

TECHNISCHE  
UNIVERSITÄT  
DRESDEN

---

# Reconstruction and Identification of Boosted Tau Pairs in the Semi-Leptonic Decay Channel at ATLAS

Master-Arbeit  
zur Erlangung des Hochschulgrades  
Master of Science  
im Master-Studiengang Physik

vorgelegt von

Christian Wiel  
geboren am 18.07.1993 in Dresden

INSTITUT FÜR KERN- UND TEILCHENPHYSIK  
FACHRICHTUNG PHYSIK  
FAKULTÄT MATHEMATIK UND NATURWISSENSCHAFTEN  
TECHNISCHE UNIVERSITÄT DRESDEN  
2018

Eingereicht am: 02.05.2018

1. Gutachter: Prof. Dr. Arno Straessner
2. Gutachter: Prof. Dr. Michael Kobel

## Abstract

Tau lepton pairs with a low angular distance are an important signature of physics scenarios extending the Standard Model of Particle Physics. The existing algorithms to reconstruct and identify these tau pairs in proton-proton collisions at the Large Hadron Collider with the ATLAS experiment are investigated. They are found to cease to work in the regime of low separation  $\Delta R < 0.4$  and alternative approaches are developed. In contrast to earlier studies which focused on the case where both tau leptons decay hadronically in this thesis algorithms for the case where one tau decays hadronically and the other one decays into lighter leptons are presented. The two leptonic decay modes are addressed separately, because the overlaps between the detector signature of an electron or a muon with the signatures of hadronic tau decay products impose different challenges. For the case where the leptonic tau decay produces a muon the existing algorithm for hadronic tau decays is corrected for the Inner Detector track of the muon. For the case where the tau decays into an electron the merging calorimeter energy depositions lead to a different approach using substructure jets. The new reconstruction and identification algorithms show promising results by achieving high signal efficiencies and background rejections down to low angular distances of  $\Delta R = 0.1$ .

## Kurzfassung

Tau-Lepton Paare mit geringen Winkelabständen sind eine wichtige Signatur für Erweiterungen des Standardmodells der Teilchenphysik. Die bestehenden Algorithmen zur Rekonstruktion und Identifikation solcher Tau Paare in Proton-Proton Kollisionsereignissen am Large Hadron Collider mit dem ATLAS Experiment werden untersucht. Es wird festgestellt, dass sie im Regime geringer Abstände  $\Delta R < 0.4$  aufhören zu funktionieren, und alternative Ansätze entwickelt werden müssen. Im Gegensatz zu früheren Untersuchungen, die den Fall zweier hadronisch zerfallender Tau Leptonen betrachtet haben, wird in dieser Arbeit der Fall, in dem eines der Taus hadronisch und das andere in leichtere Leptonen zerfällt, behandelt. Die beiden leptonischen Zerfallskanäle werden separat untersucht, da die Überlagerung der Detektorsignaturen eines Myons bzw. eines Elektron mit den Signaturen von hadronischen Tauzerfallsprodukten unterschiedliche Herausforderungen mit sich bringt. Für den Fall, dass der leptonische Zerfall in ein Myon stattfindet, können die existierenden Algorithmen für hadronische Tauzerfälle um die Spur des Myons im Inneren Detektor korrigiert werden. Zerfällt das Tau in ein Elektron, ist aufgrund der ineinander übergehenden Energiedepositionen im Kalorimeter ein anderer Ansatz notwendig der auf dem Gebrauch von Substrukturjets beruht. Die neuen Rekonstruktions- und Identifikationsalgorithmen zeigen vielversprechende Ergebnisse in Form von hohen Signaleffizienzen und Untergrundunterdrückungen bis hin zu niedrigen Winkelabständen von  $\Delta R = 0.1$ .



# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Theoretical Foundations</b>	<b>3</b>
2.1	The Standard Model of Particle Physics . . . . .	3
2.1.1	Electro-Weak Symmetry Breaking . . . . .	4
2.1.2	Properties of $\tau$ -Leptons . . . . .	5
2.1.3	Limitations of the Standard Model . . . . .	5
2.2	The Randal-Sundrum Model . . . . .	6
2.3	Boosted Topologies . . . . .	8
<b>3</b>	<b>The ATLAS Experiment at the LHC</b>	<b>13</b>
3.1	The Large Hadron Collider . . . . .	13
3.2	The ATLAS Detector . . . . .	14
3.2.1	Inner Detector . . . . .	15
3.2.2	Calorimeters . . . . .	15
3.2.3	Muon Spectrometer . . . . .	16
3.2.4	Trigger System . . . . .	18
<b>4</b>	<b>Particle Reconstruction and Identification at ATLAS</b>	<b>19</b>
4.1	Multivariate Classification Techniques . . . . .	19
4.1.1	Boosted Decision Trees . . . . .	19
4.1.2	Projective Likelihood Method . . . . .	21
4.2	Low Level Object Reconstruction . . . . .	21
4.2.1	Inner Detector Track Reconstruction . . . . .	21
4.2.2	Primary Vertex Reconstruction . . . . .	22
4.2.3	Calorimeter Cell Clustering . . . . .	22
4.2.4	The Anti- $k_t$ Jet Clustering Algorithm . . . . .	23
4.3	Electron Reconstruction and Identification . . . . .	23
4.4	Muon Reconstruction and Identification . . . . .	25
4.5	Reconstruction and Identification of Hadronic Tau Decays . . . . .	25
4.6	Reconstruction and Identification of Boosted Di-Taus in the $\tau_{\text{had}}\tau_{\text{had}}$ decay mode	28
<b>5</b>	<b>Di-Tau Reconstruction and Identification in the <math>\tau_{\text{lep}}\tau_{\text{had}}</math> Channel</b>	<b>31</b>
5.1	The $\tau_{\text{had}}\tau_e$ Channel . . . . .	32
5.1.1	Performance of Existing Algorithms . . . . .	32
5.1.2	Reconstruction . . . . .	34
5.1.3	Identification . . . . .	38
5.2	The $\tau_{\text{had}}\tau_\mu$ Channel . . . . .	56
5.2.1	Performance of Existing Algorithms . . . . .	56
5.2.2	Reconstruction . . . . .	56
5.2.3	Muon Track Removal . . . . .	58
5.2.4	Identification . . . . .	61

<b>6 Summary and Outlook</b>	<b>67</b>
<b>Bibliography</b>	<b>69</b>
<b>A Additional Tables and Figures</b>	<b>75</b>
A.1 Data Samples . . . . .	75
A.2 The $\tau_{\text{had}}\tau_e$ Channel . . . . .	77
A.3 The $\tau_{\text{had}}\tau_\mu$ Channel . . . . .	78

# 1 Introduction

The origins of particle physics go back to the discovery of the electron by J. J. Thomson and of the atomic substructure by E. Rutherford around the turn of the 19th and 20th century. This marked the beginning of over a century of ongoing research, which is characterized by a fruitful synergy between experimental and theoretical progress, where experimental findings fuel new theories and theoretical predictions motivate new experiments. For example the discovery of many different short lived particles in cosmic rays and at the upcoming synchrotron particle accelerators in the 1950s led ultimately to the development of Quantum Chromo Dynamics (QCD) [1–5] which classifies these particles by their substructure. Although the quarks which form this substructure can not be observed as free particles, they were experimentally shown to exist as pointlike partons inside of e.g. protons by deep inelastic scattering [6, 7]. In the meantime an unified theory of electroweak interactions was developed [8–10] and the Higgs mechanism [11–13] was proposed.

These theories form together the Standard Model of Particle Physics (SM), which is the most successful available theory in predicting the results of high energy experiments with high precision. In 2012, almost 50 years after its proposal, the last piece of the SM, the Higgs boson, was discovered by the ATLAS and CMS experiment at the Large Hadron Collider (LHC) [14, 15]. Although great accomplishments in the thorough determination of its properties have already been achieved [16–19], only further experimental data from the LHC or even from future colliders will reveal, if e.g. the couplings of the discovered Higgs boson coincide exactly with the predictions of the SM.

Beside all of its success the SM has also some limitations. For example it does not deliver a sufficient explanation for dark matter and is not able to unify all forces. There are many different proposed theories extending the SM to accommodate its weaknesses. Many of them predict new phenomena observable at the LHC.

The tau lepton, which was discovered in 1975 [20], is the heaviest among the three massive leptons. Because of this property it plays a central role especially in the Higgs phenomenology of many new physics scenarios. With proton-proton collisions of unprecedented luminosity and energy provided by the LHC to the ATLAS experiment searches for new physics involving high momentum particles decaying into tightly collimated tau pairs begin to be feasible. One example would be a multi TeV resonance decaying into a pair of Higgs bosons, where at least one of the Higgs bosons decays into a pair of tau leptons. A prerequisite to discover this processes is that there are algorithms able to reconstruct the two tau decays even if their detector signatures overlap. In contrast to earlier studies [21], which already showed solutions for the case where both taus decay hadronically, this thesis covers the semi-leptonic decay mode.

In Chapter 2 the SM and the new physics scenario, which is used as a benchmark, are shortly introduced. In Chapter 3 the LHC and the ATLAS detector are described. Chapter 4 covers the existing reconstruction and identification techniques for hadronic tau decays, electrons and muons at ATLAS. In Chapter 5 the existing algorithms are evaluated in the application on boosted di-tau decays and possible solutions for their shortcomings are described. The two semi-leptonic decay modes are treated separately.





## 2 Theoretical Foundations

In this chapter the theoretical foundations for this thesis are shortly summarized. In the first section the Standard Model of Particle Physics is introduced. In Section 2.2 a brief description of theories with extra dimensions with an emphasis on the Randall-Sundrum model is given. The Randall-Sundrum model is a possible extension of the Standard Model, which serves in this thesis as a benchmark for physics with boosted tau pairs. In the last section of this chapter the relativistic kinematics of boosted two particle decays are studied.

### 2.1 The Standard Model of Particle Physics

The Standard Model of Particle Physics (SM) describes elementary particles and their interactions. It combines the theory of Quantum Chromo Dynamics (QCD) [1–5] and the theory of electro-weak interactions [8–10]. The SM is very successful in predicting a wide range of phenomena, which were later on experimentally measured with a high precision.

The SM is a relativistic Quantum Field Theory (QFT). In QFTs particles are described as excitations of quantum fields. The particles are grouped into fermions, which are defined by their half-integer spin, and bosons defined by their integer spin. The SM is a relativistic theory, that means it follows a global Poincaré symmetry. This leads to conservation of four-momenta due to the Noether theorem [22, 23].

The SM also has an underlying  $SU(3)_C \times SU(2)_L \times SU(1)_Y$  local gauge symmetry, where  $Y$  stands for the electroweak hyper charge,  $L$  stands for the left-handed chirality of the weak interaction and  $C$  stands for the color charge of the strong interaction. Demanding this symmetry gives rise to the bosonic gauge fields, which mediate the interactions of the fermionic fields. The  $SU(2)_L \times SU(1)_Y$  symmetry is spontaneously broken into  $SU(1)_Q$  by a non-zero vacuum expectation value of the Higgs field (see Section 2.1.1). This leads to the massive  $W^\pm$ - and  $Z$ -Bosons, which couple to left-handed particles, and to the massless photon  $\gamma$ , which couples to particles with an electric charge. The electric charge  $Q$  is connected to the hypercharge  $Y$  and the third component of the weak isospin  $I_3^W$  by:

$$Q = I_3^W + \frac{1}{2}Y.$$

The electromagnetic interaction mediated by photons governs most of the everyday physics.  $W^\pm$ - and  $Z$ -Bosons give rise to the weak interaction, responsible for nuclear decays. Gluons  $g$  bind the color charged partons of protons and neutrons in atomic nuclei by mediating the strong force.

The fermions are grouped into leptons and quarks. An overview over all fermions in the SM is provided in Table 2.1.

The quarks carry a color charge and therefore interact through the strong force. We do not observe free quarks in nature and all composite particles are colorless. This is described as color confinement. Quarks come in 6 different flavors: up  $u$ , down  $d$ , strange  $s$ , charme  $c$ , bottom  $b$  and top  $t$ . Quarks also have an electric charge and are arranged in left-handed doublets and right-handed singlets in terms of the weak isospin.

The leptons do not carry a color charge. For each of the electrically charged leptons: electron  $e$ , muon  $\mu$ , and tau-lepton  $\tau$  there is an electrical neutral neutrino  $\nu_e, \nu_\mu, \nu_\tau$ . They are arranged in left-handed doublets and right-handed singlets. There are no left-handed neutrinos in the SM.

The fermions can be grouped into sets where the mass is the only diverging property: the charged leptons ( $e, \mu, \tau$ ), the neutrinos ( $\nu_e, \nu_\mu, \nu_\tau$ ), the up-type quarks ( $u, c, t$ ) and the down-type quarks ( $d, s, b$ ). The lightest one of each of these sets is commonly denoted as first generation. The second lightest and heaviest are denoted as second and third generation, respectively.

Neutrinos are treated as massless in the SM. This is in contradiction to the experimental observation of neutrino-oscillation [24], but the SM can be extended to accommodate this fact.

For every fermion there is an anti-fermion with quantum numbers of opposite sign.

Table 2.1: Overview of all fermions, their third component of the weak isospin  $I_3^W$  and electric charge  $Q$  in the SM:

$I_3^W$	$Q$	quarks		
$\frac{1}{2}$	$\frac{2}{3}$	$\begin{pmatrix} u_L \\ d_L \end{pmatrix}$	$\begin{pmatrix} c_L \\ s_L \end{pmatrix}$	$\begin{pmatrix} t_L \\ b_L \end{pmatrix}$
$-\frac{1}{2}$	$-\frac{1}{3}$	$u_R$	$c_R$	$t_R$
$0$	$\frac{2}{3}$	$d_R$	$s_R$	$b_R$
leptons				
$\frac{1}{2}$	$-1$	$\begin{pmatrix} e_L \\ \nu_{e,L} \end{pmatrix}$	$\begin{pmatrix} \mu_L \\ \nu_{\mu,L} \end{pmatrix}$	$\begin{pmatrix} \tau_L \\ \nu_{\tau,L} \end{pmatrix}$
$-\frac{1}{2}$	$0$	$e_R$	$\mu_R$	$\tau_R$

### 2.1.1 Electro-Weak Symmetry Breaking

To accommodate the desire for massive elementary particles without destroying the local gauge invariance the Higgs mechanism [11–13] was introduced to the SM. The Higgs mechanism consists of a scalar isospin doublet  $\Phi$  and a potential  $V$ .

$$V(\Phi) = \mu^2(\Phi^\dagger\Phi) + \lambda(\Phi^\dagger\Phi)^2 \quad (2.1)$$

For  $\mu^2 < 0$  this potential is minimized by a set of infinitely many possible vacuum states with non-zero expectation value  $v$ . The necessity to choose one unique vacuum  $\Phi_0$  then breaks the continuous symmetry  $SU(2)_L \times SU(1)_Y$  down to  $SU(1)_Q$ . This leads to massive  $W^\pm$ - and  $Z$ -bosons ( $m_W = 80.4 \text{ GeV}$ ,  $m_Z = 91.2 \text{ GeV}$ ) and retains photons massless. Small excitations around the vacuum state give rise to the massive Higgs boson  $h$  ( $m_h = 125.09(24) \text{ GeV}$  [24, 25]), its couplings to the  $W^\pm$ - and  $Z$ -bosons and its trilinear and quartic self-couplings.

An additional Yukawa coupling between the Higgs boson and the fermions is able to generate their masses. This coupling leads also to the fact, that the Higgs boson can decay into fermions and that the branching ratio of the different decay modes is enhanced for heavier fermions. The predicted branching ratios for different important Higgs decay channels are summarized in Table 2.2. With a calculated branching ratio of  $\text{BR}(h \rightarrow \tau\tau) = 6.25(10) \%$  [26]  $\tau$ -leptons are an

important signature for physics processes involving Higgs bosons.

Table 2.2: SM prediction for the branching ratio of important decay channels of the Higgs boson with  $m_h = 125.09$  GeV [26]:

Decay Channel	Branching ratio [%]
$h \rightarrow b\bar{b}$	58.1(7)
$h \rightarrow W^+W^-$	21.5(3)
$h \rightarrow gg$	8.18(4)
$h \rightarrow \tau^+\tau^-$	6.25(10)
$h \rightarrow cc$	2.9(2)
$h \rightarrow ZZ$	2.64(4)
$h \rightarrow \gamma\gamma$	0.227(5)

### 2.1.2 Properties of $\tau$ -Leptons

The tau lepton has a mass of  $m_\tau = 1776.86(12)$  MeV [24]. It is the heaviest lepton. Due to its high mass it decays after a very short mean lifetime of  $\tau = 2.903(5) \times 10^{-13}$  s into neutrinos and either hadrons ( $\tau_{\text{had}}$ ) or leptons ( $\tau_{\text{lep}}$ ). Because of the short lifetime it can usually only be detected by its decay products. In Figure 2.1 the branching ratios for the various important decay modes and their final state particles are illustrated. The decay products contain one  $\nu_\tau$  in any case.

The most common decay products for hadronic decays are charged and neutral pions. Charged pions occur in odd numbers and the sum of the charges has to add up to the  $\tau$  charge. The case with only one charged particle (charge multiplicity 1) makes up for approximately 75 % of the  $\tau_{\text{had}}$  decays and because the case with charge multiplicity 5 is extremely rare (0.1 %) the charge multiplicity 3 case makes up for almost the whole rest. In approximately 70 % of the hadronic decays there are also one or more neutral pions involved. The neutral pions decay after an even shorter mean lifetime of only  $\tau_\pi = 8.52(18) \times 10^{-17}$  s in roughly 99 % of the cases into a pair of photons.

The two leptonic decay modes  $\tau_e \rightarrow e\nu_e\nu_\tau$  and  $\tau_\mu \rightarrow \mu\nu_\mu\nu_\tau$  occur almost with the same branching ratio of 17.82(4) % and 17.39(4) % respectively.

Because of the additional neutrino compared to the hadronic decay mode, the average amount of energy of the leptonic decay, that can directly be measured in a detector, is lower.

### 2.1.3 Limitations of the Standard Model

Although the SM is very successful in predicting experimental results, it has some limitations. Cosmological observations imply that there is a large fraction of matter in our universe that seems to only interact gravitationally [27]. The SM does not contain particles which have the

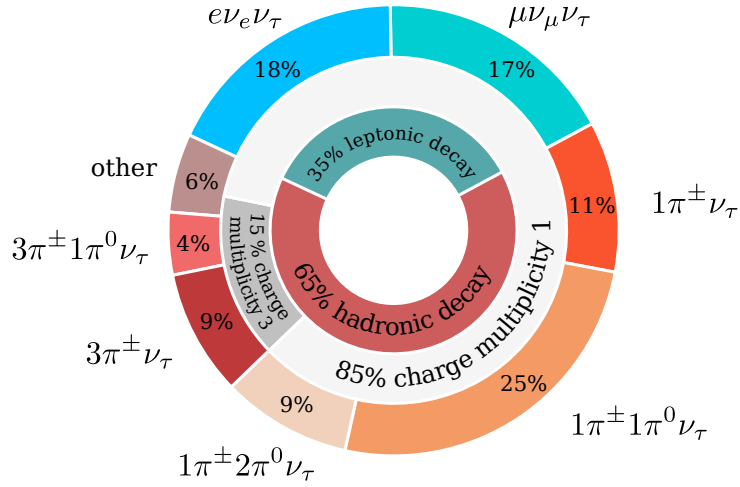


Figure 2.1: Branching ratio for the most common decay modes of the tau lepton [24].

right properties to be candidates for this so called dark matter [28].

The SM does also not describe the gravitational force. It is in principle possible to incorporate the theory of general relativity for energy scales lower than the Planck mass  $m_P = 1.221 \times 10^{19}$  GeV into the SM with an effective field theory approach [29]. Because of the tensorial structure of the space time in general relativity the resulting graviton particles from this effective field theory approach have spin-2 and because of the infinite range of gravity they have to be massless. But gravity and also the quantum corrections to gravity are so extremely small compared to the other forces, that the effects of gravity in the SM are negligible at the accessible energies of all current and proposed high energy physics experiments. This is equivalently described by the extreme difference between the Planck mass and the vacuum expectation value of the Higgs field  $v$  also known as the hierarchy problem.

The hierarchy problem is linked to the fine-tuning problem of the SM. If one assumes that the SM is valid up to the Planck mass  $m_P$ , it would be feasible to use the Planck mass as cut-off scale  $\Lambda$  for the calculation of radiative corrections to the Higgs boson mass. This leads to an incredible fine-tuning of parameters of order  $\mathcal{O}(v/M_P)$ . Supersymmetric theories provide an elegant solution to this problem. One prominent phenomenological feature of them would be, that they require at least two Higgs doublets [30].

An alternative or addition to supersymmetric theories are theories with extra dimensions. It is the aim of the next section to describe their phenomenology and how they could explain that the difference in strength between gravity and the other forces might only be an apparent one.

## 2.2 The Randall-Sundrum Model

The idea behind all of the following theories is that there are hidden space like dimensions in addition to the three we can currently observe. A figurative description of an extra dimension is, that our 4-dimensional space-time is just a membrane ('brane') in a space of higher dimensionality.

The first theory with an extra dimension was proposed by Kaluza and Klein in the 1920's [31, 32]. The extra dimension of their theory is compactified on a circle, that means it is periodical with a period of  $2\pi R$ . The smallest possible excitation in the 5th dimension would then have an energy of  $1/R$ . This means that only for energies larger than  $1/R$  the effects of the extra dimension would be apparent. Consequently this extra dimension would have to be small, as we did not observe any Kaluza-Klein excitations yet.

In 1998 Arkani-Hamed, Dimopoulos and Dvali (ADD) had a new approach to extra dimensions [33]. Their idea was to only allow gravity to propagate in the  $n$  extra compactified dimensions and to confine the other fields to the classical four dimensional space-time. Because the  $3+n$  dimensional volume accessible to gravity is much larger than the one for the other forces, gravity appears weak in our brane. This changes the classical Newtonian  $1/r^2$  law, especially when  $r$  is smaller than the size  $R$  of the extra dimension. They argue that the theory would be compatible with experimental constraints for  $n \geq 2$  extra dimensions with a size of  $R \lesssim 1$  mm.

Shortly after the ADD approach Randall and Sundrum proposed a different solution to the hierarchy problem [34, 35]. Their theory relies on only one extra dimension, which is not necessarily compactified. The core feature of the Randall-Sundrum theory is the different background metric:

$$ds^2 = e^{-2kr_c\phi} \eta_{\mu\nu} dx^\mu dx^\nu + r_c^2 d\phi^2.$$

The factor  $e^{-2kr_c\phi}$  in front of the usual Minkowski metric is named warp factor. It depends on the coordinate value of the extra dimension  $\phi$ . The brane we perceive (TeV brane) is located at  $\phi = \pi$ , while the so called Planck brane is located at  $\phi = 0$ . Consequently the warp factor leads to very short distances in the Planck brane compared to the distances in the TeV brane. The metric confines the graviton mode to be mostly located at the Planck brane. With  $kr_c \sim 11$  the theory would produce the correct relation between Planck mass and vacuum expectation value of the Higgs field on the TeV brane and one of the order of one at the Planck brane. Therefore the theory is able to solve the hierarchy problem. The massive graviton modes of this theory are predicted to have masses in the TeV range. Their main production channel at the Large Hadron Collider (LHC) would be gluon fusion. In the original proposal (RS1) all SM-fields are located on the TeV brane like in the ADD approach. In this framework the most promising decay signatures of the graviton modes would be pairs of leptons or photons. The currently strongest experimental constraints on the RS1 theory are set by ATLAS and CMS [36–39].

An alternative approach where the fermions are localized in the so called bulk between the UV- and the Planck brane has some appealing additional features. The Higgs field is still localized at the UV-brane. The fermion masses could therefore be explained by different localizations in the bulk and consequently a different overlap with the Higgs field [40]. The localization parameters are natural, that means of the order of one, hence this mechanism solves the problem of the large hierarchies between the fermion masses. There are even proposals on how to generate the flavor and neutrino mixing in the framework of warped extra-dimensions with bulk fermions [41]. The bulk RS-theory would have a different phenomenology than the RS1 model. Although the graviton modes would still have a mass in the TeV range and would be produced through gluon fusions at hadron colliders, the main decay modes would be different. Depending on the specifics of the embedding of the fermions in the bulk the gravitons would primarily decay into top pairs or pairs of  $Z$ -,  $W$ - or Higgs bosons. In Figure 2.2 branching ratios for the different

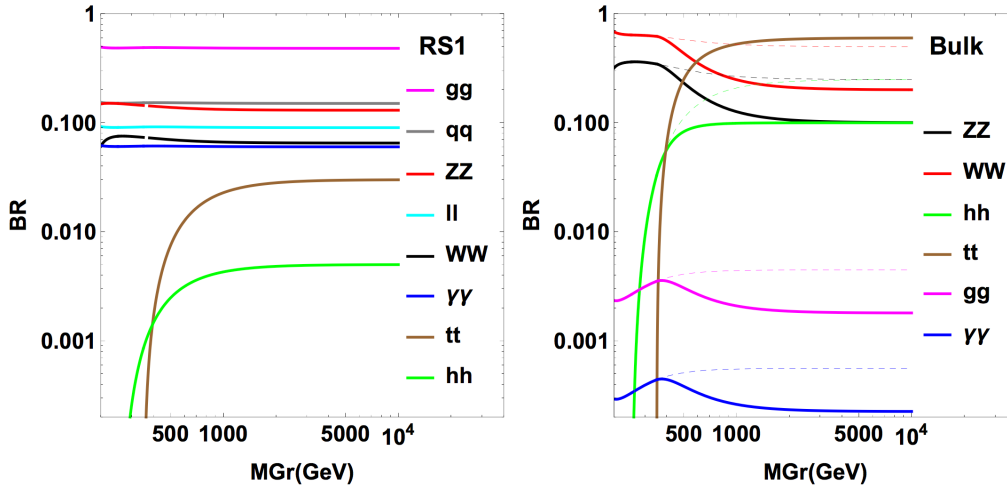


Figure 2.2: Branching ratios of the graviton mode for the RS1-(left) and the Bulk-theory(right). The symbol  $q$  stands for the sum of all quarks without the top quark, while  $l$  represents the sum of all three flavors of leptons or neutrinos. The dashed and solid line represent two extrema of fermion embedding in the bulk [45].

scenarios are compared. This thesis is motivated by the search for the graviton mode in the  $G \rightarrow hh \rightarrow \tau\tau bb$  decay mode. From the values of Table 2.2 we can derive the branching ratio of the two Higgs bosons:  $\text{BR}(hh \rightarrow bb\tau\tau) = 7.2\%$ . Although there are more frequent final states than  $bb\tau\tau$ , it is the aim of this thesis to investigate the reconstruction and identification of tau pairs, because they have a very distinctive signature. Earlier searches constrain the mass of the first graviton mode to be in the one to multi TeV range [42–44]. For gravitons with masses in this region the Higgs bosons would have a very substantial momentum. The aim of the next section is to study, what geometry is then to expect for the  $h \rightarrow \tau\tau$  decay from a standpoint of relativistic kinematics.

## 2.3 Boosted Topologies

If a particle with a high momentum decays, the angle between the momenta of the decay products is on average low. This is referred to as boosted topology. Demanding four-momentum conservation and fixing the decay axis in  $z$ -direction, one gets the following four-momenta in the

rest frame of the mother particle:

$$\mathcal{P}_M^R = \begin{pmatrix} m_M \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \mathcal{P}_{1,2}^R = \frac{m_M}{2} \begin{pmatrix} 1 \pm (2\eta - 1) \xi^2 \\ 0 \\ 0 \\ \pm \sqrt{((2\eta - 1) \xi^2 + 1)^2 - 4\eta^2 \xi^2} \end{pmatrix}$$

with

$$\eta = \frac{m_1}{m_1 + m_2} \quad \xi = \frac{m_1 + m_2}{m_M}$$

with  $m_1$  the mass of the first decay particle,  $m_2$  the mass of the second decay particle and  $m_M$  the mass of the mother particle. This parameter change is useful as the geometry should only depend on the proportion between the masses of the particles. The results in the rest frame are understood quite intuitively. The decay particle with the higher mass carries more energy and the topology is back-to-back.

The next step is to switch from the rest frame of the mother particle into the laboratory frame, where the momentum of the mother particle lies in the  $x$ - $z$ -plane and has an absolute value of the three-momentum of  $p_M^L$ . The restriction to the  $x$ - $z$ -plane is possible without loss of generality, because of the axial symmetry around  $z$ . The change of the reference frame is executed by applying the following Lorentz boost to all 4 vectors:

$$\mathcal{L} = \begin{pmatrix} \gamma & -\beta\gamma \sin \theta & 0 & -\beta\gamma \cos \theta \\ -\beta\gamma \sin \theta & (\gamma - 1) \sin^2 \theta + 1 & 0 & (\gamma - 1) \sin \theta \cos \theta \\ 0 & 0 & 1 & 0 \\ -\beta\gamma \cos \theta & (\gamma - 1) \sin \theta \cos \theta & 0 & (\gamma - 1) \cos^2 \theta + 1 \end{pmatrix}$$

with:

$$\beta\gamma = \frac{p_M^L}{m_M} \quad \gamma = \frac{E_M^L}{m_M}$$

with  $p_M^L$  the absolute value of the momentum of the mother particle in the laboratory frame and the angle  $\theta$  between  $\vec{p}_M^L$  and the  $z$ -axis. The angle between the decay products in the laboratory frame is then given by:

$$\varphi(\vec{p}_1^L, \vec{p}_2^L) = \arccos \frac{\vec{p}_1^L \cdot \vec{p}_2^L}{|\vec{p}_1^L| |\vec{p}_2^L|} = \varphi(\eta, \xi, \beta\gamma, \theta)$$

Writing down the whole formula is spared for the sake of shortness. In Figure 2.3 the function  $\varphi(\eta, \xi, \beta\gamma, \theta)$  is depicted for different interesting fixed values of  $\eta$ ,  $\xi$ , and  $\beta\gamma$ . The first interesting observation is, that for low  $\beta\gamma < 1$  the plots look similar. This is the limit of classical mechanics.

For a high boost  $\beta\gamma$  of the mother particle a large fraction of the possible  $\theta$ -values results in low angular separation  $\varphi$  between the decay products.

If the boost of the mother particle is high compared to the boost of the daughter particles in the rest frame, the direction of both daughter particles is flipped in the same direction ( $\varphi < \pi/2$ ) for any  $\theta$ . If the fraction of the mass of the decay products with respect to the mass of the mother particle  $\xi$  is high, the boost of the decay products in the rest frame is lower. Therefore a lower  $\beta\gamma$  of the mother particle is sufficient to flip the momenta of both decay products in the same direction. The value of  $\xi$  is the only difference between the cases  $h \rightarrow \tau\tau$  (see Figure 2.3 (a),  $\xi \approx 3\%$ ) and  $\Upsilon(1S) \rightarrow \tau\tau$  (see Figure 2.3,  $\xi \approx 37\%$  (b)) in this consideration.

The case where one of the decay products is much heavier than the other ( $\eta = 1$ ) is depicted on the base of the  $\tau \rightarrow \pi^\pm \nu$  decay in Figure 2.3 (c). Note that the neutrino is considered massless. The notable difference is, that even if  $\beta\gamma$  is high, it is possible that the topology between  $\nu$  and  $\pi^\pm$  in the laboratory frame is back-to-back, if the boost is in the direction of the  $\pi^\pm$  ( $\theta = \pi$ ).

For high  $\beta\gamma$  a large fraction of the  $\theta$  values yields approximately the same  $\varphi$  value than  $\varphi(\theta = \pi/2)$ . Therefore this value is a good rule of thumb for the expected angular separation between the daughter particles. With the additional restriction to the case  $\eta = 0.5$  where both daughter particles have the same mass, this leads to:

$$\varphi(\eta = 0.5, \xi, \beta\gamma, \theta = \pi/2) = \arccos\left(\frac{(\beta\gamma)^2 + \xi^2 - 1}{|(\beta\gamma)^2 - \xi^2 + 1|}\right)$$

for  $\beta\gamma > 1$  this can be written as:

$$\varphi(\eta = 0.5, \xi, \beta\gamma, \theta = \pi/2) = \arccos\left(1 - \frac{2(1 - \xi^2)}{(\beta\gamma)^2 + 1/(\beta\gamma)^2(1 - \xi^2)^2}\right)$$

using the first term of the Taylor expansion  $\arccos(1-x) \approx \sqrt{2x}$  at  $x = 0$  this can be approximated for large  $\beta\gamma$  by:

$$\varphi(\eta = 0.5, \xi, \beta\gamma \gg 1, \theta = \pi/2) \approx \frac{2\sqrt{1 - \xi^2}}{\beta\gamma}$$

This leads for  $\xi \approx 0$  to the frequently shown rule of thumb:

$$\varphi \approx \frac{2m_M}{p_M} \tag{2.2}$$

For a graviton of mass  $m_G = 1.3$  TeV the Higgs bosons would have a momentum of approximately  $p_H \approx 640$  GeV, therefore the pair of tau leptons would have an angular separation of only  $\varphi \approx 0.4$ . This value plays a crucial role in the reconstruction of hadronic tau leptons (see Section 4.5) and leads to the necessity to investigate a different approach for the reconstruction and identification of boosted tau pairs.



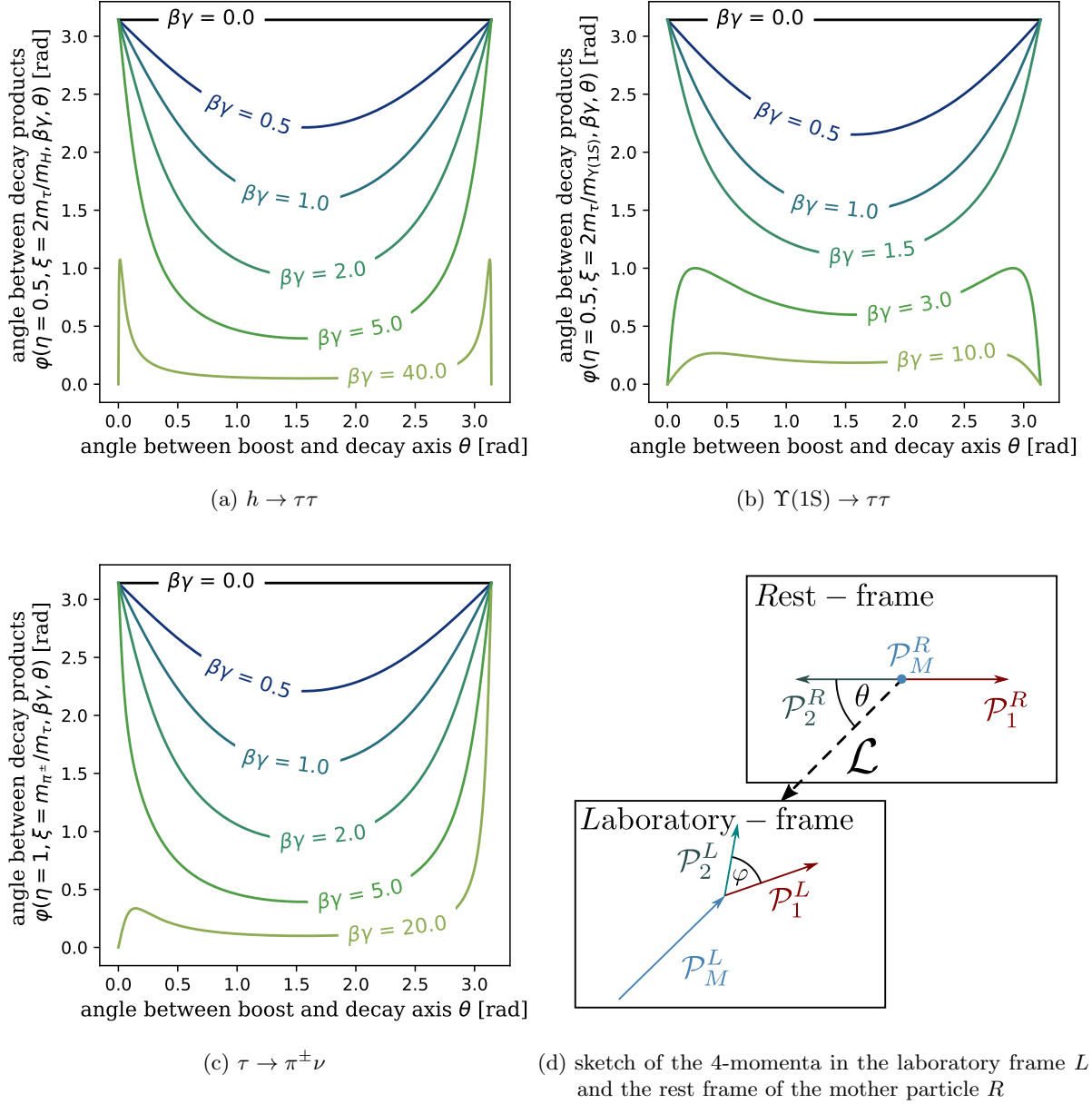


Figure 2.3: Angle  $\varphi$  between decay products for different decays in dependency on the angle  $\theta$  between the boost and the decay axis in the rest frame. The labeled lines represent different values  $\beta\gamma$  of the boost of the mother particle.



# 3 The ATLAS Experiment at the LHC

## 3.1 The Large Hadron Collider

The LHC [46] is a circular particle accelerator of the European Organization for Nuclear Research (CERN). The LHC is located near Geneva on the border between Switzerland and France. It is installed in a 26.7 km long tunnel, which previously housed the Large Electron-Positron Collider (LEP). The LHC is used to accelerate beams of protons or lead ions. The particle beams pass through a chain of pre-accelerators and are then injected into the two parallel beam pipes of the LHC. The particles of the two beam pipes circulate in opposite directions. The LHC is able to accelerate protons to energies of 6.5 TeV. To keep particles of such high energies on the intended trajectory the LHC is equipped with superconducting niobium-titanium magnets, which produce magnetic fields of up to 8 T. The LHC has to be cooled down to 2 K to achieve the superconductivity of the magnets.

The experiments CMS [47], LHCb [48], ALICE [49] and ATLAS [50] are located at the beam crossings of the LHC where the particles are brought to collision.

Besides the very high center of mass energy of the proton collisions, the LHC also achieves an unprecedented instantaneous luminosity  $\mathcal{L}$ . The integrated luminosity  $\mathcal{L}$  is connected to the expected number of events  $N_{\text{event}}$  through the cross-section of the event  $\sigma_{\text{event}}$ .

$$\mathcal{L} \sigma_{\text{event}} = \int dt \mathcal{L} \sigma_{\text{event}} = N_{\text{event}}$$

Therefore a high luminosity is crucial for the observation of events with very low cross-sections. The LHC has surpassed all expectations in producing instantaneous luminosities of over  $2 \times 10^{34} \text{ cm}^{-2} \text{ s}^{-1}$  in 2017 and in delivering a total integrated luminosity of  $93 \times 10^{39} \text{ cm}^{-2}$  to ATLAS from 2015 to 2017 during the ongoing run 2 (see Figure 3.1). Consequently even processes like  $pp \rightarrow G \rightarrow hh$  which are expected to have cross-sections lower than  $10^{-39} \text{ cm}^2$  [51] are in the reach of observability.

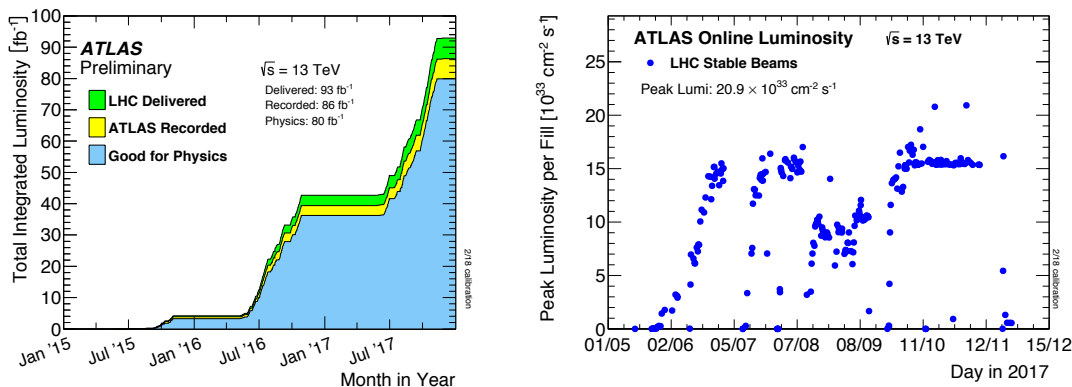


Figure 3.1: Integrated luminosity delivered to the ATLAS detector by the LHC (left). One  $\text{fb}^{-1}$  is equal to  $10^{39} \text{ cm}^{-2}$ . Peak instantaneous luminosity produced by the LHC on each day of the run in 2017 (right) [52].

## 3.2 The ATLAS Detector

The purpose of the ATLAS detector is to measure the signature of the products of the proton-proton collisions. It thereby allows for the reconstruction of the momentum, direction and type of final state particles (Chapter 4). The ATLAS detector is approximately 40 m long and 20 m high (Figure 3.2).

The origin of the ATLAS coordinate system lies in the collision point. The  $z$ -axis is aligned with the beam axis. The  $x$ -axis points to the center of the LHC and the  $y$ -axis points upwards. The azimuthal angle  $\phi$  is measured in the  $x$ - $y$ -plane starting from the  $x$ -axis. The polar angle  $\theta$  is measured from the  $z$ -axis. Instead of the polar angle  $\theta$  the pseudo rapidity  $\eta$  is commonly used.

$$\eta = -\ln \tan(\theta/2)$$

Consequently pseudo rapidity  $\eta = 0$  describes the  $x$ - $y$  plane and  $\eta \rightarrow \infty$  the  $z$ -axis. ATLAS is mirror-symmetric at the  $\eta = 0$  plane and approximately axial-symmetric around  $z$ .

The spatial distance  $\Delta R$  between two points in the angular plane is defined as:

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \quad (3.1)$$

The multiple subsystems of the ATLAS detector will be described in the following sections. The first system a particle originating from the collision point traverses is the inner detector (Section 3.2.1). The inner detector is surrounded by the calorimeters (Section 3.2.2). The muon detectors are the outermost layer of ATLAS (Section 3.2.3).

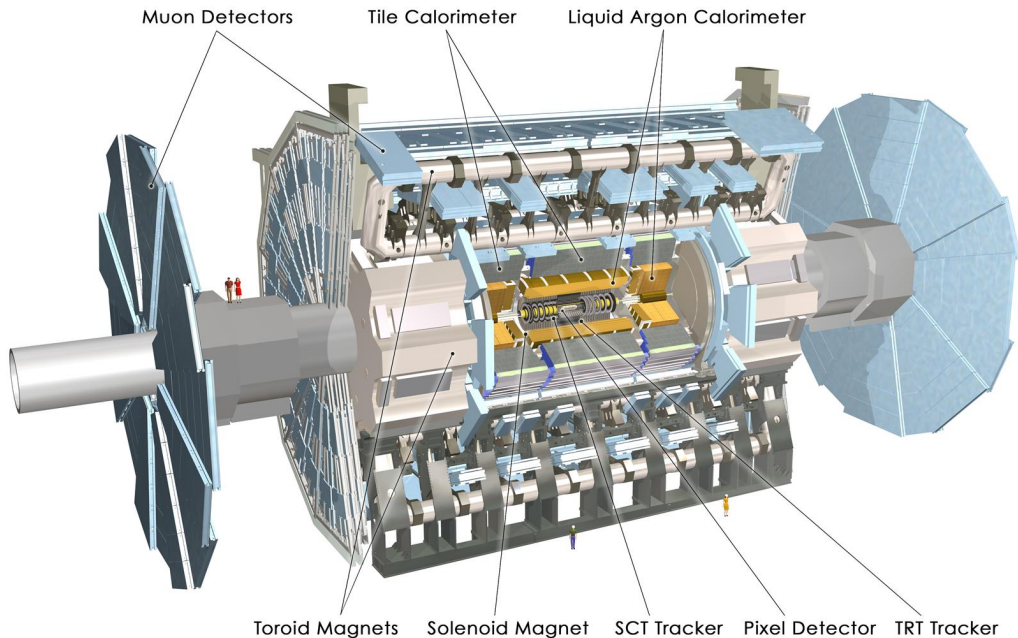


Figure 3.2: Computer generated image of the ATLAS detector [53].

### 3.2.1 Inner Detector

The main purpose of the Inner Detector (ID) is to allow for the reconstruction of trajectories of charged particles (tracks) and of production and decay vertices. The vertex reconstruction is especially useful for the detection of particles with long decay length like  $b$ -quarks and  $\tau$ -leptons. The ID is immersed in a magnetic field of 2 T, consequently the momentum and charge sign of electrical charged particles can be measured from the curvature of their tracks. In Figure 3.3 the different parts of the ID are depicted. The ID is approximately 6.2 m long and 2.1 m high.

The Pixel Detector is the subsystem nearest to the beam axis in the region where the distance to the beam is smaller than  $R < 122.5$  mm. It consists of 1744 similar modules with 46732 silicon-pixels each. The pixels have a minimum size of  $R \times (z - \phi) = 50 \times 400 \mu\text{m}^2$ . The modules are arranged in three layers. In the region  $|\eta| < 2.5$  the layers have a cylindrical shape with different  $R$  values. In the region  $|\eta| > 2.5$  there are three discs on each side located on planes defined by different  $z$  values. In May 2014 during the first long shutdown an additional innermost layer was installed, the so called Insertable B-Layer (IBL) [54]. The IBL is located in the region of  $31 \text{ mm} < R < 40 \text{ mm}$  around the beam axis. It accounts for 860000 additional pixels and is thereby allowing for a vertex reconstruction with enhanced resolution.

The Semiconductor Tracker (SCT) is the second subsystem of the ID located in the region  $299 \text{ mm} < R < 560 \text{ mm}$ . It uses modules with two silicon micro strips with a strip pitch of  $80 \mu\text{m}$  which are arranged in a stereo angle among each other. It has four coaxial cylindrical layers in the central region and  $2 \times 9$  discs in the forward and backward regions. The SCT consists of 15912 sensors.

The Transition Radiation Tracker (TRT) is the outermost layer of the ID. It is located in the region  $563 \text{ mm} < R < 1066 \text{ mm}$  around the beam axis. It consists of tubes of 4 mm diameter filled with a mixture of gas and with an anode wire in the center. In the central region the tubes are arranged parallelly to the beam axis and in the end-cap region they are arranged radially. Therefore the TRT does not allow for a direct measurement of the  $z$ -coordinate in the central region and of the  $R$ -coordinate in the end-caps. Although the information provided by the TRT is used for tracking, the main feature is, that it allows for the distinction of electrons from heavier charged particles like pions due to the different expected amount of transition radiation they produce.

### 3.2.2 Calorimeters

On their path through the calorimeters particles interact with the material and deposit energy. It is the purpose of the calorimeter system to measure this energy. Neutrinos pass the detector without interaction and muons only leave a small fraction of their energy in the calorimeters. Because of the amount, composition and arrangement of the calorimeter material all other particles deposit their whole energy in them. The detector covers the full  $2\pi$   $\phi$ -range. Therefore the negative sum of all measured momenta in the transverse plane can be used as an estimate of the sum of transverse neutrino momenta. This is referred to as missing transverse momentum  $\mathbf{p}_T^{\text{miss}}$ .

The calorimeters are sampling calorimeters. This means that they are constructed alternately from an absorber and a detection layer. The high-energy particles interact with the absorber layer and cause showers of secondary particles that ionize the active medium in the detection

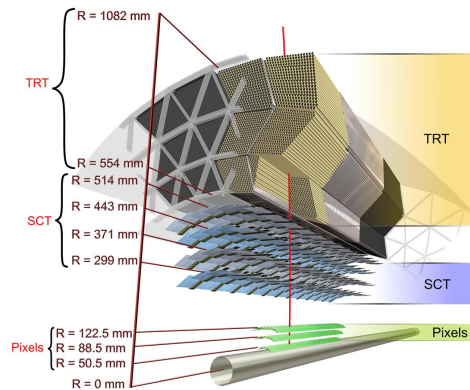


Figure 3.3: Computer generated image of the inner detector of ATLAS [55]. The innermost IBL is missing in this figure.

layer. The ions and electrons drift to the electrodes and cause a current from which the energy deposition in the cell is reconstructed.

An overview of the different calorimeter systems of the ATLAS detector is given in Figure 3.4

The liquid argon calorimeter covers the region  $|\eta| < 4.9$ . It uses liquid argon as active material and lead or copper as absorber material. It consists of the electromagnetic barrel calorimeter ( $|\eta| < 1.475$ ) in the center as well as an electromagnetic ( $1.375 < |\eta| < 3.2$ ) and a hadronic ( $1.5 < |\eta| < 3.2$ ) end-cap calorimeter at each side. The forward calorimeters ( $3.1 < |\eta| < 4.9$ ) are located between end-cap calorimeter and beam. The calorimeters have three or four layers of cells. The granularity of the cells in the middle layer of the electromagnetic calorimeter is  $\Delta\eta \times \Delta\phi = 0.025 \times 0.025$ .

The liquid argon calorimeter is surrounded by the tile barrel calorimeter ( $|\eta| < 1.0$ ) and the tile extended barrel calorimeter ( $0.8 < |\eta| < 1.7$ ). The tile calorimeter uses steel as the absorber material and scintillating tiles as the active material.

### 3.2.3 Muon Spectrometer

The muon spectrometer is the outermost system of the ATLAS detector and contributes the largest volume fraction. It consists of precision tracking chambers, trigger chambers and large magnets. The precision tracking chambers are the monitored drift tube chambers ( $|\eta| < 2.7$ ) and the cathode strip chambers ( $2.0 < |\eta| < 2.7$ ). The cathode strip chambers are used in the innermost regions to withstand the stringent conditions due to the high background rates in this region. A muon with  $\eta < 2.7$  passes on average three precision tracking chambers. To allow for a high momentum resolution even for muons of very high momenta  $p_T \approx 1$  TeV the position of the tracking chambers has to be known with a precision of less than  $30 \mu\text{m}$ , but they are installed with a precision of only  $5 \text{ mm}$ . To solve this problem, the muon spectrometer has an active optical alignment system installed to measure the relative distance between the parts with the

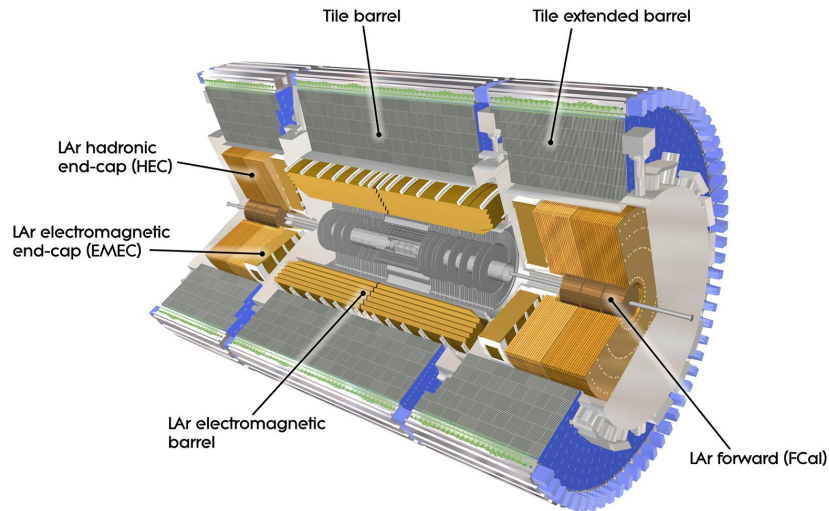


Figure 3.4: Computer generated image of the calorimeters of ATLAS [56].

necessary precision. Together with the exact measurement of the magnetic field produced by the barrel toroid and end-cap toroid magnets this ensures a momentum measurement with a precision of approximately 10% for muons of  $p_T \approx 1$  TeV.

There are two types of trigger chambers, the resistive plate chambers ( $|\eta| < 1.05$ ) and the thin gap chambers ( $1.05 < |\eta| < 2.4$ ). They allow for an additional faster but less precise track direction and momentum measurement. This is used to trigger on events containing muons above a momentum threshold.

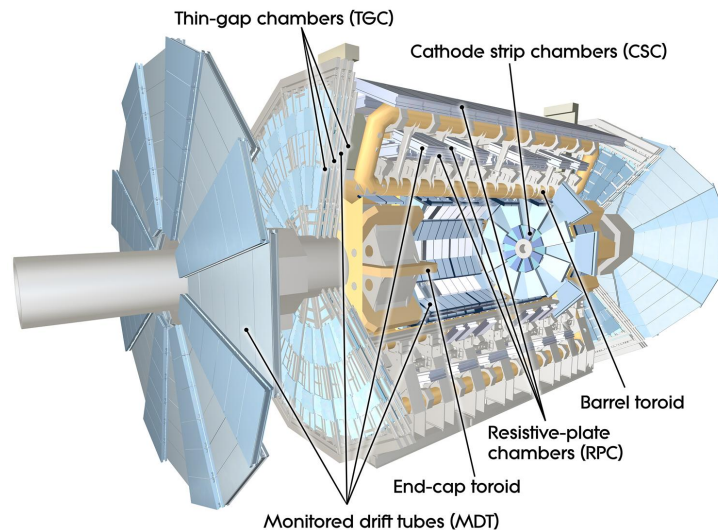


Figure 3.5: Computer generated image of the muon systems of ATLAS [57].

### 3.2.4 Trigger System

During the current run of the LHC bunches of protons collide every 25 ns. Consequently it is not feasible to store the complete readout of the ATLAS detector for every collision. Therefore an online event selection of physical important signatures is necessary. This is the task of the trigger system. It had an upgrade in 2014 during the first long shutdown [58].

The first level of the trigger system is hardware based and reduces the rate from 40 MHz down to less than 100 kHz. It consists of different subsystems optimized to select events with i.a. electrons, muons, hadronic tau leptons, jets and missing energy. These objects have to pass a certain energy threshold or are required to have certain topology among each other to activate the trigger.

The high level trigger is the second tier of the trigger system. It is software based and allows consequently for more sophisticated decisions based on reconstructed particles, which are often a simplified version of the ones created in the offline reconstruction. It is able to reduce the rate further down to 1 kHz. This leads to a total output bandwidth in the order of a few hundred MB/s.



# 4 Particle Reconstruction and Identification at ATLAS

In the following chapter an overview over various techniques to recognize different final state particles with the ATLAS detector is given.

In Section 4.2 the reconstruction of low level objects, e.g. tracks in the inner detector are described. These low level objects are used in the reconstruction of electrons (Section 4.3), muons (Section 4.4), hadronic tau decays (Section 4.5) and boosted fully hadronic di-tau decays (Section 4.6). Each of the particle reconstruction algorithms has a corresponding identification algorithm.

It is the aim of the particle reconstruction to provide candidates with maximum efficiency  $\epsilon$ . In this context efficiency is defined as the fraction of particles which pass the reconstruction:

$$\epsilon^{\text{reco}} = \frac{N_{\text{truth-match}}^{\text{reco}}}{N^{\text{truth}}},$$

where  $N^{\text{truth}}$  is the number of true particles of the type the reconstruction is intended for and  $N_{\text{truth-match}}^{\text{reco}}$  is the number of reconstructed candidates produced by the  $N^{\text{truth}}$  particles.

The particle identification algorithms are optimized to reject candidates produced by background processes and thereby increase the purity among the reconstructed particles. The purity  $p$  is the fraction of candidates originating from the  $N^{\text{truth}}$  particles with respect to the number of all candidates. For this task multivariate classification techniques play an important roll. The idea of these techniques is to combine the information of multiple properties of one candidate to decide whether it is more likely signal or background. Two of these techniques are introduced in the next sections.

## 4.1 Multivariate Classification Techniques

### 4.1.1 Boosted Decision Trees

Boosted Decision Trees (BDTs) are one frequently applied multivariate technique for classification tasks. They are used for the identification of hadronic tau decays (Section 4.5) and for the identification of boosted di-taus in the  $\tau_{\text{had}}\tau_{\text{had}}$  channel (Section 4.6). Their performance in the identification of  $\tau_{\text{lep}}\tau_{\text{had}}$  di-taus is investigated in this thesis (Chapter 5). All of these applications are currently using the implementation provided by TMVA [59].

BDTs are trained on signal and background candidates whose true classification is known. They consist of multiple decision trees (see Figure 4.1). A decision tree has multiple nodes. At every node the candidates are split into two sets depending on the value of one of their multiple classification variables. The last node a candidate passes decides whether the candidate is classified as signal or background.

The cuts at each node are optimized during training. The criterion most commonly used for this uses the Gini-impurity:

$$I_G = p(1 - p)$$

with  $p$  the fraction of true signal candidates in the set. At each node the cut which is optimizing:

$$s = I_G^P - \epsilon^1 I_G^1 - \epsilon^2 I_G^2$$

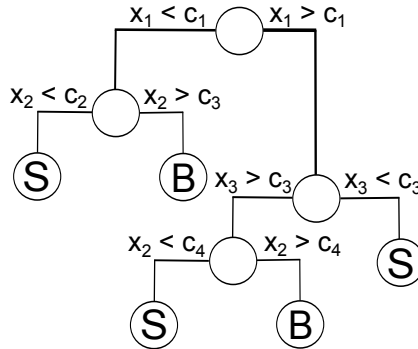


Figure 4.1: Depiction of a decision tree. The candidates are separated into two categories at each node by a cut in one variable  $x_i \geq c$ . The split in the leaf nodes determines the result of the classification.

is chosen, with  $I_G^P$  the Gini-impurity of the set before the split,  $I_G^{1,2}$  the Gini-impurity in the sets after the split and  $\epsilon^{1,2}$  the fraction of elements in the sets after the split. A single decision tree produced by this method is in most cases not optimal. One reason for this is that the algorithm is not stable against small fluctuations in the training set. Small differences in the statistical composition of the training sample might lead to a completely different cut sequences. Another reason is that for high tree depths the method will fit to the statistical noise of the sample. This is referred to as overtraining or overfitting. Boosting is used to overcome this deficit.

The idea behind boosting is to combine multiple weak classifiers into one strong classifier. In this case the weak classifiers are decision trees with a low number of nodes of approximately 7. For the BDTs in this thesis AdaBoost [60] is used. With AdaBoost the decision trees are trained one after another. The importance of candidates misclassified by previous decision trees is increased for the training of the following trees. They get a boost weight of:

$$\alpha = \frac{1 - f_{\text{err}}}{f_{\text{err}}}$$

with the misclassification rate  $f_{\text{err}}$  of the previous tree. That means trees with a low misclassification rate produce higher boost weights for wrongly classified candidates. The decision of the resulting forest is given by a weighted sum of the decisions of the single trees. The weight of a tree is determined by the logarithm of its boost weight  $\alpha$ . The result of the weighted sum is a real number between -1 and 1 referred to as BDT score. The more signal like a candidate is classified the higher its BDT score.

Because of the low tree depth of each individual tree, BDTs are very resilient against overtraining. They are also not negatively affected by strong correlations of input variables and yield in most cases a competitive performance compared to other multivariate techniques.

### 4.1.2 Projective Likelihood Method

The projective likelihood method [59] is a multivariate classification algorithm based on the estimation of the probability density functions of the input variables for signal and background. This can be done e.g. via a histogram or a kernel density estimation.

The likelihood score for a given vector of input variables  $\mathbf{x}$  is calculated with Equation 4.1, where  $P_{S(B),i}$  is the estimated signal (background) probability density function for the  $i$ -th variable. The classification is applied by a cut in the likelihood score.

$$\mathcal{L} = \frac{\mathcal{L}_S}{\mathcal{L}_S + \mathcal{L}_B} \quad \mathcal{L}_{S(B)} = \prod_{i=0}^n P_{S(B),i}(x_i) \quad (4.1)$$

The likelihood method is the optimal method if the variables are statistically independent of each other. In most cases it is difficult to find a set of independent variables which at the same time allow for a good signal-background separation. A possible solution to this problem is to use multi-dimensional probability density estimations for dependent variables.

## 4.2 Low Level Object Reconstruction

In this section the reconstruction of low level objects, which are used for the particle reconstruction later on, is summarized.

### 4.2.1 Inner Detector Track Reconstruction

The first step of inner detector track reconstruction [61, 62] is the building of three dimensional space-points from the measurements of the hits in the Pixel- and SCT detector. A hit in the pixel detector commonly results in multiple adjacent pixels exceeding the measurement threshold. These pixels are grouped into clusters. An artificial neural net is used to identify clusters, which are produced by the hits of more than one charged particle. This is for example useful to detect the highly collimated tracks of boosted hadronic tau decays.

The seeds for the track reconstruction are formed by sets of three space-points of different layers. The information of these space-points is used with the assumption of a helical trajectory to calculate a rough estimate of the momentum and the impact parameter. A large fraction of the seeds does not correspond to the trajectory of a real particle. To increase the purity of the tracks some baseline cuts on the momentum and the impact parameter are applied.

A Kalman filter [63] is used to add additional matching space-points to the seed. This is an iterative approach. The currently estimated track and its uncertainty are used to predict and add new space-points to the track. The added space-point updates the estimated track and its uncertainty.

On the resulting collection of tracks an ambiguity solver is applied to further improve purity. It is mainly based on the quality for example in terms of numbers of contained space-points. In the next step the tracks are extended with matching hits in the TRT. At last a high-precision track fit is performed.

Aside from this inside-out approach starting from SCT and Pixel tracker hits there is also an

outside-in approach starting from hits in the TRT. This is useful for the reconstruction of tracks with large energy loss or with a secondary vertex behind the first layers of the silicon detectors.

For tracks with a large displacement with respect to the interaction point there is also another algorithm, which is a slightly changed version of the standard inside-out approach [64].

### 4.2.2 Primary Vertex Reconstruction

The mean number of inelastic interactions between the protons per bunch crossing was  $\langle\mu\rangle = 37.8$  in the 2017 data taking at ATLAS [52]. It is the purpose of the primary vertex reconstruction [65] to measure the coordinates of the multiple interactions occurring at each bunch crossing. The tracks are assigned to the different vertices in this process. Consequently tracks originating from the interaction of interest can be distinguished from tracks from a different coincidental interaction (“pile-up”).

The primary vertex reconstruction in Run 2 is using a new imaging technology [66]. It uses all tracks of one event as input to find all primary vertices simultaneously.

The transverse and longitudinal impact parameters  $d_0$  and  $z_0$  can be calculated for tracks with respect to a primary vertex. The transverse impact parameter  $d_0$  is defined as the transverse distance of the closest approach of the track to the primary vertex. The longitudinal impact parameter  $z_0$  is the longitudinal distance of this closest approach to the primary vertex.

### 4.2.3 Calorimeter Cell Clustering

Adjacent cells of the calorimeter containing a significant energy deposition are grouped. This is referred to as clustering. The purpose of clustering is to isolate and measure the energy depositions of single particles. Therefore the measurements of the cells in each cluster are summed and calibrated.

At ATLAS there are two different approaches used for clustering. The topological clustering approach and the sliding window algorithm.

The sliding window algorithm [67] sums over the energy of all cells in a defined  $\Delta\eta_{\text{window}} \times \Delta\phi_{\text{window}}$  area. If a cell lies only partially in the window a fraction of its energy is added to the window. The fraction is given by the area of the cell that is contained in the window in respect to the whole area of the cell. If the energy in the window exceeds a threshold, a cluster candidate is build. For the candidate the direction is calculated as the barycenter of the energy measurements of the cells. After building all possible candidates, double counts are removed. This is done by comparing clusters whose directions lie in a  $\Delta\eta_{\text{overlap}} \times \Delta\phi_{\text{overlap}}$  window and keeping only the cluster of highest energy.

The topological clustering approach [68] uses the cell significance  $\varsigma_{\text{cell}}$ :

$$\varsigma_{\text{cell}} = \frac{E_{\text{cell}}}{\sigma_{\text{noise,cell}}}$$

with the measured energy in the cell  $E_{\text{cell}}$  and the expected noise level in the cell  $\sigma_{\text{noise,cell}}$ . The topological clusters are seeded by cells of a high  $\varsigma > \varsigma_{\text{cell}}^{\text{seed}}$ . The next step is to continuously add adjacent cells to the cluster until they fall below a certain minimal  $\varsigma_{\text{cell}}^{\text{growth}}$ . In the last step a

boundary of cells with a lower significance  $\varsigma > \varsigma_{\text{cell}}^{\text{boundary}}$  is added. At ATLAS the default values for the parameters are:

$$\varsigma_{\text{cell}}^{\text{seed}} = 4 \quad \varsigma_{\text{cell}}^{\text{growth}} = 2 \quad \varsigma_{\text{cell}}^{\text{boundary}} = 0$$

#### 4.2.4 The Anti- $k_t$ Jet Clustering Algorithm

Due to the confinement of the QCD quarks and gluons produce showers of hadrons. These showers are referred to as QCD jets. Jet clustering algorithms are used to group all calorimeter energy deposition originating from one QCD jet. The most commonly used algorithm at ATLAS is the anti- $k_t$  algorithm [69].

Although the algorithm is mainly used to combine topological clusters into jets, any set of objects can be used, if each of the objects has a defined momentum and direction. Each step of the algorithm uses the transverse momenta  $p_{T,i}$  of the objects and the angular distance between them  $\Delta R_{ij}$ . From these values the distance  $d_{ij}$  between two objects and the distance  $d_{iB}$  between an object and the beam are calculated as denoted in Equation 4.2 and 4.3. The choice of the distance parameter  $R$  depends on the use case.

$$d_{ij} = \min(1/p_{T,i}^2, 1/p_{T,j}^2) \frac{\Delta R_{ij}^2}{R^2} \quad (4.2)$$

$$d_{iB} = 1/p_{T,i}^2 \quad (4.3)$$

In each step all distances  $d_{ij}$  and  $d_{iB}$  are calculated and the smallest distance is determined. If one of the distances  $d_{ij}$  is the smallest, the  $i$ -th and  $j$ -th object are merged into one object assigned with the sum of the momenta. If one of the distances  $d_{iB}$  is the smallest, the grooming of the  $i$ -th object is complete and it is removed in further steps. This procedure is repeated until all objects are removed. The completed objects are referred to as anti- $k_t(R)$ -jets.

Because of the definition of  $d_{ij}$  and  $d_{iB}$  the algorithm starts the merging process with objects of high momentum and does not merge objects with  $\Delta R_{ij} > R$ .

### 4.3 Electron Reconstruction and Identification

The electron reconstruction [70] at ATLAS is seeded by sliding window clusters build from energy depositions in the cells of the electromagnetic calorimeter. The window has a size of  $\Delta\eta \times \Delta\phi = 0.075 \times 0.125$  and the threshold energy is 2.5 GeV. If the position of the cluster matches to an extrapolated inner detector track, the electron reconstruction is complete. The calibrated energy of the electron candidate is based on a combination of track and calorimeter information. The sliding window algorithm might be substituted with topological clustering in the future [71].

The electron identification applies the projective likelihood method to discriminate candidates originating from electrons against candidates originating from e.g. pions or photons. For this 20 variables describing different properties of the electron candidate are considered. They characterize:

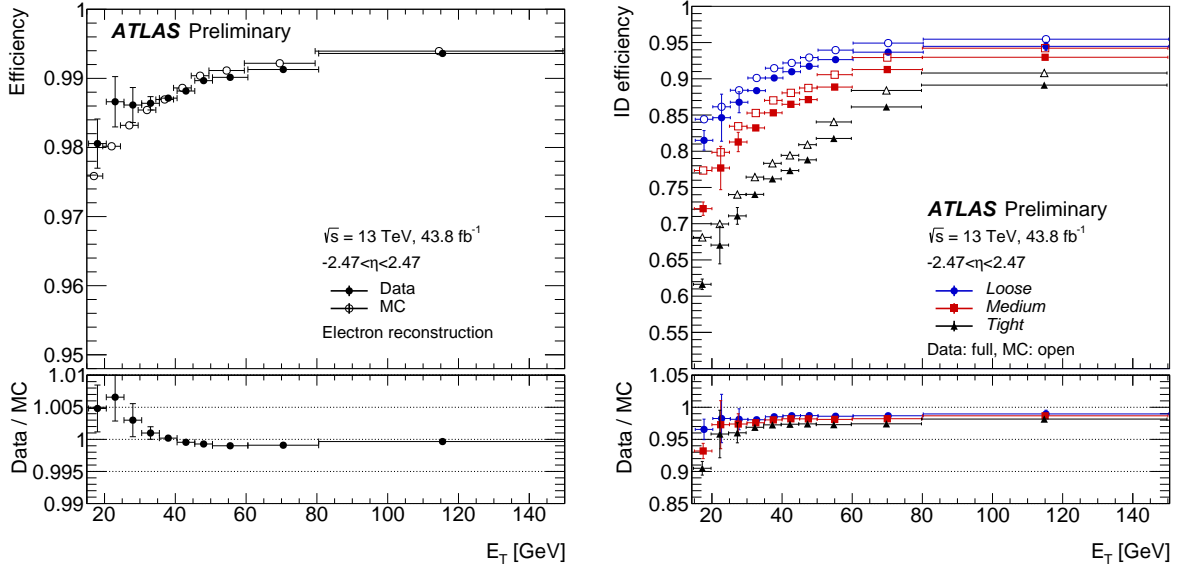


Figure 4.2: Electron reconstruction and identification efficiencies as a function of the transverse energy of the electron  $E_T$  derived from Monte Carlo simulation and a tag and probe based data analysis. The identification efficiency are derived with respect to the number of reconstructed electrons [70, 72].

- the fraction of energy deposited in the hadronic calorimeter
- the collimation and longitudinal distribution of the energy deposition in the calorimeter using information in a window of up to  $\Delta\eta \times \Delta\phi = 0.175 \times 0.175$  around the electron candidate
- the quality of the matching between the track and the cluster
- the transition radiation in the TRT
- the number of hits associated to the track in the different layers of the ID

The electron identification provides multiple working points: *Loose*, *Medium* and *Tight* corresponding to different cuts in the likelihood score. The likelihood method is optimized in bins of  $|\eta|$  and  $E_T$  of the electron candidate.

In Figure 4.2 the reconstruction and identification efficiencies are depicted as a function of the transverse momentum of the electron  $E_T$ . The efficiency of identifying candidates produced by background is estimated on Monte Carlo samples with di-jet events to be in the order of  $10^{-3}$  for all working points.

## 4.4 Muon Reconstruction and Identification

The first part of the muon reconstruction [73] is to reconstruct tracks in the muon spectrometer and the second part is to match tracks, track segments or calorimeter energy depositions among each other to build muon candidates.

The muon spectrometer tracks are reconstructed in multiple steps. At first track segments are reconstructed from the hits in the muon drift tube chambers and a nearby trigger chamber. Segments in the cathode strip chambers are reconstructed separately. Segments of different chambers are matched and a track fit is performed. Ambiguous tracks are removed while still allowing for shared segments among tracks to enable the reconstruction of highly collimated muons. Tracks with a low global track fit quality are rejected.

There are four different muon reconstruction algorithms. The *combined* muon algorithm utilizes hits from matching muon spectrometer and inner detector tracks. The hits are refitted taking i.a. energy loss into account. The *segment-tagged* muon algorithm starts from ID tracks and matches them to track segments in the muon spectrometer. This is useful, if the muon trajectory passes only one muon chamber. The *extrapolated* muon algorithm uses only the muon spectrometer information to reconstruct the muon track. This algorithm is useful in the high  $|\eta|$ -region not covered by the ID. The *calorimeter-tagged* muon algorithm looks for ID tracks, which match to calorimeter energy deposition compatible to a muon. It is independent from the muon spectrometer measurement. Ambiguities between candidates reconstructed by multiple reconstruction algorithms are resolved. Combined muons are prioritized.

The muon identification ensures the quality of the reconstruction and allows for the rejection of muons originating from in-flight kaon and pion decays. The discriminatory feature of these background muons is a kink in the track. Different variables are used to reject them. They characterize the disparity between the measured momentum and charge of the ID and muon spectrometer track as well as the global track fit quality.

The muon identification defines three working points: *Loose*, *Medium* and *Tight*. Candidates selected by a tighter working point are also included by the looser working points. The *Loose* working point uses candidates from all four muon reconstruction algorithms. *Segment-tagged* and *calorimeter-tagged* muons are only used in the region  $|\eta| < 0.1$ , where the muon spectrometer has service and cable ducts. The *Medium* working point includes only *extrapolated* and *combined* muons and the *Tight* working point selects only *combined* muons. The working points apply various different cuts in the background rejection variables to achieve different levels of purity. They also apply different quality criteria on e.g. the number of hits assigned to a track. There is also a designated working point for muons of high momentum above 100 GeV, which ensures a good momentum resolution. The muon efficiencies of the different working points are presented in Figure 4.3.

## 4.5 Reconstruction and Identification of Hadronic Tau Decays

The aim of the reconstruction of  $\tau_{\text{had}}$  decays [75] at ATLAS is to measure the summed momentum of the hadronic decay products denoted as  $\tau_{\text{had-vis}}$ . The algorithm aims for  $\tau_{\text{had}}$  decays with a transverse momentum of  $p_{\text{T}}(\tau_{\text{had}}) \gtrsim 20$  GeV where the decay products are collimated.

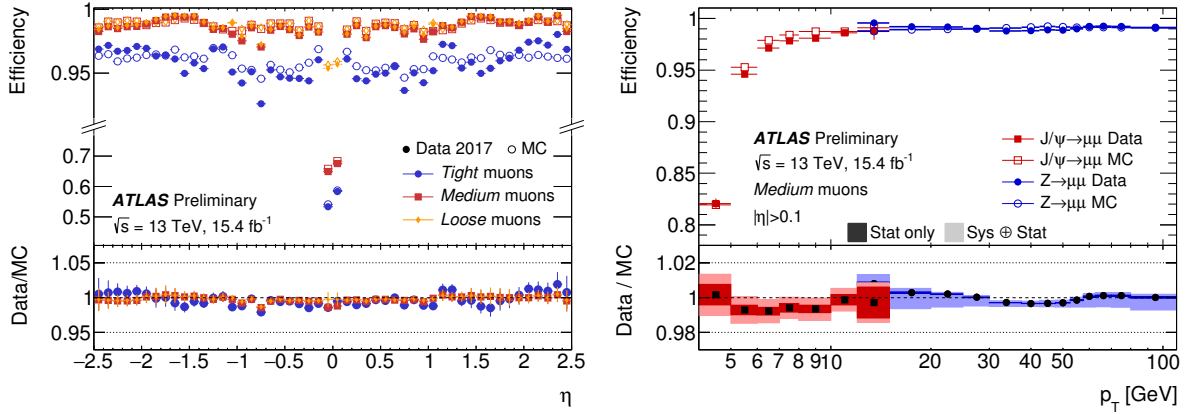


Figure 4.3: Muon identification efficiencies as a function of the pseudo-rapidity of the muon  $\eta$  for different working points (left). Muon identification efficiency for the Medium working point as function of the transverse momentum  $p_T$  (right). The values are obtained from data (full marker) and simulation (open marker) [73, 74].

The reconstruction is seeded by anti- $k_t(0.4)$ -jets with distance parameter  $R = 0.4$  build from topological calorimeter clusters, which are calibrated with the local hadronic calibration scale. Tracks in the cone of  $\Delta R < 0.2$  around the direction of the measured seed jet momentum are used to assign a primary vertex to the tau candidate.

Tracks in the  $\Delta R < 0.4$  cone around the seed-jet are considered as track candidates of the charged decay products of the  $\tau_{\text{had}}$ . Three BDTs are used to classify the tracks into four categories: *tau tracks*, *conversion tracks*, *isolation tracks* and *fake tracks* [76]. *Tau tracks* contain the tracks of the charged hadrons from the  $\tau_{\text{had}}$  decay. *Conversion tracks* are mainly tracks from electrons produced by photon conversion. *Isolation tracks* are tracks from underlying event. *Fake tracks* are tracks which do not fit in any other category and consist primarily of pile-up tracks. The track classification BDTs use variables like the angular distance  $\Delta R$  between track and seed-jet, the transverse momentum of the seed jet  $p_T^{\text{seed}}$  and the transverse impact parameter  $d_0$ .

The momentum and direction of the  $\tau_{\text{had-vis}}$  is extracted from the topological clusters associated to the jet, the *tau tracks* and the assigned primary vertex.

For the  $\tau_{\text{had}}$  identification only candidates with one or three tracks of the *tau track* class are considered. The main background for this identification algorithm are QCD jets. A BDT is used to distinguish candidates originating from true  $\tau_{\text{had}}$  decays from background candidates. It uses the following identification variables, which are also used in the identification algorithm for  $\tau_{\text{had}}\tau_{\mu}$  di-tau decays presented in this thesis:

- The **leading track impact parameter significance**  $|S_{\text{leadtrack}}|$  is the absolute value of the transverse impact parameter of the track with the highest transverse momentum divided by its estimated uncertainty.
- The **fraction of track  $p_T$  in the isolation region**  $f_{\text{iso}}^{\text{track}}$  is defined as the sum of the



transverse momenta of all *tau tracks* divided by the sum of transverse momenta of *tau tracks* and *isolation tracks*.

- The **central energy fraction**  $f_{\text{cent}}$  is calculated by dividing the transverse energy of all clusters with a barycenter in the  $\Delta R < 0.1$  region around the tau axis by the transverse energy of all clusters in the  $\Delta R < 0.2$  region.
- The **leading track momentum fraction**  $f_{\text{leadtrack}}^{-1}$  is the transverse momentum of the leading track divided by the energy of clusters in the  $\Delta R < 0.2$  region.
- The **track radius**  $R_{\text{track}}^{\text{tau}}$  is defined as:

$$R_{\text{track}}^{\text{tau}} = \frac{\sum p_{\text{T}}^{\text{track}} \Delta R_{\text{track}}}{\sum p_{\text{T}}^{\text{track}}}$$

with the sum running over all *tau tracks* and the angular distance  $\Delta R_{\text{track}}$  between the track and the  $\tau_{\text{had-vis}}$  direction.

- The **fraction of EM energy from charged pion**  $f_{\text{EM}}^{\text{track-HAD}}$  is described by the following equation:

$$f_{\text{EM}}^{\text{track-HAD}} = \frac{\sum p_{\text{T}}^{\text{track}} - E_{\text{cluster}}^{\text{Had}}}{E_{\text{cluster}}^{\text{EM}}}.$$

The sum runs over all *tau tracks*.  $E_{\text{cluster}}^{\text{Had}}$  indicates the transverse energy in the hadronic calorimeter and in the back layer of the electromagnetic calorimeter.  $E_{\text{cluster}}^{\text{EM}}$  stands for the energy in the electromagnetic calorimeter.

- The variable **ratio of EM energy to track momentum**  $f_{\text{track}}^{\text{EM}}$  is defined similar to the previous variable, but without the subtraction of the energy in the hadronic calorimeter.

$$f_{\text{track}}^{\text{EM}} = \frac{\sum p_{\text{T}}^{\text{track}}}{E_{\text{cluster}}^{\text{EM}}}$$

- The **track-plus-EM-system mass**  $m_{\text{EM+track}}$  is the invariant mass of the four-momentum acquired by summing over the four momenta of all *tau tracks* and the four momenta of the energy in the electromagnetic calorimeter of the two most energetic clusters associated to the  $\tau_{\text{had-vis}}$  candidate.
- The **ratio of track-plus-EM-system to  $p_{\text{T}}$**  denoted as  $p_{\text{T}}^{\text{EM+track}}/p_{\text{T}}$  is the transverse component of the four-momentum used to calculate  $m_{\text{EM+track}}$  divided by the transverse momentum of the  $\tau_{\text{had-vis}}$  candidate.
- To determine the **transverse flight path significance**  $S_{\text{T}}^{\text{flight}}$  the distance between the secondary vertex reconstructed from the *tau tracks* and the associated primary vertex is divided by the estimated uncertainty.
- The **track mass**  $m_{\text{track}}$  is the invariant mass of the sum of the four-momenta of all *tau tracks*.

- The **maximum angular distance**  $\Delta R_{\text{Max}}$  is defined as the maximum  $\Delta R$  between a *tau track* and the reconstructed  $\tau_{\text{had-vis}}$  direction.

Two different BDTs are deployed for the one and three *tau track* cases. The first two variables in the listing are only used for the case of one *tau track* and the last three are only used in the case of three *tau tracks*. For each working point the cut-values in the BDT score vary depending on  $p_T$  and the number of average interactions per bunch-crossing to achieve a flat signal efficiency in these variables. The identification provides three working points: *Loose*, *Medium* and *Tight*. They correspond to different signal efficiencies summarized in Table 4.1. In the future an artificial neural net might supersede the BDT in the  $\tau_{\text{had}}$  identification [77].

Table 4.1: Signal efficiency of the different  $\tau_{\text{had}}$  identification working points for one and three *tau tracks*. The values describe the efficiency to identify a candidate, which is already reconstructed with one or three *tau tracks*:

$N_{\text{tau tracks}}$	<i>Loose</i>	<i>Medium</i>	<i>Tight</i>
1	85 %	75 %	60 %
3	75 %	60 %	45 %

## 4.6 Reconstruction and Identification of Boosted Di-Taus in the $\tau_{\text{had}}\tau_{\text{had}}$ decay mode

For events with two hadronically decaying tau leptons with low angular separation  $\Delta R < 0.4$  the individual reconstruction of the  $\tau_{\text{had-vis}}$  of each  $\tau_{\text{had}}$  does not work anymore. The reason for this is the radius parameter  $R = 0.4$  used in the anti- $k_t$  algorithm producing the seed jets for the  $\tau_{\text{had-vis}}$  reconstruction. These jets can only resolve energy depositions separated by more than  $\Delta R > 0.4$ . Consequently the energy depositions of the  $\tau_{\text{had-vis}}$  of both tau leptons merge into one seed jet and the individual momenta and directions of the  $\tau_{\text{had-vis}}$  can not be measured with this jet.

A solution to this problem was introduced in [21]. The algorithm aims to reconstruct both  $\tau_{\text{had}}$  decays as one object. It uses large anti- $k_t(1.0)$ -jets build from topological clusters as a seed. This seed jet contains the whole energy deposition of the two collimated tau leptons. It is required to have a transverse momentum of  $p_T > 300$  GeV due to file size constraints. To isolate the energy deposition of each  $\tau_{\text{had-vis}}$  small anti- $k_t(0.2)$  subjets are build from the clusters in the large anti- $k_t(1.0)$ -jet. The two most energetic subjets are associated to the  $\tau_{\text{had-vis}}$  of the two individual parent particles. A primary vertex association is applied in a similar way as in the reconstruction of individual hadronic tau decays. For this all tracks in the region of the seed jet are considered, if they pass minimal quality criteria:

- $p_T > 1$  GeV,
- Number of pixel detector hits  $\geq 2$ ,

- Number of pixel detector hits + SCT hits  $\geq 7$ .

Tracks in the seed jet are assigned to subjets if their  $\Delta R$  to the subjet is smaller than 0.2. For subjets with more than one assigned track a secondary vertex is reconstructed. After the vertex association a final track selection with cuts on the impact parameters with respect to the associated primary vertex is performed:

- $d_0 \leq 1 \text{ mm}$
- $|z_0 \sin \theta| \leq 2 \text{ mm}$

The main source for di-tau candidates produced by background are QCD jets. A BDT is trained to distinguish signal from background. Some of the input variables characterize the two leading subjets individually, other combine the information of all subjets or tracks. A subset of these variables is used in the identification of tau pairs in the  $\tau_{\text{had}}\tau_e$  decay mode and is described in the corresponding section.



# 5 Di-Tau Reconstruction and Identification in the $\tau_{\text{lep}}\tau_{\text{had}}$ Channel

In this chapter the performance of the existing reconstruction and identification algorithms for electrons, muons and hadronic tau decays applied on boosted semi-leptonic tau pair decays is evaluated. Possible alternatives for the regions where the existing algorithms cease to work are presented. The  $\tau_{\text{had}}\tau_e$  and  $\tau_{\text{had}}\tau_\mu$  channels are discussed separately in Section 5.1 and 5.2. Although they both contain a hadronically decaying tau, the mutual influences between the detector signatures of the electron or muon and the  $\tau_{\text{had}}$  impose different challenges.

For all studies signal Monte Carlo samples are used, where the process  $G_{\text{RS}}^{\text{bulk}} \rightarrow hh \rightarrow \tau\tau\tau\tau$  is simulated. A listing of the samples can be found in Table A.1. Each sample is simulated with a different graviton mass between one and five TeV. This leads to tau pairs with different distributions of the visible transverse momentum of the di-tau decay products  $p_{\text{T}}(\text{di}\tau_{\text{vis}})$  and of the angular distance between them  $\Delta R(\tau_{\text{had-vis}}^{\text{truth}}, e/\mu)$  (Figure 5.1). The peaks of the  $\Delta R$  distributions correspond approximately with the values calculated with Equation 2.2. At higher graviton masses the Higgs bosons have a higher momentum and consequently the tau pairs are more collimated. For a graviton mass of  $m_G = 2$  TeV approximately 70% of the tau pairs have an angular distance in the region of  $\Delta R < 0.4$ . The distributions of  $p_{\text{T}}(\text{di}\tau_{\text{vis}})$  are shifted to higher values in the samples of higher graviton masses. Consequently  $\Delta R$  and  $p_{\text{T}}(\text{di}\tau_{\text{vis}})$  are anti-correlated in the combined sample. The samples consist of  $1.492 \times 10^6$  events in total and contain approximately  $687.4 \times 10^3$   $\tau_{\text{had}}\tau_e$  di-taus and  $672.8 \times 10^3$   $\tau_{\text{had}}\tau_\mu$  di-taus.

The events in the Monte Carlo samples are produced with Pythia 8 [78]. GEANT 4 [79] with a full model of ATLAS [80] is used to simulate the detector response. The resulting files are processed with the ATLAS reconstruction software. The Monte Carlo samples are similar to the ones obtained from real data taking, but contain additional truth information describing the four-momentum and decay chain of each particle in the event.

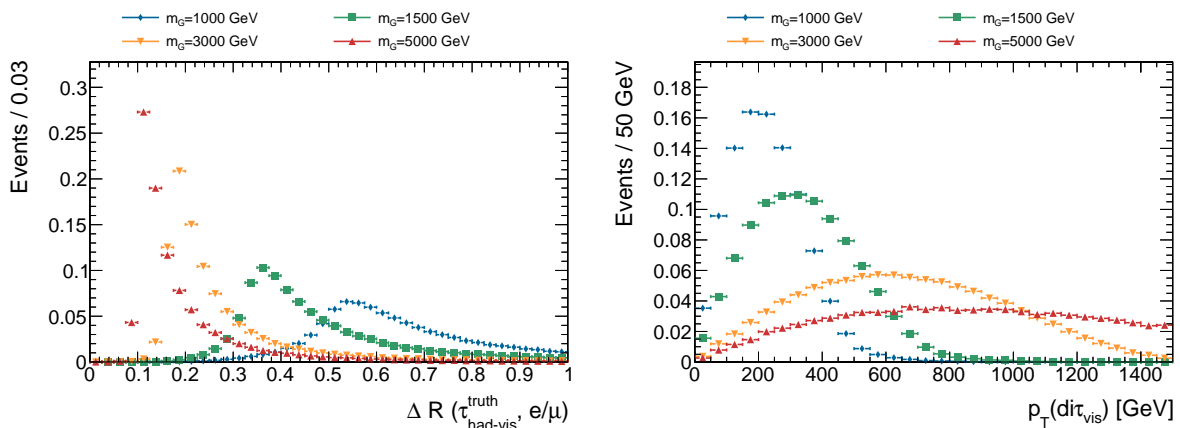


Figure 5.1: Normalized distribution of  $\Delta R$  between  $\tau_{\text{had-vis}}^{\text{truth}}$  and the electron or muon of the leptonically decaying tau (left) and of the transverse momentum of the visible decay products of the di-tau system  $p_{\text{T}}(\text{di}\tau_{\text{vis}})$  (right) in Monte Carlo sample of the  $G \rightarrow hh \rightarrow \tau\tau\tau\tau$  process.

## 5.1 The $\tau_{\text{had}}\tau_e$ Channel

The first approach for the reconstruction of  $\tau_{\text{had}}\tau_e$  decays with a low spatial separation is to use the standard algorithms described in Chapter 4. The main challenge for the standard algorithms is given by the overlap of the energy depositions of the hadronic tau decay products and of the electron in the calorimeter.

### 5.1.1 Performance of Existing Algorithms

#### Performance of the Reconstruction of Hadronic Tau Decays in the Vicinity of an Electron

To evaluate the efficiency of the  $\tau_{\text{had-vis}}$  reconstruction and identification with a nearby electron all hadronically decaying taus in the truth content of the Monte Carlo samples ( $\tau_{\text{had}}^{\text{truth}}$ ) are taken into account. If there is a reconstructed  $\tau_{\text{had-vis}}$  candidate in the  $\Delta R < 0.4$  vicinity around the momentum sum of the visible decay products of the truth tau ( $\tau_{\text{had-vis}}^{\text{truth}}$ ) the truth tau is considered as reconstructed. The  $\tau_{\text{had-vis}}$  candidates are not required to have one or three classified *tau tracks* as described in Section 4.5. The reconstruction efficiency is defined as the number of  $\tau_{\text{had}}^{\text{truth}}$  considered as reconstructed divided by the number of all  $\tau_{\text{had}}^{\text{truth}}$ . Because of the missing requirement on the number of *tau tracks* this is equivalent to the efficiency to reconstruct an anti- $k_t(0.4)$  jet in the vicinity of the  $\tau_{\text{had-vis}}^{\text{truth}}$  direction. The  $\tau_{\text{had}}^{\text{truth}}$  are considered as identified if the matched  $\tau_{\text{had-vis}}$  candidate passes the standard tau identification selection. To evaluate the influence of a nearby electron the  $\tau_{\text{had}}^{\text{truth}}$  reconstruction and identification efficiencies are depicted as a function of the  $\Delta R$  between the  $\tau_{\text{had-vis}}^{\text{truth}}$  and the nearest truth electron  $e^{\text{truth}}$  in Figure 5.2.

The  $\tau_{\text{had}}$  reconstruction is almost fully efficient independently of the angular distance to a nearby electron. For  $\Delta R(\tau_{\text{had-vis}}^{\text{truth}}, e^{\text{truth}}) > 0.4$  the identification efficiencies for all working points are approximately in the region expected from Table 4.1. The slight decrease with respect to the values in the table is caused by the cases in which the reconstruction or classification of one or more *tau tracks* fails. If the electron gets closer than  $\Delta R = 0.4$  to the  $\tau_{\text{had-vis}}^{\text{truth}}$  the identification efficiencies drop to approximately 10%. On the one hand this is due to the additional track of the electron which alters the track based identification variables of the  $\tau_{\text{had}}$  candidate (e.g.  $m_{\text{EM}+\text{track}}$ ). On the other hand the energy depositions of the  $\tau_{\text{had-vis}}^{\text{truth}}$  and the  $e^{\text{truth}}$  merge into one anti- $k_t(0.4)$  seed jet. Consequently the momentum of the  $\tau_{\text{had-vis}}^{\text{truth}}$  can not be resolved individually and the direction of the reconstructed  $\tau_{\text{had-vis}}$  candidate is drastically shifted in comparison to the  $\tau_{\text{had-vis}}^{\text{truth}}$  direction. A shift of the seed jet direction has a strong influence on identification variables which take this direction as point of reference (e.g.  $f_{\text{cent}}$ ).

In Figure 5.3 the distribution of the variable  $\Delta R_{\text{frac}}$  for the case  $0.3 < \Delta R(\tau_{\text{had-vis}}^{\text{truth}}, e^{\text{truth}}) < 0.4$  is depicted. The variable indicates the shift of the  $\tau_{\text{had-vis}}$  candidate on the axis between the  $\tau_{\text{had-vis}}^{\text{truth}}$  and the  $e^{\text{truth}}$  in the  $\eta \times \phi$ -plane. The peak at  $\Delta R_{\text{frac}} = 0$  corresponds to the case where the  $\tau_{\text{had-vis}}^{\text{truth}}$  direction is reconstructed correctly. This occurs in approximately 65% of the cases. In approximately 15% of the cases the direction of the anti- $k_t(0.4)$ -jet is in the middle between the  $e^{\text{truth}}$ - and  $\tau_{\text{had-vis}}^{\text{truth}}$ -direction ( $\Delta R_{\text{frac}} \approx 0.5$ ) and in 20% of the cases in the direction of the  $e^{\text{truth}}$  ( $\Delta R_{\text{frac}} \approx 1$ ). Although it is feasible to isolate the candidates with  $\Delta R_{\text{frac}} \approx 0$  with kinematic cuts e.g. on the ratio between the measured momenta of the  $\tau_{\text{had-vis}}$  and the electron candidates, it would come at the cost of a low reconstruction efficiency.

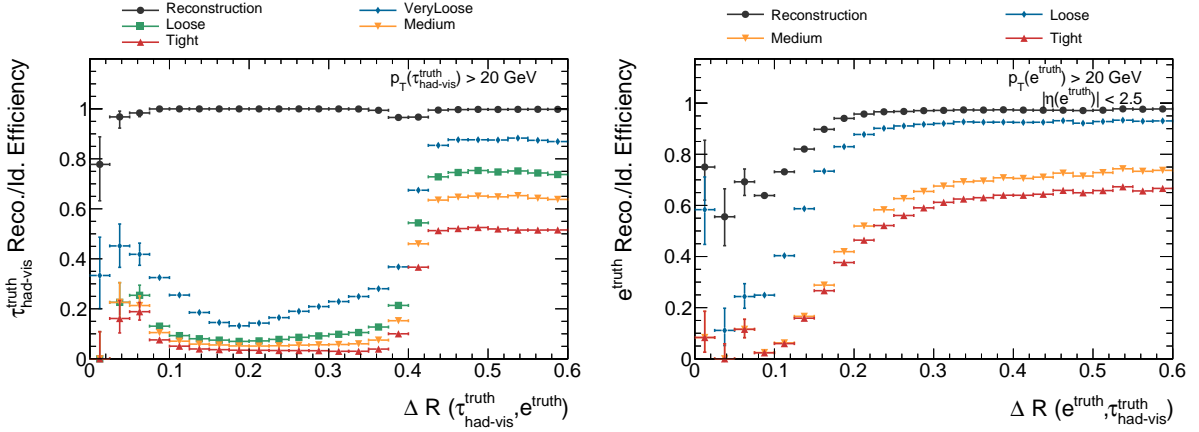


Figure 5.2: Reconstruction and identification efficiency of the standard  $\tau_{\text{had}}$  reconstruction as a function of the  $\Delta R$  between the  $\tau_{\text{had-vis}}^{\text{truth}}$  and the electron  $e^{\text{truth}}$  (left). Electron reconstruction and identification efficiency as a function of  $\Delta R$  to a nearby  $\tau_{\text{had-vis}}^{\text{truth}}$  (right). The different colors indicate the efficiencies of different working points.

Because of the directional shift, it is not easily possible to correct the identification variables for the influence of the additional electron in the  $\Delta R < 0.4$  cone. For example the information needed to recalculate  $f_{\text{cent}}$  with a different reference point is only available in earlier processing steps. Besides the problems for the identification the standard algorithm is also not able to measure the momentum of the  $e^{\text{truth}}$  and the  $\tau_{\text{had-vis}}^{\text{truth}}$  individually. This is for example necessary to calculate the invariant mass of the visible di-tau system, which is an important variable in physics analyses. Consequently a new approach for the reconstruction of a  $\tau_{\text{had}}$  decay with a nearby electron is needed.

### Performance of the Electron Reconstruction in the Vicinity of a Hadronic Tau Decay

The electron reconstruction and identification efficiencies with a nearby  $\tau_{\text{had-vis}}^{\text{truth}}$  are evaluated in a similar way as in the reversed case. All truth electrons  $e^{\text{truth}}$  originating from a leptonically decaying tau in the samples are taken into account. They are considered as reconstructed if there is a matching electron candidate. The matching is done based on the track of the electron candidate. If the  $e^{\text{truth}}$  caused the majority of the hits of the track it is associated to the electron candidate. The reconstruction and identification efficiencies as a function of the angular distance to the nearest  $\tau_{\text{had-vis}}^{\text{truth}}$  are depicted in Figure 5.2.

For  $\Delta R > 0.2$  the efficiencies of the reconstruction is flat and consequently not affected by the nearby hadronic tau. In the region  $\Delta R < 0.2$  the reconstruction efficiencies begin to decrease. A possible explanation is that, if the energy deposition of the  $\tau_{\text{had-vis}}^{\text{truth}}$  leaks into the sliding window of the electron reconstruction, its barycenter is shifted and the cluster-track matching fails. The identification efficiencies begin to decrease for values  $\Delta R < 0.3$ . This corresponds to the point where the  $\tau_{\text{had-vis}}^{\text{truth}}$  begins to deposit energy in the area around the electron, which is used for the calculation of the electron identification variables. The efficiencies for the *Medium* and *Tight*

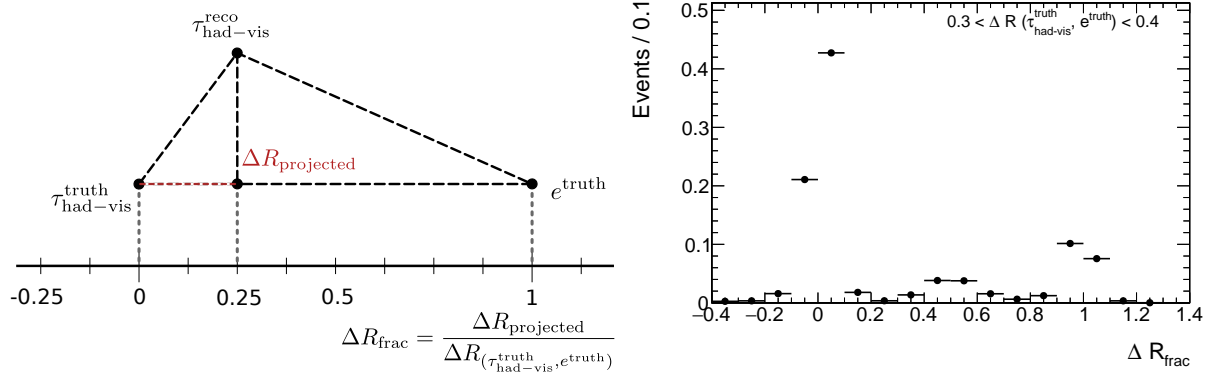


Figure 5.3: Sketch describing the variable  $\Delta R_{\text{frac}}$  (left) and the normalized distribution of  $\Delta R_{\text{frac}}$  for the case  $0.3 < \Delta R(\tau_{\text{had-vis}}^{\text{truth}}, e^{\text{truth}}) < 0.4$  (right).

working points are substantially lower than expected from Figure 4.2. The main differences of the electrons in the  $G \rightarrow hh \rightarrow \tau\tau\tau\tau$  samples in comparison to the electrons from  $Z \rightarrow ee$  events in Figure 4.2 are the more extreme kinematics of up to  $p_T \sim 1$  TeV and the higher displacement of the track originating from the decay vertex of a tau lepton.

The electron algorithms do perform sufficiently well for  $\Delta R > 0.2$  and can therefore be used to extend the sensitivity of the di-tau reconstruction to lower spatial separations.

### 5.1.2 Reconstruction

In this section a new algorithm for the reconstruction of boosted  $\tau_{\text{had}}\tau_e$  pairs is described. It uses the objects produced by the di-tau reconstruction for the fully hadronic decay mode. The algorithm is able to resolve the energy depositions of the  $\tau_{\text{had-vis}}^{\text{truth}}$  and the  $e^{\text{truth}}$  for  $\Delta R > 0.2$ . It consists of the following steps:

1. Search for a large anti- $k_t(1.0)$  seed jet with  $p_T > 300$  GeV built from topological calorimeter clusters.
2. Build small anti- $k_t(0.2)$  subjets from the clusters in the seed jet. They are required to exceed a  $p_T = 15$  GeV threshold.
3. Search for reconstructed electron candidates in the region of the seed jet.
4. Create a  $\tau_{\text{had}}\tau_e$  candidate for each subjet-electron combination with  $\Delta R(\text{subjet}, \text{electron}) > 0.1$ . The subjets are required to be among the two with highest  $p_T$ .
5. Select the tracks and associate the primary vertex in the same way as for the fully hadronic algorithm described in Chapter 4.6.

An exemplary sketch of the different reconstructed and truth objects in the seed jet region of a  $\tau_{\text{had}}\tau_e$  pair is given in Figure 5.4. The high  $R$  parameter of the seed jet ensures that it



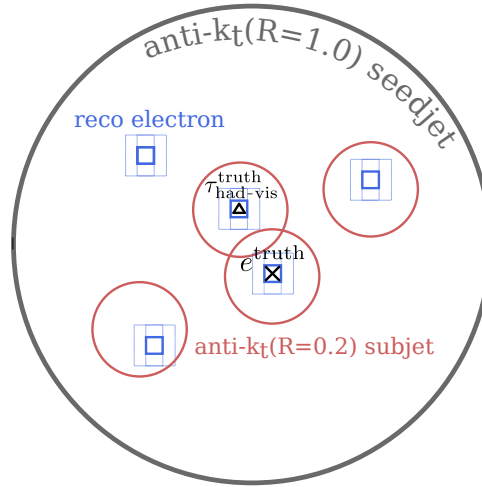


Figure 5.4: Sketch of the content of one anti- $k_t(1.0)$  seed jet reconstructed in the region of a boosted  $\tau_{\text{had}}\tau_e$  pair. The blue boxes ( $\Delta\eta \times \Delta\phi = 0.175 \times 0.175$ ) indicate the directions of reconstructed electron candidates. The red circles symbolize the anti- $k_t(0.2)$  subjets which are used to reconstruct the momentum  $\tau_{\text{had-vis}}^{\text{truth}}$ .

contains the energy depositions of the decay products of both taus completely. The large seed jet does also provide additional information about the isolation of the energy depositions, which is useful for discrimination against background from QCD jets later on. The high requirement on the transverse momentum of the seed jet of  $p_T > 300$  GeV is a remnant from the fully hadronic di-tau reconstruction.

The subjets are used to measure the energy and direction of the  $\tau_{\text{had-vis}}^{\text{truth}}$ . In most cases there are multiple subjets and electrons reconstructed in one seed jet. Most of them are caused by pile-up or underlying event. It is likely that the  $\tau_{\text{had-vis}}^{\text{truth}}$  and  $e^{\text{truth}}$  produce both a reconstructed electron and a subjet. The reconstructed electron is then most probable in the center of the subjet. This is the reason for the lower boundary on the angular separation  $\Delta R(\text{subjet}, \text{electron}) > 0.1$ . Although a higher boundary of  $\Delta R > 0.2$  would be well motivated by the distance parameter of the subjets the lower value is chosen to investigate whether the identification algorithm is able to adapt to the case where electron and subjet merge.

It would be desirable to minimize the number of candidates by applying additional requirements on the electron and subjet before matching them. The two most obvious choices would be cuts on the transverse momentum or on the working point passed by the electron. Cutting on the transverse momentum of the electron is problematic, because the two neutrinos involved in the leptonic tau decay leave in many cases only a low fraction of the transverse momentum for the electron. To exploit beneficial effects of the identification process on the purity among the candidates in the signal sample this problem was postponed to be covered after the identification process.

If the track of an electron candidate matches to a truth electron and the subjet is in the  $\Delta R < 0.2$  vicinity to a  $\tau_{\text{had-vis}}^{\text{truth}}$  the  $\tau_{\text{had}}\tau_e$  candidate is considered as truth matched. The reconstruction efficiency is defined as the number of truth matched  $\tau_{\text{had}}\tau_e$  candidates divided by

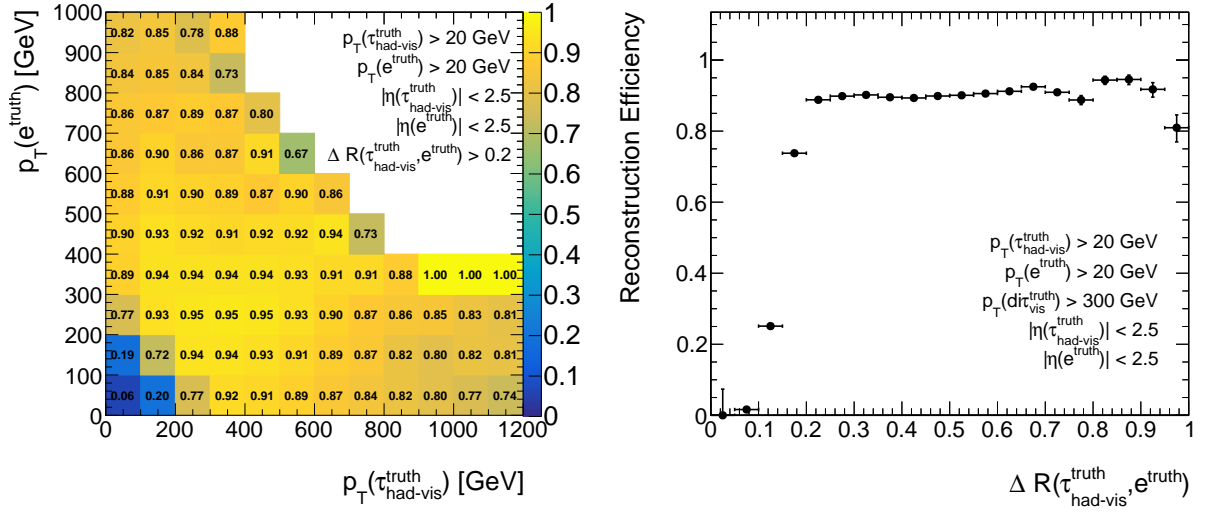


Figure 5.5: Reconstruction efficiency for  $\tau_{\text{had}}\tau_e$  pairs in 2D bins of  $p_T(e^{\text{truth}})$  and  $p_T(\tau_{\text{had-vis}}^{\text{truth}})$  (left) and in bins of  $\Delta R(\tau_{\text{had-vis}}^{\text{truth}}, e^{\text{truth}})$  (right). The efficiencies are evaluated with respect to truth  $\tau_{\text{had}}\tau_e$  pairs which fulfill the requirements indicated on the plots.

the number of truth tau pairs which share the same parent particle decaying in the  $\tau_{\text{had}}\tau_e$  channel. The reconstruction efficiency as a function of  $\Delta R(\tau_{\text{had-vis}}^{\text{truth}}, e^{\text{truth}})$ ,  $p_T(e^{\text{truth}})$  and  $p_T(\tau_{\text{had-vis}}^{\text{truth}})$  is depicted in Figure 5.5. If the visible transverse momentum of the truth di-tau system  $p_T(\text{di}\tau_{\text{vis}}^{\text{truth}})$  is lower than 300 GeV the reconstruction efficiencies suffers from the requirement on the seed jet energy. For low angular distance  $\Delta R(\tau_{\text{had-vis}}^{\text{truth}}, e^{\text{truth}}) < 0.2$  the reconstruction efficiency is decreasing because of the lower electron reconstruction efficiency in this region (see Figure 5.2) and because of the condition  $\Delta R(\text{subj}, \text{electron}) > 0.1$  the candidates have to fulfill. In the remaining kinematic regions the reconstruction algorithm is very efficient with  $\epsilon \approx 90\%$ .

In Figure 5.6 the relative deviation between corresponding quantities of the reconstructed particles and of the truth particles are presented e.g. the measured transverse momentum of the electron  $p_T(e^{\text{reco}})$  is compared to the transverse momentum of the truth electron  $p_T(e^{\text{truth}})$ . In the case where  $\Delta R(e^{\text{reco}}, \tau_{\text{had-vis}}^{\text{reco}}) > 0.2$  the absolute mean values of the relative deviations are one percent or lower. Consequently an additional energy scale factor would only allow for a marginal improvement. The standard deviations of the distributions which vary between 8 and 13% represent the resolutions of the measurements. A sophisticated multivariate energy scale might be able to improve these.

For  $\Delta R(e^{\text{reco}}, \tau_{\text{had-vis}}^{\text{reco}}) < 0.2$  the standard deviations of all distributions are larger. The mean value of the distribution of  $(p_T(e^{\text{reco}}) - p_T(e^{\text{truth}}))/p_T(e^{\text{reco}})$  has a larger off-set from zero than for  $\Delta R > 0.2$ . This means the energy deposition of the nearby  $\tau_{\text{had-vis}}^{\text{truth}}$  causes a slight increase of the reconstructed electron  $p_T(e^{\text{reco}})$  on average. The mean relative deviation of  $(p_T(\tau_{\text{had-vis}}^{\text{reco}}) - p_T(e^{\text{reco}}))$  from  $p_T(\tau_{\text{had-vis}}^{\text{truth}})$  is only 0.34(9)%. Subtracting the  $p_T(e^{\text{reco}})$  from

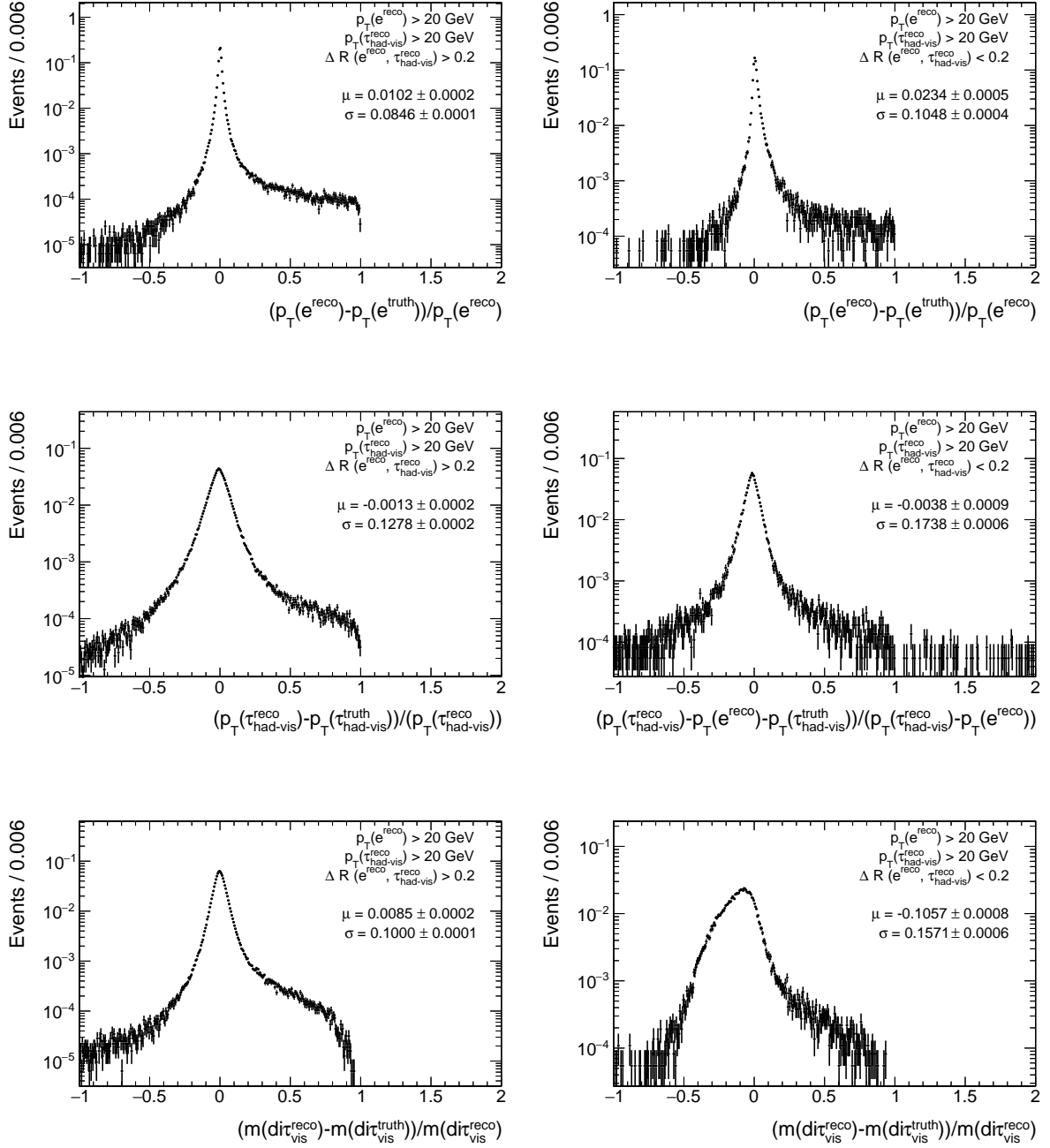


Figure 5.6: Normalized distributions of the fractional deviation of the reconstructed value from the truth value for  $p_T(e^{\text{reco}})$  (top),  $p_T(\tau_{\text{had-vis}}^{\text{reco}})$  (middle) and  $m(\text{dir}_{\text{vis}}^{\text{reco}})$  (bottom) in the region  $\Delta R(\tau_{\text{had-vis}}^{\text{reco}}, e^{\text{reco}}) > 0.2$  (left) and  $\Delta R(\tau_{\text{had-vis}}^{\text{reco}}, e^{\text{reco}}) < 0.2$  (right). The mean value  $\mu$  and standard deviation  $\sigma$  of the distribution as well as the baseline cuts on the reconstructed objects are indicated.

$p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}})$  is therefore a sufficient approximation for  $p_{\text{T}}(\tau_{\text{had-vis}}^{\text{truth}})$  in the  $\Delta R < 0.2$  regime. Nonetheless the  $(\eta, \phi)$ -direction of the measured momentum of the subject is still an insufficient approximation for the direction of the  $\tau_{\text{had-vis}}^{\text{truth}}$ . This is indicated by the large mean value and standard deviation of the distributions for the variable  $m(\text{di}\tau_{\text{vis}}^{\text{reco}})$ .

Besides a high reconstruction efficiency and a precise measurement of the momenta of the di-tau system the applicability in searches for new physics depends on a good discrimination between candidates produced from real di-taus and candidates from background. This is covered in the next section.

### 5.1.3 Identification

The aim of this section is to present a classifier, which ensures a high signal efficiency and background rejection. The background rejection is defined as the inverse background efficiency and corresponds approximately to the number of background candidates which are rejected by the classifier per accepted candidate. Ideally the signal efficiency should not depend on the production process of the di-tau system. Requiring a flat signal efficiency in kinematic variables like  $p_{\text{T}}(e^{\text{reco}})$ ,  $p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}})$ ,  $m(\text{di}\tau_{\text{vis}}^{\text{reco}})$  and  $\Delta R(e^{\text{reco}}, \tau_{\text{had-vis}}^{\text{reco}})$  is a prerequisite for this. A kinematically independent classifier allows for a flexible application in different analyses.

To train a multivariate classifier a pure signal and background sample is needed. The truth matched  $\tau_{\text{had}}\tau_e$  candidates from the simulated  $G \rightarrow hh \rightarrow \tau\tau\tau\tau$  samples are used as signal. The most common background producing  $\tau_{\text{had}}\tau_e$  candidates are QCD jets produced by quarks or gluons from the hard scattering of the protons. The number of events where a high momentum  $Z$ , Higgs or  $\Upsilon$ -meson decays into a boosted tau-pair is negligible compared to the number of candidates produced by background. This is exploited to obtain a very pure unbiased background sample by applying the reconstruction algorithm on real data with no stringent pre-selections. A fraction of the data acquired in 2015 is used. The only event selection is given by requiring large- $R$ -jet triggers to be activated. These triggers select events containing jets which exceed a certain  $p_{\text{T}}$ -threshold. The triggers with the lowest  $p_{\text{T}}$ -threshold which are not prescaled are used. Prescaled triggers only select a fraction of the events which pass the threshold. The sample and trigger names can be found in Table A.2 and Table A.3.

### Reweighting

The signal and background samples have a different kinematic composition. Especially the low  $p_{\text{T}}$  region is more abundant in the background sample. To prevent the BDT from rejecting background just by isolating different kinematic regions a weight is applied to each candidate of the background sample before the training. The weight depends on the  $p_{\text{T}}(e^{\text{reco}})$  and  $p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}})$  of the candidate. It is given by interpolating the ratio acquired by dividing the number of signal candidates by the number of background candidates in 2D bins of  $p_{\text{T}}(e^{\text{reco}})$  and  $p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}})$ . A bilinear interpolation [81] is used, which is based on solving Equation 5.1 using the values of the four nearest bin centers.

$$f(x, y) = c_1 + c_2x + c_3y + c_4xy \quad (5.1)$$

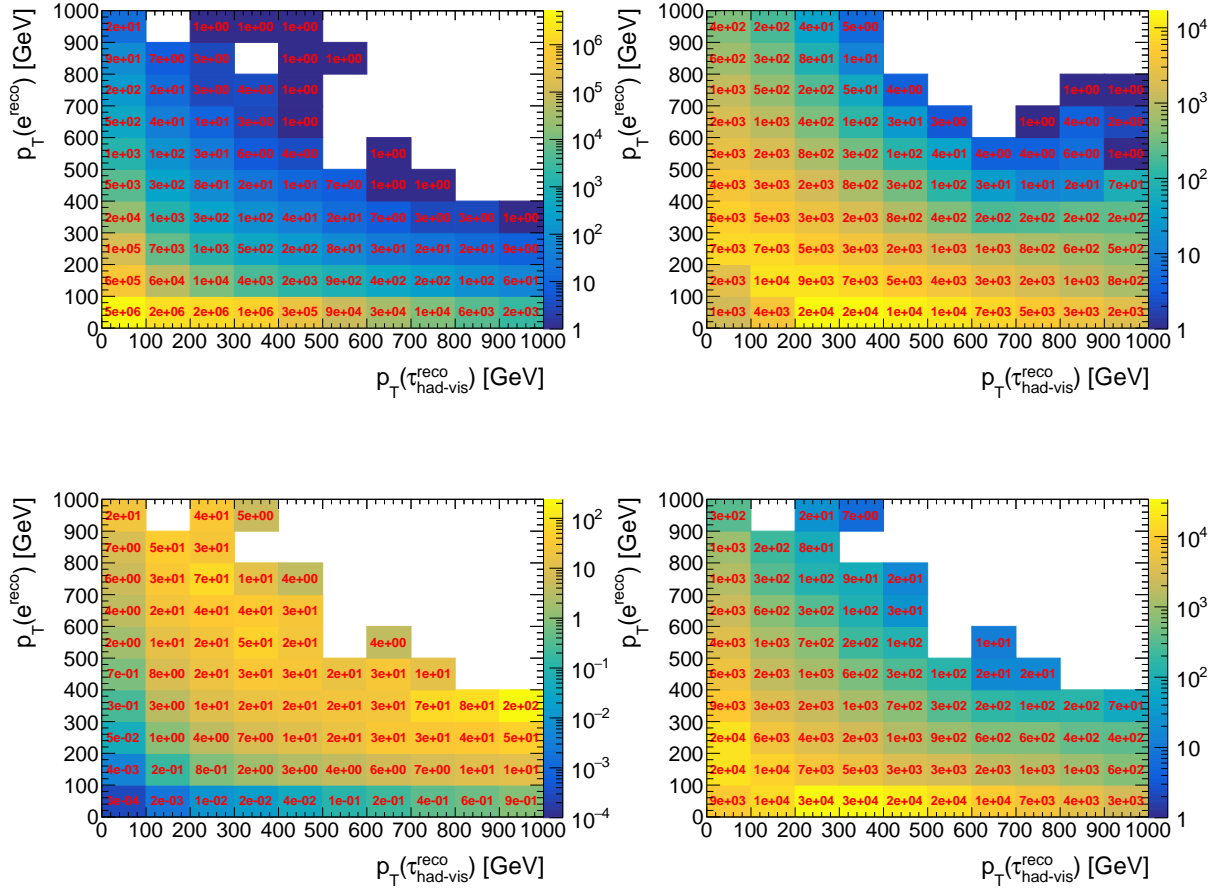


Figure 5.7: The depiction describes the reweighting applied on the background sample before the training. Number of background candidates before reweighting (top left). Weight applied on background candidates in the bin centers (bottom left). Number of signal candidates (top right). Number of background candidates after reweighting (bottom right).

The histograms of signal, background before reweighting and background after reweighting as well as the weight in the bin centers are depicted in Figure 5.7. The distribution of signal and background before reweighting are very different. Especially the bins below  $p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}}) + p_{\text{T}}(e^{\text{reco}}) < 300 \text{ GeV}$  are vastly overrepresented in the background sample. The fraction of events where either the  $\tau_{\text{had-vis}}^{\text{reco}}$  or the  $e^{\text{reco}}$  carries almost the whole momentum is also larger in the background. This leads to low weights in the region where  $p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}}) < 100 \text{ GeV}$  or  $p_{\text{T}}(e^{\text{reco}}) < 100 \text{ GeV}$  and to large weights if both objects have a high transverse momentum. The largest weights are in the order of one hundred and are only applied to a few background candidates. This can cause overfitting of the classifier to single distinctive highly weighted candidates. But the relative frequencies of candidates in bins with high weights are after the reweighting still only in the order of  $10^{-5}$ . The expected influence is therefore only small. After the reweighting the distributions of the signal and background sample have approximately the same shape. The difference in shape is caused by deviations of the true probability density from the underlying assumption of the bilinear interpolation.

### Identification Variables

In this section the identification variables are presented. Only high level variables are used. That means the variables describe physical motivated distinctive properties of the candidates. Variables which directly describe kinematic properties are avoided to foster a kinematically unbiased classifier. The variables can be grouped into three categories: variables describing properties of the  $e^{\text{reco}}$ , variables describing properties of the tau subjet and variables, which use information in the whole seed jet area.

At first variables using the information of the complete  $\tau_{\text{had}}\tau_e$  object are described. The signal and background distributions are depicted in Figure 5.8.

- The **di-tau energy fraction**  $\zeta_{\text{frac}}^{\text{seed}}$  is calculated by dividing the sum of transverse momenta of the  $e^{\text{reco}}$  and the tau subjet by the transverse momentum of the anti- $k_t(1.0)$  seed jet:

$$\zeta_{\text{frac}}^{\text{seed}} = \frac{p_{\text{T}}(e^{\text{reco}}) + p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}})}{p_{\text{T}}(\text{seed})}$$

For signal candidates most of the energy deposition in the seed jet region is caused by the di-tau. This leads to high values of  $\zeta_{\text{frac}}^{\text{seed}}$ . For background candidates the energy is mostly spread over multiple subjets leading to lower  $\zeta_{\text{frac}}^{\text{seed}}$  values.

- The **substructure scale**  $\zeta_{\text{frac}}^{\text{subsubl}}$  describes the fraction of transverse energy of the third most energetic subjet with respect to the transverse energy of the most energetic subjet:

$$\zeta_{\text{frac}}^{\text{subsubl}} = \frac{p_{\text{T}}(\text{subsubleading jet})}{p_{\text{T}}(\text{leading jet})}.$$

If there are less than three subjets the value is set to zero. For signal candidates the third subjet has in most cases only a low transverse momentum or does not exceed the 15 GeV threshold. Consequently the signal distribution is shifted to lower values compared to the background distribution.

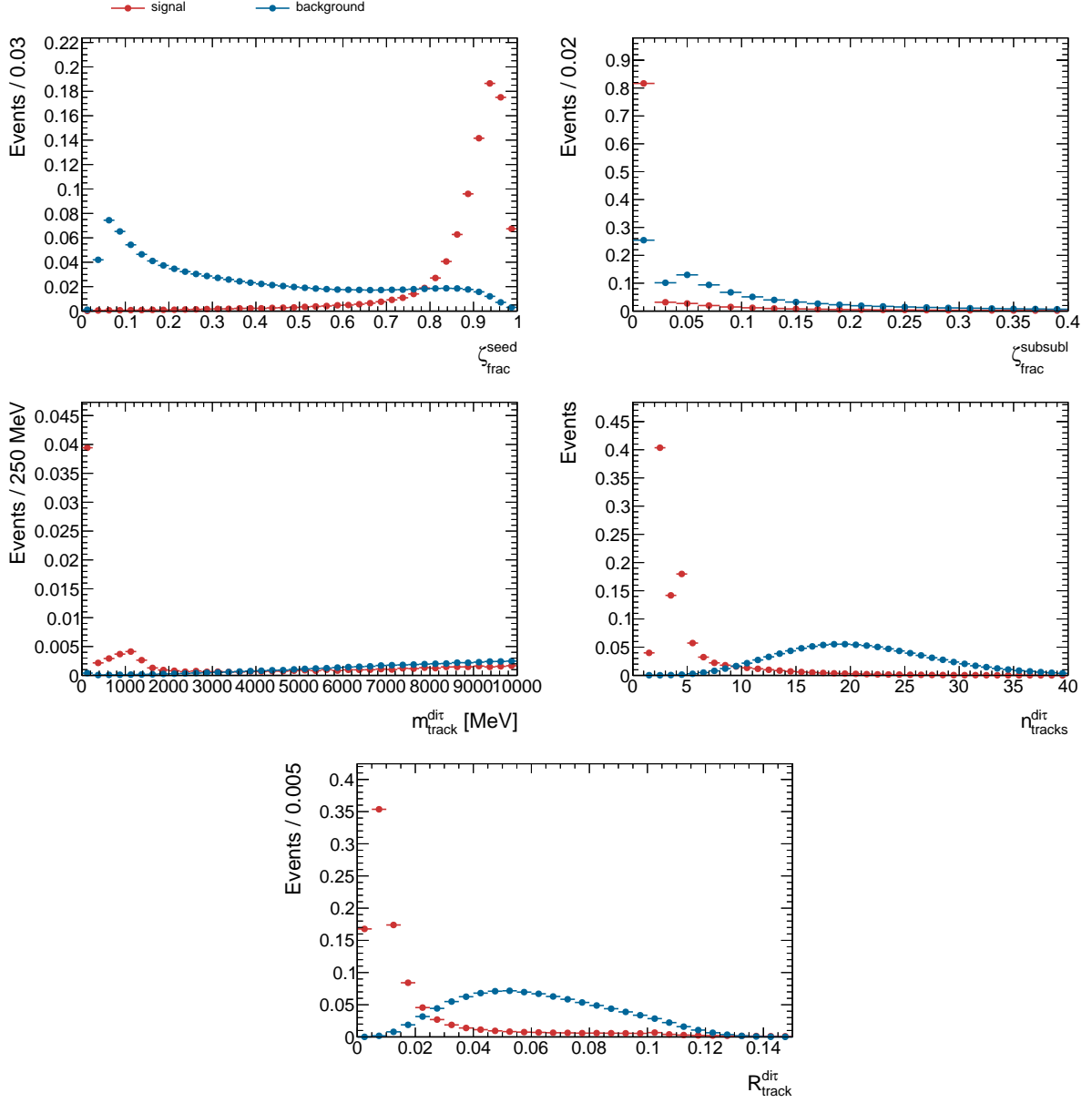


Figure 5.8: Normalized distributions of identification variables of  $\tau_{\text{had}}\tau_e$  candidates which use information in the complete anti- $k_t(1.0)$  seed jet region.

- The **invariant track mass**  $m_{\text{track}}^{\text{di}\tau}$  is the invariant mass of the sum of the four momenta of all tracks which pass the track selection and are not associated to the electron. The signal distribution has a peak at zero if only one track is associated to the candidate, a broader peak at approximately 1 GeV if multiple tracks of a tau decaying into three charged hadrons are correctly reconstructed and a large tail if there are additional fake tracks or if the electron track subtraction fails. The background distribution has only a small peak for low  $m_{\text{track}}^{\text{di}\tau}$ , but has a similar shape as the signal distribution for high  $m_{\text{track}}^{\text{di}\tau}$ .
- The **number of tracks**  $n_{\text{tracks}}^{\text{di}\tau}$  counts all tracks passing the track selection. The electron track is taken into account. The signal distribution has strong peaks at 2 and 4 corresponding to the number of charged particles from the  $\tau_{\text{had}}\tau_e$  decay. The background has on average approximately 20 tracks.
- The **average track distance**  $R_{\text{track}}^{\text{di}\tau}$  describes the transverse momentum weighted average distance of all tracks passing the selection to their nearest subjet.

$$R_{\text{track}}^{\text{di}\tau} = \frac{\sum \Delta R(\text{track}, \text{subjet}) p_{\text{T}}(\text{track})}{\sum p_{\text{T}}(\text{track})}$$

For signal the high momentum tracks collimated in the center of the tau-subjet are causing low values of  $R_{\text{track}}^{\text{di}\tau}$ . For background there are more tracks with higher angular separation leading to higher values of  $R_{\text{track}}^{\text{di}\tau}$ .

The second set of variables describes properties of the tau subjet of the  $\tau_{\text{had}}\tau_e$  candidate. The distributions are depicted in Figure 5.9 and 5.10.

- The **leading track transverse impact parameter**  $d_{0,\text{leadtrack}}^{\text{subjet}}$  is the transverse impact parameter of the track with the highest  $p_{\text{T}}$  in the  $\Delta R < 0.2$  region around the subjet axis. The signal distribution is shifted to slightly higher values than the background distribution because of the finite decay length of hadronically decaying taus.
- The **central energy fraction**  $f_{\text{core}}^{\text{subjet}}$  is calculated by dividing the sum of the transverse energy measurements of the cells in the  $\Delta R < 0.1$  region around the subjet by the sum over all cells in the  $\Delta R < 0.2$  region.

$$f_{\text{core}}^{\text{subjet}} = \frac{\sum_{\text{cells}}^{\Delta R < 0.1} E_{\text{T,cell}}}{\sum_{\text{cells}}^{\Delta R < 0.2} E_{\text{T,cell}}}$$

For signal candidates the energy is more collimated in the center of the subjet while the energy in background is more evenly spread in the subjet. This leads to a peak at high values of  $f_{\text{core}}^{\text{subjet}}$  for the signal distribution and lower values for the background.

- The **subjet energy fraction**  $f_{\text{di}\tau}^{\text{subjet}}$  describes the transverse momentum of the tau subjet with respect to the transverse momentum of the seed jet.

$$f_{\text{di}\tau}^{\text{subjet}} = \frac{p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}})}{p_{\text{T}}(\text{seed})}$$



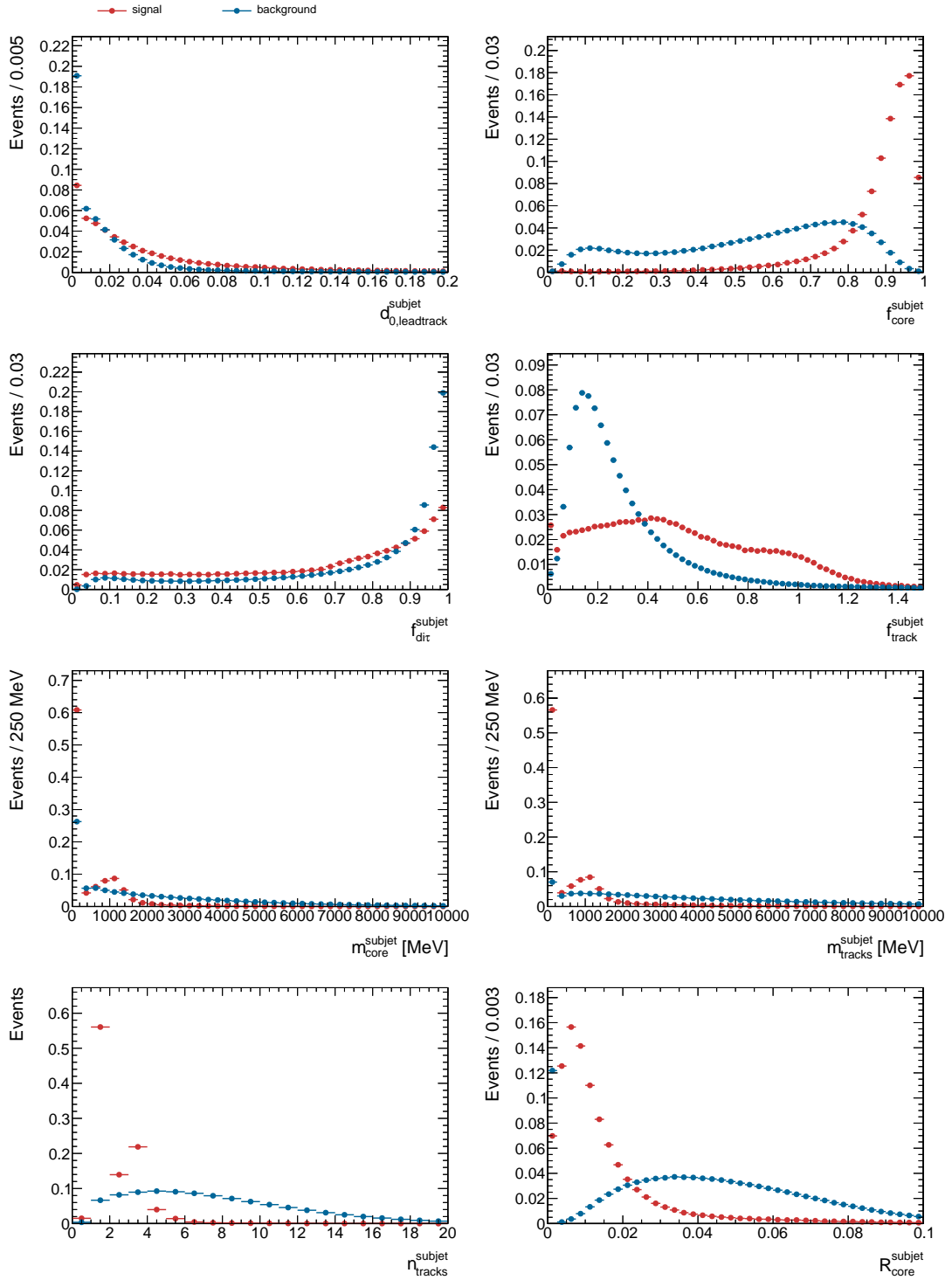


Figure 5.9: Normalized distributions of variables describing properties of the tau subject of signal and background  $\tau_{\text{had}}\tau_e$  candidates.

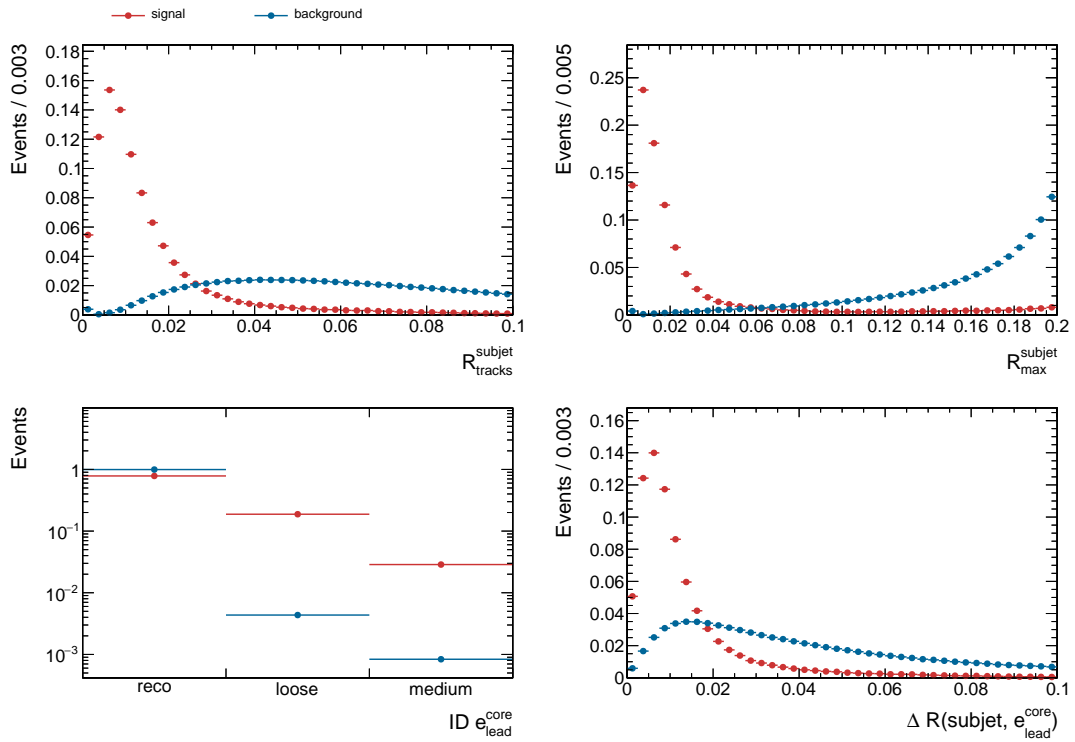


Figure 5.10: Normalized distributions of variables describing properties of the tau subject of signal and background  $\tau_{\text{had}}\tau_e$  candidates.

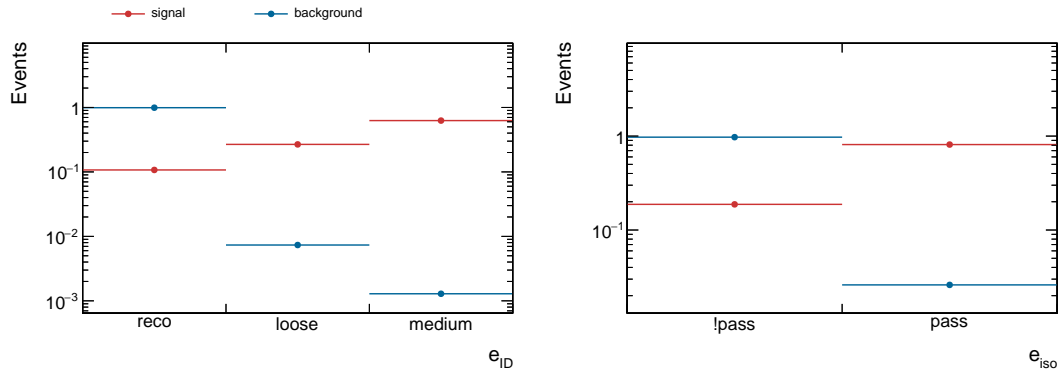


Figure 5.11: Identification variables describing the reconstructed electron.

Because of the different number of neutrinos in the leptonic and hadronic decay, the  $\tau_{\text{had-vis}}^{\text{truth}}$  carries on average more energy than the  $e^{\text{truth}}$  but any proportion between both is possible. This leads to a wide almost flat signal distribution of  $f_{\text{dir}}^{\text{subject}}$  with an increase to high values. For background candidates there is often only one hard particle involved consequently the distribution peaks stronger at high values.

- The **subject track fraction**  $f_{\text{track}}^{\text{subject}}$  is the transverse momentum of the track with the highest momentum in the region of the tau subject divided by the transverse momentum of the tau subject.

$$f_{\text{track}}^{\text{subject}} = \frac{p_{\text{T}}(\text{leading track})}{p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}})}$$

In more than 80% of all hadronic tau decays there is more than one hadron produced, but in most cases there are less than four. This leads to larger values of  $f_{\text{track}}^{\text{subject}}$  compared to background candidates where in most cases more particles are involved. The interesting shape of the signal distribution hints at the different decay modes of the hadronic tau.

- The **subject core mass**  $m_{\text{core}}^{\text{subject}}$  is determined by the invariant mass of the sum of the four momenta of the tracks in the  $\Delta R < 0.1$  region around the tau subject direction. Tracks which are associated to the  $e^{\text{reco}}$  are not taken into account. For  $m_{\text{core}}^{\text{subject}} < 3 \text{ GeV}$  the signal and background distributions are shaped and explained similarly as for  $m_{\text{track}}^{\text{dir}}$ . The larger values are prohibited because of the constrained region.
- The **subject track mass**  $m_{\text{track}}^{\text{subject}}$  is calculated similarly to  $m_{\text{core}}^{\text{subject}}$  but taking all tracks in  $\Delta R < 0.2$  into account. The signal distributions of  $m_{\text{track}}^{\text{subject}}$  and  $m_{\text{core}}^{\text{subject}}$  show almost no difference. The background distribution has slightly higher values for large  $m_{\text{track}}^{\text{subject}}$  than for  $m_{\text{core}}^{\text{subject}}$ . The reason for this behavior is that additional tracks in the  $0.1 < \Delta R < 0.2$  isolation annulus around the tau subject are uncommon for hadronic tau decays but happen more commonly for background candidates.
- The **number of subject tracks**  $n_{\text{tracks}}^{\text{subject}}$  counts tracks in  $\Delta R < 0.2$  around the tau subject axis without tracks associated to the reconstructed electron of the candidate. While the signal distribution shows a clear peak at one and three tracks the background distribution is broader with a peak at approximately 5 tracks.
- The **average core track distance**  $R_{\text{core}}^{\text{subject}}$  is calculated similarly to  $R_{\text{track}}^{\text{dir}}$  but is taking only the tracks in the  $\Delta R < 0.1$  region around the tau subject into account. The value is set to zero if there is no track in the tau subject. The one or three more collimated tracks of the hadronic tau decay cause lower values of  $R_{\text{core}}^{\text{subject}}$  compared to the multiple widespread tracks common in background candidates. The background distribution shows an interesting peak at zero caused by the case where there is no track in the core region of the tau subject.
- The **average track distance**  $R_{\text{tracks}}^{\text{subject}}$  is defined similar to  $R_{\text{core}}^{\text{subject}}$  but is taking the tracks in the  $\Delta R < 0.2$  cone around the tau subject direction into account. Because tracks in the isolation annulus are uncommon for hadronic tau decays only the background distribution is

shifted to higher values. The peak at zero for the background distribution is missing. This indicates that in approximately 12% of the cases the tracks of the background candidates are distributed around the core region  $\Delta R > 0.1$  of the subjet.

- The **maximum track distance**  $R_{\text{max}}^{\text{subjet}}$  is the maximum angular distance between a track in the  $\Delta R < 0.2$  region and the tau subjet axis. For signal this track is in most cases in the center of the core region which means the variable is smaller than  $R_{\text{max}}^{\text{subjet}} < 0.5$ . For background the outermost track is most probably in the isolation annulus  $0.1 < \Delta R < 0.2$ .
- For the **leading core electron identification ID**  $e_{\text{lead}}^{\text{core}}$  the reconstructed electron with the highest transverse momentum is searched for in the  $\Delta R < 0.1$  region around the tau subjet axis. The variable describes the standard electron identification working point this electron passes. Zero corresponds to the case where no identification is passed, one to the loose identification and two to the medium identification. The tight identification point was not used because of a software issue, but might be useful for future optimization. The hadronic tau decay is more likely to produce a fake reconstructed electron in the core region, which passes at least the loose working point, than the background. This variable can also be used to resolve the overlap between the  $\tau_{\text{had}}\tau_e$  and the  $\tau_e\tau_e$  channel.
- The **leading electron distance**  $\Delta R(\text{subjet}, e_{\text{lead}}^{\text{core}})$  is the angular distance between the electron used for the previous identification variable and the tau subjet direction. If there is no reconstructed electron the variable is set to -1. In most cases one of the charged particles in the tau subjet region is reconstructed as an electron. For signal this electron candidate is more likely to be in the center of the tau subjet.

The third set of identification variables describes the electron candidate (Figure 5.11):

- The **electron ID selection**  $e_{\text{ID}}$  describes the working point, which is passed by the  $e^{\text{reco}}$  in a similar way as ID  $e_{\text{lead}}^{\text{core}}$ . The background consist to 99% of candidates not passing any identification requirement on the electron. For signal this happens only for 10% of the candidates. A cut in this variable would therefore already be able to reduce the background significantly. But the BDT might be able to recover some of the candidates where the electron is not identified which occurs more often for low angular separation between  $\tau_{\text{had-vis}}^{\text{truth}}$  and  $e^{\text{truth}}$ . More importantly the training might benefit from the candidates where the identification is not passed. Although one of the first decision trees will use a cut in  $e_{\text{ID}}$  a later tree concentrates on features of the tau subjet to recover earlier misclassification. The decision of this tree might generalize well from the background set where the electron identification is not passed to the set where the identification passes, because the subjet and the electron are mostly independent. Consequently this is effectively increasing the background statistics in the region where the identification passes.
- The **electron isolation selection**  $e_{\text{iso}}$  describes whether the electron passes the gradient loose working point of the standard isolation selection. It is based on  $p_{\text{T}}$  dependent cuts on  $E_{\text{T}}^{\text{topcone20}}$  and  $p_{\text{T}}^{\text{varcone30}}$ . The variable  $E_{\text{T}}^{\text{topcone20}}$  is defined as the sum of the transverse energies of topological clusters in the  $0.1 < \Delta R < 0.2$  annulus around the lepton and  $p_{\text{T}}^{\text{varcone30}}$  is defined as the sum of the transverse momenta of tracks in the  $\Delta R < 10 \text{ GeV}/p_{\text{T}}$

region around the lepton excluding the tracks associated to the lepton. For signal the electron is in most of the cases isolated except when its closer than  $\Delta R < 0.2$  to the  $\tau_{\text{had-vis}}^{\text{truth}}$ . Background candidates have only in approximately 3% of all cases an isolated reconstructed electron.

The linear correlation coefficients between the different variables in the signal and in the background samples are depicted in Figure 5.12. A large fraction of the variable combinations have correlation coefficients close to zero. Some variables are highly correlated or anti-correlated, but the signal and background correlation patterns have characteristic differences. A BDT is well suited to exploit these differences to distinguish signal from background. For example for signal there is almost no correlation between  $f_{\text{di}\tau}^{\text{subject}}$  and  $n_{\text{subject}}^{\text{tracks}}$  while the correlation for background is high. This is due to the fact that for signal  $f_{\text{di}\tau}^{\text{subject}}$  is independent on the number of charged hadrons produced in the hadronic tau decay, but for background a larger number of tracks in the subjet causes on average a higher energy fraction in the jet compared to the energy in the seed jet area.

Over 20 other variables were investigated but not used for the final training. Among them were the input variables for the electron identification and e.g. variables depending on the number of tracks which do not pass the track selection. Adding these variables improves the classifier in terms of signal efficiency and background rejection, but a reduction of the number of identification variables is also desirable to increase the speed and more importantly the comprehensibility of the algorithm. Consequently the classifier presented is a compromise between both perks. The set of variables can most probably be further optimized. One of the variables  $R_{\text{tracks}}^{\text{subject}}$  and  $R_{\text{core}}^{\text{subject}}$  or  $m_{\text{tracks}}^{\text{subject}}$  and  $m_{\text{core}}^{\text{subject}}$  can be dropped if an additional variable which sufficiently describes the tracks in the isolation annulus  $0.1 < \Delta R < 0.2$  is used instead. There is also only one variable  $f_{\text{track}}^{\text{subject}}$  which has an obvious relation to the number of neutral pions involved in the hadronic tau decay. An additional variable which exploits the electromagnetic nature of the neutral pion decay might be beneficial.

## BDT Training

For the next step the signal and background samples are both split into two parts. One half is used for the training of the BDT and the other is used to evaluate the performance. The aim of this procedure is to be unbiased by possible overfitting of the classifier to features in the training sample, which are caused by statistical fluctuations and therefore do not generalize. The specifications of the BDT are summarized in Table A.4.

In Figure 5.13 the distributions of the BDT score for signal and background in the training and testing sample are depicted. The training and test distributions match very well. For values below  $-0.2$  slight signs of overtraining can be seen. In this region the distribution of the training sample is shifted to higher values compared to the distribution of the testing sample. Less than 1% of both the training and the testing signal sample fall in this area of overtraining. Consequently this has only an impact if a BDT score cut is chosen which aims for more than 99% signal efficiency. On the right hand side of Figure 5.13 the background rejection as a function of the signal efficiency is depicted. It is obtained by varying a cut on the BDT score distribution of the testing sample. The BDT allows for a background rejection of  $10^4$  at a signal efficiency of 80%.

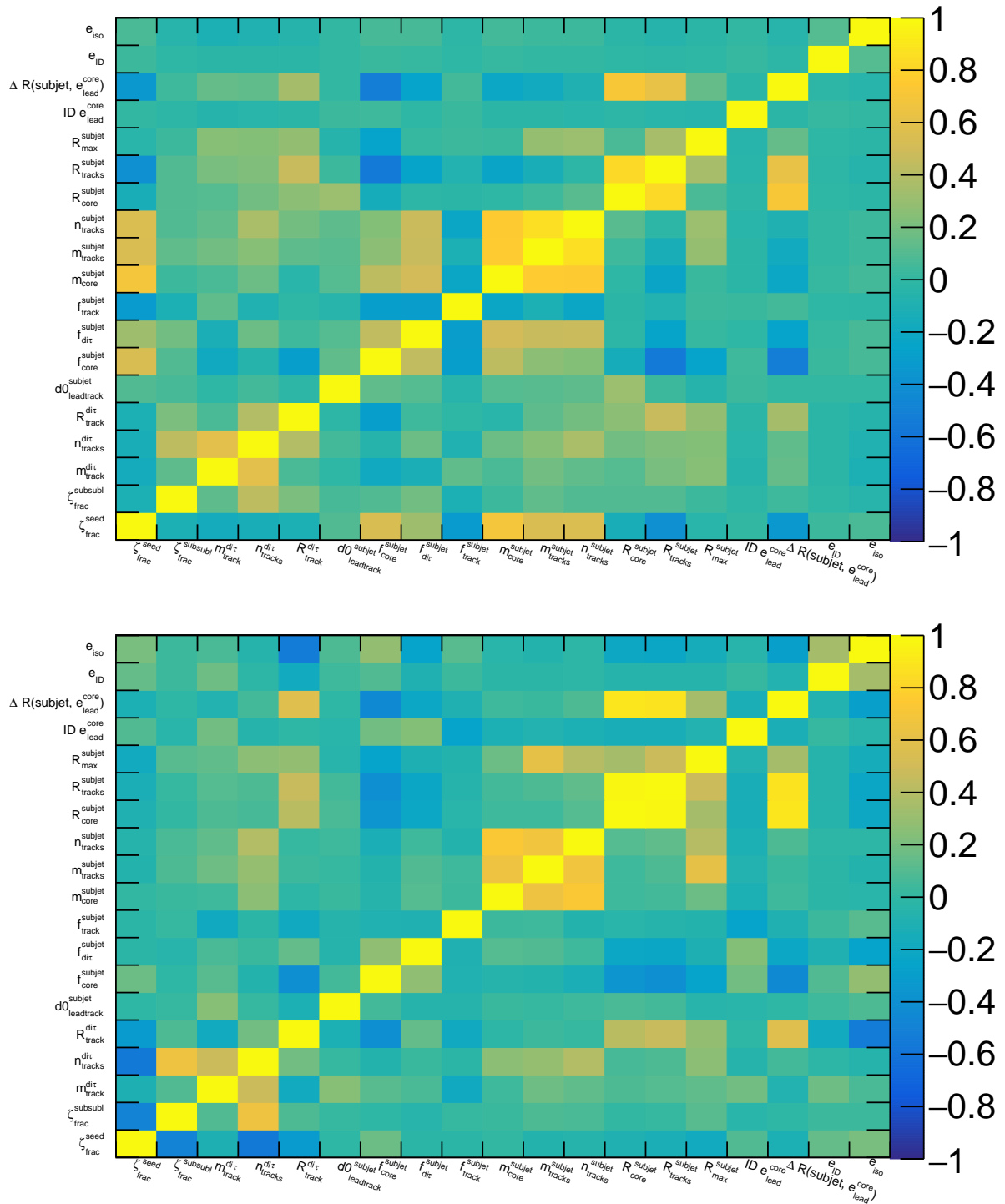


Figure 5.12: Linear correlation coefficients between the identification variables for background (top) and signal (bottom).

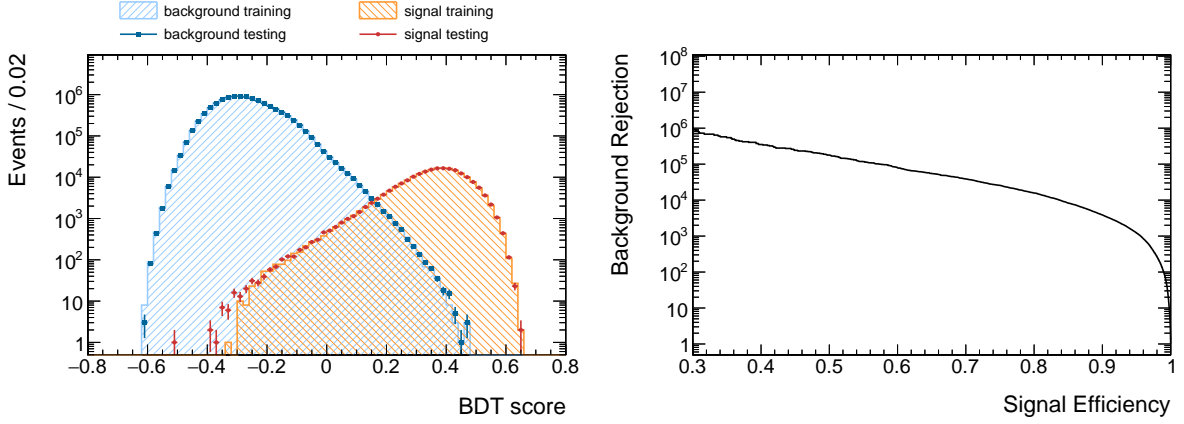


Figure 5.13: Comparison of the distribution of the BDT score in the training and testing sample for signal and background (left). Background rejection as a function of the signal efficiency in the testing sample (right).

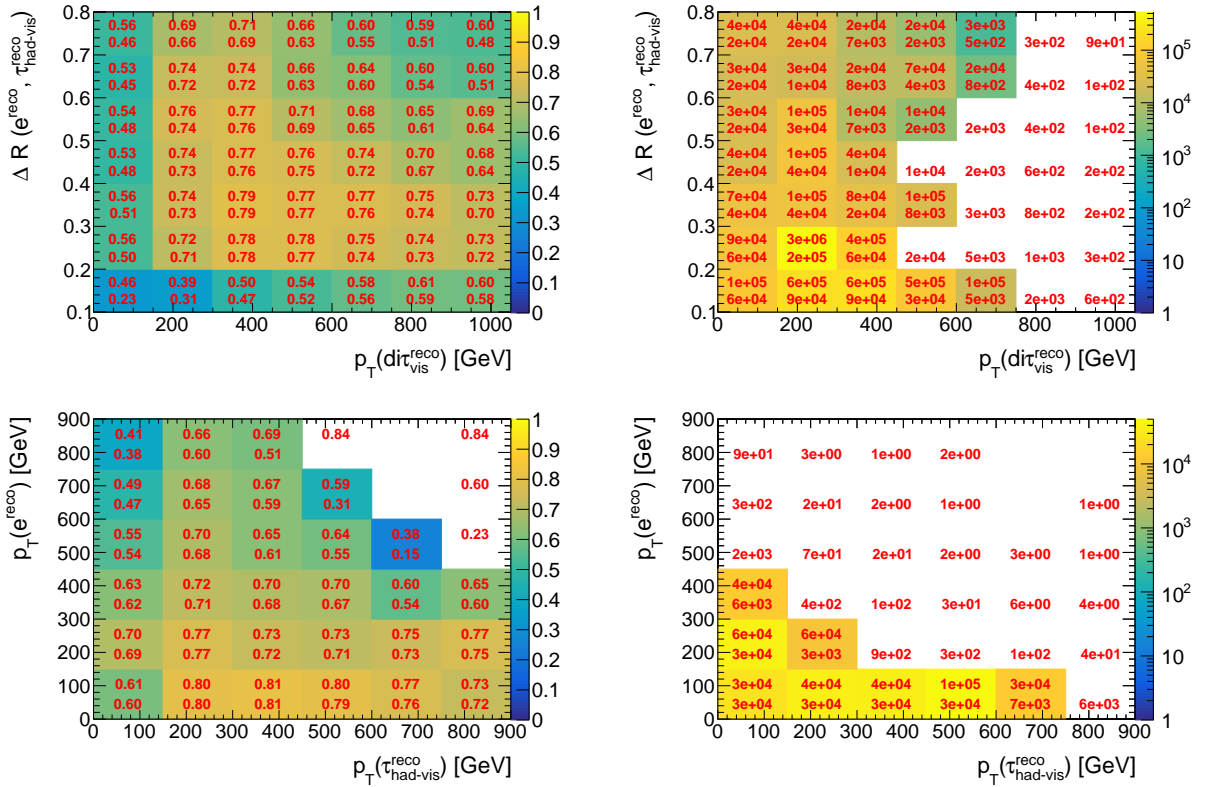


Figure 5.14: Identification efficiency (left) and background rejection (right) for BDT score  $> 0.3$  as a function of  $\Delta R(e^{\text{reco}}, \tau_{\text{had-vis}}^{\text{reco}})$  and  $p_T(\text{di}\tau_{\text{vis}}^{\text{reco}})$  (top) as well as  $p_T(e^{\text{reco}})$  and  $p_T(\tau_{\text{had-vis}}^{\text{reco}})$  (bottom). The upper and lower labels in the bins correspond to the values of the upper and lower boundaries respectively and the color indicates the expected value. In white bins no candidates pass the selection. In bins without labels there are no candidates available for selection.

In Figure 5.14 the signal efficiency and background rejection for the selection that the BDT score is greater than 0.3 are depicted as a function of different kinematic variables. The top left plot shows, that the signal efficiency for  $\Delta R < 0.2$  or  $p_{\text{T}}(\text{di}\tau_{\text{vis}}^{\text{reco}}) < 150$  GeV is in the region of approximately 50 % which is low compared to 70 to 80 % in the most efficient bins which occur for  $p_{\text{T}}(\text{di}\tau_{\text{vis}}^{\text{reco}}) > 300$  GeV and  $0.2 < \Delta R < 0.4$  where also most of the signal candidates are located. The drop in signal efficiency at  $\Delta R = 0.2$  indicates that the BDT can not completely adjust to the effect of an electron entering the tau subjet and consequently shifting the identification variables. The candidates where the  $p_{\text{T}}(\text{di}\tau_{\text{vis}}^{\text{reco}})$  is less than half of the minimum seed jet  $p_{\text{T}}$  of 300 GeV are classified as more background like, because e.g. the identification variable  $\zeta_{\text{frac}}^{\text{seed}}$  is smaller than 0.5, which is much more common for background candidates than for signal candidates. If both  $\Delta R$  and  $p_{\text{T}}$  become high, the signal efficiency is also decreasing again. Because  $\Delta R$  and  $p_{\text{T}}$  are anti-correlated in the signal sample signal candidates are not very abundant in this region.

The signal efficiency as a function of  $p_{\text{T}}(e^{\text{reco}})$  and  $p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}})$  is depicted in the lower left plot of Figure 5.14. Especially for low  $p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}})$  and high  $p_{\text{T}}(e^{\text{reco}})$  a drop in efficiency can be seen. This is also a region which is not as common as the reversed case because of the different amount of neutrinos in the leptonic and hadronic decay of the tau. This suggests that the classification of the BDT is biased towards the kinematics of the physical process which was used to produce the signal candidates. For a flexible use in different analyses in which e.g. the two taus do not share the same direct mother particle it would be desirable if this were not the case. A scheme where the  $p_{\text{T}}(e^{\text{reco}})$  versus  $p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}})$  distributions of signal and background were reweighted to be flat was investigated, but did not lead to major improvements in this regard. In the next section a method to acquire working points with a flat signal efficiency at the cost of varying background rejections is presented.

On the right hand side of Figure 5.14 the background rejections for the same selection of BDT score  $> 0.3$  are depicted. In a large region the background rejections are at least in the same order as the number of available background candidates. Consequently there are bins where no background candidate passes the selection. Because of this statistical limitation only lower bounds can be calculated for these regions. The ROOT implementation of the frequentist Clopper Pearson 68 % confidence interval is used [81, 82]. In the bins with statistical significant results the background rejection is in the order of  $10^3$  to  $10^5$  and is decreasing towards higher  $\Delta R$  and higher  $p_{\text{T}}$ .

### Working Point Tuning

To achieve working points with signal efficiencies, which are flat in certain variables, a method similar to the one used in the standard  $\tau_{\text{had}}$  identification is applied. The cut value in the BDT score is varied depending on the values of the variables in which the efficiency is desired to be flat. In this case  $p_{\text{T}}(e^{\text{reco}})$  and  $p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}})$  are used. At first the BDT Score cut as a function of the signal efficiencies is calculated in individual bins of the two variables. This is done by searching for the BDT Score cut corresponding to one to hundred percent signal efficiency in one percent steps. This model extracted from the training sample can then be used to calculate a new score for the candidates in the testing sample referred to as flat BDT score. The binned 2D representation of the BDT score cut for each of the hundred signal efficiency points is bilinearly interpolated to extract cut values for the candidate. The binning used here is summarized in



Table 5.1: Binning of the dependent variables for the working point tuning for  $\tau_{\text{had}}\tau_e$  candidates:

$p_{\text{T}}(e^{\text{reco}})$	0 GeV	50 GeV	100 GeV	150 GeV	300 GeV	350 GeV	1000 GeV
$p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}})$	0 GeV	50 GeV	150 GeV	300 GeV	400 GeV	500 GeV	1000 GeV

Table 5.1.

As an example the interpolated cut value as a function of  $p_{\text{T}}(e^{\text{reco}})$  and  $p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}})$  corresponding to 80% signal efficiency is depicted in Figure 5.15. The 2D representations where the interpolated cut value is the closest above and the closest below the BDT score of the candidate are searched for. The flat BDT score is then interpolated as following:

$$\text{flat BDT score} = 1 - \left( \epsilon_{\text{low}} + \frac{s_{\text{candidate}} - \text{cut}_{\text{low}}}{\text{cut}_{\text{high}} - \text{cut}_{\text{low}}} (\epsilon_{\text{high}} - \epsilon_{\text{low}}) \right), \quad (5.2)$$

with the interpolated cut values closest above (below)  $\text{cut}_{\text{high}}(\text{low})$ , the corresponding efficiencies  $\epsilon_{\text{high}}(\text{low})$  and the value of the BDT score of the candidate  $s_{\text{candidate}}$ . The term in brackets in Equation 5.2 corresponds approximately to the maximum signal efficiency for which the candidate can be classified as signal, when requiring the signal efficiency to be flat in the two chosen variables. Therefore a working point with a desired efficiency  $\epsilon_{\text{WP}}$  can be easily obtained by a cut on the flat BDT score  $> 1 - \epsilon_{\text{WP}}$ . The resulting distribution of the flat BDT score for signal and background is depicted in Figure 5.15. For the signal sample the distribution is flat and for the background sample it is concentrated at low values. The distributions in the training and testing samples correspond well with each other.

In Figure 5.16 the identification efficiency and background rejection for a cut flat BDT score  $> 0.2$  are depicted in the same way as for BDT score  $> 0.3$  before. The signal efficiencies are mostly flat in the corresponding variables especially compared to the results from a cut in the original BDT score in Figure 5.14. The drop in efficiency when the angular distance becomes less than  $\Delta R < 0.2$  does still occur. The increase in signal efficiency in a bin comes of course at the cost of less background rejection in the same bin. But the flat signal efficiency is still beneficial because it interferes less with the choice of the final kinematic signal selection of the physics analyses which might use the algorithm presented.

Four identification working points are defined: *VeryLoose*, *Loose*, *Medium* and *Tight* which correspond to 95%, 80%, 65% and 50% signal efficiency respectively. They are obtained by using lower boundaries on the flat BDT score of 0.05, 0.2, 0.35 and 0.5.

### Purity in the Signal Sample

The reconstruction algorithm is likely to produce more than one  $\tau_{\text{had}}\tau_e$  candidate for each seed jet that passes 300 GeV. Therefore on the one hand also tau pairs of different decay modes produce  $\tau_{\text{had}}\tau_e$  candidates and on the other hand also real  $\tau_{\text{had}}\tau_e$  decays produce multiple candidates where only one matches to the decay products and the other contain a subjet or electron, which is faked by pile-up or underlying event. The identification can be used to already reject a large fraction of these candidates. To investigate the purity among the identified  $\tau_{\text{had}}\tau_e$  candidates in

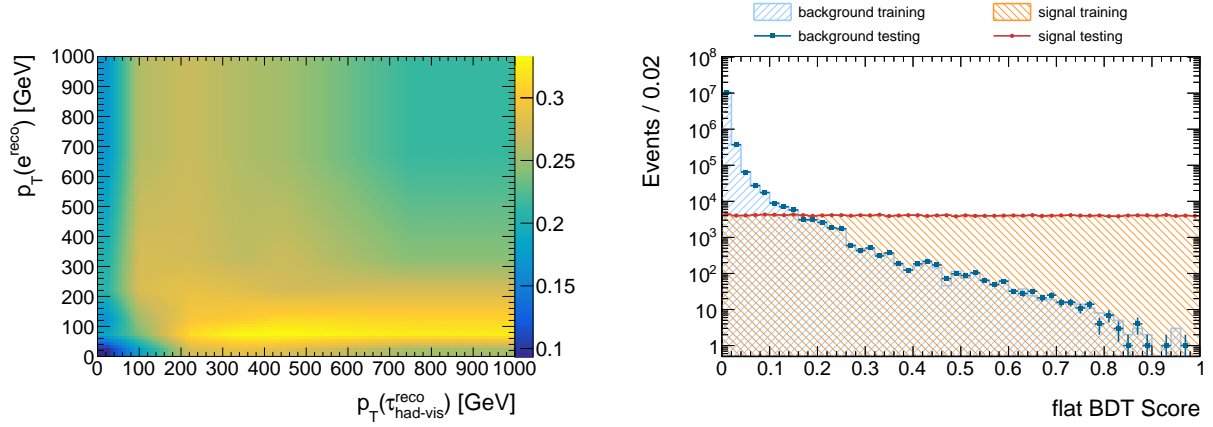


Figure 5.15: Interpolated cut value as a function of  $p_T(e^{\text{reco}})$  and  $p_T(\tau_{\text{had-vis}}^{\text{reco}})$  corresponding to 80 % signal efficiency (left). Distribution of the flat BDT score for signal and background in the training and testing sample (right).

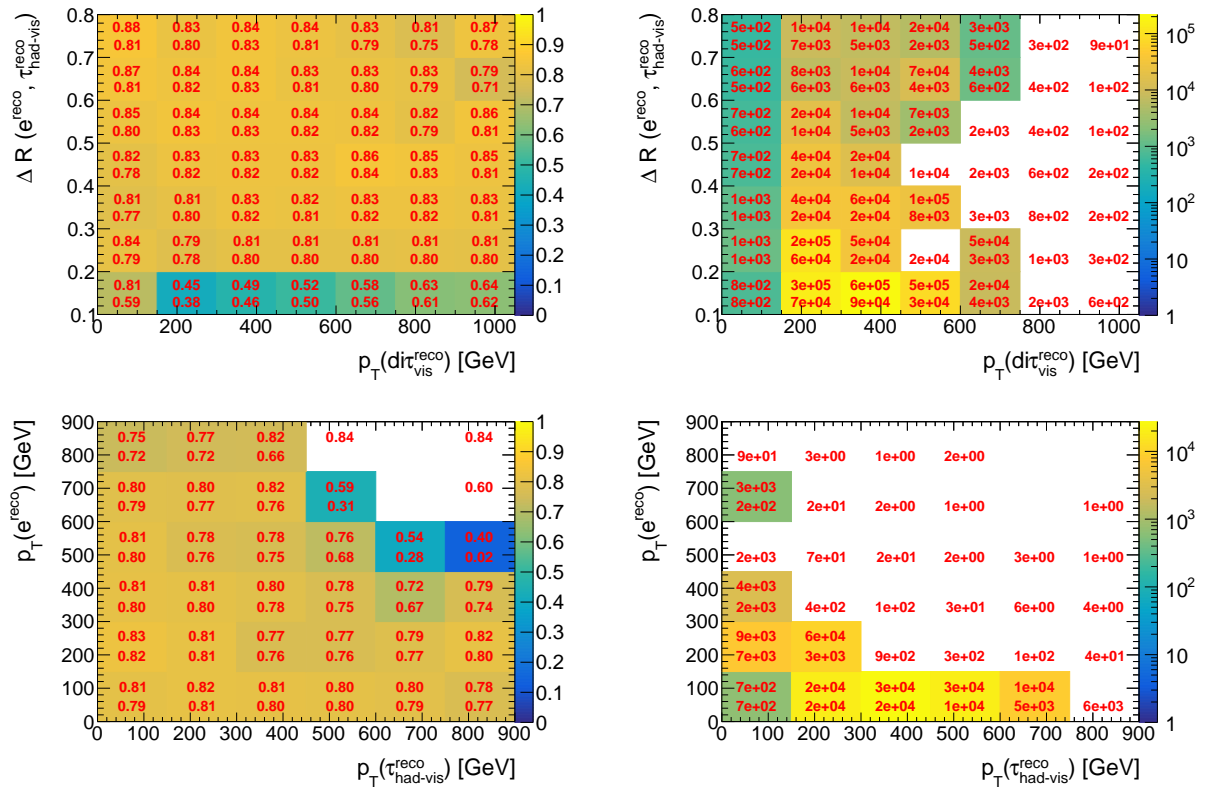


Figure 5.16: Identification efficiency (left) and background rejection (right) for flat BDT score  $> 0.2$  as a function of  $\Delta R(e^{\text{reco}}, \tau_{\text{had-vis}}^{\text{reco}})$  and  $p_T(\text{dir}\tau_{\text{vis}}^{\text{reco}})$  (top) as well as  $p_T(e^{\text{reco}})$  and  $p_T(\tau_{\text{had-vis}}^{\text{reco}})$  (bottom).

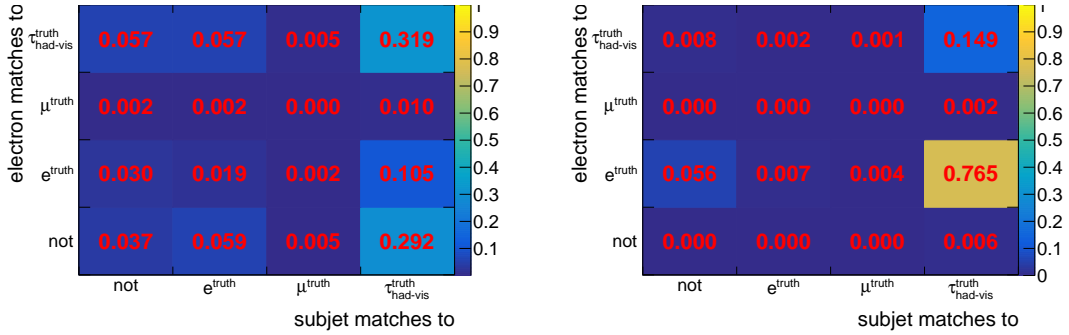


Figure 5.17: Fraction of reconstructed  $\tau_{\text{had}}\tau_e$  candidates in the signal sample where the subject and the electron match to different truth particles. The horizontal axis indicates the truth particle type the subject is matched to and the vertical axis indicates the truth particle type the reconstructed electron is matched to. In the left plot the  $\tau_{\text{had}}\tau_e$  candidates are only required to pass the very loose identification. In the plot on the right additional cuts were applied to increase the purity.

the signal sample the reconstructed electron and the subject of each candidate are matched to nearby truth particles in a vicinity of  $\Delta R < 0.1$ .

In Figure 5.17 (left) the fractions of candidates produced by different sources which pass the *VeryLoose* working point are depicted. Of only 4% of the candidates neither the reconstructed electron nor the subject matches to a  $e^{\text{truth}}$  or a  $\tau_{\text{had-vis}}^{\text{truth}}$ . The largest fraction of 32% is produced by  $\tau_{\text{had}}\tau_{\text{had}}$  decays. Taking the larger branching fraction and the on average larger visible momentum of the  $\tau_{\text{had}}\tau_{\text{had}}$  decay into account the *VeryLoose*  $\tau_{\text{had}}\tau_e$  identification seems to be approximately equally efficient for  $\tau_{\text{had}}\tau_{\text{had}}$  and  $\tau_{\text{had}}\tau_e$  decays. This has multiple reasons. The BDT was not trained to discriminate against  $\tau_{\text{had}}\tau_{\text{had}}$  decays. Furthermore the variables which depend only on the subject yield the same results for candidates from  $\tau_{\text{had}}\tau_{\text{had}}$  and  $\tau_{\text{had}}\tau_e$  decays. At last tighter working points, which oppose more stringent conditions on the electron, do reject a larger fraction of candidates from  $\tau_{\text{had}}\tau_{\text{had}}$  decays. This is also the reason for the second largest part of 29% which are candidates where the subject originates from a hadronic tau decay and no match for the electron could be found. The desired candidates originating from the  $\tau_{\text{had}}\tau_e$  decays make up only 11% of all candidates. The relevant other groups, which range between 1% and 5%, are either caused by decays where the electron is correctly matched or where the electron is faked by a hadronically decaying tau or where the tau subject is faked by an electron. This is due to the fact that the detector response of hadronic tau decays with only one charged particle and of electrons can sometimes be mistaken.

Besides the usage of tighter identification working point cuts, other variables can be used to increase the purity by removing the overlap between the decay channels. As an example the following selection was used. Especially the kinematic cuts depend on the use case.

- The transverse momentum of the subject has to be larger than  $p_T(\tau_{\text{had-vis}}^{\text{reco}}) > 80 \text{ GeV}$ .
- The transverse momentum of the electron is required to be larger than  $p_T(e^{\text{reco}}) > 10 \text{ GeV}$ .

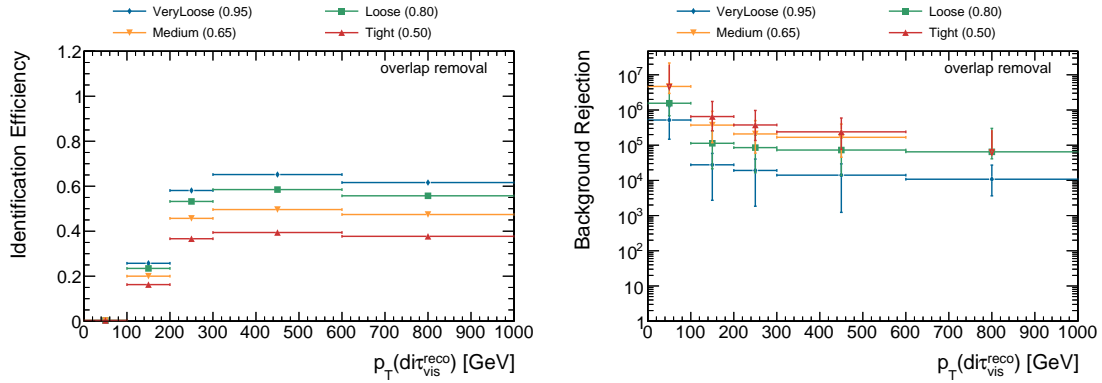


Figure 5.18: Identification efficiency (left) and background rejection (right) when the  $\tau_{\text{had}}\tau_e$  candidates are required to pass the different working points and in addition also the cuts of the overlap removal.

- The likelihood score of the standard electron identification applied on the reconstructed electron has to be larger than  $\mathcal{L}(e) > 0.5$ .
- The likelihood of the leading electron in the core region of the subjet is required to be lower than  $\mathcal{L}(e_{\text{core}}^{\text{lead}}) < 0.5$ .

The results for this selection are depicted on the right side of Figure 5.17. After applying the cuts on top of the *VeryLoose* working point the largest fraction of 77% are originating from  $\tau_{\text{had}}\tau_e$  decays. The second largest part of 15% is still originating from the fully hadronic decays and 6% come from single electrons. These results can be slightly improved by using tighter working points in addition (Figure A.1). The remaining source account for only approximately 3%. Additional cuts using the missing transverse energy, the angular distance  $\Delta R(e^{\text{reco}}, \tau_{\text{had-vis}}^{\text{reco}})$  or the visible invariant mass  $m(\text{di}\tau_{\text{vis}}^{\text{reco}})$  most probably lead to a further increase in purity.

The overlap removal also leads to a decreasing signal efficiency and an improving background rejection. The signal efficiencies and background rejections of the different working points combined with the overlap removal and as a function of the transverse momentum of the di-tau system  $p_T(\text{di}\tau_{\text{vis}}^{\text{reco}})$  are shown in Figure 5.18. The efficiency of even the loosest working point has decreased to a maximum of 60%, but the background rejection has risen to a minimum of  $10^4$  at the same time.

In summary it can be stated that the identification algorithm presented is able to distinguish candidates from  $\tau_{\text{had}}\tau_e$  decays from candidates produced by QCD multijet background. Even for low angular distance between the hadronic and leptonic decay products a good signal efficiency and background rejection can be achieved. Especially the loose working points apply no stringent selection on the reconstructed electron and focus primarily on the properties of the subjet associated to the hadronically decaying tau. Consequently the identification can not easily be used for the orthogonalization of sets of events which contain di-taus of different decay modes. The improvements with respect to the existing algorithms are mostly based on a different approach for the reconstruction of the hadronic tau decays. A method for the electron reconstruction or

identification, which differs more from the standard algorithm, would be necessary to increase the sensitivity to even lower angular distances.

## 5.2 The $\tau_{\text{had}}\tau_{\mu}$ Channel

The reconstruction and identification of boosted tau pairs where one decays into a muon and the other hadronically is less challenging than for the  $\tau_{\text{had}}\tau_e$  channel. The muon deposits only a small fraction of its energy in the calorimeter, therefore the influence of the muon on the recognition of a nearby hadronically decaying tau is mainly given by the additional track in the ID. At first the existing reconstruction and identification for muons and hadronic tau decays is evaluated when applied on  $\tau_{\text{had}}\tau_{\mu}$  decays. The identification variables for hadronic tau decays are corrected for the additional muon track and the results are used to recalculate the output of the standard identification and to train an alternative BDT using also muon variables.

### 5.2.1 Performance of Existing Algorithms

In Figure 5.19 the reconstruction and identification efficiencies of the standard algorithms are evaluated in the application on boosted  $\tau_{\text{had}}\tau_{\mu}$  decays. This was done in a similar way as in Section 5.1.1 for the  $\tau_{\text{had}}\tau_e$  channel.

The reconstruction efficiency of hadronic tau decays is, because of the missing requirement on the number of *tau tracks*, the efficiency to find an anti- $k_t(0.4)$  jet in the vicinity of the  $\tau_{\text{had-vis}}^{\text{truth}}$ , which is independent on the angular distance to a nearby muon. The direction of the anti- $k_t(0.4)$  jet also coincides well with the direction of the  $\tau_{\text{had-vis}}^{\text{truth}}$ , because of the low energy deposition of the muon in the calorimeter. For an angular distance to the nearest muon greater than  $\Delta R > 0.4$  the  $\tau_{\text{had-vis}}^{\text{truth}}$  identification efficiencies have the same values as observed in the  $\tau_{\text{had}}\tau_e$  channel before. If the muon gets nearer the identification efficiencies drop to low values of approximately 10%. This is due to the additional ID track of the muon, which is in most cases taken into account for the calculation of the identification variables.

The reconstruction and identification efficiencies for muons are very high and seem to be independent on the angular distance to a nearby hadronically decaying tau. The slight decrease for  $\Delta R < 0.2$  in the order of 10% is caused by the extreme kinematics of the 3 and 5 TeV graviton samples, which contain almost all candidates in this region. Due to the very highly boosted tau leptons in these samples, their decay vertex has more commonly a macroscopic displacement from the beam axis. The distribution of the distance of the decay vertex for the cases  $\Delta R < 0.15$  and  $\Delta R > 0.3$  can be seen on the right side of Figure 5.20. The muon identification algorithm is designed to discriminate against muons from in flight decays, consequently the efficiencies are dropping if the muon originates from a decay vertex with a large distance to the beam axis  $R$ . This can be seen on the left side of Figure 5.20. If the tau is decaying after the first layer of the pixel detector ( $R > 50.5$  mm), the muon reconstruction efficiency has already decreased by approximately 10%.

### 5.2.2 Reconstruction

Both reconstruction algorithms show a good performance when applied on boosted  $\tau_{\text{had}}\tau_{\mu}$  decays. In the following  $\tau_{\text{had}}\tau_{\mu}$  candidates are formed by combining reconstructed muons and hadronically decaying tau candidates with an angular distance of less than  $\Delta R < 1.0$ . If the  $\tau_{\text{had-vis}}^{\text{truth}}$  exceeds a momentum threshold of  $p_T(\tau_{\text{had-vis}}^{\text{truth}}) > 20$  GeV, the  $\tau_{\text{had}}\tau_{\mu}$  reconstruction efficiency is in good

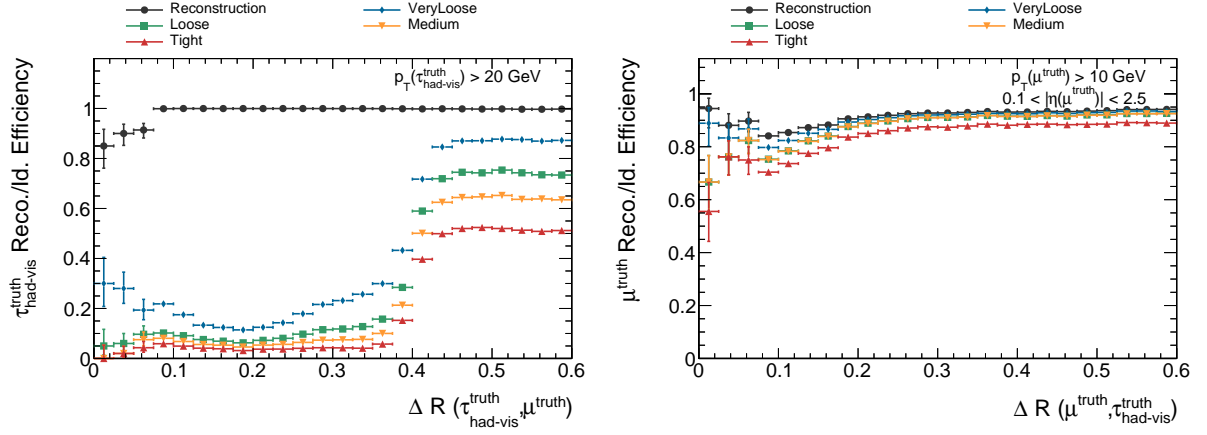


Figure 5.19: Reconstruction and identification efficiency for hadronically decaying taus as a function of the angular distance to a nearby muon  $\Delta R(\tau_{\text{had-vis}}^{\text{truth}}, \mu^{\text{truth}})$  (left). Muon reconstruction and identification efficiency as a function of the angular distance to nearby hadronically decaying tau  $\Delta R(\mu^{\text{truth}}, \tau_{\text{had-vis}}^{\text{truth}})$  (right).

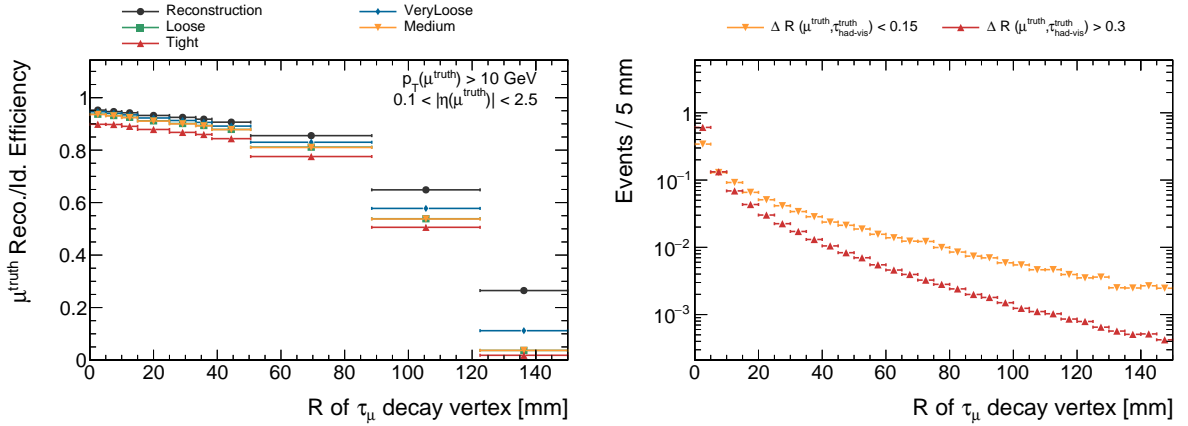


Figure 5.20: Muon reconstruction and identification efficiency as a function of the distance  $R$  of the decay vertex of the mother tau lepton to the beam axis (left). The lower boundaries of the rightmost three bins correspond to the distance of the three layers of the Pixel detector to the beam axis. Normalized distribution of  $R$  for two subsets of the muons with different requirements on the angular distance to the nearest hadronically decaying tau (right).

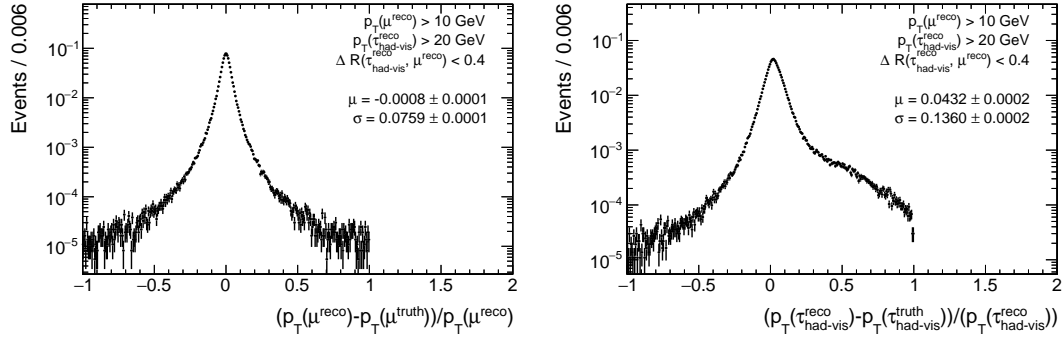


Figure 5.21: Distribution of the relative deviation of reconstructed values from the corresponding true values obtained from Monte Carlo samples. The mean value and standard deviation of the distributions are indicated on the plots.

approximation given by the muon reconstruction efficiency of 85 to 95 % depending on the position of the decay vertex of the leptonically decaying tau. The only major obstacle is the failing identification of the hadronically decaying tau caused by the additional track of the muon, which is associated to the  $\tau_{\text{had}}$  candidate. The effect of removing this track and recalculating identification variables is discussed in the next section.

The muon track has also an influence on the energy scale and vertex correction of the  $\tau_{\text{had}}$  candidate. The low mean value of the distributions of the relative deviations of the reconstructed transverse momenta from the true transverse momenta depicted in Figure 5.21 show that this influence is on average smaller than 5 %.

### 5.2.3 Muon Track Removal

In this step the tracks associated to muons which pass the requirements of the loose working point are removed from the  $\tau_{\text{had}}$  candidate. After this all standard identification variables described in Section 4.5 except for  $f_{\text{cent}}$ ,  $|S_{\text{leadtrack}}|$  and  $S_{\text{T}}^{\text{flight}}$  are recalculated for the candidate. The results for truth matched  $\tau_{\text{had}}$  candidates in the  $G \rightarrow hh \rightarrow \tau\tau\tau\tau$  samples are presented in Figure 5.22. The distributions of the remaining variables can be found in the Figure A.2 and Figure A.3 in the appendix. The distributions are shown before and after the muon track removal in the regions where a muon is nearer than  $\Delta R < 0.4$  or farther away than  $\Delta R > 0.4$  to the  $\tau_{\text{had-vis}}^{\text{truth}}$ .

Most of the identification variables use *tau tracks* as which the muon track is mostly classified. Therefore before the muon track removal these variables have a different distribution if the muon is inside or outside of  $\Delta R < 0.4$ . For example the distributions of  $R_{\text{track}}^{\text{tau}}$  and  $m_{\text{EM}+\text{track}}$  are shifted to higher values with the additional muon track. Removing the muon track restores the distributions to the same shape as in the region  $\Delta R > 0.4$ . The distribution of the number of classified *tau tracks*  $n_{\text{track}}$  also recovers the peaks at one and three after the muon track removal.

After propagating the recalculated identification variables through the standard identification of hadronically decaying taus, the flat BDT score recovers its uniform shape. The deviation from a flat distribution at low flat BDT score values is caused by the cases where the number



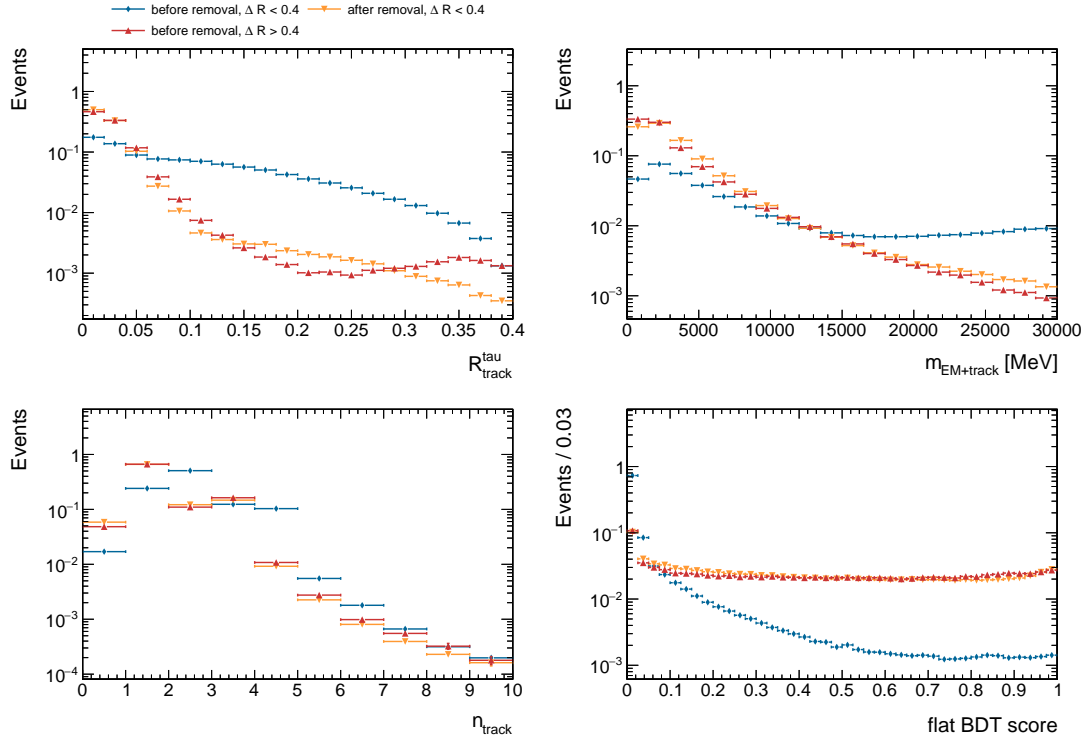


Figure 5.22: Normalized distributions of the identification variables  $R_{\text{track}}^{\text{tau}}$  and  $m_{\text{EM}+\text{track}}$ , the number of tracks  $n_{\text{track}}$  and the flat BDT score of truth matched hadronically decaying taus. The color indicates different angular distances  $\Delta R$  of the  $\tau_{\text{had-vis}}^{\text{truth}}$  to a nearby  $\mu^{\text{truth}}$  and, if the variables are calculated before or after removing tracks associated to muons.

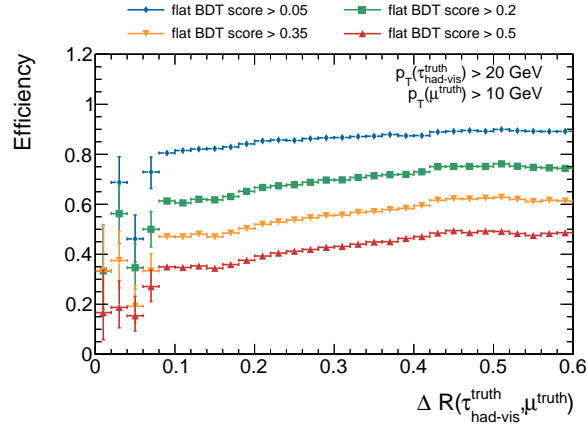


Figure 5.23: Identification efficiency for truth matched reconstructed  $\tau_{\text{had}}\tau_{\mu}$  candidates, when requiring different minimal values of the flat BDT score of the standard identification for hadronically decaying taus with recalculated identification variables.

of *tau tracks* is not one or three, because the standard  $\tau_{\text{had}}$  identification is only trained on candidates where  $n_{\text{tracks}}$  has exactly these values. The cases where the number of tracks is not one or three are caused by misidentification of fake tracks, a failing muon track removal or *tau track* reconstruction inefficiencies.

Candidates for  $\tau_{\text{had}}\tau_{\mu}$  decays where the reconstructed muon and  $\tau_{\text{had-vis}}^{\text{reco}}$  match to the corresponding truth particles are used to evaluate the identification performance. In Figure 5.23 the identification efficiencies for four cuts in the recalculated flat BDT score are depicted as a function of the angular distance  $\Delta R$  between the  $\tau_{\text{had-vis}}^{\text{truth}}$  and the  $\mu^{\text{truth}}$ . The identification efficiency is here defined as the number identified candidates divided by the number of reconstructed candidates. The efficiencies are continuously decreasing to lower values of  $\Delta R$ . The case where the muon does not pass the loose working point and the track is consequently not removed causes only a small contribution to this effect. The effects of track merging for high momentum hadronic tau decays into three charged hadrons is also not dominant. The influence of the variables  $|S_{\text{leadtrack}}|$  and  $S_{\text{T}}^{\text{flight}}$  which were not recalculated or the calorimeter energy depositions of the muon are left as possible reasons for this decrease. Unfortunately  $|S_{\text{leadtrack}}|$  and  $S_{\text{T}}^{\text{flight}}$  can not both easily be recalculated at the processing step of this analysis. To check the influence of these variables the training of the standard tau identification was reproduced leaving out  $|S_{\text{leadtrack}}|$  and  $S_{\text{T}}^{\text{flight}}$  as identification variables. Using this retrained tau identification instead leads to a signal efficiency with a smaller decrease which appears only for  $\Delta R < 0.2$ . This indicates that the influence of  $|S_{\text{leadtrack}}|$  and  $S_{\text{T}}^{\text{flight}}$  is the dominant reason for the decrease in efficiency to low  $\Delta R$ . Compared to this only a negligible influence is left for the muon energy deposition. This is reasonable because the energy loss of the muon is on average in the range of a few GeV and therefore much smaller than the energy of the tau jet.

In the next section a classifier for  $\tau_{\text{had}}\tau_{\mu}$  candidates is presented, which is trained in a similar way as for the  $\tau_{\text{had}}\tau_e$  case to use the combined information of the  $\tau_{\text{had}}$  and muon candidate in one step and the results are compared to those of the standard algorithm with recalculated variables.

### 5.2.4 Identification

The training of the classifier follows the same procedure as for the  $\tau_{\text{had}}\tau_e$  case. At first a training and testing sample is acquired from the  $G \rightarrow hh \rightarrow \tau\tau\tau\tau$  Monte Carlo samples and a part of the data taken with the ATLAS detector in 2015. The  $\tau_{\text{had}}\tau_{\mu}$  reconstruction is applied on the events by merging reconstructed hadronic tau decays and reconstructed muons with an angular distance of less than  $\Delta R < 1.0$  into  $\tau_{\text{had}}\tau_{\mu}$  candidates.

The background sample is reweighted to reassemble the  $p_T(\tau_{\text{had-vis}}^{\text{reco}})$  versus  $p_T(\mu^{\text{reco}})$  distribution of the signal sample. Two sets of variables are used for the training of the BDT. The recalculated variables of the standard  $\tau_{\text{had}}$  identification are included to describe the tau jet. The signal and background distribution as well as the correlations for these variables can be found in Figure A.5 and Figure A.4 in the appendix. In addition to the tau identification variables three variables describing the muon are used which are described in the following. Their signal and background distributions are depicted in Figure 5.24.

- The **quality of the muon**  $\mu_{\text{ID}}$  describes the working point of the standard muon selection which is passed by the reconstructed muon. Only a few percent of the background candidates pass more than the *Loose* working point, while most of the signal passes the *Tight* muon selection. Nonetheless no cut on this variable is done before hand, to profit from additional statistics for the  $\tau_{\text{had}}$  candidate as described before for the  $e_{\text{ID}}$  variable.
- The **isolation of the muon**  $\mu_{\text{iso}}$  is the result of the standard isolation selection with the gradient loose working. The selection is similar as for  $e_{\text{iso}}$  based on momentum dependent cuts on variables describing either the energy deposition in the calorimeter around the muon or the momentum of inner detector tracks around the muon. Only 3% of the background candidates have an isolated muon, while for signal only 6% are not isolated. Especially for low angular distances between the muon and the hadronic decay products  $\Delta R(\tau_{\text{had-vis}}^{\text{reco}}, \mu^{\text{reco}})$  the muon are often not isolated.
- The **energy loss fraction**  $f_{\text{E-loss}}^{\mu}$  is connected to the energy loss of the muons. The standard muon reconstruction provides a direction and momentum based parameterization motivated by the detector geometry, which describes the average energy loss  $E_{\text{avg}}$  of the muon. The measured energy loss  $E_{\text{meas}}$  is obtained from the clustered cells on the interpolated muon track through the calorimeter. These values are usually used to increase the resolution of the muon momentum measurement [73]. Here they are used to calculate the following variable:

$$f_{\text{E-loss}}^{\mu} = \frac{E_{\text{meas}} - E_{\text{avg}}}{E_{\text{meas}} + E_{\text{avg}}}. \quad (5.3)$$

In most cases high values of this variable indicate that the muon is not isolated. But they can also be caused by candidates with muon detector energy depositions higher than the expected average. For signal the distribution of  $f_{\text{E-loss}}^{\mu}$  is concentrated at around zero, where the measured values coincide with the average values. For background the distribution has peaks around zero and 0.8.

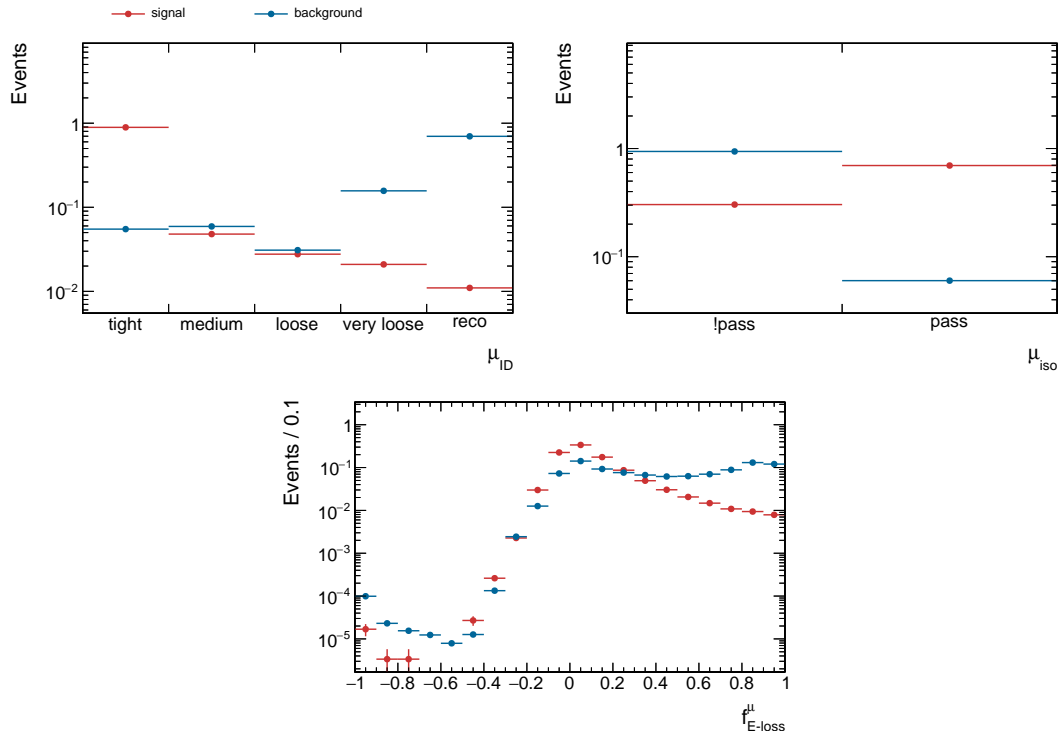


Figure 5.24: Normalized distributions of identification variables describing properties of the reconstructed muon of the  $\tau_{\text{had}}\tau_{\mu}$  candidate.

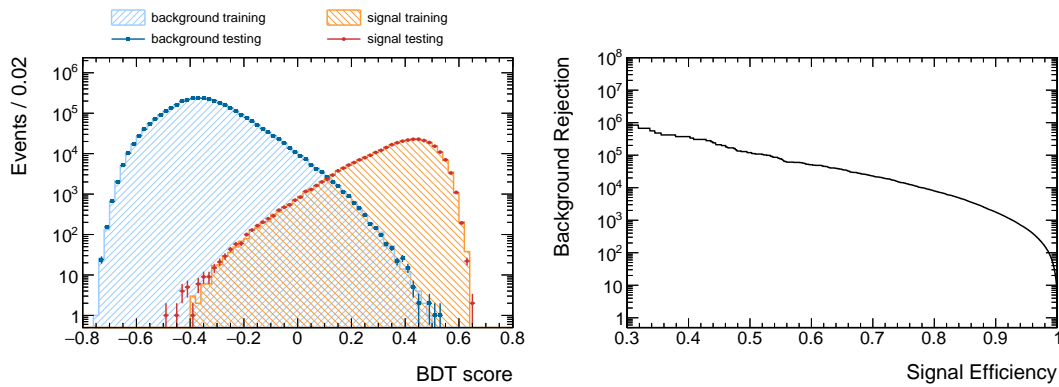


Figure 5.25: Comparison of the distribution of the BDT score in the training and testing sample for signal and background (left). Background rejection as a function of the signal efficiency in the testing sample (right).

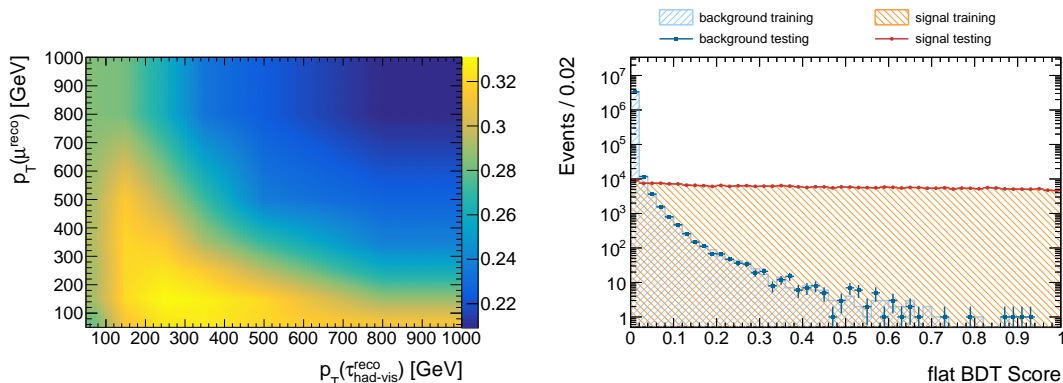


Figure 5.26: Lower boundary on the BDT score as a function of  $p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}})$  and  $p_{\text{T}}(\mu^{\text{reco}})$  corresponding to a flat BDT score greater than 0.2, which defines the *Loose* working point (left). Distribution of the flat BDT score for signal and background in the training and testing sample (right).

The resulting BDT score distribution for the training and testing sample of signal and background as well as the background rejection as a function of the signal efficiency is depicted in Figure 5.25. The distributions of the training and testing sample match very well, therefore no signs of relevant overtraining are apparent. At 80 % signal efficiency a background rejection of approximately  $10^4$  can be achieved.

The results of the BDT training are again used to tune working points, which are flat in two variables. The procedure is identical to the one for the  $\tau_{\text{had}}\tau_e$  channel described in Section 5.1.3. The transverse momentum of the muon  $p_{\text{T}}(\mu^{\text{reco}})$  and of the hadronic decay products  $p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}})$  are used. The binning of the 2D representations of the cut value can be found in Table 5.2. The resulting distribution of the flat BDT score and an exemplary representation of the BDT score cut as a function of  $p_{\text{T}}(\mu^{\text{reco}})$  and  $p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}})$  for 80 % signal efficiency are depicted in Figure 5.26. The signal distribution of the flat BDT score is uniform while the background distribution is concentrated at low values. For a flat signal efficiency the BDT score cut has to be lowered in the region where  $p_{\text{T}}(\mu^{\text{reco}})$  and  $p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}})$  are high as well as for low  $p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}})$ . Because signal candidates are less abundant in these regions, this indicates a kinematic bias of the BDT.

Four working points are defined corresponding to different cuts in the flat BDT score. They are named *VeryLoose*, *Loose*, *Medium* and *Tight* and produce approximately 95 %, 80 %, 65 % and 50 % signal efficiency respectively. The resulting signal efficiencies and background rejections for the *Loose* working point in different 2D bins of  $p_{\text{T}}(\mu^{\text{reco}})$  versus  $p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}})$  can be seen in Figure 5.27. The signal efficiencies are in most bins in the  $\pm 5\%$  region around the 80 % the tuning aimed for. Only for  $p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}}) < 100$  GeV a strong deviation to lower values can be seen. This is due to the binning of the tuning, which uses a lower boundary of 50 GeV. That means the BDT score cut is not adjusted for candidates with lower  $p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}})$ . Instead the same cut as for a candidate with  $p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}}) = 50$  GeV is applied on these candidates. It might be beneficial to adjust the binning in the future differently. The background rejection for the *Loose*

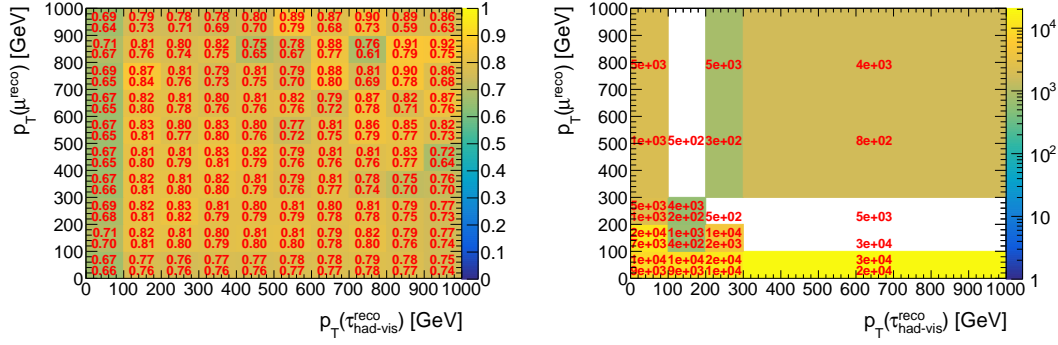


Figure 5.27: Signal efficiency (left) and background rejection (right) when requiring  $\tau_{\text{had}}\tau_{\mu}$  candidate to pass the *Loose* working point. The upper and lower numbers in each bin indicate the upper and lower boundaries calculated with the Clopper Pearson method [82] using a 68% confidence level.

Table 5.2: Binning of the two variables used for the working point tuning for  $\tau_{\text{had}}\tau_{\mu}$  candidates:

$p_{\text{T}}(\mu^{\text{reco}})$	50 GeV	100 GeV	300 GeV	400 GeV	600 GeV	1000 GeV
$p_{\text{T}}(\tau_{\text{had-vis}}^{\text{reco}})$	50 GeV	100 GeV	300 GeV	400 GeV	600 GeV	1000 GeV

working point varies between  $10^3$  and  $10^4$ . In contrast to this the *Tight* working point leads to background rejections of up to  $10^5$ .

To check the performance of the newly trained identification in comparison to results, which can be achieved by using the muon and the standard  $\tau_{\text{had}}$  identification with recalculated variables, both algorithms are applied on the candidates of the same signal and background samples. Two working point selections based on the standard algorithms with similar identification efficiencies as the newly trained *Loose* and *Tight* working point are defined. The *standard VeryLoose* selection requires the muon to pass the *Loose* selection and the  $\tau_{\text{had}}$  candidate to fulfill recalculated flat BDT score  $> 0.05$  and the *standard Medium* selection requires the muon to pass the *Medium* selection and the  $\tau_{\text{had}}$  candidate to fulfill recalculated flat BDT score  $> 0.35$ . The signal efficiencies and background rejections for these selections are depicted in Figure 5.28.

The signal efficiencies of both the newly trained and standard algorithms show a decrease to low values of  $\Delta R(\tau_{\text{had-vis}}^{\text{reco}}, \mu^{\text{reco}})$ . The new algorithm has in addition also a decreasing signal efficiency to higher values and an optimal region where  $0.2 < \Delta R(\tau_{\text{had-vis}}^{\text{reco}}, \mu^{\text{reco}}) < 0.4$ . Most of the candidates in the signal sample are in this region. For these candidates the muon is in most cases isolated and  $S_{\text{T}}^{\text{flight}}$  and  $|S_{\text{leadtrack}}|$  of the  $\tau_{\text{had}}$  candidate are altered by the muon track. Therefore the BDT is able to indirectly cut on  $\Delta R(\tau_{\text{had-vis}}^{\text{reco}}, \mu^{\text{reco}})$  which can cause this bias. Nonetheless the deviations in signal efficiency from the standard algorithms are small and the background rejections in comparable bins by a factor of approximately 10 higher.

This has to be put into perspective when considering also the  $p_{\text{T}}(\text{di}\tau_{\text{vis}}^{\text{reco}})$  dependency of the signal efficiency and background rejection on the bottom of Figure 5.28. While the standard

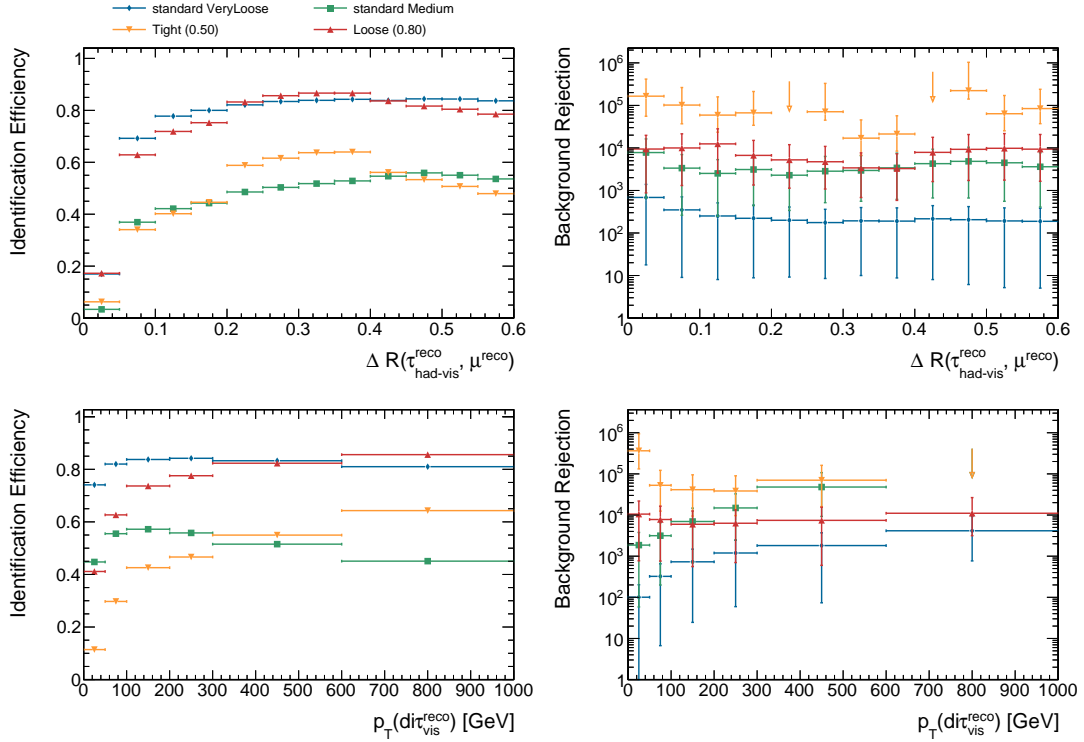


Figure 5.28: Comparison of signal efficiency (left) and background rejection (right) for different selections based on the newly trained BDT and on the existing slightly readjusted algorithms. The values are depicted as functions of the angular distance between reconstructed muon and  $\tau_{\text{had}}$  candidate  $\Delta R(\tau_{\text{had-vis}}^{\text{reco}}, \mu^{\text{reco}})$  (top) as well as of the measured visible transverse momentum of the di-tau system  $p_T(\text{di}\tau_{\text{vis}}^{\text{reco}})$  (bottom).

working points are almost fully efficient down to  $p_T(\text{di}\tau_{\text{vis}}^{\text{reco}}) = 50 \text{ GeV}$ , the newly trained algorithms show in the whole range a decreasing efficiency down to low values of the transverse momentum. This is on the one hand due to an initial bias of the BDT due to the restricted kinematics of the signal sample and on the other hand due to the incomplete working point tuning which was only done down to 50 GeV. The region for which the working points are not tuned is convoluted into the whole  $p_T(\text{di}\tau_{\text{vis}}^{\text{reco}})$  range. This indirect cut on the kinematic properties of the  $\tau_{\text{had}}\tau_{\mu}$  candidate accounts for most of the increase in background rejection which is visible in the  $\Delta R(\tau_{\text{had-vis}}^{\text{reco}}, \mu^{\text{reco}})$  plot. But even for high  $p_T(\text{di}\tau_{\text{vis}}^{\text{reco}})$  the new algorithms still lead to an improved background rejection, which can be seen best in the bin between 300 GeV and 600 GeV.

Both options, the new and the standard algorithm, have advantages. While the new algorithm is beneficial for analyses, which have a signal with similar kinematics as used here for training, the adjusted standard algorithm is more flexible and might therefore be a valid option for analyses which are looking for di-taus with a low transverse momentum. Therefore it is desirable to provide an implementation for both.

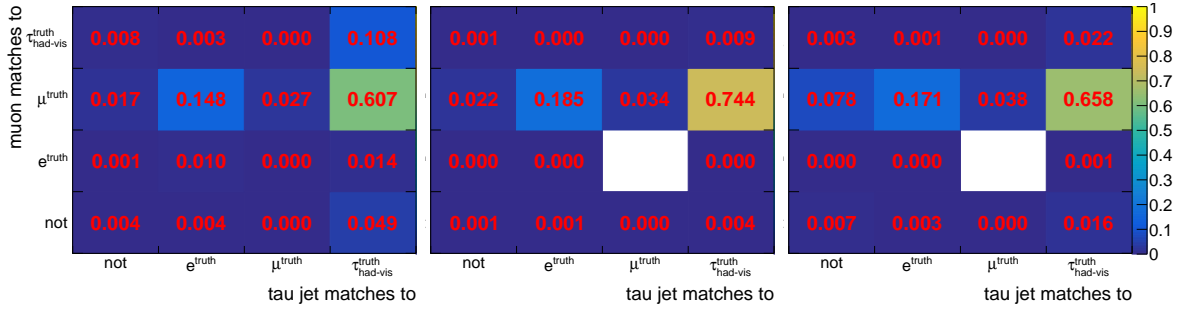


Figure 5.29: Each bin of the plots indicates the relative frequency of different categories of  $\tau_{\text{had}}\tau_{\mu}$  candidates in the  $G \rightarrow hh \rightarrow \tau\tau\tau\tau$  samples. The categories are defined by the truth particles, which produce the reconstructed muon and the tau jet of the  $\tau_{\text{had}}\tau_{\mu}$  candidate. In the left plot candidates which pass the *Loose* requirement are considered. The candidates in the middle plot are in addition required to have a muon which passes at least the *Medium* muon selection. The candidates on the right side are selected with the *standard VeryLoose* selection.

At last the proportion of  $\tau_{\text{had}}\tau_{\mu}$  candidates produced by di-taus of other decay channels in the signal sample is investigated. The procedure is similar to the  $\tau_{\text{had}}\tau_e$  channel. The truth particles which match to the reconstructed muon and  $\tau_{\text{had}}$  candidate are searched and the  $\tau_{\text{had}}\tau_{\mu}$  candidate is classified respectively. In Figure 5.29 the abundances of candidates from different sources for three distinct selections are depicted. For all selections the desired category where muon and tau jet match correctly accounts for more than 60% of the candidates. The second most common case is when the muon is matched correctly and when the hadronically decaying tau is faked by an electron. This category can only be slightly reduced by more stringent working point requirements (see Figure A.6). There is an existing technique using a BDT to reduce the overlap between electrons and hadronically decaying taus, which should be investigated for further improvements in the future. The case where the muon matches to a hadronically decaying tau occurs for the *Loose* selection in 11% of the cases. This can be almost completely prohibit when requiring the muon to pass at least the *Medium* muon selection, which also comes at the cost of only a few percent signal efficiency. Consequently this selection should be integrated into the working point.

Overall the  $\tau_{\text{had}}\tau_{\mu}$  channel shows very promising results in terms of signal efficiency, background rejection and energy resolution. The standard algorithms are well refined and tested over a long time. After the muon track removal the performance of the standard hadronic tau identification could be restored. With a new training taking also muon variables into account those results were even improved.



## 6 Summary and Outlook

Pairs of tau leptons with a low angular distance are an important signature in the search for new physics, which involves events where a high momentum particle e.g. a Higgs boson decays into two taus. The reconstruction and identification of the two semi-leptonic decay modes of such boosted di-taus were investigated in this thesis. The application of the existing standard algorithms for hadronic decaying taus, muons and electrons were evaluated. It was determined that the standard identification of hadronic tau decays with a nearby electron or muon ceases to work for low angular distances between the leptonic and hadronic decay products. Alternative approaches for both decay modes were developed, which are based on the existing algorithms including the ones for boosted di-taus in the fully hadronic decay mode. The new methods achieve a large improvement in signal efficiency in the region of very low angular distances down to  $\Delta R = 0.1$ , while at the same time allowing for a high background rejection.

The main challenge for the  $\tau_{\text{had}}\tau_e$  channel is given by the merging of the calorimeter energy depositions of the electron and the hadronic decay products into one anti- $k_t(0.4)$ -jet and therefore shifting the reconstructed direction of the momentum of the hadronic decay products. This problem was overcome by grooming small anti- $k_t(0.2)$  subjects inside of large  $R$  seed jets which contains both, electron and hadrons. Monte Carlo simulations show that with this procedure the momenta of the visible decay products of the hadronic decaying tau and of the leptonic decaying tau can be successfully reconstructed individually down to a minimal angular distance of  $\Delta R = 0.2$ . In addition to the new reconstruction an identification algorithm was developed applying Boosted Decision Trees, which ensures a high background separation against candidates produced by QCD multijet background.

The reconstruction of hadronic tau decays was shown to be almost unaffected by a nearby muon. The identification on the other hand suffers from the inner detector track of the muon if it falls in the  $\Delta R < 0.4$  region around the direction of the hadronic decay products. Two different solutions were developed and found to have comparable performance in terms of signal efficiency and background rejection. For both of them the identification variables of the standard algorithm for hadronic tau decays were recalculated leaving out tracks associated to the muon. One is based on propagating these recalculated variables through the existing identification algorithm and the other uses a newly trained BDT.

The tools, which execute the new algorithms to reconstruct and identify di-tau candidates, were implemented and are already available in the latest release of the official ATLAS software framework. With this step a major prerequisite for the application in multiple searches for new physics with boosted semi-leptonic di-taus is established. The algorithm is mainly optimized for the use in the search for a multi TeV resonance decaying into two Standard Model Higgs bosons, where at least one decays into a pair of tau leptons. A second use case is an analysis looking for decays of the Standard Model Higgs boson into two CP odd Higgs bosons with masses of less than 10 GeV and a branching ratio dominated by di-tau decays. In addition the algorithm might also be useful to extend the sensitivity of searches for R-parity violating super symmetric models to lower masses of the lightest supersymmetric particle for the case that it couples strongly to taus. For tau pairs with low spatial separation and also low transverse momentum especially the  $\tau_{\text{had}}\tau_\mu$  channel with the recalculated standard identification is promising as it shows the least kinematic bias.

The kinematic reweighting of the training sample and the momentum dependent working point

tuning ensures a classification with a flat signal efficiency in a large kinematic region. Including tau pairs from different signal samples with e.g. low momentum and low spatial separation in the training can lead to further improvements in this regard. The working point tuning, which leads to a flat signal efficiency, comes at the cost of larger fluctuations of the background rejection. An alternative approach, where the independence of the classifier is incorporated in the training, using adversarial neural nets [83] was presented recently [84] and is already showing promising results in the context of jet substructure tagging [85].

Most of the effort spend in this work was concentrated on the recovery of the reconstruction and identification of the hadronic tau decay. It was determined that the electron reconstruction and identification also suffers from nearby hadronic decay products of a tau decay. To extend the sensitivity to even lower angular distances these effects have to be studied in more detail. The standard electron and muon reconstruction are also not optimized for a large displacement of their production vertex, which occurs more often when they are originating from high momentum leptonic tau decays. Consequently a dedicated approach for these electrons and muons could lead to further small improvements.

While the problem of overlap between candidates from different tau pair decay modes was easily solved for the  $\tau_{\text{had}}\tau_{\mu}$  decay channel, the  $\tau_{\text{had}}\tau_e$  channel imposes a greater challenge. Because of the similarities between the signatures of hadronic tau decays and electrons, especially the overlap between the  $\tau_{\text{had}}\tau_{\text{had}}$  mode and the  $\tau_{\text{had}}\tau_e$  mode has to be resolved. Although cuts were found to largely separate these channels, an unified reconstruction and identification method for both channels where the decay mode classification and background separation happen simultaneously, would be desirable. This would also reduce the number of algorithms which have to be maintained.

If the muon track could be removed during an early step of the reconstruction of the hadronic tau decays, the influences of the muon on the identification, energy scale and vertex correction can be reduced even further than possible in this work. This would be very desirable, because in this case the well refined and investigated standard algorithms would be sufficient to reconstruct and identify hadronic tau decays and muons almost independently of their overlap.

The results presented in this work rely largely on the precision of the simulated samples which were used. There are ongoing efforts to determine the systematic uncertainties which are connected to this. One approach is to use varied simulation samples to estimate the dependencies of the signal efficiencies and energy measurements on changes in the simulation setup. Another method is to use Standard Model processes involving boosted tau pairs which can be measured at ATLAS as a benchmark. With the outstanding luminosity provided by the LHC there are enough high momentum  $Z$  bosons decaying into boosted tau pairs produced to utilize this process for this purpose.

# Bibliography

- [1] M. Gell-Mann, *A Schematic Model of Baryons and Mesons*, Phys. Lett. **8** (1964), 214–215.
- [2] G. Zweig, *An  $SU(3)$  model for strong interaction symmetry and its breaking. Version 2*, Developments in the Quark Theory of Hadrons, Volume 1. Edited by D. Lichtenberg and S. Rosen. pp. 22-101 (1964), 22–101.
- [3] D. J. Gross and F. Wilczek, *Ultraviolet Behavior of Nonabelian Gauge Theories*, Phys. Rev. Lett. **30** (1973), 1343–1346.
- [4] H. D. Politzer, *Reliable Perturbative Results for Strong Interactions?*, Phys. Rev. Lett. **30** (1973), 1346–1349.
- [5] H. Fritzsch, M. Gell-Mann, and H. Leutwyler, *Advantages of the Color Octet Gluon Picture*, Phys. Lett. **47B** (1973), 365–368.
- [6] E. D. Bloom *et al.*, *High-Energy Inelastic  $e p$  Scattering at 6-Degrees and 10-Degrees*, Phys. Rev. Lett. **23** (1969), 930–934.
- [7] M. Breidenbach, J. I. Friedman, H. W. Kendall, E. D. Bloom, D. H. Coward, H. C. DeStaebler, J. Drees, L. W. Mo, and R. E. Taylor, *Observed Behavior of Highly Inelastic electron-Proton Scattering*, Phys. Rev. Lett. **23** (1969), 935–939.
- [8] S. L. Glashow, *Partial Symmetries of Weak Interactions*, Nucl. Phys. **22** (1961), 579–588.
- [9] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967), 1264–1266.
- [10] A. Salam and J. C. Ward, *Weak and electromagnetic interactions*, Il Nuovo Cimento (1955-1965) **11** (Feb, 1959), 568–577.
- [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** (1964), 321–323.
- [13] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Broken symmetries and the Goldstone theorem*, Adv. Part. Phys. **2** (1968), 567–708.
- [14] ATLAS, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. **B716** (2012), 1–29, [arXiv:1207.7214 \[hep-ex\]](#).
- [15] CMS, *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 \[hep-ex\]](#).
- [16] ATLAS, *Measurement of the Higgs boson mass in the  $H \rightarrow ZZ^* \rightarrow 4\ell$  and  $H \rightarrow \gamma\gamma$  channels with  $\sqrt{s}=13$  TeV  $pp$  collisions using the ATLAS detector*, ATLAS-CONF-2017-046.

- [17] ATLAS, *Measurement of the Higgs boson coupling properties in the  $H \rightarrow ZZ^* \rightarrow 4\ell$  decay channel at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, JHEP **03** (2018), 095, [arXiv:1712.02304 \[hep-ex\]](#).
- [18] ATLAS, *Evidence for the associated production of the Higgs boson and a top quark pair with the ATLAS detector*, Phys. Rev. **D97** (2018) 7, 072003, [arXiv:1712.08891 \[hep-ex\]](#).
- [19] ATLAS, *Test of CP Invariance in vector-boson fusion production of the Higgs boson using the Optimal Observable method in the ditau decay channel with the ATLAS detector*, Eur. Phys. J. **C76** (2016) 12, 658, [arXiv:1602.04516 \[hep-ex\]](#).
- [20] M. L. Perl, *Evidence for, and Properties of, the New Charged Heavy Lepton*, in *Proceedings: 12th Rencontres de Moriond, Leptons and Multi-Leptons. Flaine-Haute Savoie, France Mar 6-Mar 18 1977. 1.*, pp. 75–97. 1977.
- [21] D. Kirchmeier, *Reconstruction and Identification of Boosted Tau Pair Topologies at ATLAS*, Master’s thesis, Dresden, Tech. U., 2015-05-18.
- [22] E. Noether, *Invarianten beliebiger Differentialausdrücke*, Gött. Nachr. (1918), 37–44.
- [23] E. Noether, *Invariante Variationsprobleme*, Gött. Nachr. (1918), 235–257.
- [24] Particle Data Group, *Review of Particle Physics*, Chin. Phys. **C40** (2016) 10, 100001.
- [25] ATLAS, CMS, *Combined Measurement of the Higgs Boson Mass in pp Collisions at  $\sqrt{s} = 7$  and 8 TeV with the ATLAS and CMS Experiments*, Phys. Rev. Lett. **114** (2015), 191803, [arXiv:1503.07589 \[hep-ex\]](#).
- [26] LHC Higgs Cross Section Working Group, *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, FERMILAB-FN-1025-T, CERN-2017-002-M, [arXiv:1610.07922 \[hep-ph\]](#).
- [27] G. Bertone, D. Hooper, and J. Silk, *Particle dark matter: Evidence, candidates and constraints*, Phys. Rept. **405** (2005), 279–390, [arXiv:hep-ph/0404175 \[hep-ph\]](#).
- [28] G. Bertone and D. Merritt, *Dark matter dynamics and indirect detection*, Mod. Phys. Lett. **A20** (2005), 1021, [arXiv:astro-ph/0504422 \[astro-ph\]](#).
- [29] J. F. Donoghue, *The effective field theory treatment of quantum gravity*, AIP Conf. Proc. **1483** (2012), 73–94, [arXiv:1209.3511 \[gr-qc\]](#).
- [30] A. Djouadi, *The Anatomy of electro-weak symmetry breaking. I: The Higgs boson in the standard model*, Phys. Rept. **457** (2008), 1–216, [arXiv:hep-ph/0503172 \[hep-ph\]](#).
- [31] T. Kaluza, *On the Problem of Unity in Physics*, Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.) **1921** (1921), 966–972.
- [32] O. Klein, *Quantentheorie und fünfdimensionale relativitätstheorie*, Z. Phys. **37** (1926), 895–906.

- 
- [33] N. Arkani-Hamed, S. Dimopoulos, and G. R. Dvali, *Phenomenology, astrophysics and cosmology of theories with submillimeter dimensions and TeV scale quantum gravity*, Phys. Rev. **D59** (1999), 086004, arXiv:hep-ph/9807344 [hep-ph].
- [34] L. Randall and R. Sundrum, *An Alternative to compactification*, Phys. Rev. Lett. **83** (1999), 4690–4693, arXiv:hep-th/9906064 [hep-th].
- [35] L. Randall and R. Sundrum, *A Large mass hierarchy from a small extra dimension*, Phys. Rev. Lett. **83** (1999), 3370–3373, arXiv:hep-ph/9905221 [hep-ph].
- [36] ATLAS, *Search for new high-mass phenomena in the dilepton final state using 36 fb<sup>-1</sup> of proton-proton collision data at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, JHEP **10** (2017), 182, arXiv:1707.02424 [hep-ex].
- [37] ATLAS, *Search for new phenomena in high-mass diphoton final states using 37 fb<sup>-1</sup> of proton-proton collisions collected at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, Phys. Lett. **B775** (2017), 105–125, arXiv:1707.04147 [hep-ex].
- [38] CMS, *Search for narrow resonances in dilepton mass spectra in proton-proton collisions at  $\sqrt{s} = 13$  TeV and combination with 8 TeV data*, Phys. Lett. **B768** (2017), 57–80, arXiv:1609.05391 [hep-ex].
- [39] CMS, *Search for high-mass diphoton resonances in proton-proton collisions at 13 TeV and combination with 8 TeV search*, Phys. Lett. **B767** (2017), 147–170, arXiv:1609.02507 [hep-ex].
- [40] S. J. Huber and Q. Shafi, *Fermion masses, mixings and proton decay in a Randall-Sundrum model*, Phys. Lett. **B498** (2001), 256–262, arXiv:hep-ph/0010195 [hep-ph].
- [41] G. Perez and L. Randall, *Natural Neutrino Masses and Mixings from Warped Geometry*, JHEP **01** (2009), 077, arXiv:0805.4652 [hep-ph].
- [42] ATLAS, *Searches for heavy ZZ and ZW resonances in the  $llqq$  and  $\nu\nu qq$  final states in pp collisions at  $\sqrt{s} = 13$  TeV with the ATLAS detector*, CERN-EP-2017-146, arXiv:1708.09638 [hep-ex].
- [43] CMS, *Search for diboson resonances in the  $2\ell 2\nu$  final state*, CMS-B2G-16-023, CERN-EP-2017-279, arXiv:1711.04370 [hep-ex].
- [44] CMS, *Search for a massive resonance decaying to a pair of Higgs bosons in the four b quark final state in proton-proton collisions at  $\sqrt{s} = 13$  TeV*, CMS-B2G-16-026, CERN-EP-2017-238, arXiv:1710.04960 [hep-ex].
- [45] A. Oliveira, *Gravity particles from Warped Extra Dimensions, predictions for LHC* arXiv:1404.0102 [hep-ph].
- [46] O. S. Bruning, P. Collier, P. Lebrun, S. Myers, R. Ostojic, J. Poole, and P. Proudlock, *LHC Design Report Vol.1: The LHC Main Ring*, CERN-2004-003-V1, CERN-2004-003, CERN-2004-003-V-1.

- [47] CMS, *The CMS Experiment at the CERN LHC*, JINST **3** (2008), S08004.
- [48] LHCb, *The LHCb Detector at the LHC*, JINST **3** (2008), S08005.
- [49] ALICE, *The ALICE experiment at the CERN LHC*, JINST **3** (2008), S08002.
- [50] ATLAS, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008), S08003.
- [51] W.-J. Zhang, W.-G. Ma, R.-Y. Zhang, X.-Z. Li, L. Guo, and C. Chen, *Double Higgs boson production and decay in Randall-Sundrum model at hadron colliders*, Phys. Rev. **D92** (2015), 116005, [arXiv:1512.01766 \[hep-ph\]](#).
- [52] *ATLAS Public Luminosity Results Run2*, <https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2>, visited on 08.03.2018.
- [53] J. Pequeno, *Computer generated image of the whole ATLAS detector*, Mar, 2008.
- [54] ATLAS, *ATLAS Insertable B-Layer Technical Design Report*, CERN-LHCC-2010-013, ATLAS-TDR-19.
- [55] J. Pequeno, *Computer generated image of the ATLAS inner detector*, Mar, 2008.
- [56] J. Pequeno, *Computer Generated image of the ATLAS calorimeter*, Mar, 2008.
- [57] J. Pequeno, *Computer generated image of the ATLAS Muons subsystem*, Mar, 2008.
- [58] Y. Nakahama, *The ATLAS Trigger System: Ready for Run-2*, J. Phys. Conf. Ser. **664** (2015) 8, 082037.
- [59] A. Hoecker, P. Speckmayer, J. Stelzer, J. Therhaag, E. von Toerne, and H. Voss, *TMVA: Toolkit for Multivariate Data Analysis*, PoS **ACAT** (2007), 040, [arXiv:physics/0703039](#).
- [60] Y. Freund and R. E. Schapire, *A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting*, Journal of Computer and System Sciences **55** (1997) 1, 119 – 139.
- [61] T. Cornelissen, M. Elsing, S. Fleischmann, W. Liebig, and E. Moyses, *Concepts, Design and Implementation of the ATLAS New Tracking (NEWT)*, ATL-SOFT-PUB-2007-007, ATL-COM-SOFT-2007-002.
- [62] ATLAS, *Performance of the ATLAS Track Reconstruction Algorithms in Dense Environments in LHC Run 2*, Eur. Phys. J. **C77** (2017) 10, 673, [arXiv:1704.07983 \[hep-ex\]](#).
- [63] R. Fruhwirth, *Application of Kalman filtering to track and vertex fitting*, Nucl. Instrum. Meth. **A262** (1987), 444–450.

- 
- [64] ATLAS Collaboration, *Performance of the reconstruction of large impact parameter tracks in the ATLAS inner detector*, ATL-PHYS-PUB-2017-014.
- [65] G. Piacquadio, K. Prokofiev, and A. Wildauer, *Primary vertex reconstruction in the ATLAS experiment at LHC*, J. Phys. Conf. Ser. **119** (2008), 032033.
- [66] *An imaging algorithm for vertex reconstruction for ATLAS Run-2*, ATL-PHYS-PUB-2015-008.
- [67] W. Lampl, S. Laplace, D. Lelas, P. Loch, H. Ma, S. Menke, S. Rajagopalan, D. Rousseau, S. Snyder, and G. Unal, *Calorimeter clustering algorithms: Description and performance*, ATL-LARG-PUB-2008-002, ATL-COM-LARG-2008-003.
- [68] ATLAS, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. **C77** (2017), 490, [arXiv:1603.02934 \[hep-ex\]](#).
- [69] M. Cacciari, G. P. Salam, and G. Soyez, *The Anti- $k(t)$  jet clustering algorithm*, JHEP **04** (2008), 063, [arXiv:0802.1189 \[hep-ph\]](#).
- [70] ATLAS, *Electron efficiency measurements with the ATLAS detector using the 2015 LHC proton-proton collision data*, ATLAS-CONF-2016-024.
- [71] ATLAS Collaboration, *Electron and photon reconstruction and performance in ATLAS using a dynamical, topological cell clustering-based approach*, ATL-PHYS-PUB-2017-022.
- [72] *ATLAS Electron Public Results*, <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/EGAM-2018-002/index.html>, visited on 11.03.2018.
- [73] ATLAS, *Muon reconstruction performance of the ATLAS detector in proton-proton collision data at  $\sqrt{s} = 13$  TeV*, Eur. Phys. J. **C76** (2016) 5, 292, [arXiv:1603.05598 \[hep-ex\]](#).
- [74] *ATLAS Muon Public Results*, <https://atlas.web.cern.ch/Atlas/GROUPS/PHYSICS/PLOTS/MUON-2017-002/index.html>, visited on 13.03.2018.
- [75] *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.
- [76] D. Duschinger, *Search for neutral bosons beyond the Standard Model decaying into tau leptons using the ATLAS detector at the LHC*. PhD thesis, Dresden, Tech. U., In preparation, 2018.
- [77] C. Deutsch, *Identification and Classification of Hadronic Tau Lepton Decays in the ATLAS Experiment*, Master's thesis, Bonn U., 2017.

- [78] T. Sjostrand, S. Mrenna, and P. Z. Skands, *A Brief Introduction to PYTHIA 8.1*, Comput. Phys. Commun. **178** (2008), 852–867, [arXiv:0710.3820](#) [hep-ph].
- [79] GEANT4, *GEANT4: A Simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003), 250–303.
- [80] ATLAS, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. **C70** (2010), 823–874, [arXiv:1005.4568](#) [physics.ins-det].
- [81] R. Brun and F. Rademakers, *ROOT: An object oriented data analysis framework*, Nucl. Instrum. Meth. **A389** (1997), 81–86.
- [82] C. J. Clopper and E. S. Pearson, *The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial*, Biometrika **26** (1934) 4, 404–413.
- [83] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, *Generative Adversarial Networks*, ArXiv e-prints (June, 2014), [arXiv:1406.2661](#) [stat.ML].
- [84] G. Louppe, M. Kagan, and K. Cranmer, *Learning to Pivot with Adversarial Networks* [arXiv:1611.01046](#) [stat.ME].
- [85] C. Shimmin, P. Sadowski, P. Baldi, E. Weik, D. Whiteson, E. Goul, and A. Sogaard, *Decorrelated Jet Substructure Tagging using Adversarial Neural Networks*, Phys. Rev. **D96** (2017) 7, 074034, [arXiv:1703.03507](#) [hep-ex].



# A Additional Tables and Figures

## A.1 Data Samples

Table A.1:  $G \rightarrow hh \rightarrow \tau\tau\tau\tau$  Monte Carlo samples with different graviton mass:

Sample name	number of events
mc16_13TeV:mc16_13TeV.425108.MadGraphPythia8EvtGen_A14NNPDF23LO_RS_G_hh_4tau.c10_M1000.merge.AOD.e6072.s3126_r9364_r9315	193000
mc16_13TeV:mc16_13TeV.425100.MadGraphPythia8EvtGen_A14NNPDF23LO_RS_G_hh_4tau.c10_M1500.merge.AOD.e5485.s3126_r9364_r9315	200000
mc16_13TeV:mc16_13TeV.425101.MadGraphPythia8EvtGen_A14NNPDF23LO_RS_G_hh_4tau.c10_M1750.merge.AOD.e5485.s3126_r9364_r9315	200000
mc16_13TeV:mc16_13TeV.425102.MadGraphPythia8EvtGen_A14NNPDF23LO_RS_G_hh_4tau.c10_M2000.merge.AOD.e5485.s3126_r9364_r9315	200000
mc16_13TeV:mc16_13TeV.425103.MadGraphPythia8EvtGen_A14NNPDF23LO_RS_G_hh_4tau.c10_M2250.merge.AOD.e5485.s3126_r9364_r9315	200000
mc16_13TeV:mc16_13TeV.425104.MadGraphPythia8EvtGen_A14NNPDF23LO_RS_G_hh_4tau.c10_M2500.merge.AOD.e5485.s3126_r9364_r9315	200000
mc16_13TeV:mc16_13TeV.425105.MadGraphPythia8EvtGen_A14NNPDF23LO_RS_G_hh_4tau.c10_M3000.merge.AOD.e6072.s3126_r9364_r9315	100000
mc16_13TeV:mc16_13TeV.425106.MadGraphPythia8EvtGen_A14NNPDF23LO_RS_G_hh_4tau.c10_M4000.merge.AOD.e6072.s3126_r9364_r9315	100000
mc16_13TeV:mc16_13TeV.425107.MadGraphPythia8EvtGen_A14NNPDF23LO_RS_G_hh_4tau.c10_M5000.merge.AOD.e6072.s3126_r9364_r9315	99000

Table A.2: Data sample names used to acquire background candidates for training of the classifiers:

Sample name
data15_13TeV:data15_13TeV.00284484.physics_Main.merge.AOD.r9264_p3083
data15_13TeV:data15_13TeV.00284427.physics_Main.merge.AOD.r9264_p3083
data15_13TeV:data15_13TeV.00284420.physics_Main.merge.AOD.r9264_p3083
data15_13TeV:data15_13TeV.00284285.physics_Main.merge.AOD.r9264_p3083
data15_13TeV:data15_13TeV.00284213.physics_Main.merge.AOD.r9264_p3083
data15_13TeV:data15_13TeV.00284154.physics_Main.merge.AOD.r9264_p3083

Table A.3: Names of triggers used to select the events for the background samples:

Trigger name
HLT_j360_a10r_L1J100
HLT_j360_a10_lcw_L1J100
HLT_j400_a10r_L1J100
HLT_j400_a10_lcw_L1J100
HLT_j420_a10_lcw_L1J100
HLT_j420_a10r_L1J100
HLT_j390_a10t_lcw_jes_30smcINF_L1J100
HLT_j420_a10r_L1J100
HLT_j420_a10t_lcw_jes_L1J100
HLT_j420_a10_lcw_subjes_L1J100
HLT_j420_a10t_lcw_jes_40smcINF_L1J100
HLT_j440_a10r_L1J100
HLT_j440_a10t_lcw_jes_L1J100
HLT_j440_a10_lcw_subjes_L1J100
HLT_j440_a10t_lcw_jes_40smcINF_L1J100
HLT_j460_a10r_L1J100
HLT_j460_a10t_lcw_jes_L1J100
HLT_j460_a10_lcw_subjes_L1J100
HLT_j480_a10r_L1J100
HLT_j480_a10t_lcw_jes_L1J100
HLT_j480_a10_lcw_subjes_L1J100

Table A.4: BDT-Parameters used for the classification of semi-leptonic decaying tau pairs:

parameter name	description	value
<b>NTrees</b>	Number of decision trees	400
<b>BoostType</b>	Boosting Algorithm	AdaBoost
<b>MinNodeSize</b>	Minimum required number of candidates in a leaf node in percent	0.1

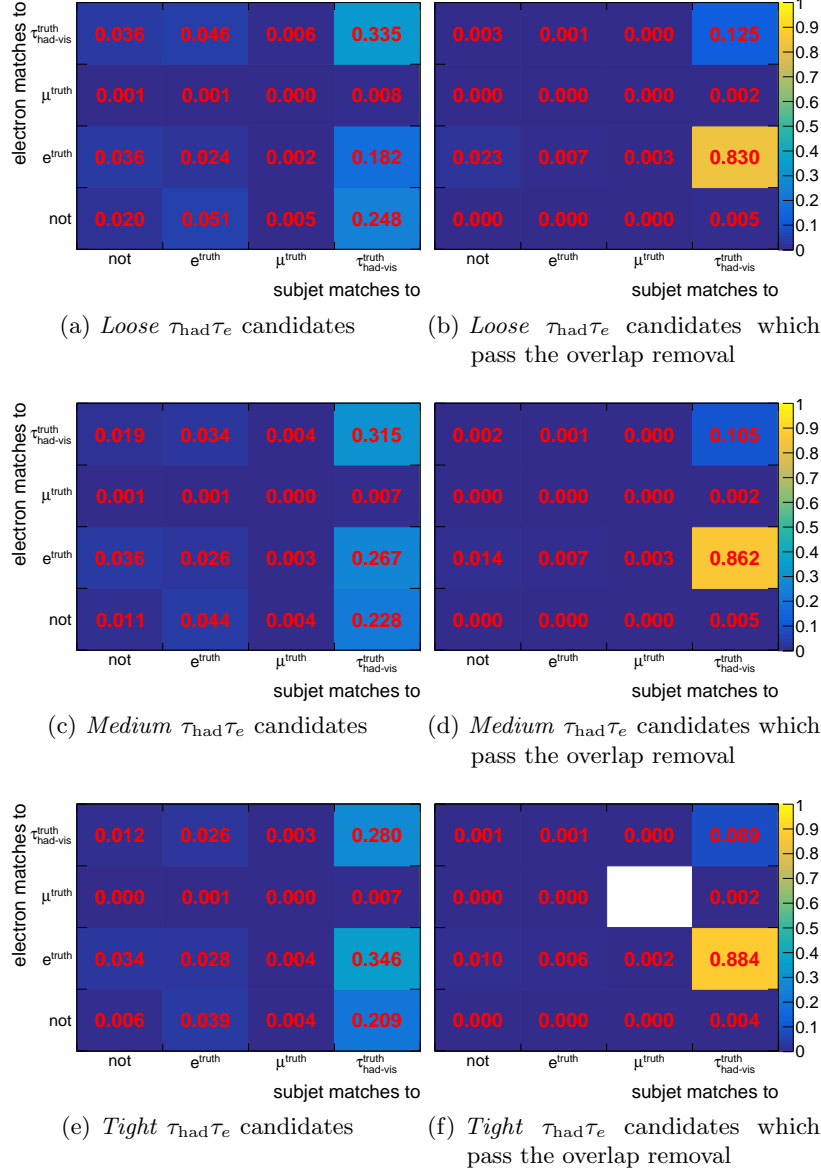
A.2 The  $\tau_{\text{had}}\tau_e$  Channel

Figure A.1: Each bin of the plots indicates the relative frequency of different categories of  $\tau_{\text{had}}\tau_e$  candidates in the  $G \rightarrow hh \rightarrow \tau\tau\tau\tau$  samples. The categories are defined by the truth particles, which produce the reconstructed electron and the tau subjet. In the left column the candidates have to pass different working points of the  $\tau_{\text{had}}\tau_e$  identification and in the right they have to additionally pass the overlap removal requirements presented in section 5.1.3.

### A.3 The $\tau_{\text{had}}\tau_{\mu}$ Channel

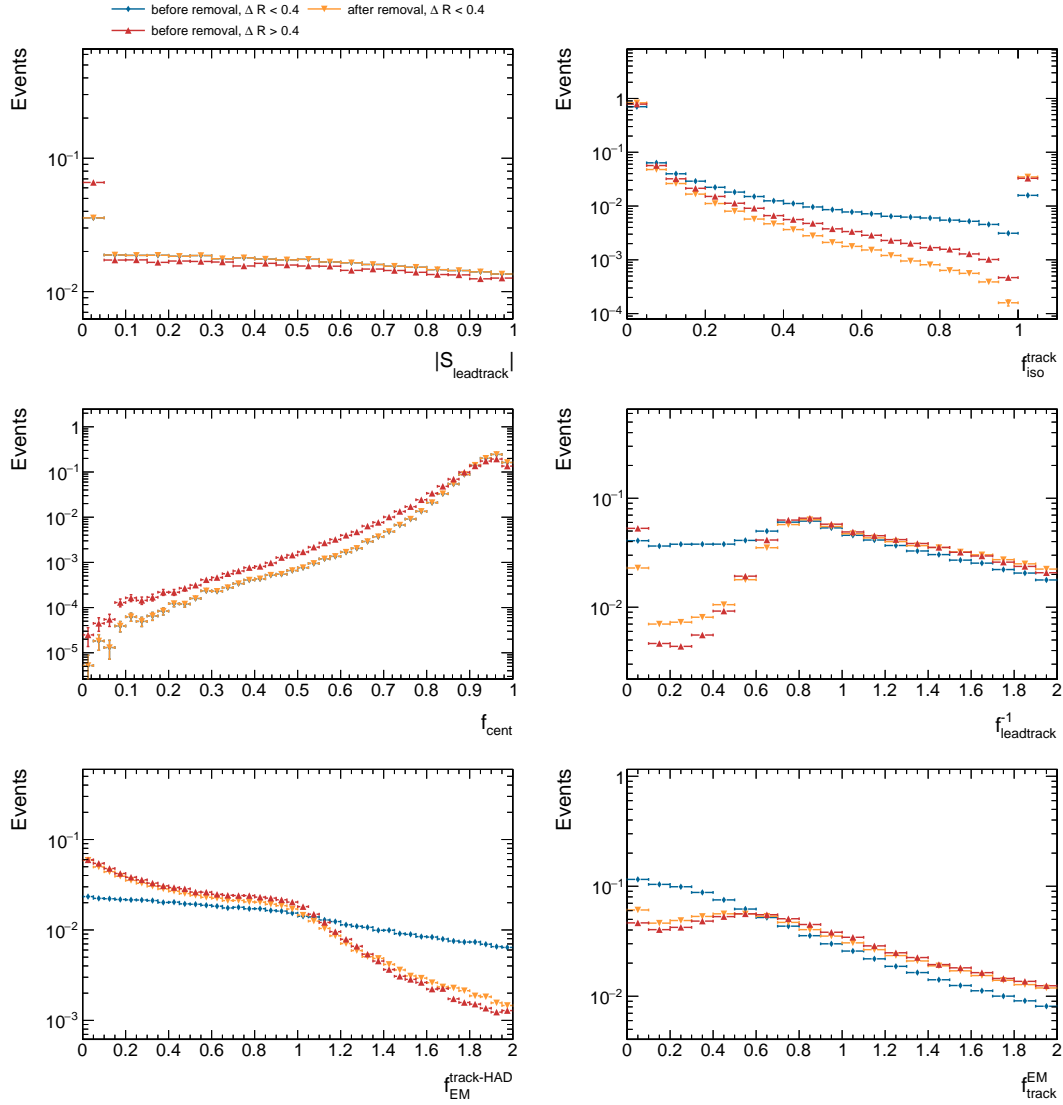


Figure A.2: Normalized distributions of identification variables for hadronically decaying taus. The color indicate different angular distances  $\Delta R$  of the  $\tau_{\text{had-vis}}^{\text{truth}}$  to a nearby  $\mu^{\text{truth}}$  and if the variables are calculated before or after removing tracks associated to muons.

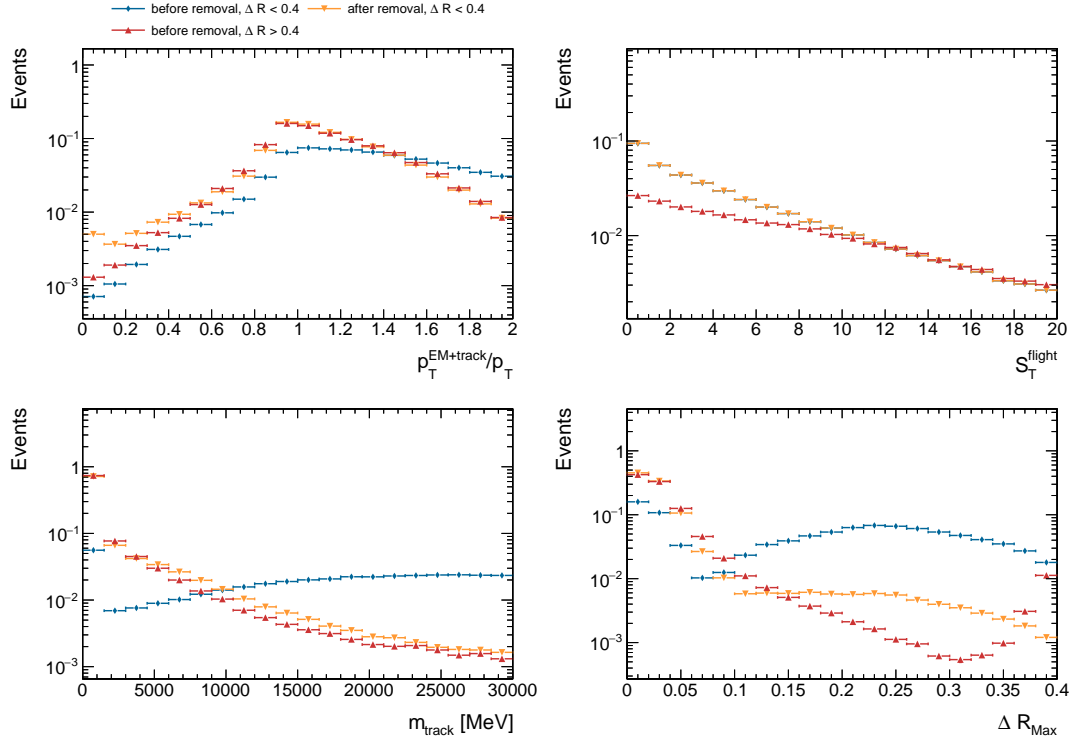


Figure A.3: Normalized distributions of identification variables for hadronically decaying taus. The color indicate different angular distances  $\Delta R$  of the  $\tau_{\text{had-vis}}^{\text{truth}}$  to a nearby  $\mu^{\text{truth}}$  and if the variables are calculated before or after removing tracks associated to muons.

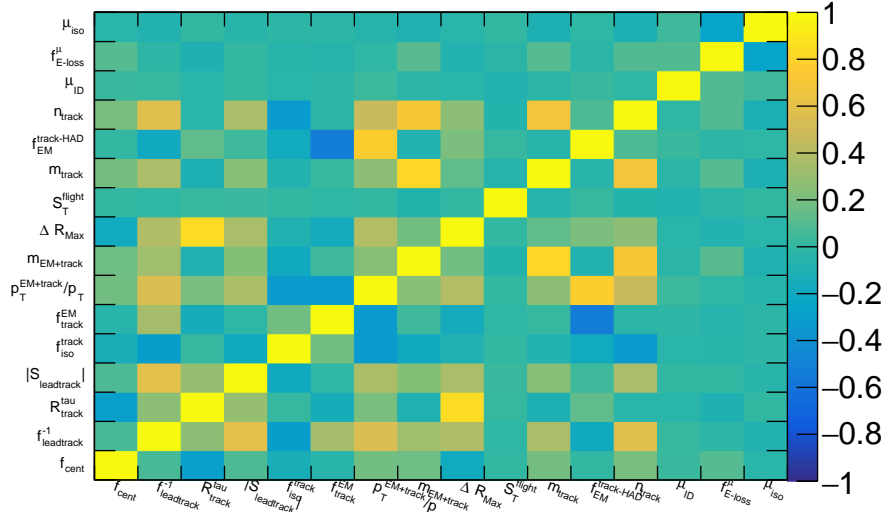


Figure A.4: Correlation between identification variables for  $\tau_{\text{had}}\tau_{\mu}$  candidates in the background sample.

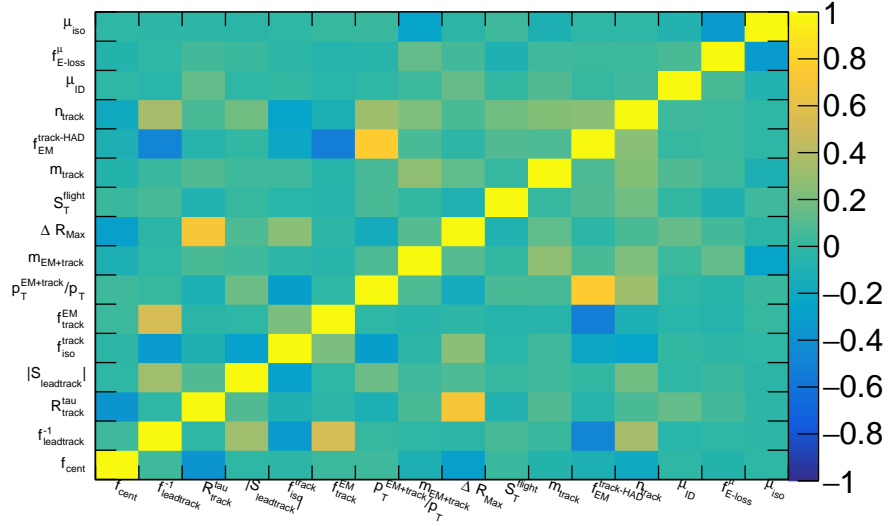


Figure A.4: Correlation between identification variables for  $\tau_{\text{had}}\tau_{\mu}$  candidates in the signal sample.

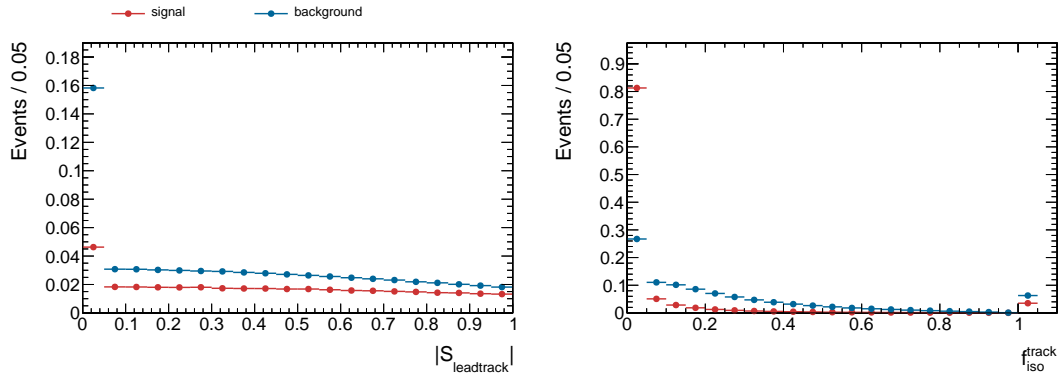


Figure A.5: Normalized distributions of identification variables for signal and background used for the  $\tau_{\text{had}}\tau_{\mu}$  identification.

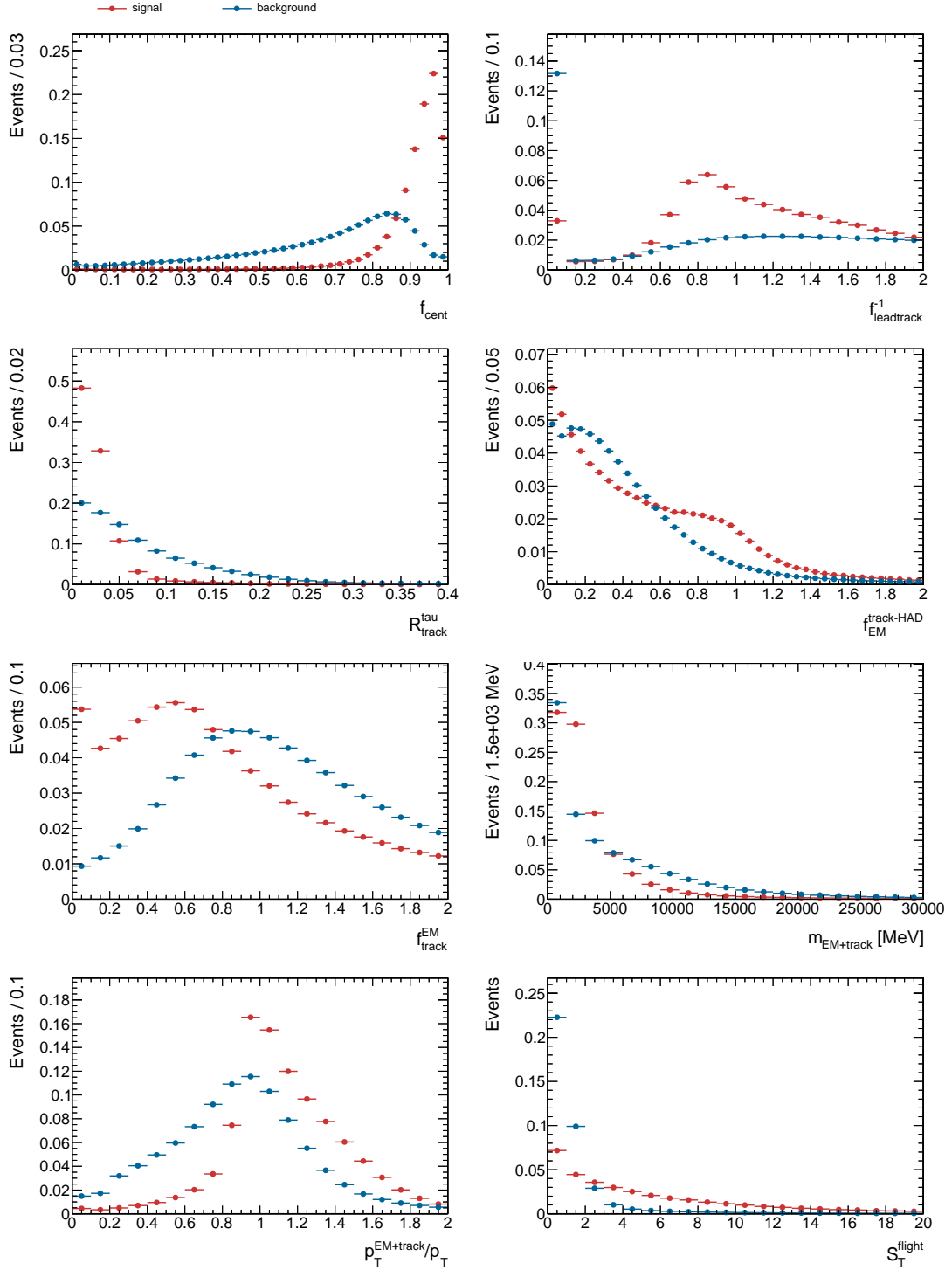


Figure A.5: Normalized distributions of identification variables for signal and background used for the  $\tau_{\text{had}}\tau_{\mu}$  identification.

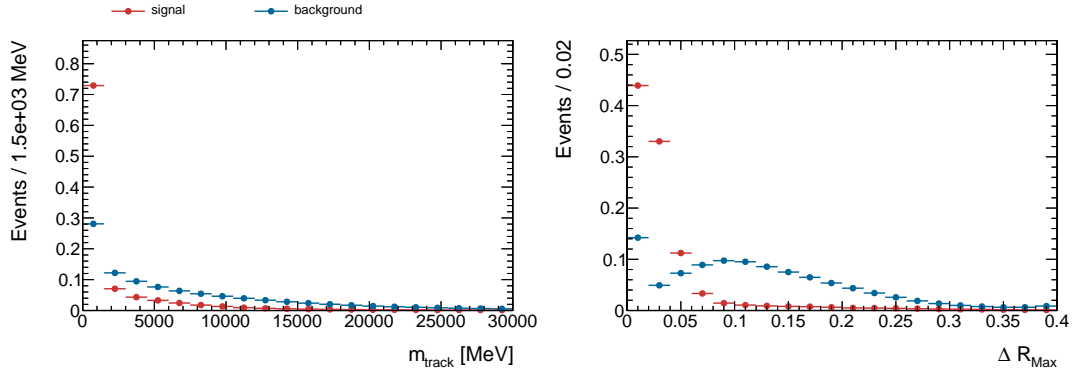


Figure A.5: Normalized distributions of identification variables for signal and background used for the  $\tau_{\text{had}}\tau_{\mu}$  identification.

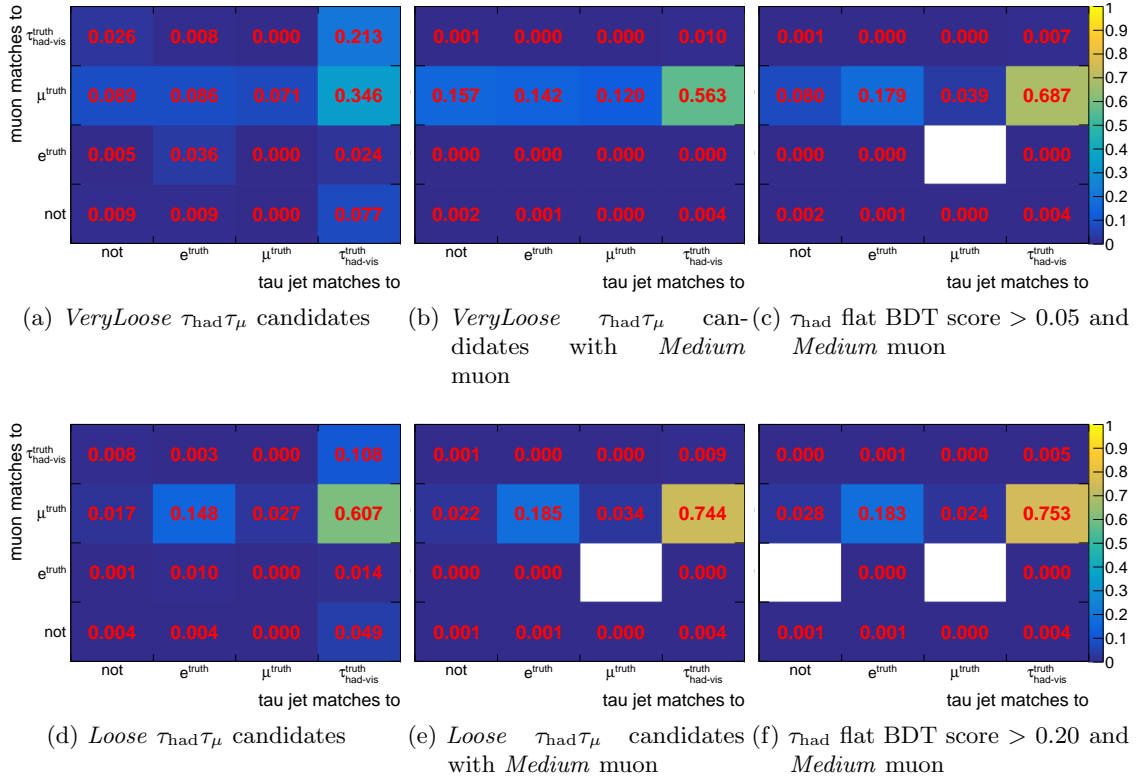


Figure A.6: Each bin of the plots indicates the relative frequency of different categories of  $\tau_{\text{had}}\tau_{\mu}$  candidates in the  $G \rightarrow hh \rightarrow \tau\tau\tau\tau$  samples. The categories are defined by the truth particles which produce the reconstructed muon and the tau jet. The results for various candidate selections are presented. In the left column candidates which pass different working points of the newly trained  $\tau_{\text{had}}\tau_{\mu}$  identification are used. In the middle an additional requirement on the reconstructed muon is applied and on the right the standard  $\tau_{\text{had}}$  identification with recalculated variables is used.



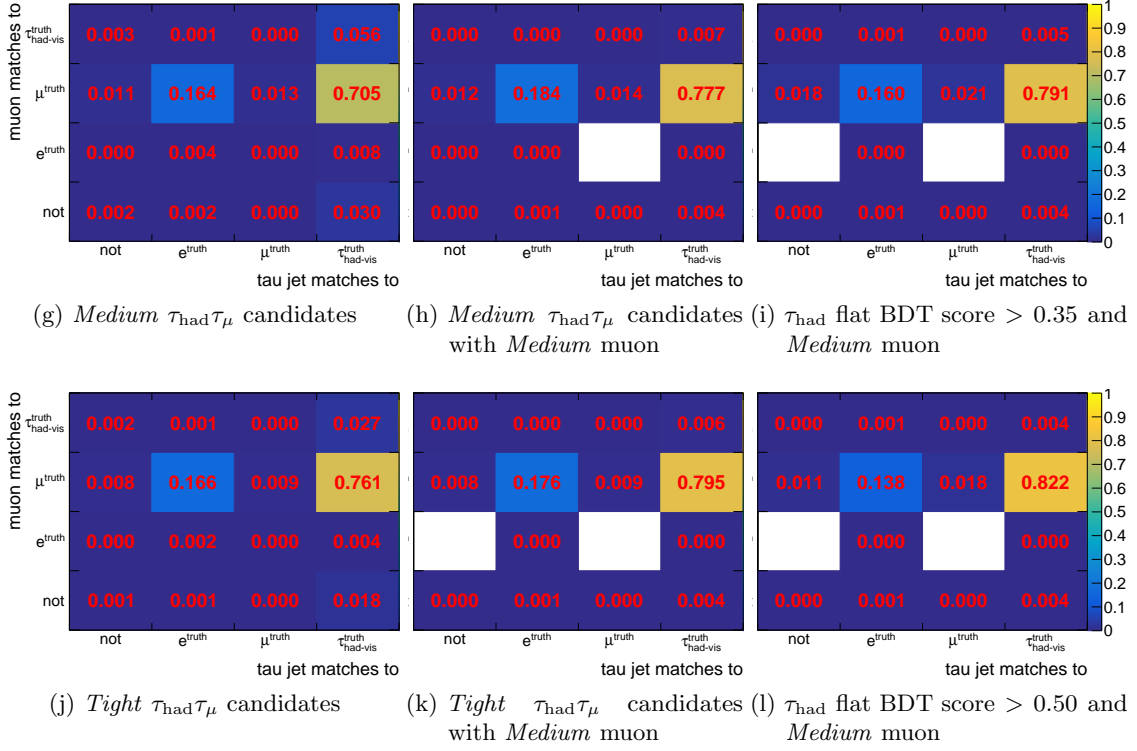


Figure A.6: Each bin of the plots indicates the relative frequency of different categories of  $\tau_{\text{had}}\tau_{\mu}$  candidates in the  $G \rightarrow hh \rightarrow \tau\tau\tau\tau$  samples. The categories are defined by the truth particles which produce the reconstructed muon and the tau jet. The results for various candidate selections are presented. In the left column candidates which pass different working points of the newly trained  $\tau_{\text{had}}\tau_{\mu}$  identification are used. In the middle an additional requirement on the reconstructed muon is applied and on the right the standard  $\tau_{\text{had}}$  identification with recalculated variables is used.

## List of Abbreviations

**SM** Standard Model of Particle Physics

**QFT** Quantum Field Theory

**RS** Randall-Sundrum

**LHC** Large Hadron Collider

**QCD** Quantum Chromo Dynamics

**CERN** European Organization for Nuclear Research

**LEP** Large Electron-Positron Collider

**IBL** Insertable B-Layer

**TRT** Transition Radiation Tracker

**SCT** Semiconductor Tracker

**ID** Inner Detector

**BDT** Boosted Decision Tree

## Danksagung

Ich danke allen die diese Arbeit ermöglicht haben. Das spannende und vielfältige Thema, welches ich bearbeiten durfte, hat mir sehr gefallen. Dafür und für die ausgezeichnete Unterstützung gilt mein Dank Prof. Dr. Arno Straessner. Für die wichtigen Hinweise zu meiner Arbeit und die offene Diskussion neuer Ansätze danke ich David Kirchmeier und Dr. Wolfgang Mader. Ohne die Vorarbeit und Hilfe von David Kirchmeier hätte es nicht zu dieser Arbeit kommen können. Dirk Duschinger danke ich für die Entwicklung und Erklärung des hilfreichen *THOR* Frameworks, welches meinen Fortschritt beschleunigt und einen schnellen Einblick in die Tau Performance Arbeit ermöglicht hat. Max Märker danke ich für die spannenden Ausführungen zu Neuheiten im Bereich neuronaler Netze. Der gesamten Tau Gruppe in Dresden danke ich für die erfrischenden Gespräche und Diskussionen in zahlreichen Meetings oder im Büro. Bei allen Beschäftigten des IKTPs bedanke ich mich für die angenehme kollegiale Arbeitsatmosphäre. Meiner Familie danke ich für ihren uneingeschränkten Rückhalt und besonders Julia für ihre liebevolle Unterstützung.



## **Selbständigkeitserklärung Statement of Authorship**

Ich versichere, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Ich reiche sie erstmals als Prüfungsleistung ein. Mir ist bekannt, dass ein Betrugsversuch mit der Note "nicht ausreichend" (5,0) geahndet wird und im Wiederholungsfall zum Ausschluss von der Erbringung weiterer Prüfungsleistungen führen kann.

I hereby declare that I have written this Master's thesis independently and have listed all the used sources and means. I understand that attempted fraud will result in the failing grade "not sufficient" (5,0) and in case of recurrence in exclusion from completing of any further examinations and assessments.

Christian Wiel  
Dresden, den 2.5.2018