \left( \tan \frac{\theta}{2} \right). \quad (3.1)$$

Angular distances are measured in the  $(\eta, \varphi)$ -plane by

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\varphi^2}. \quad (3.2)$$

The transverse plane plays an important role, because the beams are supposed to have no momentum and energy components orthogonal to the  $z$ -axis. Therefore, not measured, “missing”, transverse energy and momentum (marked by index T) can be calculated.

### 3.2.2. The Inner Detector

The innermost part of the ATLAS detector is the Inner Detector (ID), whose purpose is to measure the primary and secondary vertices of the collisions and to provide a high momentum resolution of charged particles within a pseudorapidity range  $|\eta| < 2.5$ . An electron identification is provided in  $|\eta| < 2.0$ . The ID is about 7 m long and reaches about 1.1 m in radial expanse. A sketch of the radial distances is shown in Figure 3.3. The whole block is immersed in a solenoidal magnetic field of 2 T, forcing charged particles on curved tracks, and thus, giving the possibility to derive information on momentum and charge. The ID is divided in three parts with different sensor types, the Pixel Detector, the Semiconductor Tracker (SCT) and the Transition Radiation Tracker (TRT).

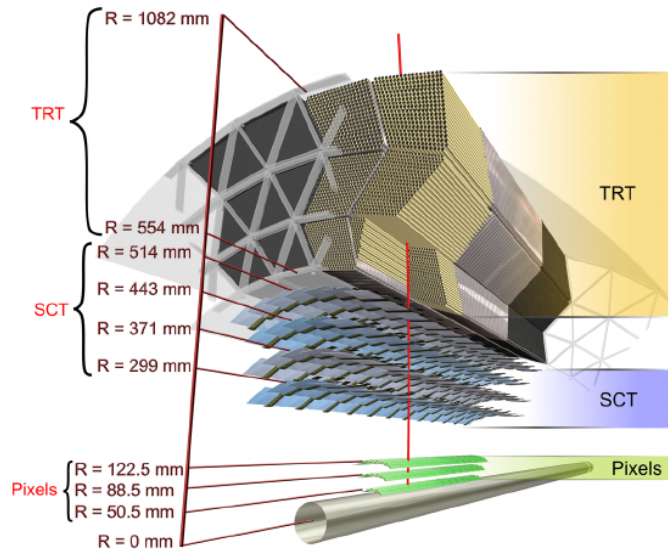


Figure 3.3.: Sketch of the barrel region of the Inner Detector. [26]

#### The Pixel Detector

The Pixel Detector directly borders the beam pipe with a minimum distance to the beam axis of 3.27 cm [27]. Four layers of pixel modules in the barrel region and four disks in the end-caps give a high tracking precision and impact parameter resolution. The pixel modules consist of silicon sensors, which are divided in pixels of typically  $2000 \mu\text{m}^2$  in size. With over 92 million read-out channels, the measurement points

reach a resolution of  $10\ \mu\text{m}$  in the  $r\text{-}\varphi$  plane and  $110\ \mu\text{m}$  in  $z$ . The innermost layer of the ID is special to the other three layers, since it consists of extra 3D silicon sensors, which further improve performance.

#### **The Semiconductor Tracker**

Similar to the Pixel Detector, the SCT also uses silicon sensors, but instead of pixel matrices, the sensors consist of wider, layered semiconductor strips. The accuracy of these sensors is  $17\ \mu\text{m}$  in the  $r\text{-}\varphi$  plane and  $580\ \mu\text{m}$  in  $z$ . The SCT has four layers in the barrel and nine disks in each end-cap, with overall about 6.3 million read-out channels.

#### **The Transition Radiation Tracker**

The outer part of the Inner Detector is the TRT, which uses wire chambers as sensors that pick up transition radiation produced by traversing charged particles. It has not the layered structure like the silicon detectors, but is filled with straw tubes of 4 mm in diameter, which provides a higher number of hits per track (approximately 36). The straws are 144 cm long in the barrel region and 37 cm long in end-caps and always aligned in a plane with the beam pipe, so only measurements in  $r\varphi$  are possible, with a precision of 130 mm. Since the TRT end-caps don't reach to the beam pipe, but enclose the end-caps of the SCT, the coverage ends with  $|\eta| < 2$ . In this range, the TRT provides a good electron identification, since the light electrons are more relativistic than e.g. hadrons and thus, produces more transition radiation.

### **3.2.3. The Calorimeter System**

The Inner Detector is surrounded by a sampling calorimeter with the purpose to absorb particles to give energy and position information. Only neutrinos and muons are supposed to pass the calorimeter system without detection. The calorimeter consists of three distinct detectors (see Figure 3.4): an electromagnetic calorimeter (EMC), a hadronic calorimeter (HC) and a forward calorimeter.

The EMC uses liquid argon (LAr) as active sampling material and lead as an absorber, which is ideally suited for measuring electrons and photons. The coverage

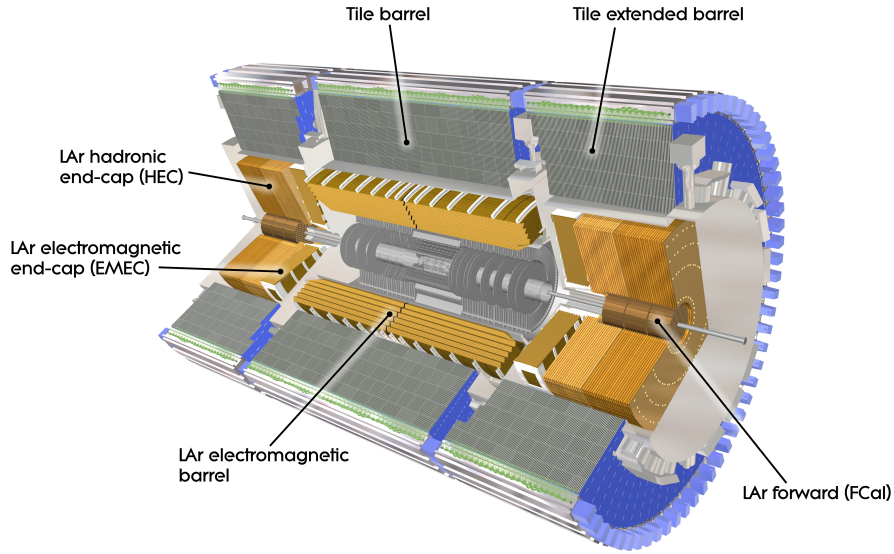


Figure 3.4.: Sketch of the Calorimeter System.[28]

of the pseudorapidity range is split up by the barrel section ( $|\eta| < 1.475$ ) and end-cap sections ( $1.375 < |\eta| < 3.2$ ).

In the end-caps of the HC ( $1.5 < |\eta| < 3.2$ ), also LAr technology is used in combination with copper absorbers. The HC barrel, with a central part and two extended barrel cylinders on each side, is made of steel/scintillator tiles and provides hadronic calorimetry up to  $|\eta| < 1.7$ .

Finally the LAr forward calorimeter extends the pseudorapidity range to  $|\eta| < 4.9$  with both electromagnetic and hadronic energy measurements, based on copper and tungsten absorbers.



# 4. Reconstruction and Identification of Tau Leptons

## 4.1. Tau Lepton Related Physics

The tau lepton is the heaviest lepton with a mass of 1776.86(12) MeV and a mean lifetime of 290.3(5) fs[8]. The corresponding proper decay length amounts to 87.03  $\mu\text{m}$  and therefore tau leptons decay within the LHC beam pipe, making the reconstruction only possible by their decay products. Due to the high mass, tau leptons cannot only decay into lighter leptons as muons and electrons but also into hadrons with a branching ratio of 65 %. These two main decay channels with their branching ratios are listed below.

$$\tau \rightarrow W + \nu_\tau \tag{4.1}$$

$$\hookrightarrow l + \nu_l + \nu_\tau \qquad 35 \% \tag{4.2}$$

$$\hookrightarrow had + \nu_\tau \qquad 65 \% \tag{4.3}$$

$l$  stands for electrons and muons,  $had$  refers generally to hadrons. Since neutrinos pass the detector nearly without interactions, the remaining lepton of a leptonic tau decay is hard to distinguish from primary leptons, which makes it difficult to use in physics analyses. This thesis shall focus on the hadronic decay of the tau leptons, and in the following, the reconstruction only considers the “visual” part of the decay, leaving the neutrinos aside.

The hadronic decay products move away from the interaction point in a conical shape and form a so-called jet (see Figure 4.1), which consists of various hadrons, mostly pions but also kaons. While the number of neutral particles is not fixed, the charge conservation requires an odd numbers of charged particles in the final state.

#### 4. Reconstruction and Identification of Tau Leptons

---

The two important decay modes are those with one charged hadron (77.2%) and with three charged hadrons (22.8%). These modes are referred to as “1-prong” and “3-prong”.

When identifying particles by their signature in the detector, a major challenge is to reject similar looking background processes. In the case of a hadron collider like the LHC, the dominant background are quark- or gluon-initiated (so-called QCD) jets. Their cross sections are many orders of magnitude greater than the cross sections for weak interactions involving leptons like the tau. The most important difference that jets of tau decays feature compared to QCD background is the mentioned 1- or 3-prong characteristic. But also, the tau jet has a more collimated shape, due to a boosting from the intermediate  $W$  boson.

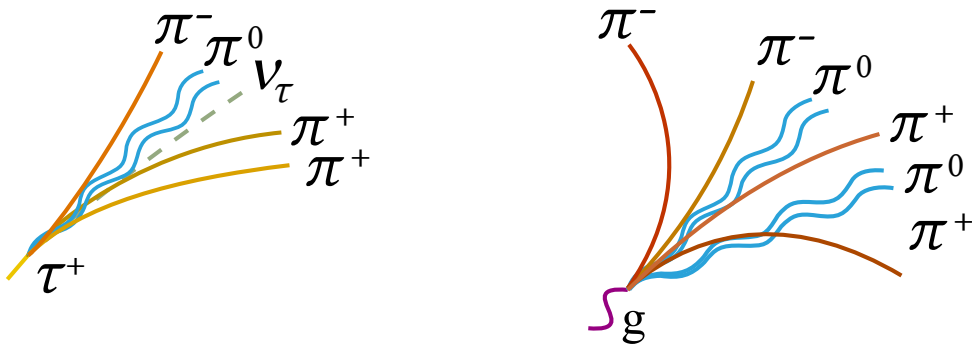


Figure 4.1.: Sketch of a possible tau jet (left) and QCD jet (right).

The reconstruction and identification of hadronic tau decays at ATLAS is accomplished with the measurements given by the Inner Detector and the calorimeter. Jet seeds, reconstructed in the calorimeter, are associated with tracks from the ID through a dedicated track selection. Since the LHC aims for high luminosities a special concern here goes to dealing with pile up. A list of identification variables is calculated including tracking and energy information. Based on multivariate analyses, the variables are used to discriminate tau jets against QCD background. The following sections further describe the reconstruction and identification.



## 4.2. General Track Reconstruction

The reconstruction of charged particles[29] starts with the information given by the Pixel and SCT detectors. A *connected component analysis*[30] builds clusters by grouping connected pixels and strips, where a sufficient amount of energy is deposited, together. These clusters, referred to as “hits”, form three-dimensional space-points, which indicate where the particle traversed the detector. When two particles occur with a spatial distance of less than a few pixels, only one merged cluster can be constructed.

In the next step, track seeds are formed by combining space points to a trajectory. At first, sets of three points are used, which provides the minimum number of necessary points to reconstruct a circular track, and also gives the maximum number of possible space point combinations. The application of various criteria on the trajectories, such as requirements on momentum and impact parameter, reduces the number of track seeds to achieve only tracks considered as good-quality. The *combinatorial Kalman filter*[31] then builds track candidates based on the seeds by including more space points into the trajectory.

Within these candidates are still tracks, which share space points, so an ambiguity-solving is necessary. The algorithm of the ambiguity solver has the following features:

**Track Score** Every track is assigned with a track score, based on how many clusters or holes are associated to the track. Holes are intersections on the trajectory, where an active sensor is located without a related cluster.

**Shared Cluster Solving** An artificial neural network (NN), specially trained to identify merged clusters, decides if a shared cluster is incorrectly assigned to multiple track candidates or if it is a merged cluster, correctly build on different tracks.

**Track Fit** On track candidates that pass basics quality criteria, a high-resolution fit is performed with all available information from the Inner Detector.

Sorted by the track score, the ambiguity solver iteratively accepts or rejects track candidates with the help of the NN and the fit. In the end, no cluster is shared by more than two tracks, and no track has more than two shared clusters. Fitted tracks that pass the ambiguity solver without modification are added to the final track collection.

### 4.3. Reconstruction of Hadronic Tau Decays

The reconstruction process of hadronic tau decays begins with the information given by the calorimeter. Topological clusters[32, 33], made of grouped calorimeter cells, serve as input for the anti- $k_t$  jet algorithm[34, 35]. Jets with a distance parameter  $R = 0.4$  are passed as tau candidates to the reconstruction algorithm, if they satisfy a transverse momentum  $p_T > 10 \text{ GeV}$  and a pseudorapidity  $|\eta| < 2.5$ , which corresponds to the coverage of the ID tracking.

The  $p_T$  of the tau candidates is set to the transverse energy  $E_T$  of the jet seed, assuming a mass of zero. The tau energy scale is determined by the energy deposition in the clusters within  $\Delta R > 0.2$  to the jet seed axis, and is also used as input for the tau identification. After the reconstruction, the energy is finally calibrated to compensate for detector defects and to correct for pile up contributions [36]. The  $\eta$  and  $\varphi$  variables of the tau candidates are also derived from the jet seeds.

#### 4.3.1. Vertex Association

In dense environments with multiple simultaneous interactions, the calculated primary vertex of the jet seed does not necessarily correspond to the actual tau decay vertex. The *tau jet vertex association* (TJVA) is the algorithm used to improve the vertex finding by choosing among track vertices corresponding to the jet. The vertex candidate with the highest  $p_T$  fraction of a set of selected tracks is set to the new jet seed vertex. The used track selection is optimised for taus and features the same criteria as described in Section 4.3.2, excluding cuts on impact parameters. The impact parameters from the tracks are recalculated with respect to the new vertex.

#### 4.3.2. Track Selection

Tracks are associated to the tau candidate if they are within the *core cone*  $\Delta R < 0.2$  around the jet axis and satisfy the following set of quality criteria.

- $p_T > 1 \text{ GeV}$ ,
- At least two hits in the Pixel Detector,
- At least seven hits in both the Pixel Detector and the SCT,

- $|d_0| < 1.0$  mm,
- $|z_0 \sin \theta| < 1.5$  mm.

$d_0$  is the distance of closest approach of the track to the jet vertex in the transverse plane, while  $z_0$  is the longitudinal distance of closest approach. Both,  $d_0$  and  $z_0$ , are recalculated by the TJVA. Tracks, which pass the criteria but lay in the *isolation annulus*  $0.2 < \Delta R < 0.4$ , are separately used for discriminating variables in the tau identification.

## 4.4. Identification of Hadronic Tau Decays

The tau reconstruction process provides a set of tau candidates that not only contains true tau jets but also fake objects. The biggest fraction of fake taus consists of QCD jets, but also electrons and muons can mimic the signature of tau jets, especially those of 1-prong taus. In the tau identification, different techniques are deployed to discriminate the tau decays from background processes. In the case of muon rejection, a cut based algorithm is used, while both the electron veto and the hadronic jet rejection apply multivariate methods, commonly *boosted decision trees* (BDT), which are described in Section 5.1. To improve the performance for hadronic jet rejection, two separate BDTs are applied for 1-prong and 3-prong decays. The BDTs are trained with Monte Carlo (MC) simulated events, using Drell-Yan  $\tau^- \tau^+$  samples for signal and di-jet samples for background. The discriminating variables[36] are derived from the tracks and the calorimeter clusters.



# 5. Multivariate Tau Track Classification

The track selection in the reconstruction process plays a crucial role, because the number of associated tracks determines the track multiplicity of the decay, and thus, which BDT is applied in the tau identification. Also, since QCD jets, unlike tau jets, tend to higher numbers of tracks, they are an important discriminating variable. Even if the tau candidate is a real tau decay, the set of tracks given by the jet seed consists of possibly more than the direct decay tracks. One group of tracks emerges through hadronic interactions in the Inner Detector, like photon conversions in decays of neutral pions. Another large contribution of tracks comes from pile up, which mainly arise from two sources: In-time pile up refers to underlying events from the same bunch crossing, while out-of-time pile up occurs when interactions from a previous or following bunch crossing overlap with the current one. The average number of interactions per crossing increased over the years with the technical improvements of the LHC, as shown in Figure 5.1, which increases the challenge on the track selection.

The current method, described in Section 4.3.2, is optimised to reconstruct tau decays with one or three charged particles. The major causes of inefficiency in reconstructing the correct track multiplicity come from merged tracks and missing hits in the ID. Both can result in not reconstructed tracks and thus in a lower number of associated tracks. On the other hand, the track multiplicity can be shifted to higher values if tracks are falsely assigned to the tau candidate. Since the tau track association takes reconstructed tracks as input, the association performance is limited by the track reconstruction efficiency. Figure 5.2 shows the efficiency to reconstruct the correct track multiplicity on simulated tau events, in red only with the track reconstruction inefficiency and in black with the track selection inefficiency on top of the reconstruction efficiency.

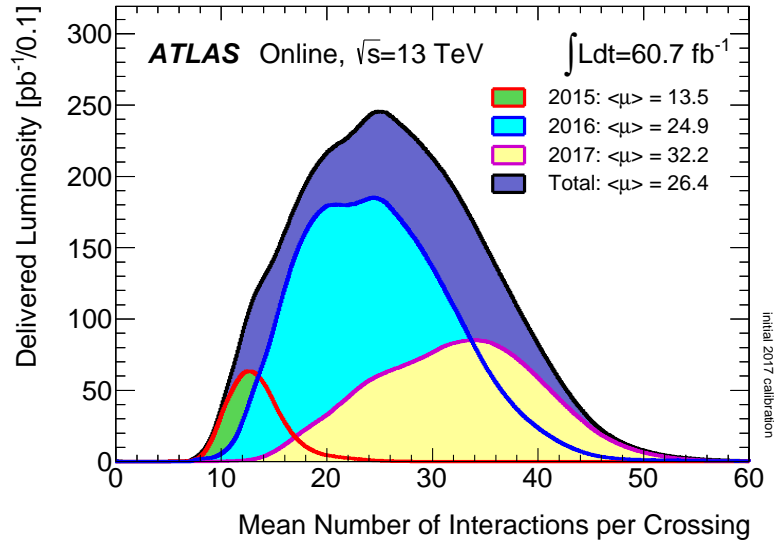


Figure 5.1.: Shown is the luminosity-weighted distribution of the mean number of interactions per crossing,  $\langle\mu\rangle$ , for pp collision data. The luminosity represents the initial estimate and includes data recorded through 22nd August 2017.[37]

The association of tracks to a tau candidate, based on the six observables mentioned in section 4.3.2, surely can be improved by taking more track relevant variables into account. For purposes like this, special multivariate computing algorithms are developed, which gained more and more attention with recent increase in computing powers. The cut based method is a fast and simple solution, but can be outperformed by more complex techniques, like Boosted Decision Trees (BDT) or neural networks. Especially when the variables are non-linearly correlated, consecutively applied cuts lack in discriminating capability.

So this work studies the implementation of BDTs in the track selection. The optimisation of this algorithm is done on simulated Monte Carlo[38] samples, with dedicated  $\gamma \rightarrow \tau\tau$  samples for training and testing. Mass-sliced Drell-Yan- $\tau\tau$  samples in combination with an inclusive low-mass  $Z \rightarrow \tau\tau$  sample are used for further testing and validation. More information on the samples can be found in Appendix A. To obtain tau events with sufficient track information, a pre-selection is applied, that requires  $p_T > 20$  GeV,  $|\eta| < 2.5$  and a matching to a true hadronic tau decay. The next section describes the BDT algorithm and then the implementation is presented.

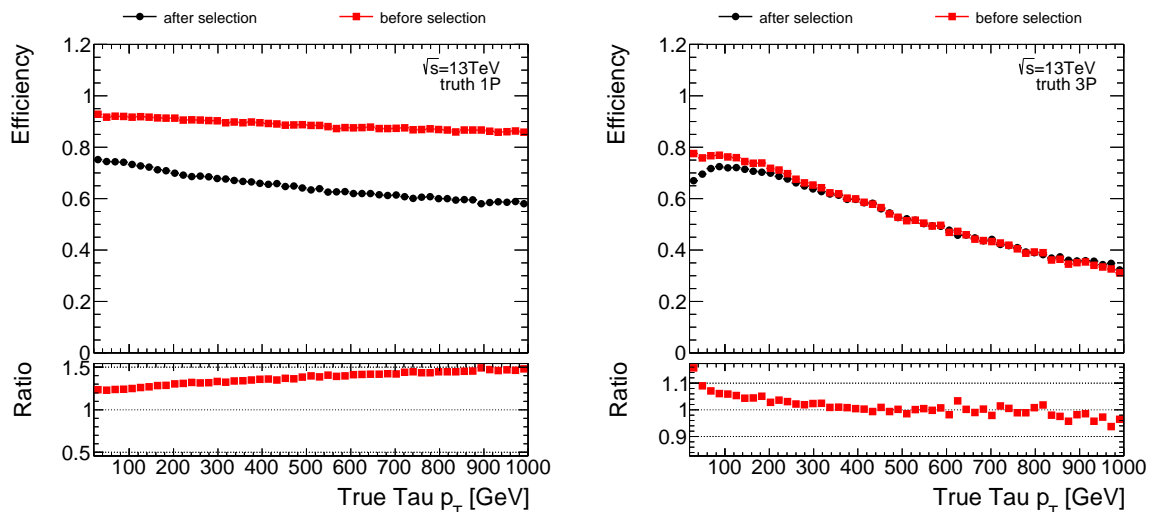


Figure 5.2.: Efficiency to reconstruct 1-prong (left) and 3-prong (right) tau decays in simulated tau events, plotted against the true tau transverse momentum. The efficiency shown in black refers to the step after the track selection, while the red dots correspond to the track reconstruction efficiency (which is equivalent to a perfect track selection).

## 5.1. Boosted Decision Trees

Boosted decision trees denotes a machine learning algorithm used for classification and decision problems. Based on a classifier, which is trained on training data, the algorithm predicts a desired response on events, that is mainly a decision between two states, for example signal and background. For this purpose, the algorithm combines several discriminating variables into one final discriminator, the BDT score. The higher the score of a particular event, the more likely it is a signal event, and the lower the score, the more background-like it appears.

The BDT algorithm, used in this study, is provided by the Toolkit for Multivariate Analysis (TMVA)[39], which itself is implemented in the scientific software framework ROOT. The available binary decision trees are structured as shown in Sketch 5.3 and define the basic routine. To classify the events of a data set as signal or background, the algorithm starts at the root node and proceeds down the tree by splitting the the data set in two branches at each internal node, whereby the splitting is achieved through a cut on one of the given variables. At the end nodes, the subsets consist, at best, of either signal or background events.

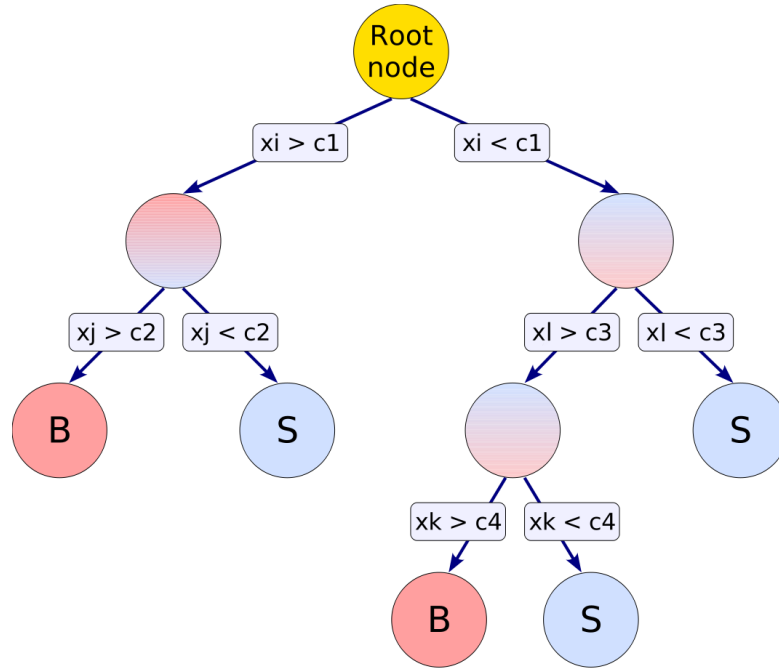


Figure 5.3.: Sketch of a binary decision tree with the root node at the top, internal nodes and the end nodes (leaves). Each node corresponds to a cut,  $c$ , on a discriminating variable,  $x$ . [39]

In the training process, the variables and cuts, on which the separation is done, are determined by maximising the separation power at each node, which is defined by the Gini Index,  $p(1 - p)$ , with the purity  $p$ , given by the ratio of signal events to all events in one node. The best separation is achieved, when the difference of the Gini Indices from the parent node to the sum of the daughter nodes is maximised. While calculating the Gini Index, the variables are scanned in discrete steps, whose maximum number is set by the parameter  $nCuts$ . The depth of a decision tree is directly limited by a  $MaxDepth$  value, but a branch also can end before this depth, when the node size falls below a minimum percentage of training events ( $MinNodeSize$ ).

The separation power of a single decision tree is rather small and they also tend to overly adapt to statistical fluctuations in the training set. The solution is to assemble many dependent decision trees in one classifier. The forest of decision trees is incrementally build with a boosting of misclassified events in between. While the initial tree sees the raw input sample, the following trees are trained on samples with adapted event weights. Misclassified events get a higher weight and correct classified



events a lower weight, so that the total weight of the sample remains the same. This emphasises the misclassified events in the growing of the forest and increases the overall separation power. Two of the available boosting algorithms are Adaptive Boosting (*AdaBoost*) and Gradient Boosting (*GradBoost*). They are defined by their loss function,  $L$ , which measures the deviation of the classifier,  $F$ , to the true value,  $y$ . Adaptive Boosting uses an exponential loss,  $L(F, y) = e^{-F \cdot y}$ , which leads to the classifier deviation

$$F_{AdaBoost}(\mathbf{x}) = \frac{1}{N_{trees}} \sum_i^{N_{trees}} \ln(\alpha_i) \cdot f_i(\mathbf{x}) , \quad (5.1)$$

with the set of input variables,  $\mathbf{x}$ , the boost weights,  $\alpha_i$ , and the single tree classifier,  $f_i(\mathbf{x})$  (with  $-1 < f_i < 1$ ). Gradient Boosting applies a more robust loss function – with respect to outliers in the data set – by using a binomial log-likelihood loss,

$$L(F, y) = \ln \left( 1 + e^{-2F \cdot y} \right) . \quad (5.2)$$

The combined classifier then results in

$$F_{GradBoost}(\mathbf{x}) = 2 \left( 1 + \exp \left\{ -2 \sum_i^{N_{trees}} f_i(\mathbf{x}) \right\} \right)^{-1} - 1 . \quad (5.3)$$

Both boosting algorithms harmonise well with small individual decision trees of depths about 2 to 4. Such small trees also provide a good robustness against over-training, which occurs when the classifier specialises on statistical fluctuations of the training sample, leading to a decreasing performance on test samples.

## 5.2. BDT in Track Classification

The track selection in the tau reconstruction is equivalent to a classification problem, where the current cut based method knows of three different track types. The main class, supposed to collect the charged tracks of the tau decay, is defined by the core cone region ( $\Delta R < 0.2$ ) and the track quality criteria. Similarly, the tracks from the outer cone must pass the selection cuts, but lie in the isolation annulus of the jet seed ( $0.2 < \Delta R < 0.4$ ). The third class consists of the tracks, that don't pass the quality criteria and are rejected in the selection process.

With the implementation of a new multivariate classification, the track classes are redefined and an additional class is introduced: the conversion tracks. While the cut based categories primarily depend on the distance parameter  $\Delta R$ , the new classes are rather build on truth information and the tracks origin, with the truth particle type and the particles barcode as main parameters. The list of track classes is described in the following list:

**Tau Tracks (TT)** are the charged tracks from the direct tau decay. They are associated to the tau candidate and serve as major input for the tracking information used in the tau identification. This class corresponds to the “core tracks” of the former cut based track selection.

**Conversion Tracks (CT)** This class contains pure photon conversion tracks, meaning electron tracks that originate from a photon conversion process. Since photon conversions in a tau decay are highly correlated with the occurrence of neutral pions, these tracks could be useful in calculations, where the neutral tau decay products are considered, like the tau energy calibration or the determination of the exact decay mode.

**Isolation Tracks (IT)** are tracks from underlying events and correspond to the tracks from the isolation annulus of the cut based selection. This class provides further information to the tau identification, since QCD jets are supposed to have more contributions in this category.

**Fake Tracks (FT)** is a combined class and includes all the other tracks that do not fit in the classes above. Considerable contributions come from pile up and secondary interactions. Reconstructed tracks, which could not be matched to a true generated track in the MC, are labelled as “unclassified” and are also put in this class.

The tracks of a tau candidate shall now be classified into the four track categories through the application of a BDT algorithm. To make use of the TMVA package’s binary BDTs, the task must be reduced to a two class problem, which requires the implementation of multiple classifiers. An approach with the least number of BDTs is a decision tree like procedure, where in a first step the tracks are divided in two intermediate classes and then in two separate steps, the final track classes are

discriminated. The choice of all possible intermediate class combinations falls to the combination of **tau tracks** with **conversion tracks** (TT+CT) and **isolation tracks** with **fake tracks** (IT+FT). This configuration, shown in Figure 5.4, yields the best performance.

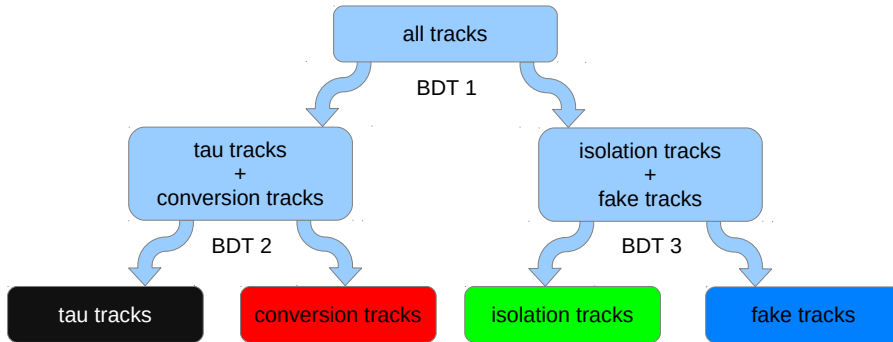


Figure 5.4.: Sketch of the track classification strategy with three BDTs.

The training process of the BDTs proceeds, in the same way as the BDTs' application, according to the following routine: All tracks are used for the training of the first BDT (BDT1), while the secondary BDTs are fed only with tracks, that are classified by BDT1 to the corresponding branch. In particular, BDT1 is trained with the combined samples TT+CT and IT+FT and then for BDT2, which distinguishes **tau tracks** and **conversion tracks**, only tracks are used that have the corresponding truth category (TT and CT) *and* are classified as TT+CT. The same applies for BDT3, respectively.

The optimisation of the BDT algorithms is done with focus on the combined performance, which is represented by the final track selection efficiency

$$\frac{\# \text{ of tau decays with matching reconstructed \& truth track multiplicity}}{\# \text{ of tau decays}}, \quad (5.4)$$

where the reconstructed track multiplicity equals the number of tracks classified as **tau tracks**. Furthermore, the actual track classification efficiency is to be maximised, which is illustrated by migration plots like Figure 5.10. They show the percentage contributions of the classified track classes to the truth track classes. To reduce complexity, the tuning for all three BDTs is done at once with the same pa-

parameters for each BDT. The choice of the discriminating variables is explained in Section 5.3 and the BDT parameter tuning in Section 5.4. The cuts on the BDT scores,  $F$ , are individually determined by maximising the score function

$$\frac{S(F)}{\sqrt{S(F) + B(F)}}, \quad (5.5)$$

with the signal efficiency,  $S$ , and the background efficiency,  $B$ . This function is a valid representation of the number of standard deviations away from zero for the signal, in case of large MC statistics, and shows the best performance among various score functions.

### 5.3. Discriminating Variables

The BDT algorithm is very robust against correlations, so several variables with also possibly redundant information can be applied as input. The Inner Detector provides a wide range of track information, including all kinds of sensor hits. The amount of variables, used for the classification, is kept at a moderate size, to achieve a feasible computing time during the training process. In order to evaluate and compare different classifiers, TMVA provides various benchmark quantities: Independent of the classifier, a separation score for each variable is calculated, which quantifies the amount of overlap of the variable distributions of signal and background. Another ranking refers to the variable importance, which measures how often a variable is used for node splitting, weighted with the separation power of the specific nodes.

Since the BDT algorithm bins the variables for the splitting procedure, it is disadvantageous when the variable distribution peaks in a narrow range, respectively in one bin. To avoid this, variable transformations can be applied. Here, variables that peak at low values are transformed with a logarithm function to stretch out the important region.

The following list introduces a set of possible discriminating variables. For a few very interesting ones, the histograms with the distributions for each tracks class are shown here, the rest of the variable distributions can be found in Appendix B. The rankings for variable importance and separation score are presented in Table 5.3, as

well as in Tables B.1 and B.2. In Figure 5.7, a matrix with the linear correlation coefficients of the variables for the `tau` tracks is shown.

### Distance Parameter: $\Delta R$

The distance of tracks to the jet seed axis,  $\Delta R$ , is an important variable for distinguishing tracks of a tau decay from those of background jets, since the tau jet cone is assumed to have a tighter shape, resulting in smaller  $\Delta R$  for the tau tracks. The distribution is shown in Figure 5.5.

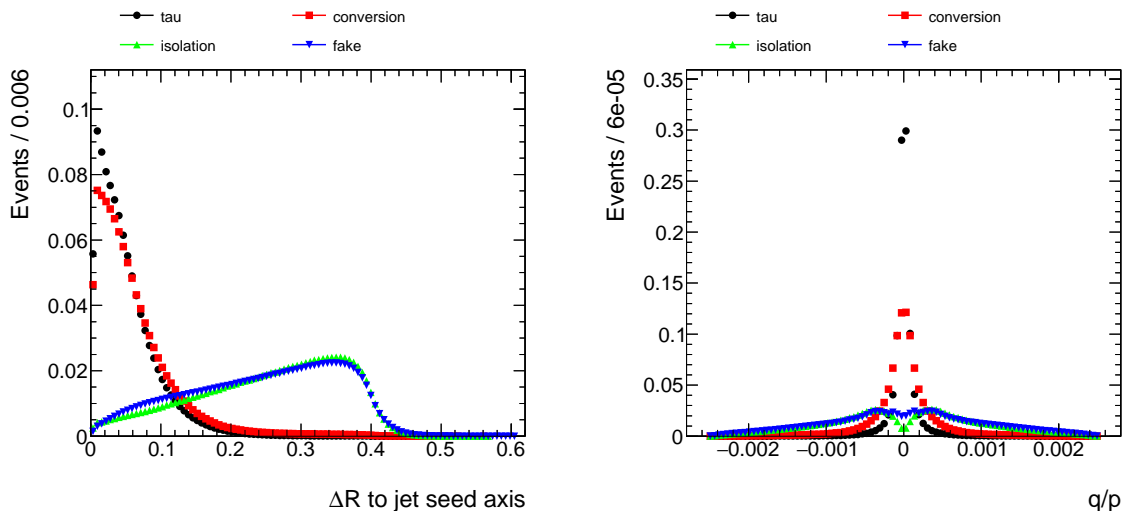


Figure 5.5.: Normalised distribution of the distance of tracks to the jet seed axis,  $\Delta R$ , shown on the left. The right plot shows the distributions for  $q/p$ . The different colours correspond to the track classes: tau tracks (black), conversion tracks (red), isolation tracks (green) and fake tracks (blue).

### Inverse Track Momentum: $q/p$

The charge of a track,  $q$ , is either  $\pm 1$ , so  $q/p$  equals to the inverse track momentum with the sign of the charge. The distribution is shown in Figure 5.5. As can be seen in Table B.2, this variable provide good separation power in BDT3, which distinguishes isolation tracks and fake tracks.

### Impact Parameter: $\ln |d_0|$

$d_0$  is the distance of closest approach of the track to the primary vertex in the transverse plane. The distribution is shown in Figure 5.6. Secondary interactions in the tau decay generally appear after the particles flew some distance,

## 5. Multivariate Tau Track Classification

which applies for `conversion tracks` and parts of the `fake tracks`. The right peak in the `fake tracks`' distribution most likely comes from secondary hadronic interactions of the tau decay products in the detector.

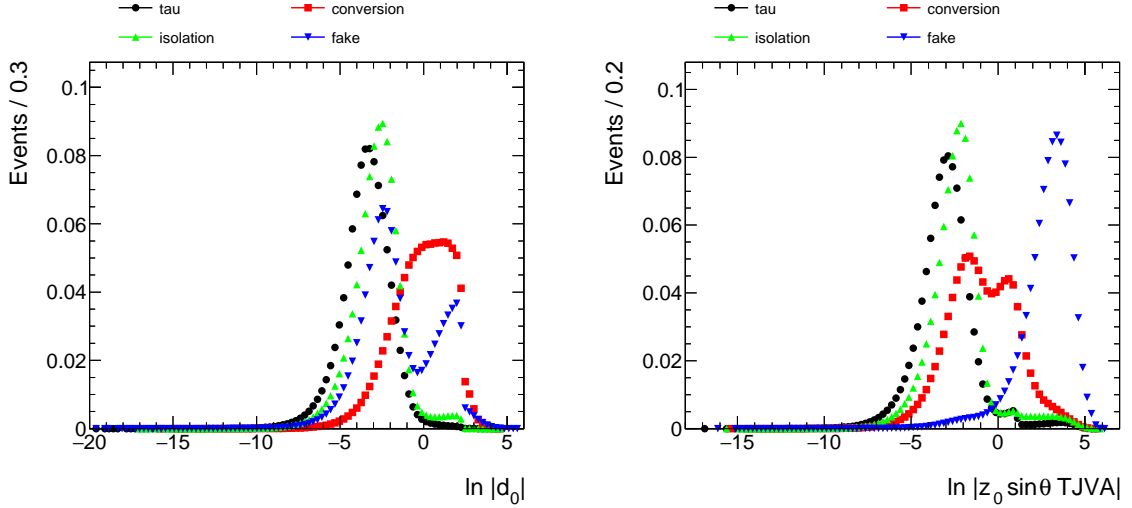


Figure 5.6.: Normalised distributions of the impact parameters,  $\ln |d_0|$  (left) and  $\ln |z_0^{\text{TJVA}} \sin \theta|$  (right). The different colours correspond to the track classes: `tau tracks` (black), `conversion tracks` (red), `isolation tracks` (green) and `fake tracks` (blue).

### Impact Paramter: $\ln |z_0^{\text{TJVA}} \sin \theta|$

$z_0$  is the longitudinal distance of closest approach of the track to the primary vertex. Since this variable is very sensitive to the correct association of the tau decay vertex, the recalculated  $z_0^{\text{TJVA}}$  is used (for TJVA see Section 4.3.1). The distribution is shown in Figure 5.6. The fact that `fake tracks` don't originate from the tau lepton is descriptively reflected in this distribution, since these tracks have a greater distance to the tau vertex.

### Pseudorapidity: $\eta$

The pseudorapidity shows good separation power for `tau tracks` and `conversion tracks`, as shown in Table B.1. Since with higher pseudorapidity, and respectively lower angle to the beam pipe, the neutral pions transverse more material, the possibility of photon conversion increases. This leads to the tendency of `conversion tracks` to higher values for  $\eta$ .

**Conversion Radius I:  $R_{\text{conv}}^{\text{I}}$** 

The conversion radius[40] is intended to classify `conversion tracks` and calculates as:

$$R_{\text{conv}} = \sqrt{\frac{|d_0| \cdot p_{\text{T}}}{0.3}}. \quad (5.6)$$

However, a better separation power provides the transformed variable

$$R_{\text{conv}}^{\text{I}} = \ln R_{\text{conv}}.$$

**Conversion Radius II:  $R_{\text{conv}}^{\text{II}}$** 

$$R_{\text{conv}}^{\text{II}} = \begin{cases} R_{\text{conv}}^{\text{I}} & \text{for } d_0 \cdot q > 0 \\ -R_{\text{conv}}^{\text{I}} & \text{otherwise} \end{cases} \quad (5.7)$$

Just like the  $R_{\text{conv}}^{\text{I}}$ , this variable is considered to classify `conversion tracks`, but yields only a moderate separation power.

**Transverse Momentum of the Jet Seed:  $p_{\text{T}}^{\text{jetSeed}}$  and  $\ln p_{\text{T}}^{\text{jetSeed}}$** 

These are the only non-track-specific variables and although they have small separation scores, their distributions differ slightly for `conversion tracks`, giving separation power in BDT2. The main purpose of this variables is however to provide correlations between the discriminating variables. A tau jet with higher momentum is supposed to have a narrower cone, leading to smaller  $\Delta R$  of the tracks. Another consequence of an increasing tau lepton energy is the merging of tracks, which results in a decrease of hits in the ID, shown in Figure 5.7. In some cases, the tau lepton decays even after passing the first layer of the Pixel Detector. For those events, the decay tracks don't have hits in the Innermost Pixel Layer.

**Electron Probability  $P_e$** 

Based on the information from the TRT, this variable is calculated to give the probability, that the track comes from an electron, and so it is predestined for the classification of `conversion tracks`.

**Number of Expected Hits in the Pixel Detector:  $n_{\text{PixHits}}$** 

This variable represents the number of hits, that a particle would leave in the

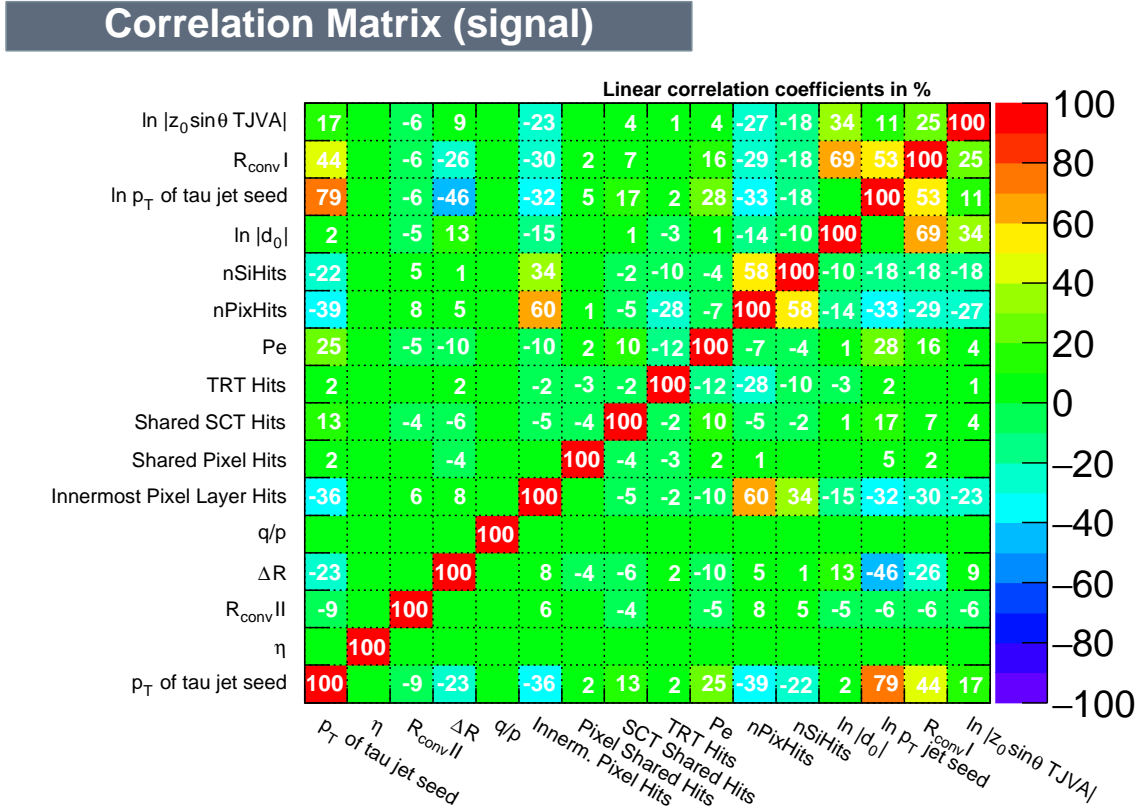


Figure 5.7.: Shown are the linear correlations for the variables of the tau tracks. The plot is provided by TMVA.

Pixel Detector, when all sensors worked properly. However, some sensors loose functionality and are labelled as dead sensors.  $nPixHits$  is the sum of hits in the pixel layers and the number of crossed dead sensors.

**Number of Expected Hits in the Silicon Detectors:  $nSiHits$**

$nSiHits$  is the sum of hits from the Pixel Detector and the SCT, plus the crossed dead sensors from both sub-detectors. Similar to  $d_0$ , this variables indicates tracks, that originate from further away of the tau decay vertex, probably outside the SCT. So a considerable amount of fake tracks have no hits the silicon sensors.



Variable	Separation Score	Importance
$\Delta R$	0.6743	0.13640
$R_{\text{conv}}^{\text{I}}$	0.378	0.09596
$\ln  d_0 $	0.045 93	0.08122
$\ln \left  z_0^{\text{TJVA}} \sin \theta \right $	0.2209	0.08084
$\ln p_{\text{T}}^{\text{jetSeed}}$	0.003 716	0.07063
$q/p$	0.4964	0.06911
$n\text{SiHits}$	0.075 25	0.06762
$P_e$	0.1385	0.06196
$\eta$	0.032 21	0.05998
Number of Hits in the TRT	0.025 84	0.05141
$R_{\text{conv}}^{\text{II}}$	0.1308	0.05124
$n\text{PixHits}$	0.074 11	0.04944
Number Of Innermost Pixel Layer Hits	0.095 88	0.04633
$p_{\text{T}}^{\text{jetSeed}}$	0.002 709	0.02814
Number of Shared Hits in the Pixel Detector	0.013 29	0.02613
Number of Shared Hits in the SCT	0.016 78	0.02356

Table 5.1.: Shown are the separation scores and variable importances for each discriminating variable, provided by TMVA. The values belong to BDT1 (TT+CT vs. IT+FT) of the track classification presented in Section 5.5; for BDT2 and BDT3 see Appendix B. The list is ranked according to the importance in descending order.

The following variables count specified hits in particular sub-detectors. Compared to the other variables, they have the lowest separation power but are still useful.

- Number of Hits in the Innermost Pixel Layer
- Number of Shared Hits in the Pixel Detector  
Shared hits correspond to sensors, which are triggered by several tracks (see Section 4.2).
- Number of Shared Hits in the SCT
- Number of Hits in the TRT

tau tracks	25 014 795
conversion tracks	8 845 009
isolation tracks	11 290 383
fake tracks	165 368 923

Table 5.2.: Number of tracks in the training data sets.

## 5.4. Tuning the BDT Algorithm

Boosted Decision Trees possess a variety of parameters, which have to be adapted to the specific problem. The main parameters are already introduced in Section 5.1 and concern the size of the trees and the forest, the boosting and the node splitting. The parameters that are tuned for the track selection are described in the following. Other noteworthy algorithm options, that are used with their default values are listed in Table C.1 in the appendix and the complete list is documented in the TMVA's user guide[39].

The sizes of the training data sets for the different track classes are listed in Table 5.2 and indicate the rate of their occurrence in the tau decay. Tests show, that the discriminating features in the variable space are already sufficiently covered, when using smaller sample sizes than the maximum available. Trainings with 2.5 million tracks per class provide similar results as trainings with more tracks and therefore the BDT options are optimised with samples of 2.5 million tracks in size.

The size of the decision trees is kept low in combination with a large tree collection of several hundred trees. Under this condition, *MinNodeSize*, *NTrees* and *MaxDepth* are optimised. The node splitting in all set-ups is done with the Gini Index as the separation criterion, and the granularity with which the variables are scanned, given by *nCuts*, is adjusted from the default, 20, to higher values. This should improve especially the separation power in the root node, since here the entire variable range is considered. It also helps dealing with variables that have narrow peaks in their distribution.

As an example for the optimisation, Figure 5.8 shows the tau reconstruction efficiencies for different BDT configurations. On the left side, the number of trees are varied in the steps 200, 850, 1000 and 1500. The set-up with 1000 trees yields the best performance, considering the complete tau  $p_T$  range up to 2 TeV. The maximum variation in efficiency only slightly exceeds one percent. The improvement from

the baseline, set at 850 trees, to the set-up with 1000 trees is about 0.1–0.7%. For the *MaxDepth* parameter, the value 5 is preferable to 3, since the former improves reconstruction in a wider momentum range.

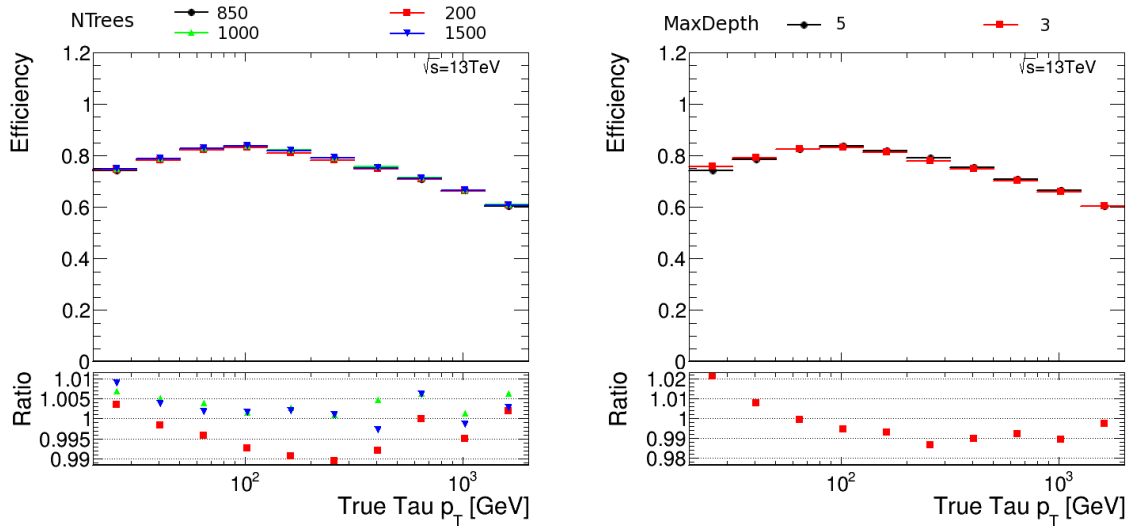


Figure 5.8.: Efficiency to reconstruct tau decays with the correct track multiplicity, plotted against the true tau transverse momentum in a logarithmic scale. Different colours correspond to different BDT configurations: on the left with varied *NTrees*, and on the right with varied *MaxDepth*.

Finally, the BDT configuration with the best performance is chosen and the BDT options are listed in Table 5.3. The results for the track selection using this set-up is presented in the next section.

## 5.5. Results

In this section, the results are presented for the the track selection using three BDTs as illustrated in Figure 5.4. The plots in Figure 5.9 show the efficiency to reconstruct tau decays with the track multiplicity matching the truth value. For tau decays with one charged track, an improvement is achieved over the entire spectrum of the tau transverse momentum, which extends up to 40% for high  $p_T$ . The rather small improvement in the very low  $p_T$  region, is mainly caused by an inefficient tau vertex association. This method even has to suffer a loss in efficiency for 3-prong tau decays. Here the efficiency drop in the mid- $p_T$  region is around 8%. Considering that 1-prong

Option	Value
AdaBoostBeta	0.5
BaggedSampleFraction	0.5
BoostType	AdaBoost
MaxDepth	5
MinNodeSize	0.75
NTrees	1000
NegWeightTreatment	IgnoreNegWeightsInTraining
UseBaggedBoost	true
nCuts	100

Table 5.3.: BDT algorithm options.

tau decays appear more than three times as often as 3-prong decays, the combined improvement is predominantly positive, with an increase in efficiency in the range of 3 %.

The performance of the track classification is illustrated in the migration matrices in Figure 5.10. The overall reconstruction efficiency, defined by the combination of both diagonals for 1-prong and 3-prong, evaluates to 86.8 %. The important part, the correct classification of `tau tracks`, has a maximum in 1-prong tau decays with 94.6 %.

## 5.6. Multi-Classification BDT

The TMVA framework offers a special multi-classification option for Boosted Decision Trees, which provides several training input channels, as well as a classifier with multiple output scores. The algorithm is technically based on binary classification trees, but combines the training of multiple binary classifications in one process, by applying a “one against the rest” classification for each input class. As a consequence, the classifier consists of a number of trees, that is the product of the *NTrees* parameter and the number of classes. In a consecutive manner, the training first builds a tree for each class, then calculates the boosting and then builds the following trees. Since Adaptive Boosting only supports real binary classifications, the AdaBoost option is not available and the Gradient Boosting is used.

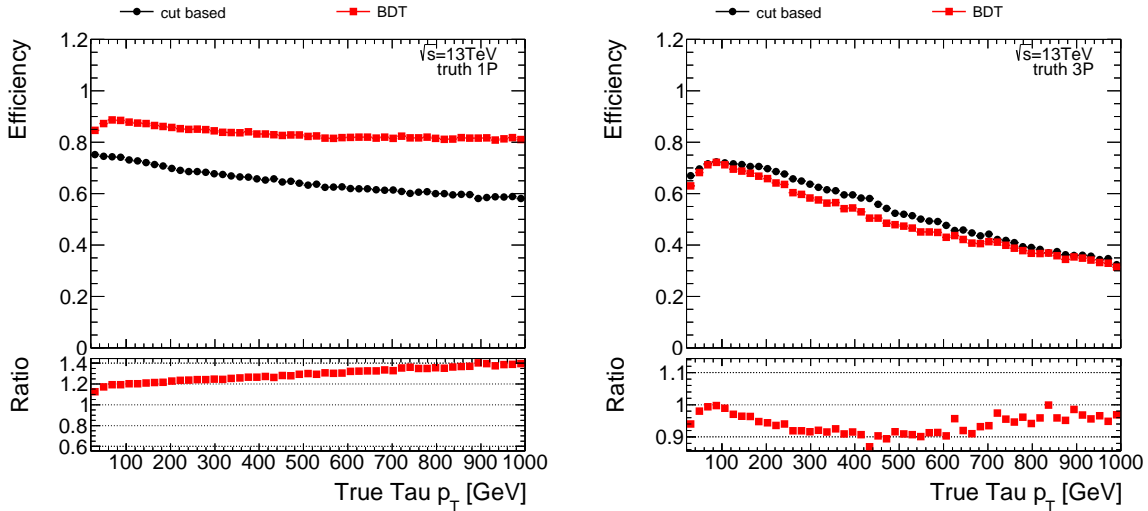


Figure 5.9.: Efficiency to reconstruct 1-prong (left) and 3-prong (right) tau decays on simulated tau events, plotted against the true tau transverse momentum. The efficiency shown in black refers to the cut based method, while the red dots correspond to the track selection with the BDT set-up.

The track selection with the multi-classification BDT leads to the results shown in Figure 5.11 and 5.12. Here, besides the boosting, the same BDT parameters are used as in the two-tiered BDT set-up. The parameters are listed in Table 5.4.

The multi-classification shows similar improvements in the 1-prong decay mode as the two-tiered method, but moreover also yields a higher efficiency in the 3-prong reconstruction than the cut-based method. This is also reflected in the migration matrices, which show a correct classification of tau tracks with 96.9% in 1-prong and 95.9% in 3-prong.

## 5.7. Performance Uncertainty

When tuning the BDTs, the difference in the efficiencies between two set-ups may only exceed some tenth of a percent. In order to validate the results, it is necessary to check, if these improvements are covered by uncertainties. The statistical uncertainties are negligible and in fact the statistical error bars in the efficiency plots are smaller than the data points. For the systematic uncertainties, this thesis investigates the influence of the uncertainties of the tracks' impact parameters,  $d_0$  and  $z_0$ ,

## 5. Multivariate Tau Track Classification

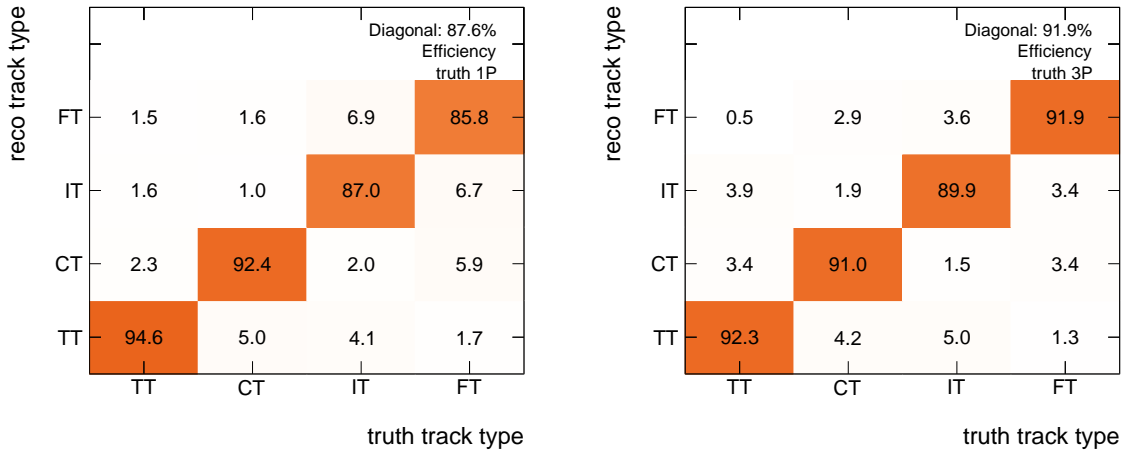


Figure 5.10.: The migration matrices show the fractions of the truth track classes to the reconstructed track categories, classified with the two-tiered BDT set-up.

on the track selection efficiency. This is accomplished by applying the track selection on tracks, which are smeared in their impact parameters.

The smearing recalculates the variables by taking a Gaussian random value with the given variable value as mean and an uncertainty as standard deviation, which is derived from uncertainty systematics for detector material and dead sensors. The systematics for dead sensors are dependent on the tracks' transverse momentum, while the material systematics are dependent on the transverse momentum and the

Option	Value
BaggedSampleFraction	0.5
BoostType	Grad
MaxDepth	5
MinNodeSize	0.75
NTrees	1000
NegWeightTreatment	IgnoreNegWeightsInTraining
UseBaggedBoost	true
nCuts	100

Table 5.4.: BDT Algorithm Options.

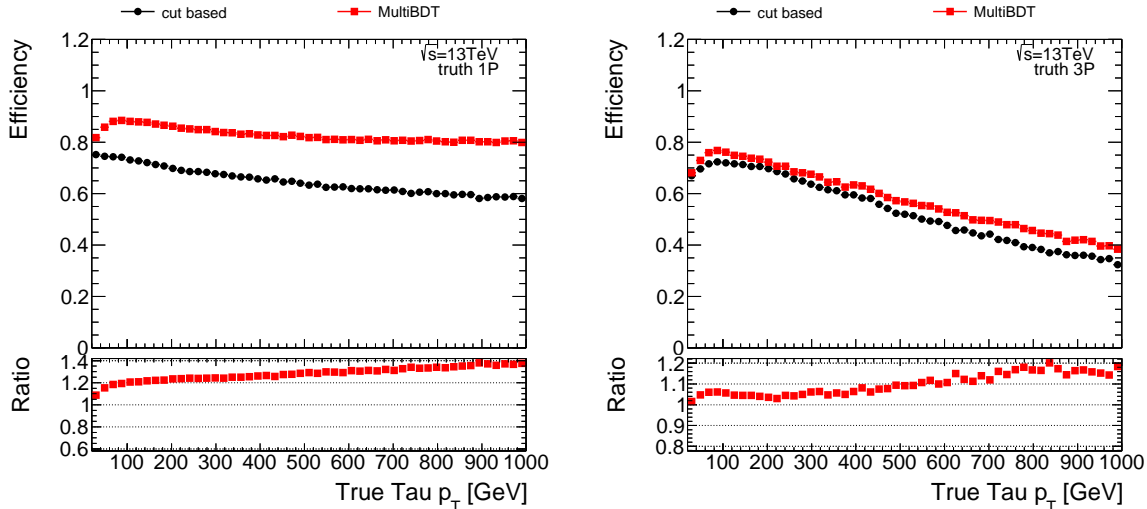


Figure 5.11.: Efficiency to reconstruct 1-prong (left) and 3-prong (right) tau decays on simulated tau events, plotted against the true tau transverse momentum. The efficiency shown in black refers to the cut based method, while the red dots correspond to the track selection with the multi-classification BDT.

pseudorapidity. To get an impression of their sizes, Table 5.5 lists a few examples of the standard deviations.

	$p_T = 1 \text{ GeV}$	$p_T = 20 \text{ GeV}$
$\sigma_{z0\_dead}$	0,0035	0,0009
$\sigma_{z0\_meas}$	0,0498	0,0370
$\sigma_{d0\_dead}$	0,0013	0,0003
$\sigma_{d0\_meas}$	0,0340	0,0059

Table 5.5.: This table shows chosen examples of the standard deviations  $\sigma$  for the track resolution smearing, derived from detector material uncertainties,  $\sigma_{meas}(p_T, \eta = 0)$ , and uncertainties through dead sensors,  $\sigma_{dead}(p_T)$ .

The smearing of the impact parameters affects the following discriminating variables, which also have to be recalculated:

- $\ln |d_0|$ ,
- $\ln |z_0^{\text{TJVA}} \sin \theta|$ ,
- $R_{\text{conv}}^I$ ,

## 5. Multivariate Tau Track Classification

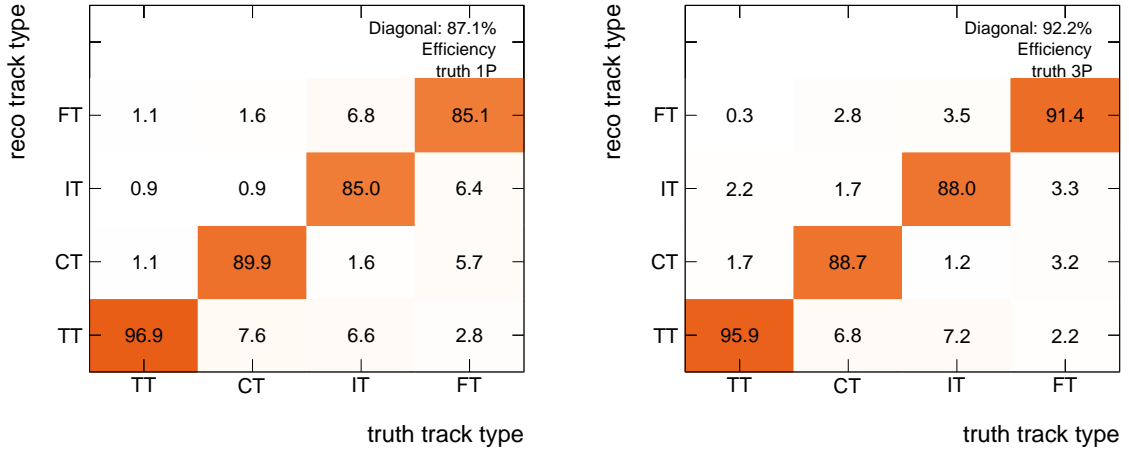


Figure 5.12.: The migration matrices show the fractions of the truth track classes to the reconstructed track categories, classified with the multi-classification BDT.

- $R_{\text{conv}}^{\text{II}}$  .

The track selection efficiency with smeared tracks and the BDT set-up, described in Section 5.6, is shown in Figure 5.13. The difference between the efficiencies for the original tracks and the smeared tracks is in the range of permille. So based on the uncertainties on the track's impact parameters, the improvements are in a relevant scale.



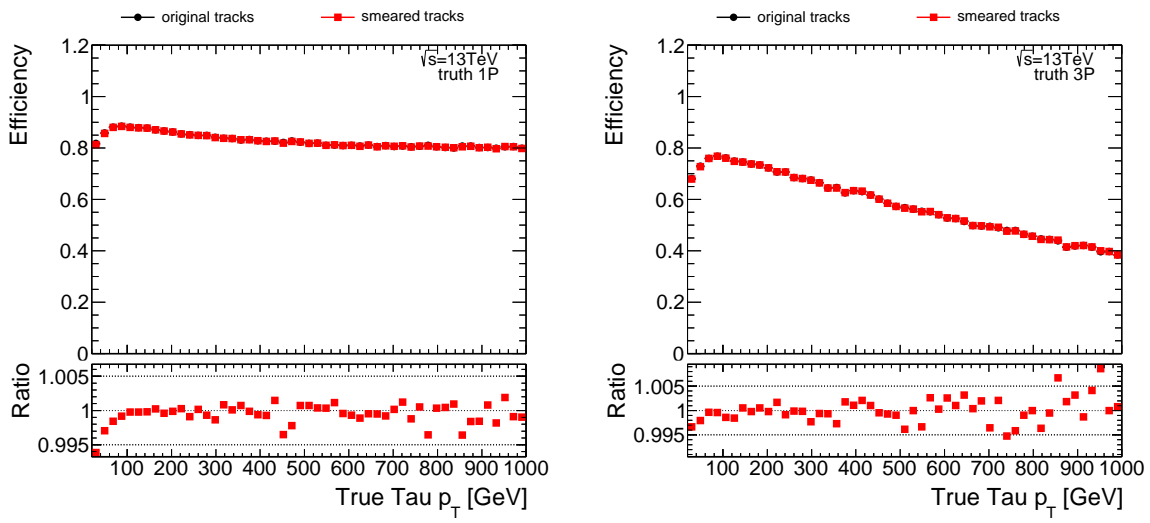


Figure 5.13.: Efficiency to reconstruct 1-prong (left) and 3-prong (right) tau decays on simulated tau events, plotted against the true tau transverse momentum. The efficiency shown in black refers to the track selection with the original track samples, while the red dots correspond to the track selection with smeared tracks.



## 6. Summary and Outlook

An efficient reconstruction of particles by their signatures in the detector is a major key point in particle physics analyses. The efficiency to reconstruct tau leptons is mostly limited by the track reconstruction, but is also reduced by the inefficiency of the track selection. To improve the tau lepton reconstruction, this thesis investigates the track selection in the reconstruction process with Boosted Decision Trees (BDT). BDTs are machine learning algorithms that need, for the training, truth information about the categories, which they have to classify. In particle physics, a common way to acquire such data samples are so called Monte Carlo Simulations. For the tau decay reconstruction, dedicated samples provided by the ATLAS collaboration are used for training and testing of the classifiers.

The purpose of the track selection is the association of tau decay tracks to the tau candidate while rejecting tracks from underlying events. Therefore, the cut based track selection currently classifies tracks in **tau tracks**, **isolation tracks** and **fake tracks**. In a new approach, BDT based classifiers are applied to categorise the following four track classes: **Tau tracks** are assigned to the tau candidate and are used together with **isolation tracks** as input for the tau identification. The new class of **conversion tracks** can also be used for that purpose. **Fake tracks** are rejected and play no further role. Two different set-ups for the track classification are presented in this thesis. The first one is a two-tiered classification utilizing three simple BDTs. The second one uses a single multi-classification BDT. Both set-ups are optimised for tau track selection and show improvements in the reconstruction efficiency for tau decays with one charged particle (1-prong). Compared to the cut based method, the efficiency increases by 11%. In the 3-prong decay mode, the two-tiered set-up loses efficiency, while the multi-classification BDT shows a promising improvement of 3.5% increase in tau reconstruction efficiency. Overall the classification of **tau tracks** yields good results with efficiencies of about 95%.

When comparing multivariate analyses, the more complex algorithms like Boosted Decision Trees or neural networks may lack in transparency and the physical motivation for the input variables appears not that clear. But exactly because of the more complex structure, the performance usually exceeds those of simple cut methods, especially when dealing with non-linear correlations between input variables. Ongoing and future investigation with neural networks in the track selection may even outperform the results shown in this thesis.

## A. Data Sets

The Boosted Decision Trees are trained and tested on Monte Carlo[38, 41] generated events samples. For the training  $\gamma \rightarrow \tau\tau$  samples are used and the testing is done on  $Z \rightarrow \tau\tau$  and Drell-Yan- $\tau\tau$  samples. The exact sample names are shown in Table A.1.

---

mc16_13TeV.361108.PowhegPythia8EvtGen_AZNLOCTEQ6L1_Ztautau.merge.AOD.e3601_s2997_r8903_r8906
mc16_13TeV.301040.PowhegPythia8EvtGen_AZNLOCTEQ6L1_DYtautau_120M180.merge.AOD.e3649_s2997_r8903_r8906
mc16_13TeV.301041.PowhegPythia8EvtGen_AZNLOCTEQ6L1_DYtautau_180M250.merge.AOD.e3649_s2997_r8903_r8906
mc16_13TeV.301042.PowhegPythia8EvtGen_AZNLOCTEQ6L1_DYtautau_250M400.merge.AOD.e3649_s2997_r8903_r8906
mc16_13TeV.301043.PowhegPythia8EvtGen_AZNLOCTEQ6L1_DYtautau_400M600.merge.AOD.e3649_s2997_r8903_r8906
mc16_13TeV.301044.PowhegPythia8EvtGen_AZNLOCTEQ6L1_DYtautau_600M800.merge.AOD.e3649_s2997_r8903_r8906
mc16_13TeV.301045.PowhegPythia8EvtGen_AZNLOCTEQ6L1_DYtautau_800M1000.merge.AOD.e3649_s2997_r8903_r8906
mc16_13TeV.301046.PowhegPythia8EvtGen_AZNLOCTEQ6L1_DYtautau_1000M1250.merge.AOD.e3649_s2997_r8903_r8906
mc16_13TeV.301047.PowhegPythia8EvtGen_AZNLOCTEQ6L1_DYtautau_1250M1500.merge.AOD.e3649_s2997_r8903_r8906
mc16_13TeV.301048.PowhegPythia8EvtGen_AZNLOCTEQ6L1_DYtautau_1500M1750.merge.AOD.e3649_s2997_r8903_r8906
mc16_13TeV.301049.PowhegPythia8EvtGen_AZNLOCTEQ6L1_DYtautau_1750M2000.merge.AOD.e3649_s2997_r8903_r8906
mc16_13TeV.301050.PowhegPythia8EvtGen_AZNLOCTEQ6L1_DYtautau_2000M2250.merge.AOD.e3649_s2997_r8903_r8906
mc16_13TeV.301051.PowhegPythia8EvtGen_AZNLOCTEQ6L1_DYtautau_2250M2500.merge.AOD.e3649_s2997_r8903_r8906
mc16_13TeV.301052.PowhegPythia8EvtGen_AZNLOCTEQ6L1_DYtautau_2500M2750.merge.AOD.e3649_s2997_r8903_r8906
mc16_13TeV.301053.PowhegPythia8EvtGen_AZNLOCTEQ6L1_DYtautau_2750M3000.merge.AOD.e3649_s2997_r8903_r8906
mc16_13TeV.301054.PowhegPythia8EvtGen_AZNLOCTEQ6L1_DYtautau_3000M3500.merge.AOD.e3649_s2997_r8903_r8906
mc16_13TeV.301055.PowhegPythia8EvtGen_AZNLOCTEQ6L1_DYtautau_3500M4000.merge.AOD.e3649_s2997_r8903_r8906
mc16_13TeV.301056.PowhegPythia8EvtGen_AZNLOCTEQ6L1_DYtautau_4000M4500.merge.AOD.e3649_s2997_r8903_r8906
mc16_13TeV.301057.PowhegPythia8EvtGen_AZNLOCTEQ6L1_DYtautau_4500M5000.merge.AOD.e3649_s2997_r8903_r8906
mc16_13TeV.301058.PowhegPythia8EvtGen_AZNLOCTEQ6L1_DYtautau_5000M.merge.AOD.e3649_s2997_r8903_r8906
mc16_13TeV.425200.Pythia8EvtGen_A14NNPDF23LO_Gammatautau_MassWeight.merge.AOD.e5468_s2997_r8903_r8906

---

Table A.1.: Listed are the training and testing samples. The Drell-Yan samples are sliced in subsamples according to the truth resonance mass. From the  $Z \rightarrow \tau\tau$  sample only events are used that satisfy a truth resonance mass less than 120 GeV.

## B. Discriminating Variables

This chapter contains distribution plots of the variables and the tables with variable importance for BDT2 (Table B.1), BDT3 (Table B.2) and multi-classification BDT (Table B.3).

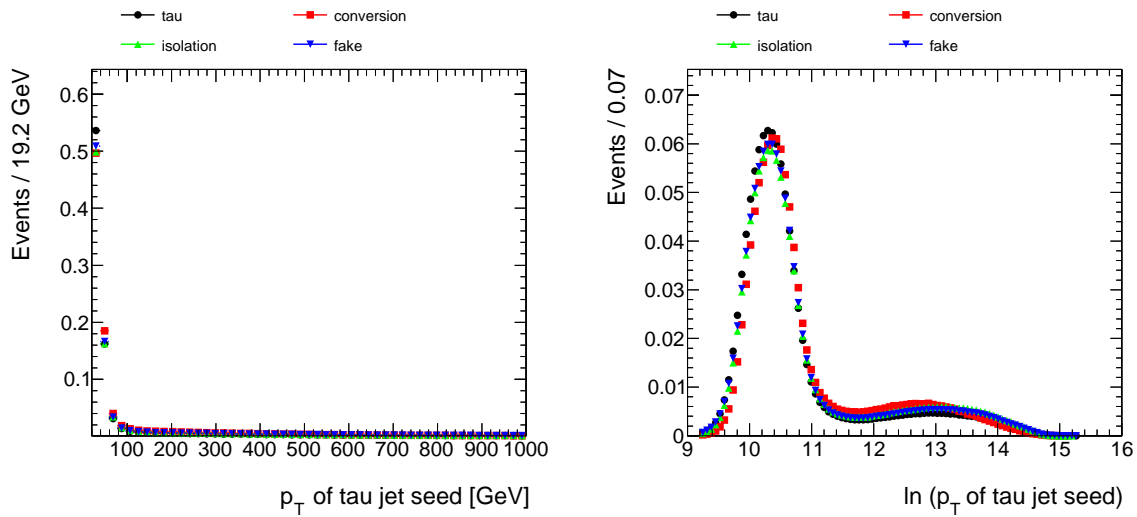


Figure B.1.: Normalised distributions of the momentum variables  $p_T^{\text{jetSeed}}$  (left) and  $\ln p_T^{\text{jetSeed}}$  (right) for the track classes tau tracks, conversion tracks, isolation tracks and fake tracks.

B. Discriminating Variables

---

BDT2		
Variable	Separation Score	Importance
$\ln  d_0 $	0.5828	0.10330
$P_e$	0.2245	0.08471
$R_{\text{conv}}^{\text{I}}$	0.2232	0.08069
$\eta$	0.064 28	0.07414
$\ln p_{\text{T}}^{\text{jetSeed}}$	0.013 47	0.07367
Number of Innermost Pixel Layer Hits	0.604	0.07078
$\ln  z_0^{\text{TJVA}} \sin \theta $	0.3105	0.07066
$q/p$	0.201	0.06618
$\Delta R$	0.027 21	0.06361
$n\text{PixHits}$	0.527	0.06124
Number of Hits in the TRT	0.048 24	0.05938
$R_{\text{conv}}^{\text{II}}$	0.3241	0.05597
$n\text{SiHits}$	0.3104	0.04714
$p_{\text{T}}^{\text{jetSeed}}$	0.006 927	0.03500
Number of Shared Hits in the Pixel Detector	0.024 57	0.02762
Number of Shared Hits in the SCT	0.018 25	0.02589

Table B.1.: Shown are the separation scores and variable importances for each discriminating variable, provided by TMVA. The values belong to BDT2, which separates `tau tracks` and `conversion tracks`. The list is ranked after the importance in descending order.



---

BDT3

Variable	Separation Score	Importance
$\ln  z_0^{\text{TJVA}} \sin \theta $	0.8227	0.13580
$nSiHits$	0.1052	0.11460
$q/p$	0.0122	0.08101
$R_{\text{conv}}^I$	0.1342	0.07657
$\Delta R$	0.006 715	0.07554
$\eta$	0.005 797	0.07295
$P_e$	0.004 294	0.07165
$\ln  d_0 $	0.1318	0.07066
Number of Hits in the TRT	0.002 571	0.06884
$\ln p_T^{\text{jetSeed}}$	0.000 607 5	0.06754
$nPixHits$	0.1194	0.04370
$R_{\text{conv}}^{II}$	0.1286	0.03579
$p_T^{\text{jetSeed}}$	0.000 461 9	0.03529
Number of Shared Hits in the SCT	0.003 959	0.02193
Number Of Innermost Pixel Layer Hits	0.1026	0.01653
Number of Shared Hits in the Pixel Detector	0.000 686 8	0.01153

Table B.2.: Shown are the separation scores and variable importances for each discriminating variable, provided by TMVA. The values belong to BDT3, which separates **isolation tracks** and **fake tracks**. The list is ranked after the importance in descending order.

multi-classification BDT	
Variable	Importance
$\Delta R$	0.1202
$\ln p_T^{\text{jetSeed}}$	0.8787
$\ln  d_0 $	0.8017
$R_{\text{conv}}^I$	0.7852
$\eta$	0.7652
$q/p$	0.7272
$P_e$	0.7229
$\ln  z_0^{\text{TJVA}} \sin \theta $	0.7163
Number of Hits in the TRT	0.5286
$R_{\text{conv}}^{II}$	0.5029
$nPixHits$	0.4841
$nSiHits$	0.4476
$p_T^{\text{jetSeed}}$	0.4289
Number Of Innermost Pixel Layer Hits	0.4077
Number of Shared Hits in the Pixel Detector	0.3060
Number Of Innermost Pixel Layer Hits	0.2948

Table B.3.: Shown are the variable importances for each discriminating variable sorted in descending order. The values belong to the multi-classification BDT presented in Section 5.6.

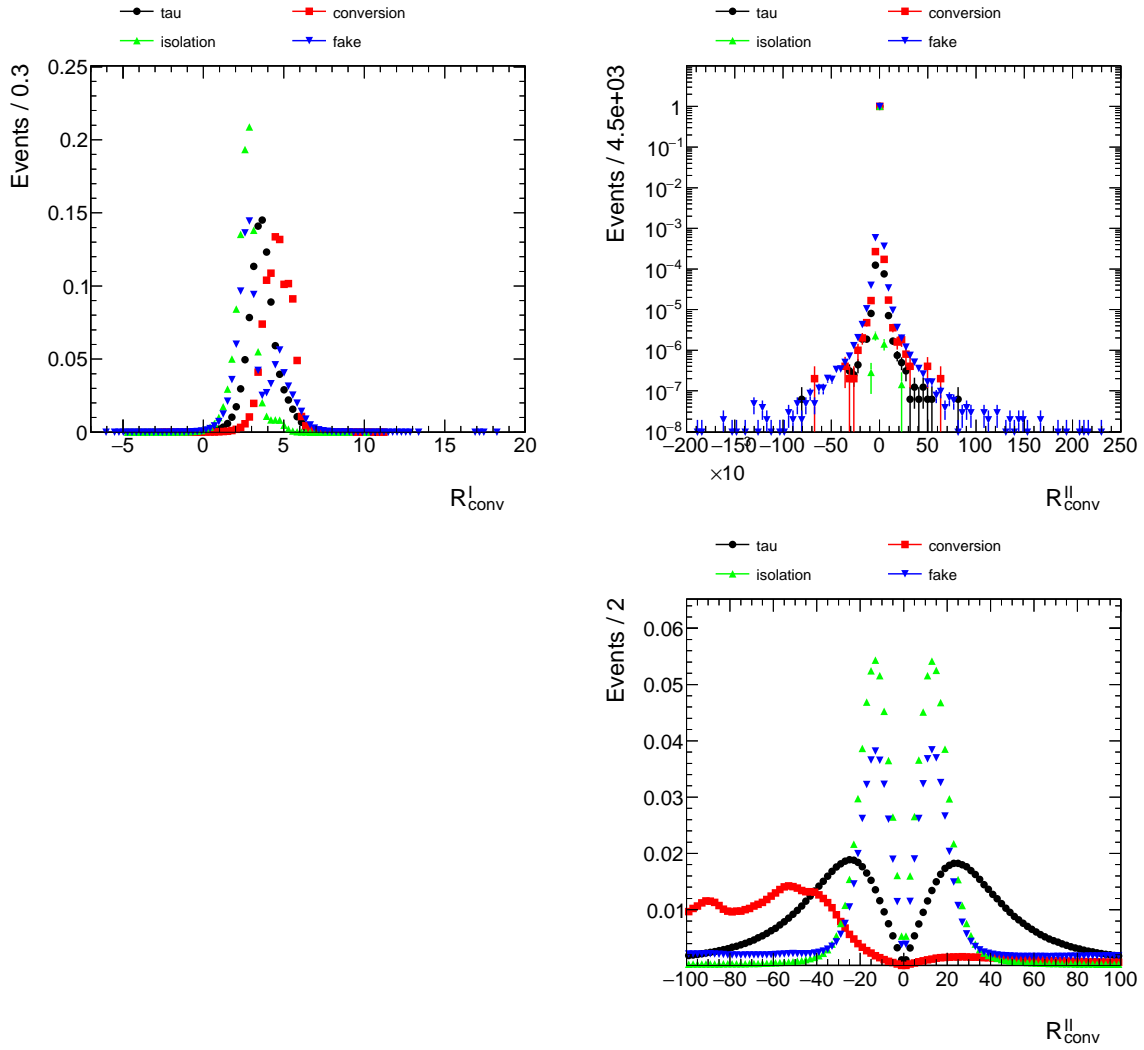


Figure B.2.: Normalised distributions of the variables  $R_{\text{conv}}^I$  (upper left) and  $R_{\text{conv}}^{II}$  (upper right). The distribution in the lower left shows  $R_{\text{conv}}^{II}$  in a smaller range. The different colours correspond to the track classes tau tracks, conversion tracks, isolation tracks and fake tracks.

## B. Discriminating Variables

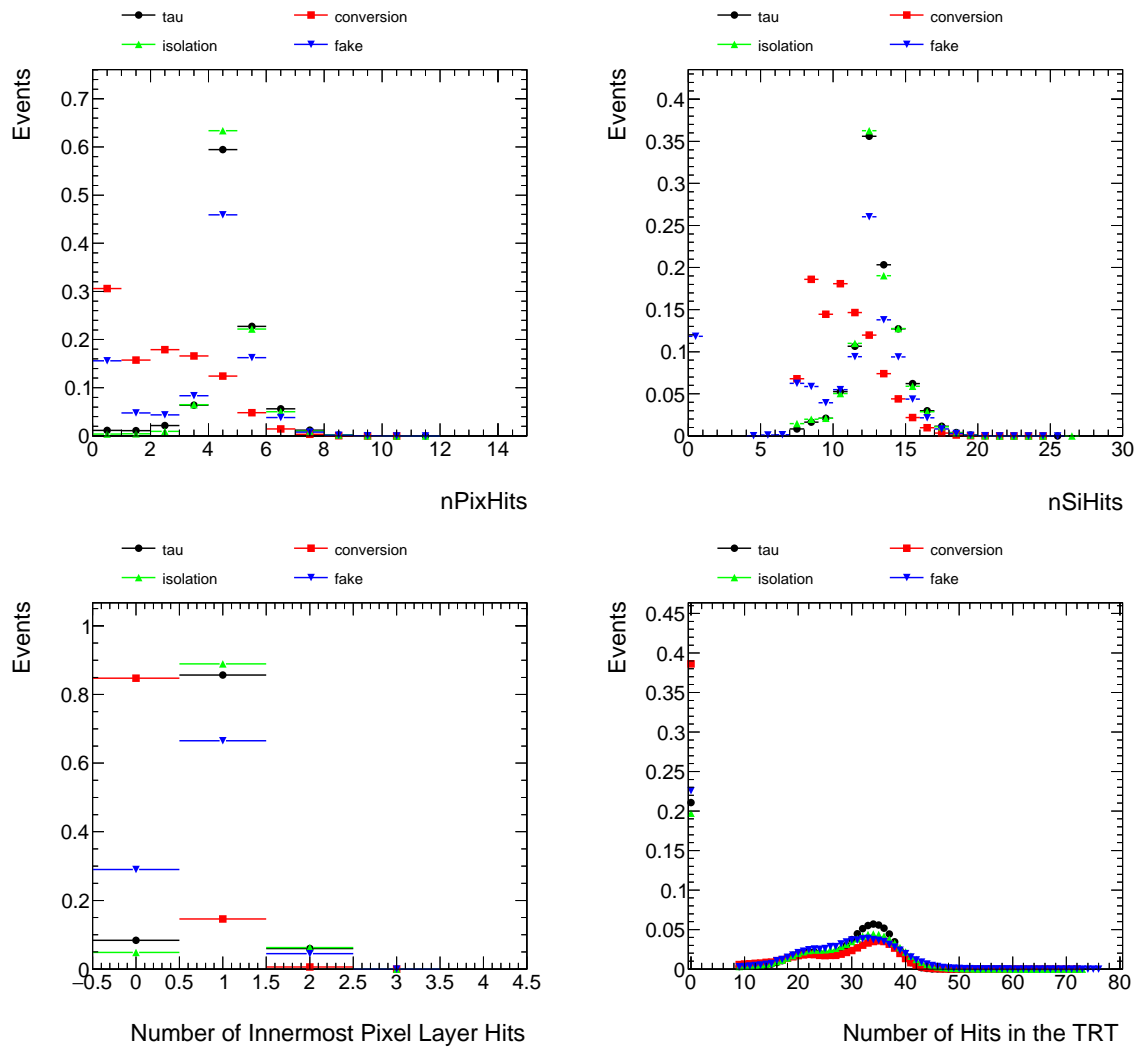


Figure B.3.: Normalised distributions of the variables  $nPixHits$  (upper left),  $nSiHits$  (upper right), Number of Innermost Pixel Layer Hits (lower left) and Number of Hits in the TRT (lower right). The different colours correspond to the track classes tau tracks, conversion tracks, isolation tracks and fake tracks.

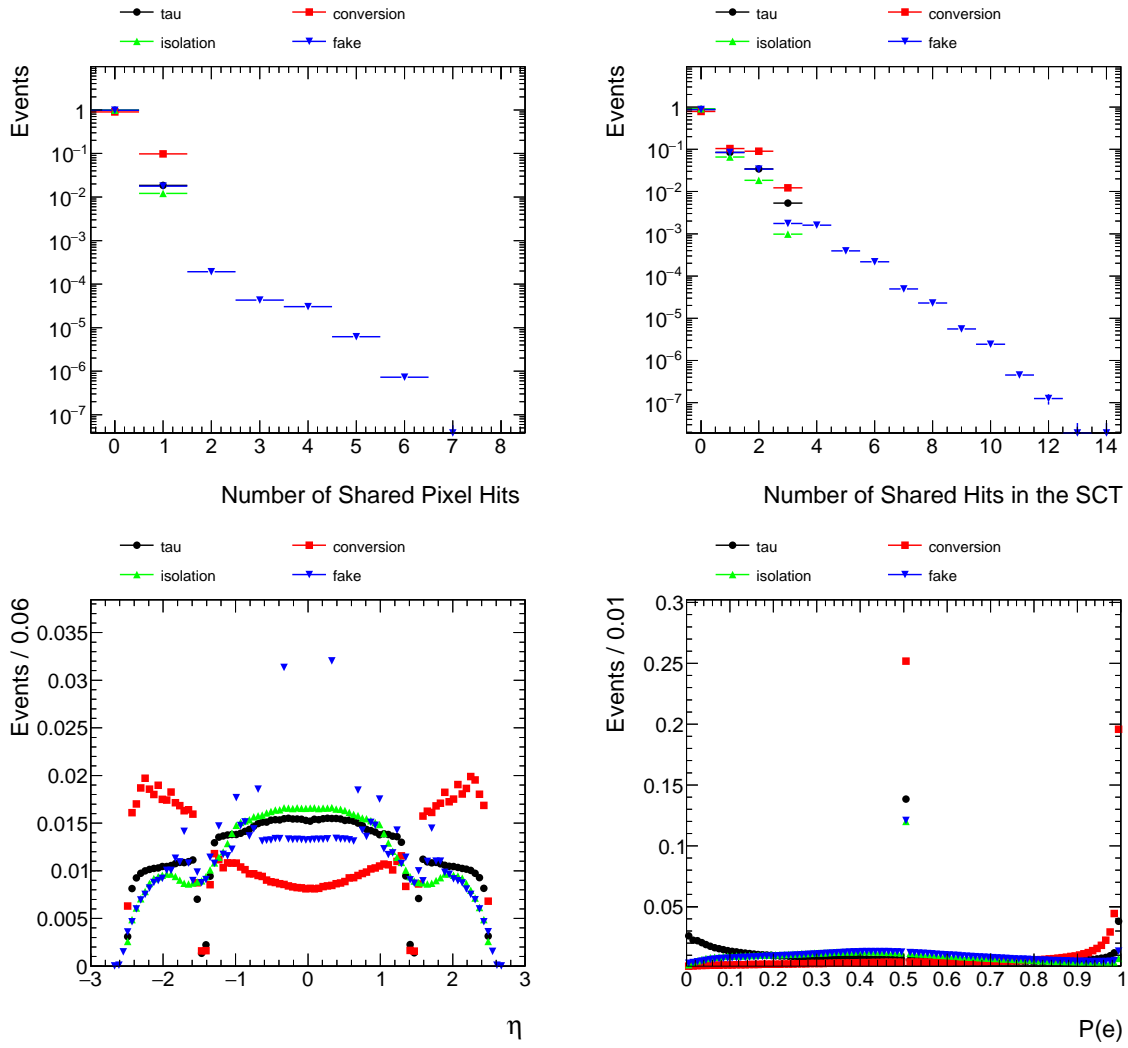


Figure B.4.: Normalised distributions of the Number of Shared Hits in the Pixel Detector (upper left), the Number of Shared Hits in the SCT (upper right),  $\eta$  (lower left) and  $P_e$  (lower right). The different colours correspond to the track classes tau tracks, conversion tracks, isolation tracks and fake tracks.



## C. BDT Algorithm Options

The Table C.1 shows the BDT options, that are used with their default values.

Option	Default	Description
Shrinkage	1	Learning rate for GradBoost algorithm.
SeparationType	GiniIndex	Separation criterion for node splitting.
PruneMethod	NoPruning	Remove nodes of a tree from the bottom up.
NodePurityLimit	0.5	In boosting, nodes with purity $>$ NodePurityLimit are signal; background otherwise.
BaggedSampleFraction	0.6	Relative size of bagged event sample to original size of the data sample.
UseYesNoLeaf	True	Use Sig or Bkg categories as classification of the leaf node.

Table C.1.: BDT Algorithm Options.[39]





# Bibliography

- [1] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett.* **B716** (2012), 1–29, [arXiv:1207.7214 \[hep-ex\]](#).
- [2] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, *Phys. Lett.* **B716** (2012), 30–61, [arXiv:1207.7235 \[hep-ex\]](#).
- [3] J. J. Thomson, *Cathode Rays*, *Philosophical Magazine* **44** (1897) 269, 293–316.
- [4] J. Chadwick, *The Existence of a Neutron*, *Proc. Roy. Soc. Lond.* **A136** (1932) 830, 692–708.
- [5] S. Fukuda et al., *The Super-Kamiokande detector*, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **501** (2003) 2, 418–462.
- [6] F. Halzen and S. R. Klein, *IceCube: An instrument for neutrino astronomy*, *Review of Scientific Instruments* **81** (2010).
- [7] L. Evans and P. Bryant, *LHC Machine*, *Journal of Instrumentation* **3** (2008) 08, S08001.
- [8] Particle Data Group, *Review of Particle Physics*, *Chin. Phys.* **C40** (2016/2017) 10, 100001.
- [9] P. Langacker, *Introduction to the Standard Model and Electroweak Physics* [arXiv:0901.0241 \[hep-ph\]](#).
- [10] S. L. Glashow, *Partial Symmetries of Weak Interactions*, *Nucl. Phys.* **22** (1961), 579–588.
- [11] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964), 508–509.
- [12] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, *Phys. Rev. Lett.* **13** (Aug, 1964), 321–323.

- [13] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, *Phys. Rev. Lett.* **13** (Nov, 1964), 585–587.
- [14] M. Czakon, M. Zralek, and J. Gluza, *Are neutrinos Dirac or Majorana particles?*, *Acta Phys. Polon.* **B30** (1999), 3121–3138, [arXiv:hep-ph/9910357](#).
- [15] S. Hughes, *Lecture 1*, *Massachusetts Institute of Technology*, 8.022 (Spring 2005).
- [16] G. Bertone, D. Hooper, and J. Silk, *Particle dark matter: Evidence, candidates and constraints*, *Phys. Rept.* **405** (2005), 279–390, [arXiv:hep-ph/0404175](#).
- [17] S. P. Martin, *A Supersymmetry primer*, [FERMILAB-PUB-97-425-T](#), [arXiv:hep-ph/9709356](#) [hep-ph].
- [18] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, *JINST* **3** (2008), S08003.
- [19] CMS Collaboration, *The CMS experiment at the CERN LHC*, *JINST* **3** (2008), S08004.
- [20] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, *JINST* **3** (2008), S08002.
- [21] LHCb Collaboration, *The LHCb Detector at the LHC*, *JINST* **3** (2008), S08005.
- [22] E. Mobs, *The CERN accelerator complex*, [OPEN-PHO-ACCEL-2016-009](#) (Jul, 2016).
- [23] C. Pralavorio, *The LHC racks up records*, *CERN* (Jun, 2017).
- [24] ATLAS Collaboration, *The Run-2 ATLAS Trigger System*, [ATL-DAQ-PROC-2016-003](#) (Feb, 2016).
- [25] J. Pequenao, *Computer generated image of the whole ATLAS detector*, [CERN-GE-0803012](#) (Mar, 2008).
- [26] K. Potamianos, *The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider*, *PoS EPS-HEP2015* (2015), 261, [arXiv:1608.07850](#) [physics.ins-det].
- [27] ATLAS Collaboration, *The Pixel Detector of the ATLAS Experiment for the Run 2 at the Large Hadron Collider*, *Nucl. Part. Phys. Proc.* **273-275** (2016), 1166–1172.

- 
- [28] J. Pequenao, *Computer Generated image of the ATLAS calorimeter*, [CERN-GE-0803015 \(Mar, 2008\)](#).
- [29] ATLAS Collaboration, *Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC run 2*, [CERN-EP-2017-045](#), [arXiv:1704.07983 \[hep-ex\]](#).
- [30] A. Rosenfeld and J. Pfaltz, *Sequential Operations in Digital Picture Processing*, [Journal of the ACM \(JACM\) 13 \(1966\) 4](#), 471–494.
- [31] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, [Nucl. Instrum. Meth. A262 \(1987\)](#), 444–450.
- [32] W. Lampl et al., *Calorimeter clustering algorithms: Description and performance*, [ATL-LARG-PUB-2008-002](#), [ATL-COM-LARG-2008-003](#) .
- [33] ATLAS Collaboration, *Local hadronic calibration*, [ATL-LARG-PUB-2009-001-2](#), [ATL-COM-LARG-2008-006 \(2009\)](#).
- [34] M. Cacciari, G. P. Salam, and G. Soyez, *The anti- $k_t$  jet clustering algorithm*, [JHEP 04 \(2008\)](#), 063, [arXiv:0802.1189 \[hep-ph\]](#).
- [35] M. Cacciari, G. P. Salam, and G. Soyez, *FastJet User Manual*, [Eur. Phys. J. C72 \(2012\)](#), 1896, [arXiv:1111.6097 \[hep-ph\]](#).
- [36] ATLAS Collaboration, *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*, [ATL-PHYS-PUB-2015-045 \(Nov, 2015\)](#).
- [37] ATLAS Collaboration, *Number of Interactions per Crossing*, [Public Results - Run 2 \(2017\)](#).
- [38] M. H. Seymour and M. Marx, *Monte Carlo Event Generators*, [MCNET-13-05](#), [arXiv:1304.6677 \[hep-ph\]](#).
- [39] A. Hoecker et al., *TMVA - Toolkit for Multivariate Data Analysis*, [PoS ACAT \(2007\)](#), 040, [arXiv:physics/0703039](#).
- [40] CDF Collaboration, *Two-particle momentum correlations in jets produced in  $p\bar{p}$  collisions at  $\sqrt{s} = 1.96$  TeV*, [Phys. Rev. D 77 \(May, 2008\)](#), 092001.
- [41] T. Sjostrand et al., *An Introduction to PYTHIA 8.2*, [Comput. Phys. Commun. 191 \(2015\)](#), 159–177, [arXiv:1410.3012 \[hep-ph\]](#).

## **Acknowledgements**

I would like to thank all those, who have contributed to the success of this work. First of all, my gratitude goes to Prof. Dr. Arno Straessner and Dr. Wolfgang Mader for the excellent supervision. I deeply appreciate their support and commitment. A special thanks goes to Dirk Duschinger for his support and advice. His guidance was essential for my research. Furthermore, I express my gratitude to my colleagues Lorenz Hauswald, Sebastian Wahrmund, Max Märker and David Kirchmeier for their helpful advice. I would also like to extend my appreciation to all the members of the IKTP. The atmosphere and working conditions were great.

Finally, I sincerely thank my family for their support and sympathy over the entire time. Without their help, my studies haven't been possible.

## **Selbstständigkeitserklärung**

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher weder im Inland noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Richard Hartmann  
Dresden, 16. Oktober 2017