TECHNISCHE UNIVERSITÄT DRESDEN

INSTITUT FÜR KERN- UND TEILCHENPHYSIK
FACHRICHTUNG PHYSIK
FAKULTÄT FÜR MATHEMATIK UND NATURWISSENSCHAFTEN

Diplomarbeit

zur Erlangung des akademischen Grades
Diplom-Physiker

Determination of the misidentification rate of hadronically decaying tau leptons using $Z \to ee + jet$ events with ATLAS

Felix Friedrich

geboren am 05.10.1981 in Schwerin

Dresden 2009

1. Gutachter: Jun.-Prof. Dr. Arno Straessner

2. Gutachter: Dr. Wolfgang Mader

eingereicht am 15.12.2009

Abstract

In searches for the Higgs boson and physics beyond the Standard Model, the presence of tau leptons in the final state is an important signature at the ATLAS detector of the Large Hadron Collider. Hadronically decaying tau leptons can be reconstructed and identified over a broad kinematic range. Their signature in the detector is similar to other processes involving hadronic interactions. Therefore it is important to determine the misidentified tau lepton candidates in the detector. An identification method use cuts based on calorimeter variables, which are presumed to be well understood for the first data. In this thesis a method is described, to estimate the number of misidentified hadronically decaying tau leptons candidates caused by QCD dijet background events using $Z \rightarrow ee + jet$ events.

Kurzfassung

Die Anwesenheit von Tau Leptonen im Endzustand ist eine wichtige Signatur bei der Suche nach dem Higgs Boson oder nach neuer Physik jenseits des Standard Modells am ATLAS Detektor des LHCs. Hadronisch zerfallende Tau Leptonen können im ATLAS Detektor über einen grossen kinematischen Bereich rekonstruiert und identifiziert werden.Ihre Signatur im Detektor ähnelt dabei jener anderer hadronischer Prozesse, wie dem QCD Dijet Untergrund. Daher ist es wichtig die Anzahl der im Detektor falsch rekonstruierten und falsch identifizierten Tau-Leptonen-Kandidaten zu bestimmen. Eine wichtige Identifizierungsmethode nutzt Schnitte auf Kalorimetervariablen, welche in den ersten Daten als besonders gut verstanden gelten. In dieser Studie wird eine Methode vorgestellt, die, ausgehend von diesen Kalorimetervariablen, die Anzahl von fehlidentifizierten, hadronisch zerfallenden Tau-Lepton-Kandidaten in QCD Untergrundereignissen mit Hilfe von $Z \rightarrow ee + jet$ Ereignissen abschätzt.

Contents

Chapter 1

Introduction

When the tau lepton was observed in 1974 at the Stanford electron-positron collider, particle physics was nearly 80 years old. Since the discovery of the electron as the first particle of the leptons by Thompson in 1897, particle physics has made a big development. The most important part of particle physics is represented by the Standard Model which describes all known particles and their interactions. Over the decades it is proven by many experiments and particle colliders. But the development is not finished. Thus, the last lepton, the tau-neutrino, was observed in the year 2000. There are various studies ongoing, concentrating on new approaches beyond the current limits of the Standard Model. Many extensions of the Standard Model predict new particles with high masses. One of them is the Higgs boson, which is the quanta of the scalar Higgs field. This field is predicted by the Standard Model, to acquire the mass of the particles. If this field exist, the Higgs boson has to be observable. More Higgs bosons are predicted by the MSSM Model, an extension of the Standard Model, which includes Supersymmetry. To discover the Higgs boson the development of particle colliders with high collision energy and high efficient detectors are necessary. The recently started Large Hadron Collider and with it the ATLAS detector provide the environment to start a new era of particle physics. Having a collision energy of 14 TeV, new developed analyzing techniques and state-of-the-art computational infrastructure, it will be possible to measure and analyze known particles, like the tau lepton, with high precision and to study their decay products. Also there is the potential to discover the Higgs boson and new physics beyond the Standard Model. At many process possible at the LHC tau leptons are present in the final state and therefore an important signature for discovering new physics. Hence, a good tau lepton identification efficiency is essential.

Tau leptons have a short life time and decay as well as leptons and hadrons. While electrons and muons can be detected directly with specialized detector components, various analyzing techniques are necessary for the reconstruction of tau leptons. A suitable method of tau lepton reconstruction is based on the hadronic decay products. But this opens a challenge, to distinguish between the tau decay products and other processes involving hadronic interactions, like QCD dijets. The latter are produced at the LHC with very high rates being many orders of magnitude larger than tau leptons. Because QCD dijets are similar to hadronically decaying tau leptons, QCD dijets can also be identified as tau lepton candidates. Therefore it is necessary to measure the misidentification rate of tau lepton candidates caused by QCD dijets. This thesis is focusing on the determination of these misidentification rate (fake rate) and describes a method, to estimate the misidentification rate caused by QCD jets using jets from $Z \to ee + jet$ events. In $Z \to ee + jet$ processes no tau leptons are involved and therefore this process is suited to study the tau misidentification rate. At this, first a clean $Z \to ee + jet$ event is selected. With the additional hadronic jets the tau reconstruction and identification algorithm is tested, determine a misidentification rate depending on the hadronic jet variables. After that, the expected tau misidentification rate caused by QCD jets is compared with the estimated fake rate. The purpose is to find a set of variables, that a good agreement between estimated and expected tau fake rate is realized.

After this first chapter, in Chapter 2 the theory of Standard Model, this thesis based on, and important processes involving tau leptons are described. An overview of the Large Hadron Collider and the ATLAS experiment is given in Chapter 3. Chapter 4 characterized the tau reconstruction and identification process. In Chapter 5 the determination method of the misidentification rate is described and results are presented. A summary and an outlook on further developments is given in Chapter 6.

Chapter 2

Theoretical Foundations

2.1 The Standard Model

2.1.1 Overview

The tau particle is one of the elementary particles whose interactions are described by the theory of the Standard Model (SM). This theory was tested in many experiments and found to describe nature at the level of quantum corrections. There are two types of matter particles, leptons and quarks, both carrying a spin of $1/2$ in units of $\hbar$ and therefore called fermions [1]. As currently known they are structureless at smallest distances. Leptons and quarks are arranged in three generations. The first generation consist of the electron and the electron-neutrino on the lepton side and the up- and down-quark on the quark side. The second generation are the muon with the muon-neutrino for the leptons and charm- and strange-quark for quarks. The tau particle belongs to the third generation of leptons together with the tau-neutrino. The third generation quarks are the bottom- and top-quark. The fermions cover a wide mass spectrum, beginning with the almost massless neutrinos. The top quark is the heaviest fermion with nearly two hundred GeV reaching the mass of big atoms. All fermions are listed in Table 2.1 and every particle has a corresponding anti-particle, having an opposite electric charge. The atoms and therefore all the visible matter in the universe are only made up of the first generation of leptons and quarks. The second and third generations can only be produced in cosmic radiation and particle accelerators. The interaction between the particles are described by four forces, three of them are included in the Standard Model. These are the electromagnetic interaction, the weak interaction and the strong interaction. The gravitation is not included and described by the Standard Model. The combination with the other forces is still a challenge for building a unified theory. At distances of ordinary particle physics the gravitation does not play a role, because its effect is many orders of magnitude smaller than the others. Hence it is neglected in this study. Leptons interact electromagnetically and weakly if they are charged and participate only in weak interaction if they are neutral. Quarks can interact in all three ways, strongly, weakly and electromagnetically. All interactions have associated particles, which act as mediator between the particles [1]. These particles are called gauge bosons and have a spin of 1. The interactions and their particles can be found in Table 2.2.

2.1.2 Interactions

The interactions in the Standard Model are described as quantum field theories with local gauge symmetries. In this theory a free fermion is described by a Lagrangian, which is invariant under gauge transformation of the fermion field. A global gauge transformation leads to conservation of quantum numbers, while local gauge transformation introduce new quantum fields, which can be identified with the gauge bosons.

	leptons			quarks		
Generation	particle	el. charge/e	mass/MeV	particle	el. charge/e	mass/GeV
first	ν_e	0	$< 2 \times 10^{-6}$	u	$2/3$	$1.5 - 3.3 \times 10^{-3}$
	e^-	-1	0.511	d	$-1/3$	$3.5 - 6.0 \times 10^{-3}$
second	ν_μ	0	< 0.19	c	$2/3$	1.27
	μ^-	-1	105.66	s	$-1/3$	104×10^{-3}
third	ν_τ	0	< 18.2	t	$2/3$	171.2
	τ^-	-1	1776.84	b	$-1/3$	4.2

Table 2.1: Fundamental fermions [2]

interaction	range/m	particle	mass/GeV
electromagnetic	∞	γ	$< 10^{-27}$
weak	10^{-18}	W^+/W^-	80.40
		Z	91.19
strong	$< 10^{-15}$	g	0

Table 2.2: Fundamental interactions [2]

The **electromagnetic interaction** is described by quantum electrodynamics (QED) and is based on the $U(1)$ symmetry group. The gauge boson is the photon and couples to the electric charge, but carries no electric charge. The photon can not self-interact and to preserve local gauge symmetry it has no mass. Both properties are resulting in an infinite range of the electromagnetic interaction.

The **strong interaction** is described by quantum chromodynamics (QCD), which is constructed in analogy to QED and is based on a $SU(3)_C$ symmetry group. The strong charge is the color charge, with three types: “red”, “green” and “blue”. Therefore every quark exists in these three colors. The force carrier is the massless gluon and carries also a color charge. Due to this charge the gluon can interact with other gluons, leading to the finite range of the strong interaction. Color charged particles can not exist as free particles (confinement). Hence, bound states of quarks – mesons, consisting of a quark-anti-quark pair and hadrons, consisting of three quarks – have no color charge. If two quarks are separated from each other, their mutual potential energy increase so much that the energy stored in the potential converts into a new $q\bar{q}$-pair. The quarks bind into hadrons, which mostly emerge in quite well-collimated jets. In contrast, at small distances of about nucleon diameters quarks can be seen as non-interacting particles (asymptotic freedom).

The **weak interaction** can not be described alone as a gauge theory and must be combined with the electromagnetic interaction, resulting in the electroweak interaction (often called GSW theory[1] [3]). The gauge bosons of the weak interaction are two charged bosons, $W^{\pm}$, and one neutral boson, Z, which are very massive. This leads to a short-ranged interaction. The weak interaction is realized as a chiral gauge theory, which have different couplings to left-handed and right-handed particles. The left-handed particles are arranged as charged doublets, while the right-handed particles are chargeless singlets. The underlying symmetry group $SU(2)_L \times (U)_Y$ reflects the unification of the two theories.

In the weak theory and in QCD, the coupling strength decreases with increasing energy, because their gauge bosons can interact with themselves. Unlike, in QED the coupling increases due to the non-interacting photon.

[1] named after the authors of the theory, Glashow, Salam, Weinberg

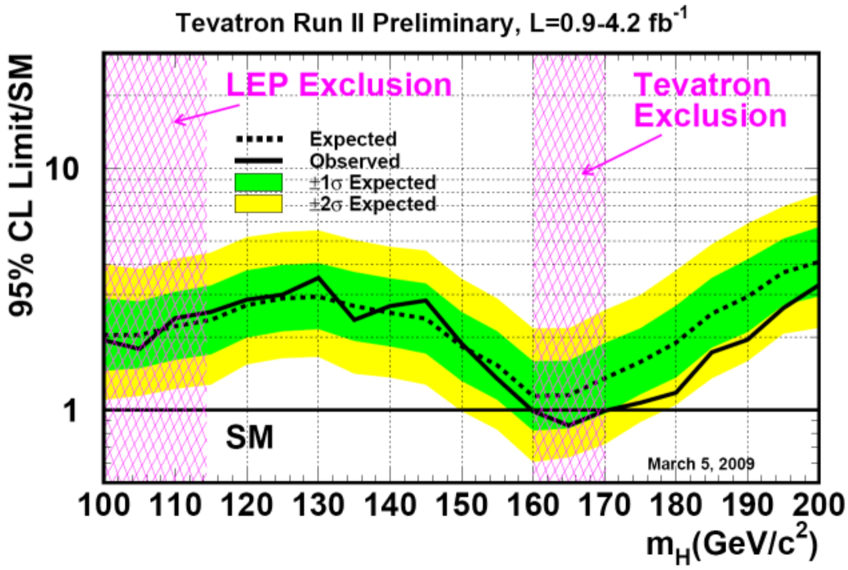


Figure 2.1: Observed and expected 95% C.L. upper limits on the ratios to the SM cross section, as functions of the Higgs boson mass for the combined CDF and DØ analyses. [7]

2.1.3 Higgs mechanism

In the Standard Model all particles are massless, but obviously in nature they are not. The particles must acquire a mass in some way that does not break the gauge symmetries [3]. It is assumed, that a suitable scalar "Higgs"-field exists, following ideas of Higgs and others [4]. This Higgs field has a potential which at the ground state breaks the symmetry spontaneously. The fermion masses are put in "by hand" [3] via Yukawa-like couplings to the Higgs field. Introducing a scalar field, composed as a complex isospin doublet of scalar fields, the symmetry is broken by giving a non zero vacuum expectation value to the neutral part of the doublet. Then it turns out that three parts of the scalar field are eliminated by the gauge bosons to acquire mass. The remaining neutral scalar field is physical and should be observed. This field can be identified with the Higgs boson [5, ch.7/17].

Up to now the Higgs boson was not discovered. But nevertheless, theoretical considerations concerning the Higgs mass can be given and lead to an upper bound of $m_H \approx 1$ TeV by unitarity arguments. Further constraints can be derived under the assumption that the Standard Model is valid only to a cutoff energy scale λ. Beyond this scale new physics is expected. For a cutoff scale of the order of the Planck mass[2], the Standard Model Higgs boson is required to be in the range of $130 < m_H < 180$ GeV. If new physics appear at $\lambda = 1$ TeV, the Higgs boson mass is expected to be $50 < m_H < 800$ GeV. The limits do not have the intention that there must be a higgs boson within these limits, but rather that perturbation theory will fail or new physics will enter [3], if a Higgs boson is founded beyond these limits. There are also limits from direct searches. A lower bound is given by the LEP experiments at CERN [6], excluding Higgs mass region below 114.4 GeV at 95% confidence level. Experiments at Tevatron exclude an additional region between 160 GeV and 170 GeV [7]. The observed and expected upper limits from the Tevatron experiments are shown in Figure 2.1.

If the Higgs boson is not of the simple form of a scalar explained above, another mechanism must be involved, providing a more complicated structure. Supersymmetry can be one of these extensions. In Supersymmetry all standard model particles have supersymmetric partners with

[2] $m_{\text{planck}} = 1.22 \times 10^{19}$ GeV

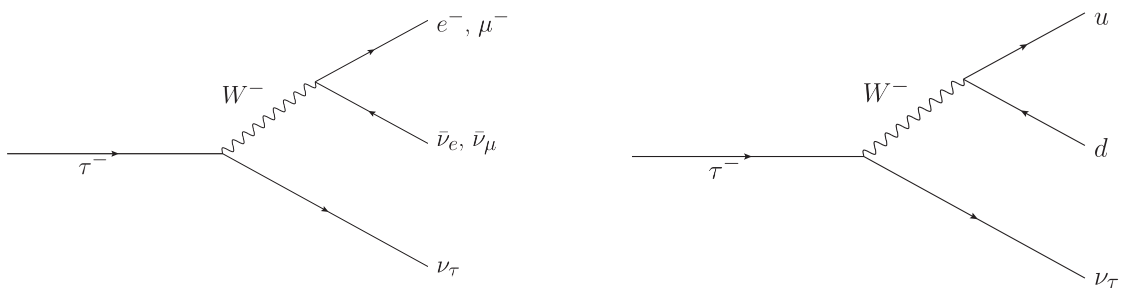


Figure 2.2: Feynman diagrams of leptonic (left) and hadronic (right) τ^- decays. The τ^+ decays are charged conjugated.

spin differing by $1/2$. The Minimal Supersymmetric Standard Model (MSSM) is the simplest supersymmetric extension of the Standard Model and introduces two Higgs doublets whose neutral components have two vacuum expectation values. This leads to five physical Higgs bosons: a charged boson pair ($H^\pm$), two neutral scalar bosons (h,H) and one pseudoscalar (A). The ratio of the vacuum expectation values of the Higgs doublets is called $\tan\beta$. Together with the mass of the pseudoscalar A, both parameters define the main properties of this model. In the MSSM the couplings of Higgs bosons to fermions and bosons are different from those in the Standard Model resulting in different production cross-sections and decay rates. For high values of $\tan\beta$ Supersymmetry favors[3] tau leptons in final states of Higgs boson decays.

2.2 Physics with Tau Leptons

2.2.1 The Tau Lepton

The tau (or τ) lepton is part of the third fermion generation and with a mass of $m_\tau = 1776.84 \pm 0.17$ MeV [2] it is the heaviest lepton. It is unstable having a lifetime of 290.6 fs, being $c\tau \approx 87\,\mu$m, which has to be multiplied by $\beta\gamma$ to obtain the actual decay length. A tau lepton with $E = 40$ GeV traverse about of 2 mm through the detector. The τ lepton is the only lepton heavy enough to decay leptonically and hadronically. In 35.2% of the time it decays leptonically and in 64.8% of the time into one or more hadrons. Figure 2.2 shows the Feynman diagrams of both decay modes. Considering only hadronically decaying τ leptons (τ_{had}), decays with only one charged particle (or 1-prong) occur in about 72% of the time and with three charged particles (or 3-prong) in about 23% of the time. The 5-prong decay has only a fraction of about 0.1%. Some important τ^- branching fractions are listed in Table 2.3. The τ^+ modes are charge conjugated. The hadronic decay products of τ leptons are dominated by $\pi^\pm$ and π^0 mesons in the final state, but there are also small fractions of decays containing $K^\pm$ and K^0 mesons. In decays of high-momentum tau leptons the τ flight direction is quite well reproduced by the direction of the leading particle (particle with highest momentum). A dedicated reconstruction and identification method of τ leptons is essential for hadronically decaying tau leptons and described in Chapter 4. The leptonic decay mode is measured as electrons and muons in the final state but hard to be separate from direct electron and muon production. Due to the large fraction of hadrons in τ_{had} final states it is a challenge to distinguish hadronic tau decays from hadronic jets, which are produced at the LHC in processes with very high cross sections (see Figure 3.1).

[3] in some Supersymmetry parameter models for high $\tan\beta$ the Next-to-LSP (Lightest Supersymmetric Particle) is the stau, which is the supersymmetric partner of the SM τ. The stau decays to τ-lepton + LSP.

τ^- decay mode	BR %	
$e^-\bar{\nu}_e\bar{\nu}_\tau$	17.85	leptonic
$\mu^-\bar{\nu}_\mu\bar{\nu}_\tau$	17.36	
$\pi^-\bar{\nu}_\tau$	10.91	hadronic
$\pi^-\pi^0\bar{\nu}_\tau$	25.52	1-prong
$\pi^-2\pi^0\bar{\nu}_\tau$	9.27	
$\pi^-3\pi^0\bar{\nu}_\tau$	1.04	
$K^-\bar{\nu}_\tau$ + Neutrals	1.57	
$\pi^-\pi^+\pi^-\bar{\nu}_\tau$	9.32	hadronic
$\pi^-\pi^+\pi^-\pi^0\bar{\nu}_\tau$	4.61	3-prong
$K^-\pi^+\pi^-\bar{\nu}_\tau$ + Neutrals	0.48	

Table 2.3: τ^- main decay modes [2]

2.2.2 Standard Model Processes with Tau Final States

The reconstruction and identification of tau leptons play an import role for measuring Standard Model processes [8]. The most important source for τ leptons is the $W \to \tau\nu$ process (Figure 2.3a) with a cross section of 1.7×10^4 pb at center-of-mass energy $\sqrt{s} = 14$ TeV. Having one tau lepton and a neutrino in the final state, this process requires a good tau identification and missing energy reconstruction due to the neutrino. Hence, the selection of this channel is difficult and will require a τ and missing transverse energy, $E_{\mathrm{T}}^{\mathrm{miss}}$, trigger. This decay channel can be used to study the leptonic branching ratio of the W boson and the cross section of W production.

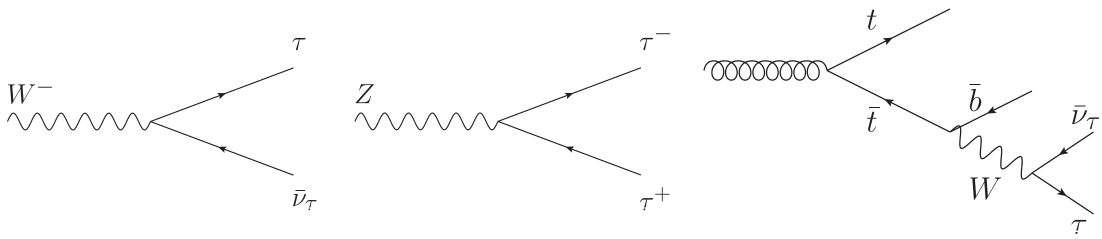
The $Z \to \tau\tau$ channel (Figure 2.3b) has a cross section of one order of magnitude less than $W \to \tau\nu$, but two tau leptons in the final state, with a invariant mass near the Z mass. Due to no additional missing energy this channel is called the “golden channel” for the detection of τ leptons. The case, one τ lepton decaying hadronically and the other leptonically gives the possibility to use the electron or muon trigger without to bias the selection of the event. This channel can be used to understand the efficiency of tau identification and τ-trigger as well as reconstruction methods for visible and invariant masses $m_{\tau\tau}$ of both tau leptons. The latter can be reconstructed using, for example, a collinear approximation method [8, p. 1280ff]. The dominant background for both processes is QCD dijet production, which has a cross section that is approximately six orders of magnitude higher than for $W \to \tau\nu$.

Top pair production, where one or both W bosons decaying leptonically, is also a process with tau leptons in the final state (Figure 2.3c). This process has more jets in the event, coming from hadronic W boson decays and b-quarks. This gives a different environment compared to $W \to \tau\nu$ and $Z \to \tau\tau$, but leads also to a more complex and difficult reconstruction. All Standard Model processes having tau leptons in final states can be considered as backgrounds for Higgs boson and new physics searches using tau leptons.

2.2.3 Standard Model Higgs Boson Searches with Tau Final States

The production cross section of the Standard Model Higgs boson at the LHC depends on the Higgs mass and the center-of-mass energy (Figure 2.4). The dominant process is the production via gluon fusion with a fraction of about 90%, followed by the Vector Boson Fusion (VBF, qqH), associated production with a Vector Boson (WH,ZH) and $t\bar{t}$-pair production.

The search for the Higgs Boson is performed [8] in various decay channels, depending on the Higgs mass. Figure 2.5 shows the branching ratios for the important decay modes of the SM Higgs boson as a function of its mass. The combined discovery significance for the Higgs boson is shown in Fig. 2.6.



(a) (b) (c)

Figure 2.3: Some Standard Model processes with tau leptons in final state

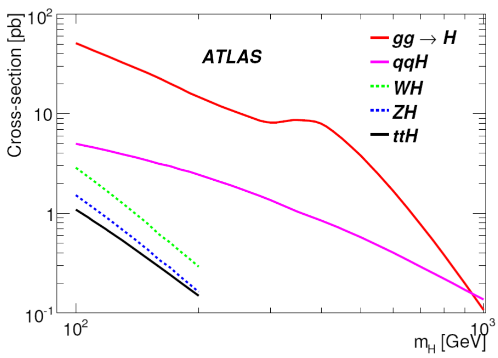


Figure 2.4: cross-sections for the production channels of SM Higgs boson at the LHC at 14 TeV [8]

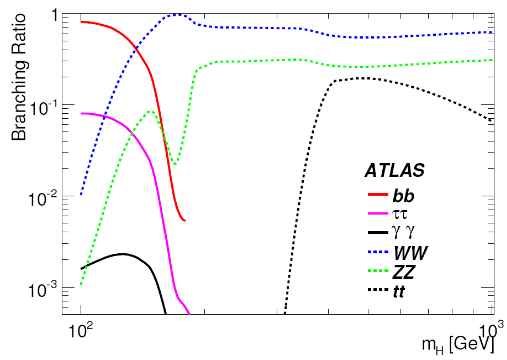


Figure 2.5: Branching ratios for important decay modes of the SM Higgs boson as function of its mass [8]

At low masses the decay of the Standard Model Higgs boson into a $\tau^+\tau^-$ pair plays an important role, representing about 10% of the branching ratio. The full-hadronic decay appears in about 12% of the time, the semi-leptonic in about 46% and the full-leptonic in about 42%. The Higgs boson mass can be reconstructed using a collinear approximation for the τ deacys, assuming that the decay products are emitted in the same direction as the τ lepton. Backgrounds for this process are reducible W+jets and $t\bar{t}$+jets events and irreducible $Z \to \tau\tau$ events. In case of this decay the Higgs VBF production channel gives a good detection signature with two high p_{T} forwards tagging jets and no jet activity between them (rapidity gap) in addition to the $H \to \tau\tau$ signature. In Figure 2.7 the Feynman diagram of this process is shown. In the low mass region this combination has the highest sensitivity of decting the Standard Model Higgs boson. Considering intermediate and higher mass regions, $H \to WW^{(*)} \to \ell\nu\ell\nu$ is the dominant decay. But because there are neutrinos in the final state, a direct reconstruction of the Higgs boson mass is not possible. The $H \to ZZ^{(*)} \to 4\ell$ decay has a smaller branching ratio than the WW channel, but leads to a clean experimental signature with four leptons in the final state. Therefore Higgs boson mass can be fully reconstructed. Both Z bosons are on-shell if $m_H > 2m_Z$. Studies for the final states $4e$, 4μ and $2e2\mu$ can be found in Reference [8, p. 1243], while for the $2\mu 2\tau$ and $2e2\tau$ final states studies have been recently started.

2.2.4 Searching for New Physics with Tau Final States

Large parts of the MSSM parameter space predict decays of the MSSM Higgs bosons to τ leptons. At high values of $\tan\beta$, associated neutral Higgs production with bottom quarks is dominant. Searches for light neutral Higgs bosons are based on the same channels considered for the

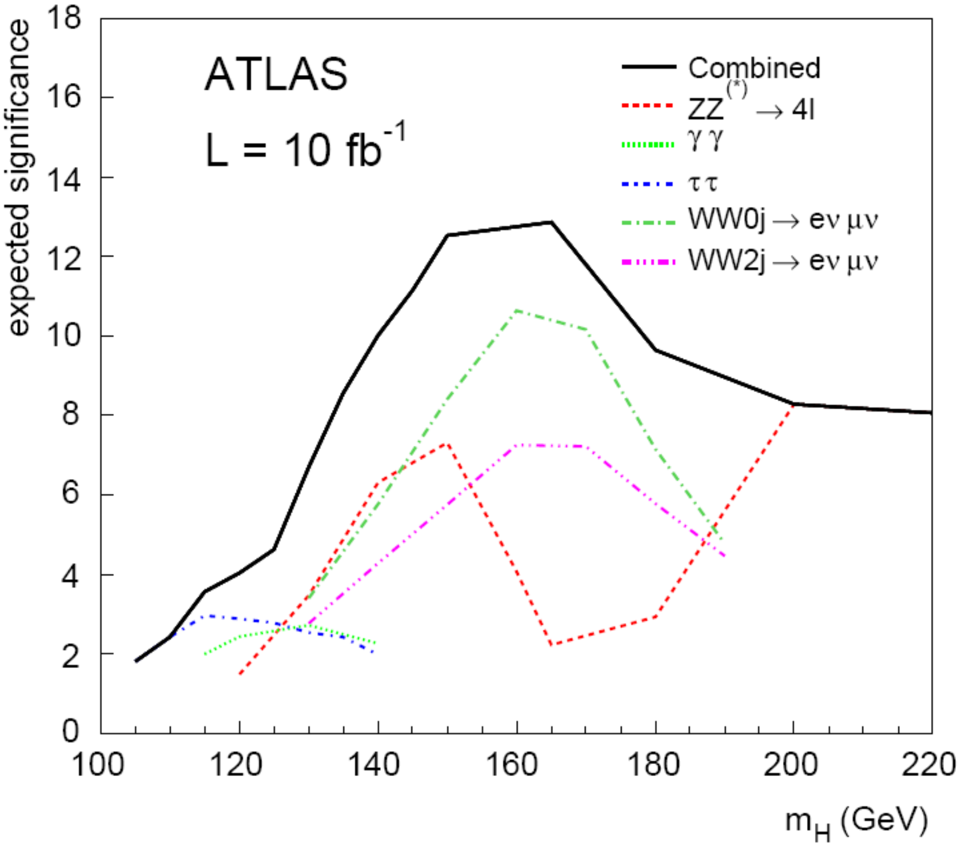


Figure 2.6: Higgs discovery significance of the ATLAS detector for different channels and the combination for an integrated luminosity of L=10 fb^{-1}[8].

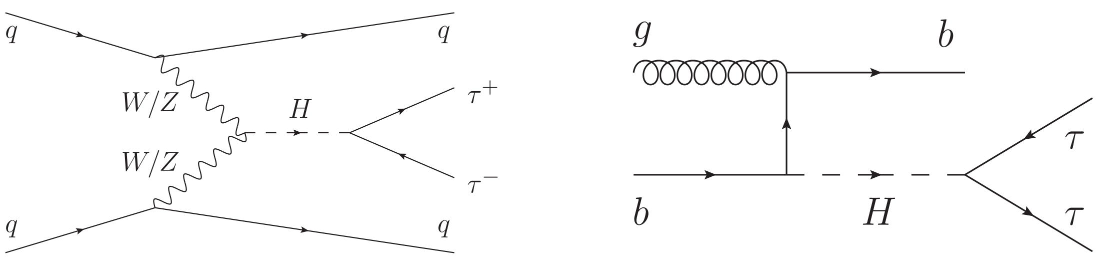


Figure 2.7: VBF production channel of SM Higgs boson with decay $H \to \tau\tau$ (left) and MSSM gb fusion of a heavy Higgs boson decaying into $\tau\nu$ (right)

SM Higgs boson. Additional decay channels become accessible in certain regions of the MSSM parameter space if searching for heavier Higgs bosons.

All three neutral Higgs bosons can decay into $\tau^+\tau^-$ pairs with a branching fraction of about 10 %. The decay modes $H \rightarrow \tau\tau$ and $A \rightarrow \tau\tau$ are important discovery channels for heavy Higgs bosons at the ATLAS detector [9]. The highest sensitivy is expected in the lepton-hadron final state of tau decays. Heavy Higgs bosons are produced via gluon-fusion or in association with b quarks. In Figure 2.7 the Feynman diagram of this process is shown.

The decay modes for charged Higgs bosons are depending on the Higgs mass. For a Higgs boson mass below the top-quark mass the main decay mode is into $\tau\nu$. In this mass region charged Higgs bosons are predominantly produced in top-quark decays. For heavy charged Higgs boson with $m_{H^\pm} > m_t$ the main production process will be gb-fusion. The dominant decay mode is to tb, but according to current expectations the discovery potential will be mainly for the $H^\pm \rightarrow \tau\nu$ decay channels [9].

There are also models which predict heavy resonances of W and Z bosons, hence referred as W' and Z', coupling preferentially to third generation of leptons. This gives also the motivation to search for processes like $Z' \rightarrow \tau\tau$ or $W' \rightarrow \tau\nu$, with large Z' and W' masses ($O(1\ \text{TeV})$).

Chapter 3

The ATLAS Detector at the Large Hadron Collider

3.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [10] is a particle accelerator at the European Laboratory for Particle Physics (CERN[1]) close to Geneva, Switzerland. The LHC is a proton-proton collider and was designed for an energy of 7 TeV per beam giving a nominal center-of-mass energy of 14 TeV. It is also possible to accelerate heavy ions instead of protons. Installed in the former LEP (Large Electron-Positron Collider) tunnel up to 100 m underground the LHC has a circumference of 27 km and was completed in 2008 after 7 years of construction. The design luminosity is 10^{34} cm^{-2} s^{-1}. Further machine design parameters are listed in Reference [11, S.291, Table 9.6].

To investigate particle collisions at the LHC, several detectors were built: ATLAS [11] and CMS [12] as detectors for multi purpose physics analysis, ALICE [13] for heavy ion collisions, LHCb [14] to investigate CP-violation and properties of the bottom quark and LHCf [15] and TOTEM [16] to study particle productions, elastic scatterings and total cross sections of pp collisions.

At high design luminosity and energy, bunches of up to 10^{11} protons will collide about 40 million times per second leading to a very high flux of particles and radiation doses. This opens a wide range of measurements in particle physics, like searching for new physics phenomena and precise measurements of the Standard Model. Figure 3.1 shows the predictions of cross sections of important Standard Model processes for the LHC ($\sqrt{s} = 14$ TeV) compared to the Tevatron Collider ($\sqrt{s} = 1.96$ TeV). Most of the interesting processes like Higgs boson production are dominated by quark and jet productions and are therefore difficult to detect.

The LHC started with first beam on 12th September 2008, but had to be stopped some days later due to an incident [17]. A faulty electrical connection between two magnets led to an electric arc and the release of some tons of liquid helium, which damaged several magnets of the LHC. After repair, the LHC restarted successfully with low beam energy in November 2009. For 2010 it is planned to run with 3.5 TeV per beam with options to ramp up to 5 TeV [18] leading to a peak luminosity of 10^{32} cm^{-2} s^{-1} and an integrated luminosity of 200–300 pb^{-1}.

[1] Conseil Européen pour la Recherche Nucléaire, (European Organization for Nuclear Physics)

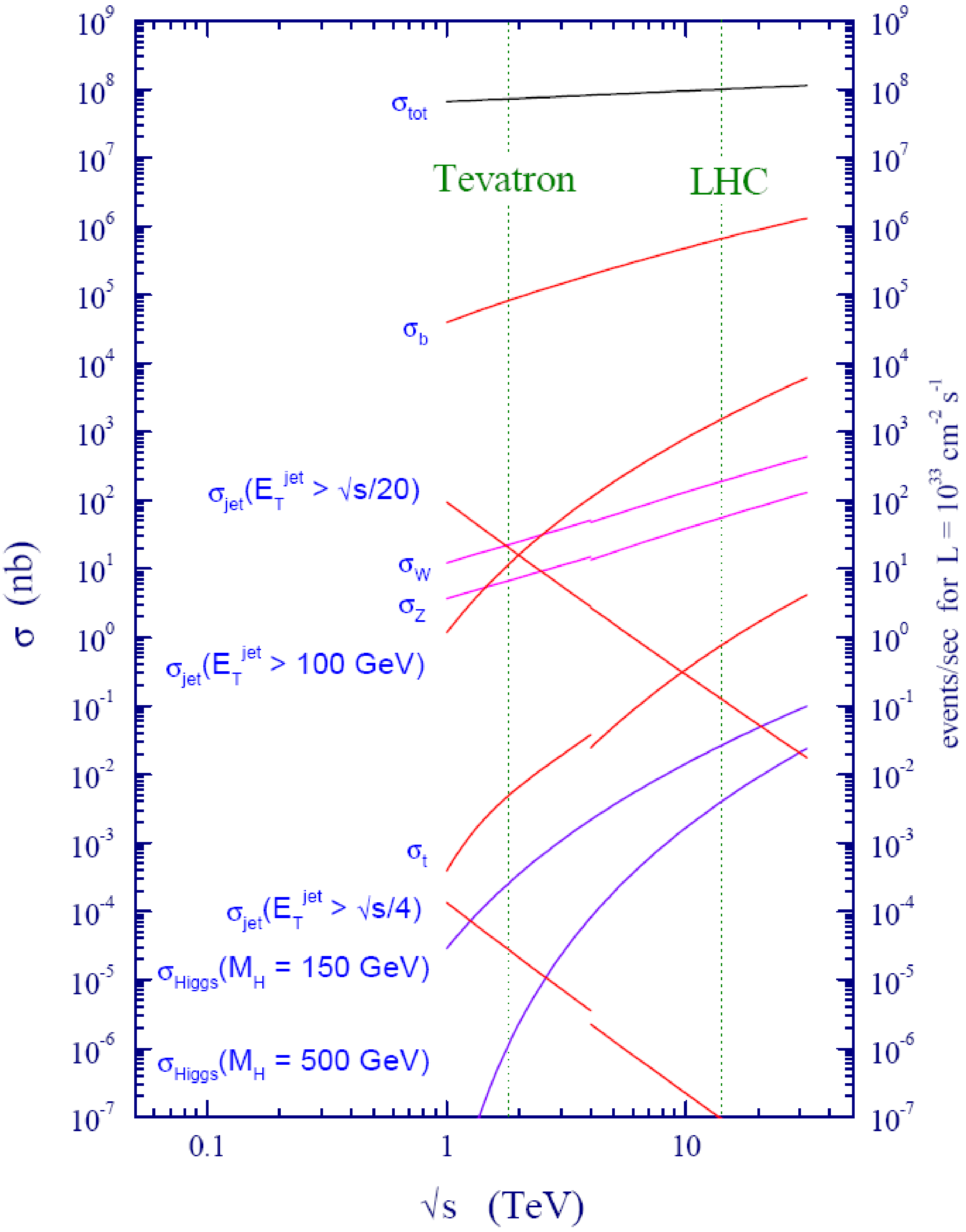


Figure 3.1: Standard Model cross sections at the LHC and Tevatron [19].
The discontinuities in the lines rely on changes from $p\bar{p}$ to pp colliders

3.2 The ATLAS Detector

This study was done within the ATLAS[2] collaboration. ATLAS [11] is one of the two detectors to study a wide range of physics at the TeV scale accessible at the LHC. The main tasks [8] are the search for new physics like the Higgs boson or Supersymmetry and the validation and the measurement of the parameters of the Standard Model. The dimensions of the detector are 25 m in height, 44 m in length and a overall weight of approximately 7000 metric tons. A layout structure with inner detectors, calorimeters and muon spectrometers is shown in Figure 3.2.

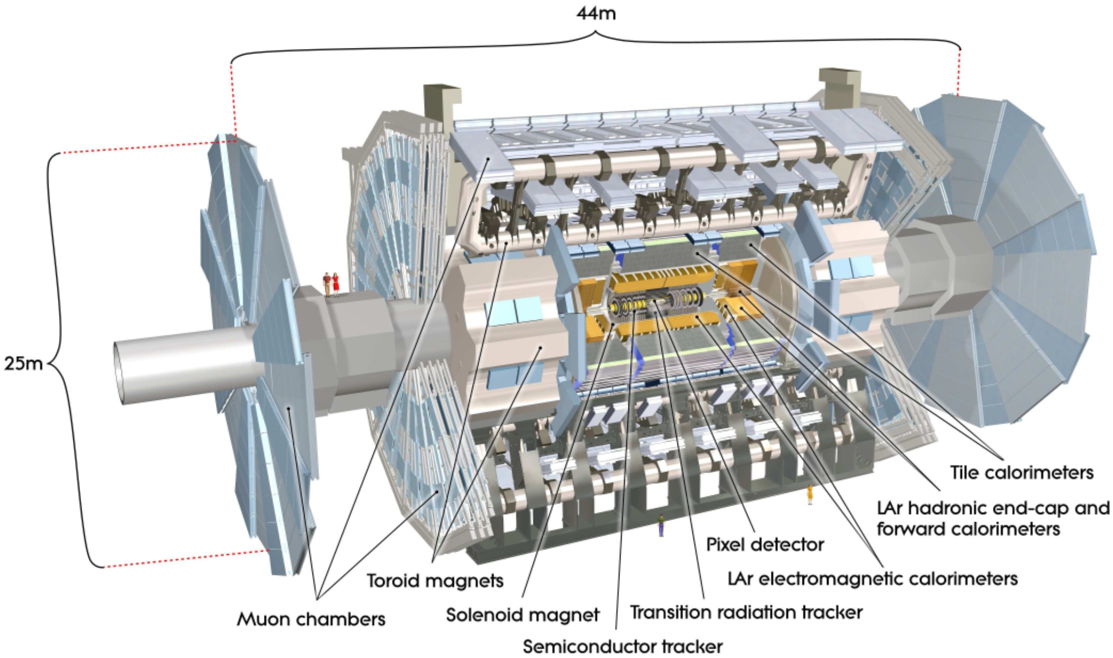


Figure 3.2: Overview of the ATLAS detector [11].

In ATLAS a specialized coordinate system is used, where the interaction point is the origin of the coordinate system. The z-axis is defined by the beam line, while the $x-y$ plane is transverse to the beam direction in the way that the x-axis point towards the center of the LHC ring and the y-axis upwards. The azimuthal angle ϕ is defined around beam axis and the polar angle θ is the angle from the beam axis. It can be replaced by the definition of pseudorapidity $\eta = -\ln(\tan(\theta/2))$. Particles traveling along the beam direction correspond to a infinite pseudorapidity, while particles produced perpendicular to the beam have a pseudorapidity of zero. The geometrical distance of two objects in the pseudorapidity-azimuthal plane is called ΔR and defined as $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. Transverse variables like the transverse momentum p_T are defined in the $x-y$ plane as $p_T = \sqrt{p_x^2 + p_y^2}$.

The ATLAS detector consists of many sub-detectors in a shell-like construction around the beam line and has a forward-backwards symmetry. The inner part closest to the interaction point is a central tracking detector in a solenoidal magnetic field. It is surrounded by self-contained electromagnetic and hadronic calorimeters and a muon spectrometer. A superconducting toroid system forms the outer part of the detector. The whole detector including all subsystems are divided in three parts, one barrel part and two end-cap parts. The energy resolution and the coverage in η for the sub-detectors can be found in Figure 3.3.

[2] A Toroidal LHC ApparatuS

Detector component	Required resolution	η coverage	
		Measurement	Trigger
Tracking	$\sigma_{p_T}/p_T = 0.05\%\ p_T \oplus 1\%$	± 2.5	
EM calorimetry	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	± 3.2	± 2.5
Hadronic calorimetry (jets)			
barrel and end-cap	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	± 3.2	± 3.2
forward	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$	$3.1 < \lvert\eta\rvert < 4.9$
Muon spectrometer	σ_{p_T}/p_T=10% at p_T = 1 TeV	± 2.7	± 2.4

Figure 3.3: Energy resolution and coverage in η for the ATLAS detector [11]. The units for E and p_{T} are in GeV.

Inner Detector

The Inner Detector (ID) is a tracking detector designed to reconstruct tracks of charged particles. It is designed to give a good resolution of momentum measurements and particle reconstruction efficiencies. For tagging tau leptons and b-quark decays the observation of secondary vertices is necessary. The Inner Detector is contained within a solenoidal magnetic field of 2 Tesla covering a pseudorapidity of $|\eta| < 2.5$. It consists of 3 independent subsystems. An outline of the ID is shown in Figure 3.4. The Inner Detector is exposed to a high particle flux density, because of its position closest to the beam line and interaction point. Thus, a high granularity and a fast readout system is required.

At inner radii a **pixel detector** is used providing approximately 80 million readout channels. The pixel detector has a very high degree of segmentation and use silicon diodes as detection technology. This is necessary to cope with the high track density expected and to find short lived particles like τ leptons. The pixels are arranged in three layers with an intrinsic accuracy of 10 µm in the Rϕ direction and 115 µm in the beam direction. The innermost layer, called B-layer, improves the performance of the secondary vertex measurements, necessary for τ_{had} reconstruction.

The next sub-detector is the **Semiconductor Tracker** (SCT) using silicon strips. The barrel layer covers $|\eta| < 1.4$ and the end-caps $|\eta| < 2.5$. The SCT provides around 6.3 million readout channels in a more granular segmentation than the pixel detector and provides up to 8 hits per track. The resolution is 17 µm (Rϕ) and 580 µm (z).

The third part of the Inner Detector is the **Transition Radiation Tracker** (TRT), which is composed of gaseous straw tubes. In the barrel part the straws are arranged parallel to the beam axis, while in the end-caps a radial arrangement is used, giving 351000 readout channels. The TRT contributes only information from the Rϕ plane with an resolution of 130 µm per straw. With around 36 hits per track, the TRT provides a continuous tracking and more space point measurements for a track at a less precision than the pixel and SCT detectors. This leads to an improvement of the momentum resolution at small pseudorapidities, $|\eta| < 2.0$. The TRT is not only designed for tracking measurements, but also for simple particle identification. The Transition Radiation, which occurs when a charged particle with high velocity traverse two media with different dielectric constants, is also used to discriminte between electrons and pions.

Calorimeter

The adjacent part are the calorimeters, designed to detect particles with electromagnetic and hadronic interaction in a range of $|\eta| < 4.9$. The calorimeters provide the measurement of energies of electrons, photons and jets and the identification of the particle species. The granularity

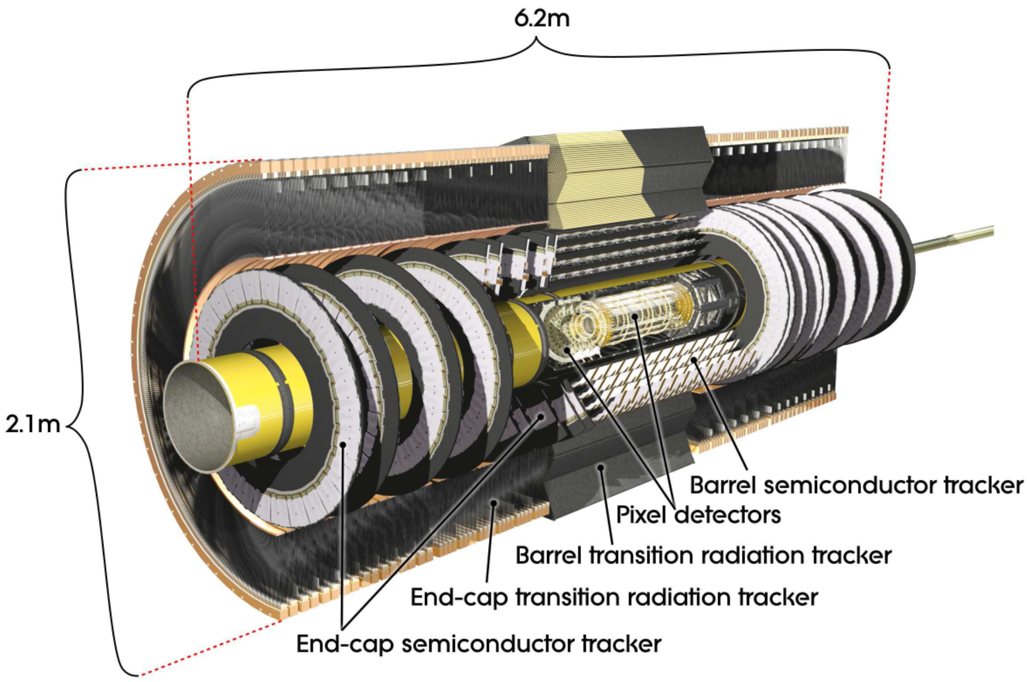


Figure 3.4: Overview of the inner detector [11].

of the calorimeter varies from a fine grained structure at the region, which overlap with the ID, and a more coarser structure at the rest. Due to high homogeneity and a wide range of acceptance the calorimeters allow to reconstruct missing transverse energy. The main parameters of the calorimeter system are listed in Reference [11, Table 1.3]. A view of the three parts of the calorimeters is shown in Figure 3.5. They are described in the following.

The **Electromagnetic Calorimeter** (EM) measures electrons and photons. The EM is a sampling calorimeter and uses liquid argon as active material and lead as absorber. Particles traveling through the high-density absorber initiate a shower cascade, depositing all energy in the calorimeter by ionization of the active material. The EM is longitudinally divided into three samplings with different granularity in η and ϕ, consisting of a barrel part ($|\eta| < 1.475$) and two end-cap parts ($1.375 < |\eta| < 3.2$). In regions where many support and structure material is in front of the calorimeter, a presampler is added to correct for the energy loss in this regions. The segmentation of the EM can be found in Figure 3.7.

The **Hadronic Calorimeter** measures strongly interacting particles, forming jets. It consists of a scintillator tile calorimeter (TileCal) in the barrel part, which covers $|\eta| < 1.7$. The TileCal use iron as absorber and plastic scintillator plates as active material. It provides a good resolution and prevents also leaks of non-muon particles into the muon system. The end-cap calorimeters ($1.5 < |\eta| < 4.9$) consists of the Hadronic End-Cap Calorimeter (HEC) and a high-density Forward Calorimeter (FCal). Due to the high radiation density in the end-cap and forward region a radiation-hard material is necessary. Therefore a copper-liquid argon mix is used in the HEC and in the first part of the FCal, while tungsten-liquid argon is used in the second and third part of FCal. Due to the large coverage in η, a good missing transverse energy measurement is possible, which is necessary for many processes involving τ leptons.

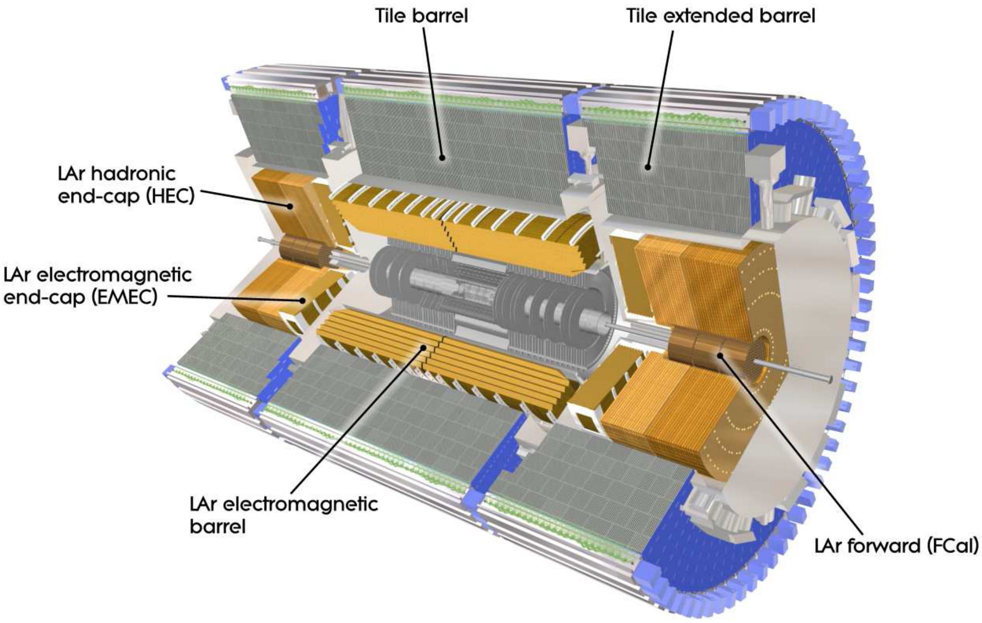


Figure 3.5: Overview of the calorimeters [11]

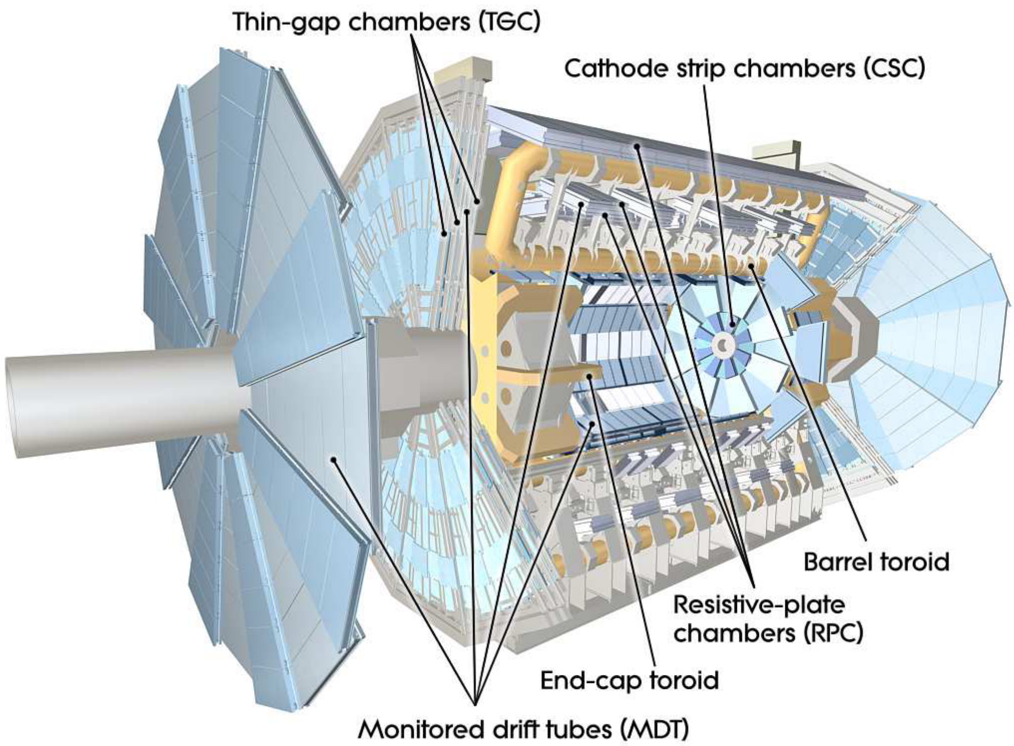


Figure 3.6: Overview of the muon spectrometer [11]

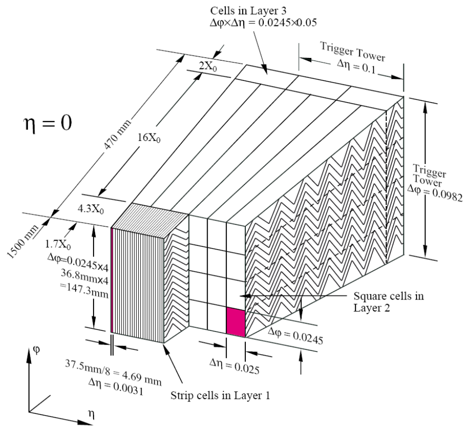


Figure 3.7: Structure of the LAr EM calorimeter in the barrel region [11]

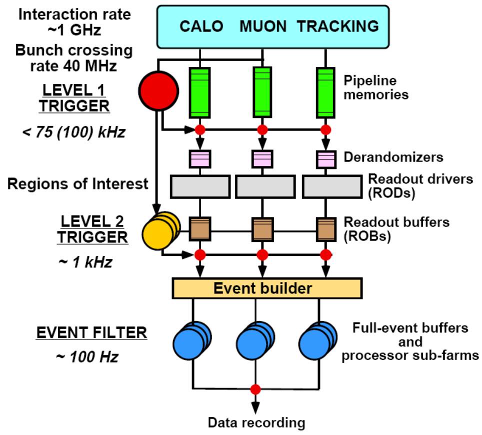


Figure 3.8: Overview of the trigger system [20]

Muon Spectrometer

The Muon Spectrometer is designed for high precision measurements and identification of muons. Because of their higher mass in comparison to an electron, muons do not initiate a particle shower in the electromagnetic calorimeter. The Muon Spectrometer has its own magnetic field of 0.5–1 Tesla, providing measurements of the muon transverse momentum independently of the inner detector. The Muon Spectrometer uses four different chamber systems: Monitored Drift Tubes (MDT) and Cathode Strip Chambers (CSC) designed for measurements of track coordinates, Resistive Plate Cambers (RPC) and Thin Gap Chambers (TGC) having fast drift times and are used for triggering. A layout structure of the muon spectrometer is shown in Figure 3.6.

Trigger

To handle the high interaction rates of up to 10^9 Hz at the LHC a efficient trigger system is essential. To storage interesting physics events the rate must be reduced to ~ 200 Hz, leading to trigger only on new physics and important standard processes. The trigger system is arranged in a chain of three steps, each depends on decisions made before and applying additional criteria. The first step is called Level 1 (L1) and uses limited detector information from the muon chambers and the calorimeter to define a Region-of-Interest (RoI). L1 is realized in hardware and reduce the event rate to ~ 100 kHz. The RoI are used as seed for the Level 2 Trigger (L2), which has full access to calorimeter information and lowers the event rate to 3.5 kHz. The reduction to 200 Hz is realized by the Event Filter (EF), which is an offline analysis step, due to the high decision time of ~ 4 s. The Event Filter has access to informations about magnetic field, alignment and calibration constants. After this step the event data will be written to mass storage for later analysis. A processing scheme is shown in Figure 3.8. Even though using this trigger chain, some trigger tend to have a too high event rate, due to high particle flux mainly coming from QCD interactions. In this case only a fraction of events triggered will be accepted. The inverse of the fraction is called pre-scale factor.

3.3 Simulation and Reconstruction of Physics Events

To understand the real data seen by the detector and to compare this with theoretical predictions, the simulation of particle interactions is necessary. The whole process of data generating and reconstruction is done in ATHENA, the ATLAS offline software framework. The full event simulation is organized in different steps:

Event Generation: Monte Carlo (MC) techniques are used to simulate the probability statistics of particle interactions. In this study datasets are used generated with PYTHIA, a multipurpose Monte Carlo generator, which is commonly used for the simulation of hadronic processes. PYTHIA [21] is a leading-order generator, calculating all events in leading and lowest order. Higher corrections are realized at parton shower level. Other Monte Carlo generators using Feynman rules to calculate the matrix elements of the particle interactions and match these with the parton showers.

Simulation: The generated final state particles are passed through a detector simulation, to track the way of particles through the detector and simulate the interaction of the particles with the detector material and the magnetic fields.

Digitization: This part provides the response of the electronic readout system on the particle interactions. The output is digitized in a way, that it looks like real data.

Reconstruction: The particle tracks and calorimeter clusters are reconstructed to create physics objects, such as electrons, tau leptons or jets. While a track is reconstructed by hits in the inner detector, a cell is the smallest reconstructible object in the ATLAS calorimeter and provides mainly the reconstructed raw energy. A topological cluster (*TopoCluster*) is formed by a group of cells around a seed cell, above some noise energy thresholds. As described in Chapter 2, quarks and gluons can radiate other partons and due to the color confinement they bind into hadrons. As in QCD, the radiation is in general in the flight direction and therefore the hadrons have the same direction of the initial partons. A collimated *jet* of these hadrons can be seen in the detector. There are different jet algorithms used to identify these jets. In this study *Cone4TopoJets* are used, which are based on a clustering algorithm. The algorithm tries to group particles, that are close together in a cone of $\Delta R < 0.4$ [11, ch. 10].

The results of the full simulation are stored in various formats, depending on which information should be to include. In this study the Analysis Object Data (AOD) format is used, that provides sufficient information of the reconstructed objects for this analysis.

Chapter 4

The Reconstruction of Hadronically Decaying Tau Leptons at ATLAS

4.1 Tau Reconstruction Process

The reconstruction of τ leptons in ATLAS is a challenge, because it is hard to distinguish τ leptons from hadronic jets, which are produced with very large cross sections at the LHC. Due to the difficulty of discriminating between leptonic decay modes from primary electrons and muons, the reconstruction of τ leptons is understood as a reconstruction of hadronic decay modes. Hadronically decaying τ leptons have a well known signature. They decay in 72% of the cases with one charged particle (1-prong) and to 23% with three charged particles (3-prong). This leads to a low track mulitplicity. The branching ratio for the 5-prong decay is 0.1%, but can be detected hardly in an environment with many hadronic jets, like the LHC. The decay products are well collimated and the invariant mass is smaller than the τ lepton mass, due to the not detectable neutrino. With a significant life time $(ct_{\tau} \approx 87\,\mu\text{m})$ the decay vertex can be resolved from the primary interaction vertex in the silicon tracker. A hadronically decaying tau lepton deposits an extensive amount of energy in the calorimeter.

The Tau Reconstruction Algorithm

The τ reconstruction algorithm implemented in ATLAS is a higher level reconstruction method. It uses components already provided by other algorithms specific to different subdetectors, such as topological clustering of the energy deposits in the calorimeter or track reconstruction in the Inner Detector. The τ reconstruction algorithm uses two types of methods to find a τ candidate. A tracking-based (or track-based) method, which uses good quality tracks as seed and a calorimetry-based (or calo-based) method, which runs on clusters reconstructed in the hadronic and electromagnetic calorimeter. At the first step, the two seeds – tracks and clusters – will be identified.

To reconstruct a track-seed candidate, a track with transverse momentum of $p_T > 6\,\text{GeV}$ and satisfied quality criteria is required and will be called a seed track. The quality criteria are the transverse impact parameter with respect to the interaction vertex, a chi-square of the track fit, the number of hits in the Transition Radiation Tracker (TRT), the ratio of high and low threshold TRT-hits and the number of hits in the silicon detector, in the pixel detector and in the B–layer of the pixel detector. Then the algorithm looks for other tracks in a cone of $\Delta R < 0.2$ around the seed track passing particular quality criteria. If a second track is found, the algorithm looks for a third track with loose requirements. All quality criteria are listed in Table 4.1. Tau candidates with more than eight tracks found are not treated as a valid track-based candidate.

Track Criteria		Seed Track	Assoc. Track	3rd Track	Loose-Track
p_T in GeV	$>$	6	1	1	1
$\|\eta\|$	$<$	2.5	2.5	2.5	2.5
Impact parameter d_0 in mm	$<$	1	1	1	1
Silicon hits	$\geq$	8	8	8	6
TRT hits	$\geq$	10	no cut	no cut	no cut
normalized χ^2	$<$	1.7	1.7	no cut	1.7
Pixel hits	$\geq$	no cut	1	1	1
B layer hits	$\geq$	no cut	1	0	1
High/low threshold hits ratio	$<$	no cut	0.2	no cut	no cut

Table 4.1: Track quality criteria for seed tracks and default associated tracks (track-based candidates) and loose tracks (calo-based candidates) [22]

For a reconstruction of a calo-seeded candidate, a TopoJet with $E_T > 10$ GeVand $|\eta| < 2.5$ is required. The algorithm associates tracks within a cone radius of ΔR$<$ 0.3 around the TopoJet center passing quality criteria listed in Table 4.1 to the calo-seeded candidate. The track is referred as loose track. Concerning both methods, a reconstructed tau candidate with n-prongs means that there are n tracks associated to it.

After identification of the τ candidates, track-seeds and calo-seeds will be matched using a cone radius of ΔR$<$ 0.2 around the TopoJet center. In case of matching only one τ candidate will be build, having two valid seeds. Regarding reconstructed τ leptons from $Z \to \tau\tau$ events with $E_{\mathrm{T}} > 10$ GeV and $|\eta| < 2.5$, about 70% of τ candidates have two valid seeds, 25% have only a calorimeter seed, and 5% have only a track seed [23].

The energy of τ candidates with a calo-seed is calculated using the H1 weighting method [8, p.242 and 298ff] by summing over cells associated to a TopoJet within ΔR$<$ 0.4. The cells are weighted according to a function dependent on η, ϕ and the calorimeter layer of the cell. Since this method was originally developed for QCD jets it overestimates the energy and a τ specific calibration has to be applied to correct for this effect. An energy flow algorithm is used for candidates having a track seed. The approach based on dividing the energy deposits in cells to pure, charged and neutral electromagnetic energy as well as charged hadronic energy, replacing the charged energies by track momenta and include contribution from π^0 mesons. Also a correction for energy leakage coming from charged particles outside the narrow cone is applied. The algorithm is described in detail in Reference [8, p. 246]. For candidates with two valid seeds, the cell-weighting method is considered as the default energy calibration.

In case of track-seeded tau candidates the η and ϕ direction is taken from the direction of the track at the primary vertex in case of 1-prong candidates and from the p_T-weighted track barycenter in case of multi-prong candidates. This direction is used to calculate the four-momentum vector for candidates with track-seeds and both seeds. For candidates with a calo-seed, η and ϕ are calculated based on E_T-weighted barycenter of calorimeter cells. In case of tau candidates with only a calo-seed it is also used for building the four-momentum vector. The charge of the τ lepton is determined by the sum of the charges of the reconstructed tracks. The charge misidentification of ΔQ= 1 is about $2-4$%, mainly coming from single-prong decays, which migrate to three-prong decays due to tracks from photon conversion or three-prong decays reconstructed as a single-prong decay.

To include also contributions from π^0 mesons in τ lepton decays, TopoClusters found in the electromagnetic calorimeter with $E_{\mathrm{T}} > 1$ GeV and isolated from tracks, are interpreted as energy deposits from π^0 mesons. This methods finds 50% of the π^0 clusters in tau decays with one or two π^0 mesons, while approximately 66% of $\tau \to \pi\nu$ decays are reconstructed correctly with zero π^0

clusters [22].

4.2 Tau Identification Process

The τ reconstruction algorithm tries to find many candidates to be suitable for a τ lepton, but do not distinguish between real tau leptons and tau candidates originate from hadronic jets, electrons or muons. Therefore different τ identification algorithm are implemented in ATLAS.

To separate tau leptons and electrons, tracks are divided into categories [22]. This categories are based on informations from the energy deposit in the hadronic calorimeter and from energy that not associated with a charged track in the strip part of the electromagnetic calorimeter. On these tracks, fixed cuts are placed on the ratio of E_T from the electromagnetic calorimeter to the track p_T and the ratio of high threshold to low threshold hits in the TRT. Both ratios tends to be higher for electrons than for charged hadrons. This method is used for $|\eta| \leq 1.7$ and ${p_\mathrm{T}}^{trk} \leq 50$ GeV, where the TRT provides good discrimination between electrons and hadrons. Outside this region, a simplified method based only on energy from the hadronic calorimeter and $E_\mathrm{T}/p_\mathrm{T}$-ratio is used. This procedure suppresses electrons by a factor of 60, while retaining 95% of hadronically decaying tau leptons. To veto against muons a cut on the τ candidate energy deposit in the calorimeter with $E_\mathrm{T} > 5$ GeV can be applied.

Because of the overwhelming numbers of hadronic jets coming from QCD interactions the separation of τ candidates coming from true τ leptons and candidates originate from hadronic jets is more difficult. For this case rectangular cut-based approaches are implemented, using information from track and calorimeter seeds. These methods are suitable for an integrated luminosity of $\sim 10-200\,\mathrm{pb}^{-1}$, which covers the early data taking period. It is expected that the calorimeter is one of the detector components that is well understood and easy to calibrate at the beginning phase. Therefore the cut-based methods are based mainly on variables from the calorimeter. These so called “safe variables” avoid informations from τ specific energy calibrations, energy flow calculations, π^0 cluster reconstructions or precision tracking informations. Currently there are two cut-based approaches: the calorimeter-based approach use only variables derived from calorimeter deposits. The calorimeter+track-based methods extends the former by including information from tracking. The intention for two methods is to avoid to use informations from the tracking detector, in case this component is not well aligned or understood. Both methods have three threshold levels, called loose, medium and tight, which correspond to reconstruction efficiencies of $\sim$30%, $\sim$50% and $\sim$70%. The reconstruction efficiency is defined as the ratio of numbers of reconstructed τ candidates with correct track multiplicity to numbers of true hadronic decaying τ leptons, with cuts $E_\mathrm{T} > 10$ GeV and $|\eta| < 2.5$ applied. The variables, where the approaches are based on, are separately optimized for the three thresholds, single-prong (including zero-prong) and multi-prong (more than one prong) as well as several E_T bins. The following five E_T bins are defined: $10-25$ GeV, $25-45$ GeV, $45-70$ GeV, $70-100$ GeV and > 100 GeV.

The Calorimeter-Based Approach

This method can be used for calo-seeded τ candidates and is based on four calorimeter variables, following listed:

- **The Electromagnetic Radius:**

$$R_\mathrm{EM} = \frac{\sum_{i=1}^{\Delta R<0.4} E_{\mathrm{T},i}\sqrt{(\eta_i - \eta_\mathrm{caloseed})^2 + (\phi_i - \phi_\mathrm{caloseed})^2}}{\sum_{i=1}^{\Delta R<0.4} E_{\mathrm{T},i}} \tag{4.1}$$

The electromagnetic radius represents the transverse shower profile and is defined in Eq. 4.1. i runs over all cells in the electromagnetic calorimeter that are associated to the τ candidate within a cone radius of ΔR$<$ 0.4 around the seed cluster. The quantities η_i, ϕ_i and $E_{\mathrm{T},i}$ denotes the position and energy of the cell i, while η_{caloseed} and ϕ_{caloseed} are the coordinates for the τ candidate calorimeter seed. The shower profile for τ decays is narrower and smaller than for hadronic jets.

- **Transverse Energy Width in the η Strip Layer:**

$$W_{\mathrm{strip}} = \sqrt{\frac{\sum_{i=1}^{\Delta R<0.4} E_{\mathrm{T},i}^{\mathrm{strip}} (\eta_i - \eta_{\mathrm{caloseed}})^2}{\sum_{i=1}^{\Delta R<0.4} E_{\mathrm{T},i}^{\mathrm{strip}}}} \tag{4.2}$$

where i runs over all cells in the electromagnetic calorimeter that are associated to the τ candidate within a cone radius of ΔR$<$ 0.4 around the seed cluster. $E_{\mathrm{T},i}^{\mathrm{strip}}$ is the transverse energy in the cell of the strip layer. Due to highly collimated decay products it is expected that entries from true τ decays leads to a maximum near zero, while hadronic jets have a higher W_{strip}.

- **Isolation in the Calorimeter:**

$$f_{\mathrm{iso}} = \frac{\sum_{i}^{0.1<\Delta R<0.2} E_{\mathrm{T},i}}{\sum_{j}^{\Delta R<0.4} E_{\mathrm{T},j}} \tag{4.3}$$

Most of the energy from the collimated τ decay products is collected in a small cone around the calorimeter seed and therefore lower the fraction of the energy in an outer cone ring. In Eq. 4.3 i runs over EM calorimeter cells associated with τ candidate in a cone ring around the seed with $0.1 < \Delta$R$<$ 0.2 and j runs over cells in a cone with ΔR$<$ 0.4.

- **Ratio of EM Energy and Total Energy:**

$$\frac{E_{\mathrm{T}}^{EM}}{E_{\mathrm{T}}^{total}} = \frac{\sum_{i}^{\Delta R<0.4} E_{\mathrm{T},i}^{EM}}{\sum_{i}^{\Delta R<0.4} E_{\mathrm{T},i}^{EM} + \sum_{j}^{\Delta R<0.4} E_{\mathrm{T},j}^{Had}} \tag{4.4}$$

Most of the energy of a τ lepton is deposit in the electromagnetic part of the calorimeter while QCD jets have a broader distribution and leave more energy in the hadronic part. In Eq. 4.4 the sums run over all cells associated to τ candidate within a cone of ΔR$<$ 0.4 with respect to the seed. Both, electromagnetic energy and hadronic energy are taken after H1 style calibration.

The cuts on these variables are listed in Appendix A. The distribution of all four variables is shown in Figure 4.1 separately for 1-prong and 3-prong τ candidates for a signal and a background sample. The efficiency for τ_{had} and the rejection rates against hadronic jets are shown in Figure 4.2. If a τ candidate with a calorimeter seed passes the cuts placed on these variables, it is classified as a tau having a safe-calo tau id.

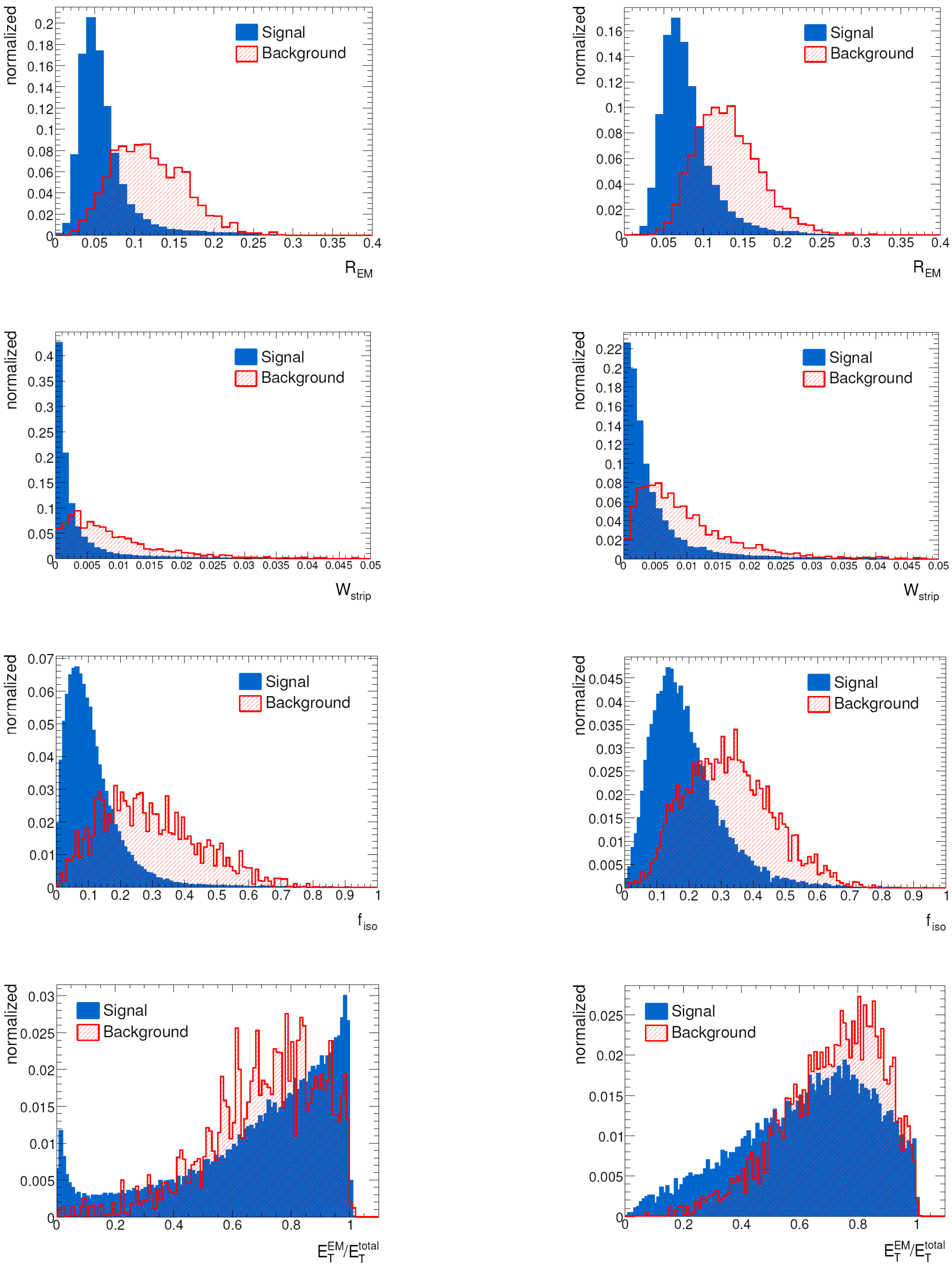


Figure 4.1: Distribution of calorimeter variables for 1-prong (left) and 3-prong (right) τ candidates within a $E_{\mathrm{T}}^{\mathrm{vis}}$ range of 25–45 GeV. Signal is from $Z \to \tau\tau$ and $A \to \tau\tau$ events matched to truth hadronically decaying tau leptons. Background is from QCD dijets events. [23]

		1-prong		3-prong	
Selection	E_T-range (GeV)	efficiency	rejection	efficiency	rejection
Loose	10-25	0.764 ± 0.002	17 ± 0.1	0.610 ± 0.003	4.3 ± 0.2
	25-45	0.773 ± 0.002	32 ± 0.2	0.746 ± 0.003	4.7 ± 0.4
	45-70	0.780 ± 0.003	45 ± 0.02	0.775 ± 0.005	7.8 ± 0.04
	70-100	0.849 ± 0.003	32 ± 0.005	0.867 ± 0.007	4.6 ± 0.01
	>100	0.937 ± 0.001	43 ± 0.007	0.701 ± 0.003	14 ± 0.01
Medium	10-25	0.515 ± 0.002	71 ± 0.06	0.553 ± 0.003	9.8 ± 0.2
	25-45	0.550 ± 0.004	130 ± 0.09	0.550 ± 0.002	28 ± 0.2
	45-70	0.582 ± 0.003	190 ± 0.01	0.597 ± 0.005	31 ± 0.02
	70-100	0.593 ± 0.004	290 ± 0.002	0.592 ± 0.008	36 ± 0.005
	>100	0.630 ± 0.002	780 ± 0.002	0.665 ± 0.003	22 ± 0.01
Tight	10-25	0.315 ± 0.001	190 ± 0.02	0.318 ± 0.002	62 ± 0.06
	25-45	0.351 ± 0.001	290 ± 0.06	0.339 ± 0.002	110 ± 0.09
	45-70	0.327 ± 0.002	600 ± 0.005	0.337 ± 0.004	140 ± 0.01
	70-100	0.368 ± 0.003	700 ± 0.001	0.401 ± 0.006	120 ± 0.003
	>100	0.369 ± 0.001	2300 ± 0.001	0.355 ± 0.002	320 ± 0.003

Figure 4.2: Efficiency for hadronically decaying τ leptons and rejection rates against hadronic jets for cut-based methods. Uncertainties shown are due to Monte Carlo statistics. (from [23])

The Calorimeter + Track-Based Approach

This approach can be used for τ candidates having both seeds. It combines the four variables from the calo-based methods and five variables coming from tracking. Providing more information it is expected to have a better performance with this approach than only using calorimeter informations. The additional variables are: the ratio E_T over p_T of the leading track, the ratio of EM transverse energy to sum of p_T of the three leading tracks, the same ratio but using the hadronic transverse energy, the ratio of the sum of p_T of the three leading tracks to total calorimeter cell energy and the width of the track momenta. Because this thesis includes only the calo-based approach, the variables are not further explained here. If a τ candidate passes the cuts that are placed on these variables, it is classified as a tau candidate having a safe tau id. Detail descriptions and variables distributions can be found in Reference [23].

Projective Likelihood

Beside the rectangular cut-based methods also more complex algorithms are implemented. One of these is a logarithmetic likelihood function, which uses information from calorimeter as well as the tracker. Separate probability distributions functions (pdfs) have been used for single and multi prong events made in p_T bin separations at 0, 10, 20, 30, 45, 70, 100, 150, 220, 500, 999 GeVand only one η bin. The likelihood method needs well understood variables and sufficient data, therefore the cut-based methods are preferably used. In comparison to the projective likelihood method, the cut-based approaches perform quite well [23, Sec. 4].

4.3 Methods for Fake Rate Estimations

Because tau leptons play an important role for detecting processes involving Higgs bosons and new physics, the measurement of the correct number of tau leptons in final states is essential. Tau candidates not being real tau leptons are called fake taus and have different sources. Since the expected rate of QCD jets exceeds the rate of hadronically decaying tau leptons (τ_{had}), and since the fake rates are expected to be in the range of 10^{-3} to few 10^{-2} for low p_{T} tau leptons [8, p. 251ff], precise measurements of the tau fake rate caused by QCD jets are necessary.

To estimate the τ_{had} fake rate various methods are studied. Besides the large amount of dijets, studies are performed to estimate the fake rate directly using QCD dijets. The method presented in Reference [8, p. 251ff] uses a very clean sample of QCD jets with no significant contamination from true tau leptons. This is achieved by selecting dijet events with two jets having similar p_{T} and being back-to-back in ϕ. One jet has to fulfill certain criteria on the track multiplicity, to ensure that it is not a true hadronic τ decay. If the criteria are fulfilled the second jet is used as a probe jet to test if it had passed the tau reconstruction and identification. Since a prescaled low threshold single jet trigger is used in the method, the probe jet should be required to not have caused the event be triggered. This avoids direct dependence on the trigger. The fake rate is determined as the number of probe jets identified as τ_{had} ($\Delta R < 0.15$) divided by the number of probe jets. In recent studies [24] the fake rate is observed at about 5–7% for an p_{T} range from 15 GeV to 120 GeV for the calorimeter-based approach with medium threshold. For the calo+track-based approach the fake rate is about 10% for the low p_{T} range up to 40 GeV but decreases rapidly to about 1% for the high p_{T} range. Also it is observed, that the fake rate is dependent on the jet origin. The fake rate is higher for QCD jets coming from quarks, than for QCD jets coming from gluons.

Due to the large cross section of the QCD dijet events, they are produced in a very high rate at the LHC. Because of the difficult trigger aspects, dijet events are not always an optimal sample to study tau misidentification rates. In particular jet based triggers, which are usually pre-scaled, may introduce a bias in the measurement. Therefore, in this thesis a method is described to estimated the misidentification rate caused by QCD dijets by using $Z \to ee + jet$ events. The τ_{had} misidentification rate (or fake rate) is determined and parametrized from additional jets in $Z \to ee + jet$ events. Then the fake rate is applied on QCD dijet events to estimated the number of fake taus coming from QCD dijets. In the following chapter, the whole method is described and the results are discussed.

Chapter 5

Analysis

5.1 Introduction

In this chapter a method is described to estimate the number of misidentified tau leptons (also referred to as fake taus) from QCD dijet events with a parametrized tau misidentification rate (also referred to as tau fake rate)determined in $Z \to ee + jet$ events.

Events without real tau leptons are a good base to study fake taus. Because there are no real tau leptons involved in $Z \to ee + jet$ events, this process is used as a τ lepton free control sample. The two electrons in the final state can be detected with high efficiency and low background and can be taken to tag the event. The additional hadronic jets are tested if they passed the tau reconstruction and identification algorithm described in Chapter 4. Using this sample, the tau misidentification rate is calculated as a function of different variables expecting to describe the event topology sufficiently. At the next step the number of misidentified tau leptons in a QCD dijet sample is estimated, using the fake rate calculated as a function of sensitive variables. The ratio between the number of estimated fake taus and expected fake taus in QCD dijet events is calculated. This step is called "tau fake rate validation". The best results are expected if both samples, $Z \to ee + jet$ and QCD dijets, have similar event topologies. In this case, the tau misidentification rates achieved from $Z \to ee + jet$ events can be projected to QCD dijet events.

The tau fake rate is divided into two parts, according to the process of tau reconstruction and identification, to be not affect by the differences in topology of $Z \to ee + jet$ and QCD dijet. The parts are the reconstruction fake rate (also referred to as reco fake rate or f_{reco}) and the identification fake rate (also referred to as id fake rate or f_{id}), as defined in the following:

- **Fake Tau**

$$\text{fake tau} = \tau \text{ candidates matched to a jet within } \Delta\text{R} < 0.2 \tag{5.1}$$

  A fake tau is defined as a τ candidate, that is matched to a Cone4TopoJet with ΔR< 0.2. Both, tau candidates and jets are required to pass certain cuts listed in Sec. 5.2.

- **Reconstruction Fake Rate**

$$f_{\text{reco}} = \frac{\text{number of fake taus}}{\text{all jets}} \tag{5.2}$$

  The reconstruction fake rate is the ratio of the number of fake taus defined in Eq. 5.1 to the number of all jets. If the fake rate is calculated for 1-prong (3-prongs) fake taus, only the number of jets having 1-prong (3-prongs) is used. The jets are required to pass the cuts listed in Sec. 5.2.

- **Identification Fake Rate**

$$f_{\mathrm{id}} = \frac{\text{number of fake taus passing } \tau \text{ identification}}{\text{number of fake taus}} \tag{5.3}$$

The identification fake rate is the ratio of fake taus passing the calorimeter cut-based tau identification with one of the threshold levels loose, medium, tight to the number of all fake taus. If the fake rate is calculated for 1-prong (3-prongs) fake taus passing the tau identification, only the number of fake taus having 1-prong (3-prongs) is used as denominator.

It is expected, that f_{reco} is dominated by the quantities and methods used at the reconstruction step of a τ candidate. Since fake tau candidates are originally gluon- or quark-jets passing the tau reconstruction, the reconstruction fake rate is influenced by the origin of the fake tau candidates. By contrast, f_{id} is influenced by the differences between applying the tau identification on a fake tau candidate. They have already passed the tau reconstruction algorithm and therefore fulfill the requirements to be a tau candidate. Thus, the influences on f_{id} are caused by the additional four calorimeter variables in the tau identification.

5.2 Event Selection

5.2.1 Event Samples

In Table 5.1 all event samples used for this study are listed with their cross section σ, generator event filter $\varepsilon_{\mathrm{filter}}$ efficiency and their expected number of events in $10\,\mathrm{pb}^{-1}$ of data. The latter was calculated with

$$\mathrm{N}_{exp}^{10pb^{-1}} = 10\,\mathrm{pb}^{-1} \times \sigma \times \varepsilon_{\mathrm{filter}}. \tag{5.4}$$

The QCD dijet samples are split into sub-samples J0–J5 corresponding to p_{T} ranges of 8–17 GeV, $17-35$ GeV, $35-70$ GeV, $70-140$ GeV, $140-280$ GeV and $280-560$ GeV. For the $Z \to ee$ sample, a generator event filter requiring one high p_{T} lepton within $|\eta| < 2.8$. All samples are generated with the Monte Carlo generator PYTHIA. The complete dataset names of the samples are listed in Appendix B.

Sample	σ in pb	$\varepsilon_{\mathrm{filter}}$ Efficiency	expected events in $10\,\mathrm{pb}^{-1}$data
J0	1.17×10^{10}	1	1.17×10^{11}
J1	8.67×10^{8}	1	8.67×10^{9}
J2	5.60×10^{7}	1	5.60×10^{8}
J3	3.28×10^{6}	1	3.28×10^{7}
J4	1.52×10^{5}	1	1.52×10^{6}
J5	5112	1	5.11×10^{4}
$Z \to ee$	1143.96	0.96	1.10×10^{4}

Table 5.1: Event samples used in the analysis, with cross section σ, generator event filter ($\varepsilon_{\mathrm{filter}}$) efficiency and the expected events at the LHC for an integrated luminosity of $10\,\mathrm{pb}^{-1}$

5.2.2 Fake Rate Determination

While this study was performed on generated events, the method is designed to be used with real data. For the tau fake rate determination step, a pure $Z \to ee + jet$ sample is necessary. Therefore different cuts were applied to select $Z \to ee + jet$ events out of the large QCD dijet background in data. Only jets from events reconstructed as to be a $Z \to ee$ process are used to determine the tau fake rate. To test the rejection of the QCD background, the same cuts are applied on the QCD dijet samples J0–J5. An overview of the generated events used for the fake rate determination ($Z \to ee$) and QCD background rejection study (J0–J5) is shown in Table 5.2. Because of computational resources there are only a limited samples of Monte Carlo events useable. Therefore the Monte Carlo events have to be weighted, to reflect the the numbers of events, which are expected in data. The weights can be found in Table 5.1.

Sample	MC events	weight for expected events in 10 pb^{-1}data
J0	999266	117086
J1	997956	8686
J2	979316	572
J3	980000	33
J4	998120	1.52
J5	872972	0.06
$Z \to ee$	128366	0.09

Table 5.2: Generated Monte Carlo events used for the fake rate determination ($Z \to ee$) and for study the QCD rejection rates.

Selection of $Z \to ee$ events

- Events with a trigger signal *EF_e10_medium* are required. *EF_e10_medium* is implemented as a high level trigger (EF = Event Filter) and requires at least one electron with $E_T >$ 10 GeV and medium quality criteria based on calorimeter informations. The trigger is defined for the beginning luminosity of 10^{31} cm^{-2} s^{-1}. The QCD background is reduced by this trigger up to three orders of magnitude over a broad energy range. The distribution of p_T of all electrons and the invariant mass of the tag electrons are shown in Figure 5.1 and Figure 5.2 .

- Electrons are selected, which are reconstructed by a cluster-based algorithm (*ElectronAuthor*) and passed a medium quality threshold (*medium*-electron). These cuts select high p_T electrons and prevent selecting photons or conversions.

- The electron selection requires $p_T > 15$ GeV and $|\eta| < 2.5$, excluding the detector crack-region $1.37 < |\eta| < 1.52$. This reduce nearly completely the QCD background except for the p_T region of $35 - 70$ GeV.

- At least two electrons with opposite charge are required, which have an invariant mass in the range of 70 GeV $< m_{\text{ee}} <$ 100 GeV. The range is indicated in Figure 5.2.

- If more than one pair of electrons passing the cuts, the pair with highest scalar sum of p_T is chosen as tag-electrons assuming coming from the Z boson.

The efficiency of the signal selection is defined as

$$\varepsilon = \frac{\text{MC events after the last cut}}{\text{input MC events}} \times 100\% \; , \tag{5.5}$$

while the rejection is defined as $r = (1-\varepsilon) \times 100\%$. The input Monte Carlo events are taken from Table 5.2. This leads to an efficiency of about $24.5\% \pm 0.12\%$ for the $Z \to ee$ signal process. The

QCD background is fully rejected ($r = 100\%$). Also the purity of the $Z \to ee$ event selection is checked, found to be $100\%^{+0\%}_{0.0036\%}$, using

$$purity = \frac{\text{events with both tag-electrons matched with truth electrons } (\Delta R < 0.1)}{\text{all selected events}} \times 100\%. \quad (5.6)$$

A cut flow overview for the signal sample and QCD background is presented in Table 5.3, showing the MC events which passed the certain cuts. The invariant mass of the tag electrons as shown in Figure 5.2 is peaking at the Z boson mass.

Sample	Trigger	*EleAuthor*	"medium"-Ele	p_T ,$\|\eta\|$	oppos. charge	mass window
J0	34	2	0	0	0	0
J1	799	259	16	0	0	0
J2	3369	2388	47	47	1	0
J3	6900	6232	77	4	2	0
J4	5743	5480	59	3	0	0
J5	6372	6186	68	1	0	0
$Z \to ee$	101089	62642	40026	34137	33846	31553

Table 5.3: Cutflow: Numbers in columns are Monte Carlo events after the cut printed in the column head. The QCD background (J0–J5) is fully rejected. The $Z \to ee$ selection efficiency is about of 31%.

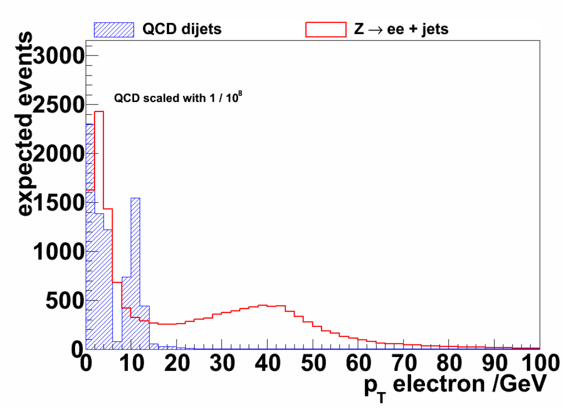


Figure 5.1: p_T distribution of electrons in the $Z \to ee + jet$ sample and QCD dijet sample, both after the trigger cut. The QCD dijet distribution is downscaled by a factor of 10^8.

Jets

As jets, *Cone4TopoJets*[1] are used, with $p_T > 10$ GeV and $|\eta| < 2.5$. Since electrons can be reconstructed as QCD jets, it is necessary to avoid the case, where tag electrons and probe jets are the same object. Therefore in case of the $Z \to ee + jet$ sample, jets are removed which overlap with the tag electrons within a cone of $\Delta R < 0.4$. Further electrons are not removed, because TopoJets are used as seed for tau candidates by the tau reconstruction algorithm as described in 4. As a cross-check it is shown that no jet is matched with truth electrons in a cone of $\Delta R < 0.2$. Since at least one jet is required to pass the cuts for calculating tau fake rates, $Z \to ee$ events with less than one jet are not considered. Comparing to Table 5.3, this reduces the number of events from 31553 to 14969 MC events. Using the latter, the expected number of events for $L = 10\,\text{pb}^{-1}$ data is 1282.

[1] described in Chapter 3.3

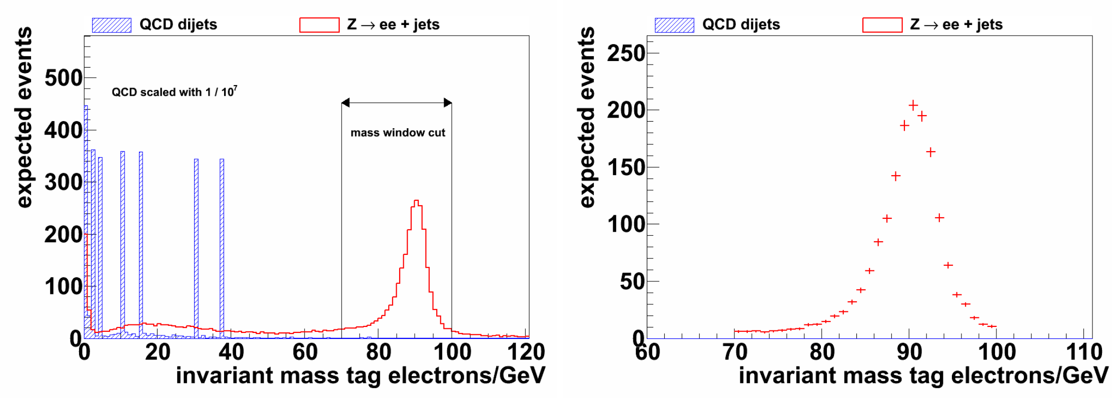


Figure 5.2: The invariant mass of the tag electrons is shown. Left: Only the trigger cut is applied on both samples and the QCD dijet events are downscaled by 10^5. The mass windows cut range is indicated by the arrow. Right: After all cuts the QCD dijet background is fully rejected.

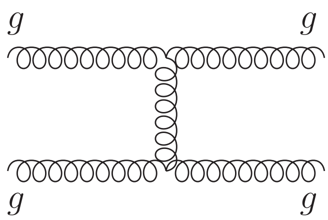


Figure 5.3: QCD dijet production at the LHC

Fake Tau

Tau candidates are used with $p_T > 10$ GeV, $|\eta| < 2.5$ and having both a calorimeter seed and a track seed. If a tau candidate is matched with a jet within $\Delta R < 0.2$, the tau candidate is called fake tau candidate. The separation value $\Delta R < 0.2$ is motivated by the distribution of minimal ΔR between τ candidates and QCD jets, both after passing the cuts. Figure 5.5 shows that most of the tau candidates are very close to QCD jets. It is being caused by QCD jets are reconstructed as a tau candidate as therefore they are the same objects.

5.2.3 Fake Rate Validation

To validate the fake rate determined in $Z \to ee + jet$ events, only QCD dijet events are used. In order to have unbiased QCD dijet events, no trigger cut is applied in the event selection. As **Jets**, *Cone4TopoJets* are used, with $p_T > 10$ GeV and $|\eta| < 2.5$. As **Fake Taus**, τ candidates are used, with $p_T > 10$ GeV, $|\eta| < 2.5$ and having both, a calorimeter seed and a track seed. If a tau candidate is matched with a jet within $\Delta R < 0.2$ (see Figure 5.5), the tau candidate is called fake tau candidate. Table 5.4 shows the generated MC events and the corresponding weights for $10\,\mathrm{pb}^{-1}$data.

5.3 Differences between $Z \to ee + jet$ and QCD dijet events

Due to the high energy, the production process at the LHC is dominated by gluon-gluon fusion. In a $Z \to ee$ environment hadronic jets originate mainly from the hard scattering process and therefore are dominated by quark-jets (diagrams (d) and (f) in Figure 5.4). In QCD dijet events the jets are

Sample	MC events	weight for expected events in 10 pb^{-1}data
J0	499859	234066
J1	339437	25536
J2	333203	1681
J3	400000	82
J4	319912	4.74
J5	200000	0.26

Table 5.4: Generated Monte Carlo events used for the fake rate validation.

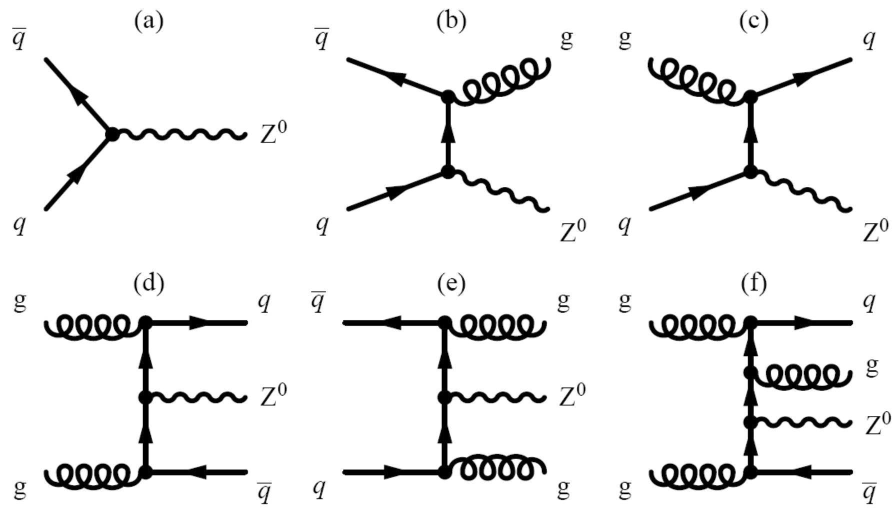


Figure 5.4: Feynman diagrams for the production of Z boson together with 0,1,2,3 jets (from [26, p.120])

mainly gluon-jets. A Feynman diagram of the dominant process is shown in Figure 5.3. A detail description of the production of jets in association with Z bosons at the LHC and ATLAS detector can be found inf Reference [8, p. 777ff] and more general in Reference [25]. The different origins of the hadronic jets in $Z \to ee + jet$ and QCD dijet events are reflected by the distributions of E_{T}, η and the track multiplicity. Figure 5.6 shows the distribution of these quantities for fake τ candidates for both samples. In Figure 5.7 the four calorimeter variables, which are used in the calo-based tau identification, are plotted for both, 1-prong and 3-prong fake taus and can be compared with distributions coming from real tau leptons as seen in Figure 4.1.

5.4 Details of the Measurement Method

Fake Rate Determination

To calculate the tau misidentification rate in $Z \to ee + jet$ events, a "tag-and-probe" method is used. The electrons coming from the Z boson are used for the event selection and therefore called tag electrons. The additional jets are used for probing the tau reconstruction and calculating the misidentification rate. These jets are called probe jets. The determination of the tau misidentification rate in only performed on $Z \to ee + jet$ events. In the following a common description of the method is given with technical details. The method can be applied for up to 3 variables used for the fake rate parametrization ($f(x_1)$, $f(x_1,x_2)$ or $f(x_1,x_2,x_3)$). The main objects working with are

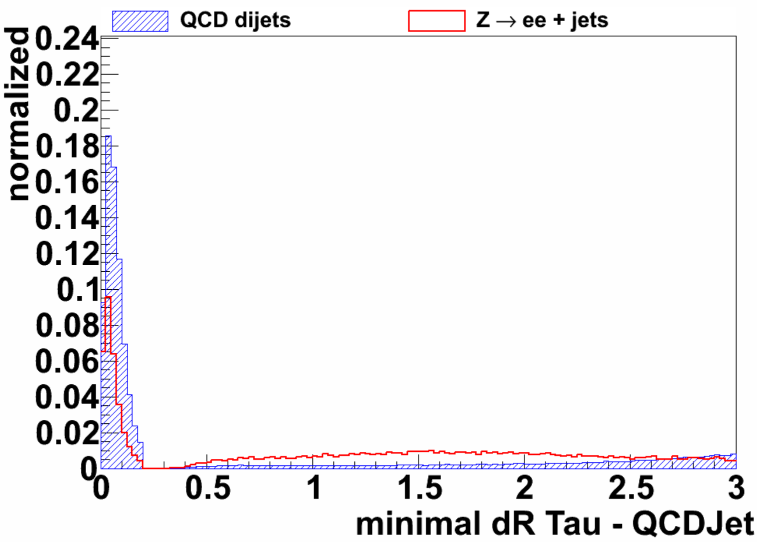


Figure 5.5: minimal ΔR between τ candidates and QCD jets, both after cuts ($p_T > 10$ GeV, $|\eta| < 2.5$). Regarding the QCD dijet events, the cuts listed at fake rate validation are applied.

histograms. Their dimension depends on the number of variables the misidentification rate should be parametrized. As an example, the identification fake rate parametrized with three variables is used. Also from now on, only jets and τ candidates already passing these cuts listed in Sec. 5.2 are considered. The analyis steps for the fake rate determination are:

1. **Binning:** Creation of a three dimensional[2] template histogram, representing on the axis the binning of the three variables. The numerator, denominator and fake rate histograms are derived from this template histogram.

2. **Numerator Histogram:** For every fake tau candidate passing the calo-based tau identification an entry is added to a bin, according to its values in the three variables. This histogram is the numerator in Eq. 5.3. For each threshold loose, medium, tight, a own histogram is filled.

3. **Denominator Histogram:** For every fake tau candidate an entry is add to a bin, according to its values in the three variables. This histogram is the denominator in Eq. 5.2.

4. **identification fake rate:** Calculate the identification fake rate by dividing the according histograms (2) and (3) bin-per-bin and store the resulting fake rate into a histogram. The histogram contains now the identification fake rate parametrized in three variables, represented by the individual bins.

If the fake rate is calculated for 1-prong (3-prong) fake τ candidates, in step (2) and (3) every fake tau candidate has to fulfill this prong requirement. The procedure for f_{id} is shown in Figure 5.8, while for f_{reco} the method performs analogue.

[2] a so called 3-dimensional histogram has 3 dimension for the physical variables and a “fourth” dimension to store the histogram entries.

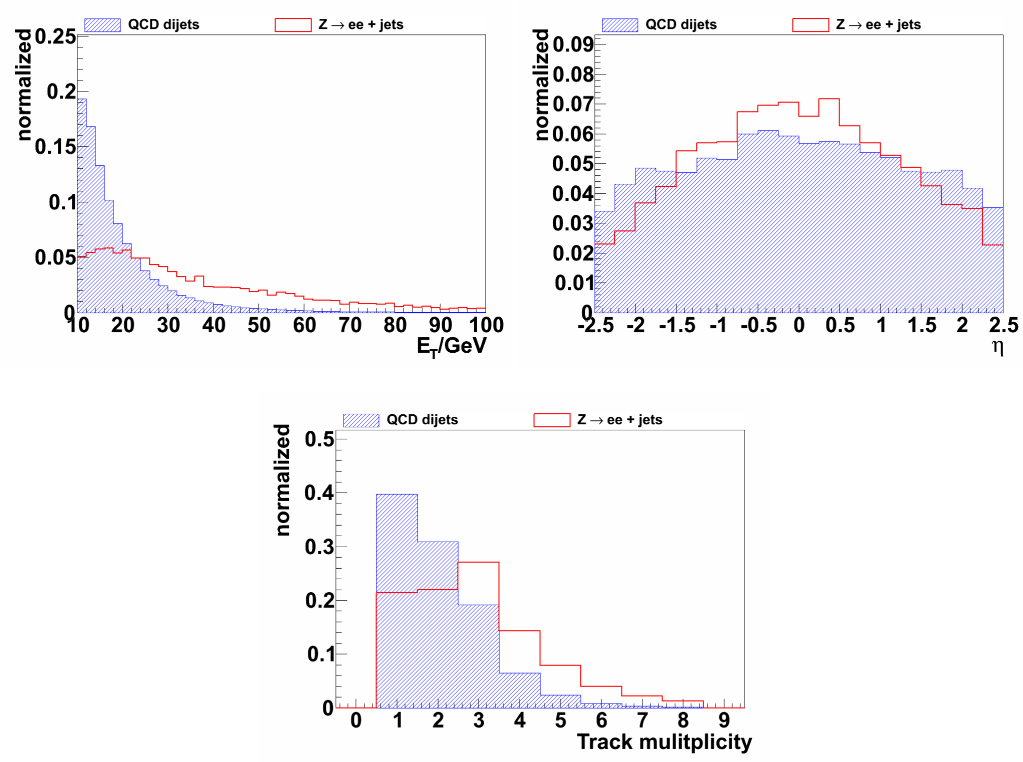


Figure 5.6: Distribution of E_T, η, numbers of tracks and EM Radius achieved from $Z \to ee + jet$ and QCD dijet events for tau candidates identified as fake taus

Fake Rate Validation

To estimate the number of misidentified taus in QCD dijet events, the stored parametrized fake rate from $Z \to ee$ events is used. In the following the method is described, using again f_id parametrized in three variables ($f(x_1, x_2, x_3)$) as an example.

1. Each fake tau candidate is associated with a identification fake rate f_id, depending on the variables x_1, x_2, x_3. The identification fake rate can be seen as a possibility that a fake tau candidate with certain properties will pass the calo-based tau identification. The estimated number of fake taus candidates passing the tau identification is calculated as

$$\mathrm{N}^{\mathrm{est}} = \sum_{i=1}^{\text{fake taus}} f_\mathrm{id}(x_i, y_i, z_i) \times w_\mathrm{sample} \tag{5.7}$$

   where the sum runs over all fake taus, f_id is the identification fake rate depending on the variables of the fake tau candidate and w_sample is the weight, according to the sample running on. The distribution of estimated fake taus is filled in a histogram over E_T.

2. The expected number of fake tau candidates which passed the tau identification is directly achieved as

$$\mathrm{N}^{\mathrm{exp}} = \sum_{i=1}^{\text{fake taus}} \left(w_\mathrm{sample} \times \begin{cases} 0 & : \text{ has not passed tau identification} \\ 1 & : \text{ has passed tau identification} \end{cases} \right) \tag{5.8}$$

   where the sum runs over all fake tau candidates. w_sample is the weight, according to the sample running on. The distribution of expected fake taus is filled in a histogram over E_T.

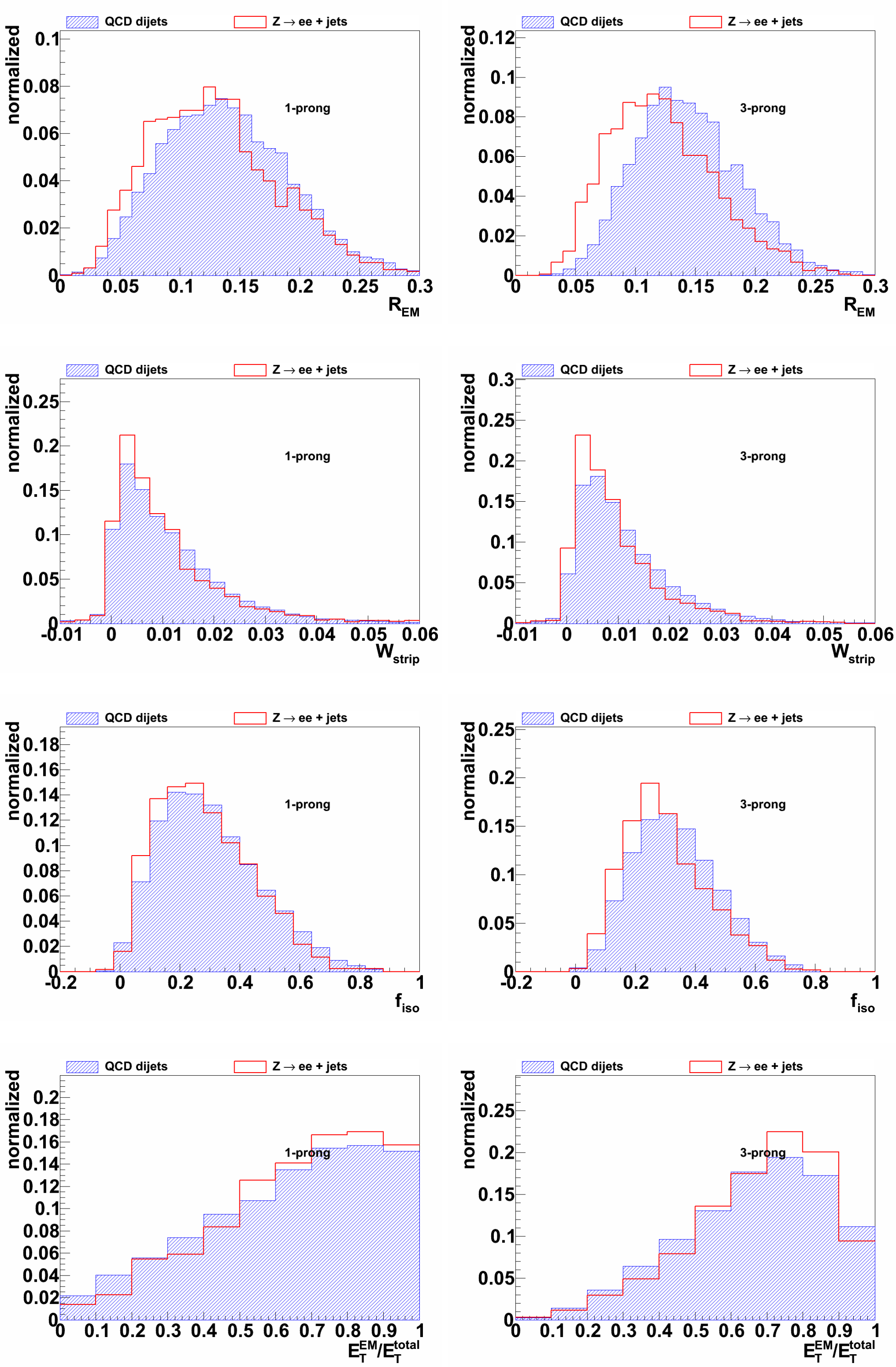


Figure 5.7: Distribution of the calorimeter variables for 1-prong (left) and 3-prong (right) fake taus.

3. To check the goodness of the parametrization, both distributions (5.7 and 5.8) are compared by calculating the ratio for each E_T bin:

$$\mathrm{r} = \frac{\text{expected number of fake tau candidates passing the tau identification}(\mathrm{N}^{\mathrm{exp}})}{\text{estimated number of fake tau candidates passing the tau identification}(\mathrm{N}^{\mathrm{est}})} \tag{5.9}$$

If the parametrization works fine, the ratio should be 1 over a wide range of E_T.

The procedure for f_{id} is shown in Figure 5.8, while for f_{reco} the method performs analogue.

Calculation of Uncertainties

Since this method is foreseen to be applied to data, the uncertainties expected in real data events are estimated here. The following descriptions concentrate on the identification fake rate as defined in Eq. 5.3 but can also be assigned analogous to the reconstruction fake rate (Eq. 5.2). The identification fake rate is a ratio of passed to total events, where the passed events p are a sub-sample of the total events t. Thus, the uncertainties for the fake rate can be handled like for efficiencies. Assuming Gaussian uncertainties, the error is $\delta p = \sqrt{p}$ for passed events and $\delta t = \sqrt{t}$ for total events. Then, using standard error propagation this leads to

$$\delta f_{\mathrm{id}} = \sqrt{\frac{p^2(t+p)}{t^3}} = \sqrt{f_{\mathrm{id}}(1+f_{\mathrm{id}})} \tag{5.10}$$

In some cases this error calculation will give unphysical results [27]. If there are 0 passed events the error is founded to be 0. If $p = t$, which is equal to $f_{\mathrm{id}} = 1$, the error range will span the range of the fake rate from negative values to larger than 1. Therefore a Bayesian method is used described in Reference [27]. The model is the binomial distribution. The prior is the flat distribution from 0 to 1, which reflects the range of the fake rate. To assign the errors a 68.3% confidence belt is used. The method varies in Eq. 5.11 the lower border a and upper border b as long as the integral reaching the confidence belt of 0.683.

$$\mathrm{frac} = \int_a^b x^p(1-x)^{t-p}\mathrm{d}x \stackrel{!}{=} 0.683 \tag{5.11}$$

The lower and upper error for f_{id} are calculated by $f_{\mathrm{id}} - a$ and $b - f_{\mathrm{id}}$. The lower border a is set to zero and only the upper border is calculated in Eq. 5.11, if there are no passed events ($f_{\mathrm{id}} = 0$). Likewise, if $f_{\mathrm{id}} = 1$, the upper border b is set to 1, giving an upper error of 0, and only the lower border is calculated in Eq. 5.11. This prevents to have errors reaching into the negative region or in the region bigger than 1.

In the validation step, the uncertainty for the expected number of id fake taus is $\delta\mathrm{N}^{\mathrm{exp}} = \sqrt{\mathrm{N}^{\mathrm{exp}}}$ for each bin, using normal Gaussian uncertainties. If the error of the estimated number of fake taus passing the tau identification is calculated, the uncertainty of the fake rate has to take into account. According to Eq. 5.7, the error per bin is

$$\delta\mathrm{N}^{\mathrm{est}} = \sqrt{\sum_{i=1}^{\text{fake taus}} (\delta f_{\mathrm{id}}(x_i,y_i,z_i)^2) \times w_{\mathrm{sample}}^2} \tag{5.12}$$

where the sum runs over all fake tau candidates, $\delta f_{\mathrm{id}}(x_i,y_i,z_i)$ is the error for the parametrized identification fake rate of Eq. 5.7 and w_{sample} is the weight, according to the sample running on. This uncertainty is mainly dependent on the distribution and flatness of the error of f_{id}. The errors of the identification fake rate depends on the numbers of passed events in Eq. 5.11. The number

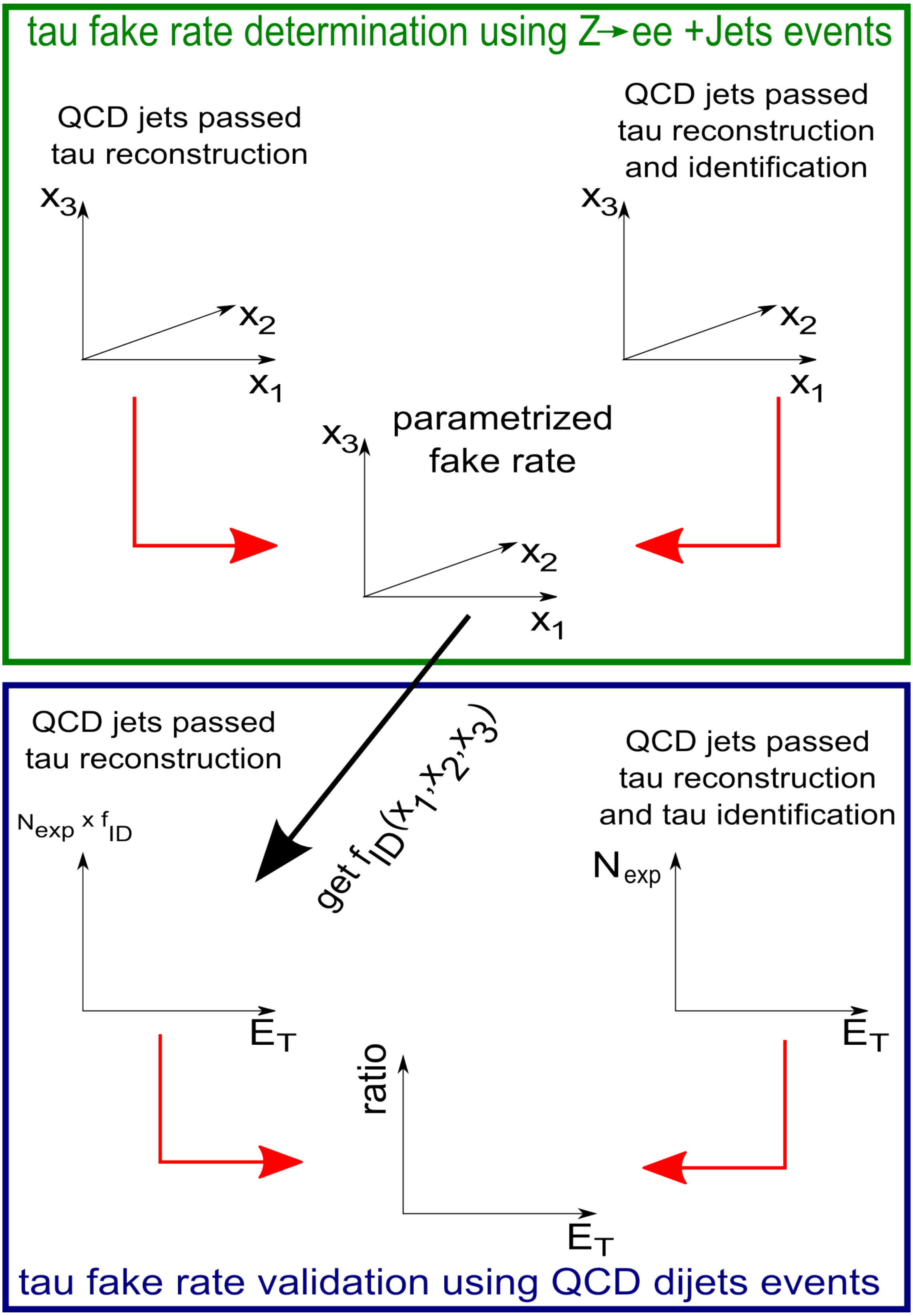


Figure 5.8: Work flow of the determination and validation step of the tau identification fake rate parametrized in 3 variables x_1, x_2, x_3. $f_{\mathrm{id}}$is determined with jets from $Z \to ee + jet$ events and validated with QCD dijet events.

of fake tau candidates used for determine the identification fake rate varies for every variable and bin, which the fake rate is parametrized on. Therefore every fake rate has its own error and the distribution of the errors are not flat. If a variable or bin is favored, the associated fake rate has a low error. When this error is multiplied with a very a weight and the larger errors are multiplied with a lower weight, the overall error $\delta \mathrm{N}^{\mathrm{est}}$for one bin can be get small.

The ratio in Eq. 5.9 can range from 0 to ∞, and has not the meaning of an efficiency. Therefore normal error propagation can be applied:

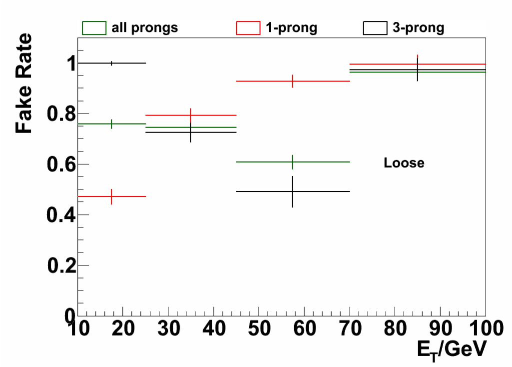
$$\delta r = r \times \sqrt{\left(\frac{\delta \mathrm{N}^{\mathrm{exp}}}{\mathrm{N}^{\mathrm{exp}}}\right)^2 + \left(\frac{\delta \mathrm{N}^{\mathrm{est}}}{\mathrm{N}^{\mathrm{est}}}\right)^2} \tag{5.13}$$

5.5 Identification Fake Rate

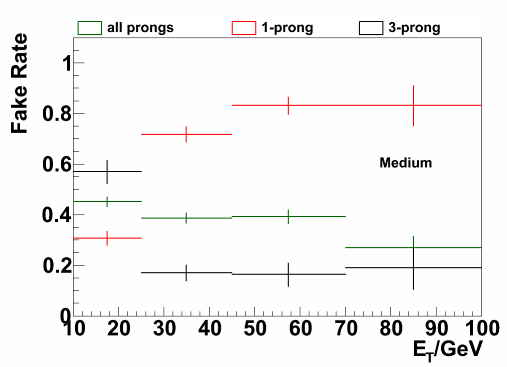
The identification fake rate is defined in Eq. 5.3 and is be studied only for the calorimeter-based tau identification, described in Sec. 4.2.

1-dimensional: E_T

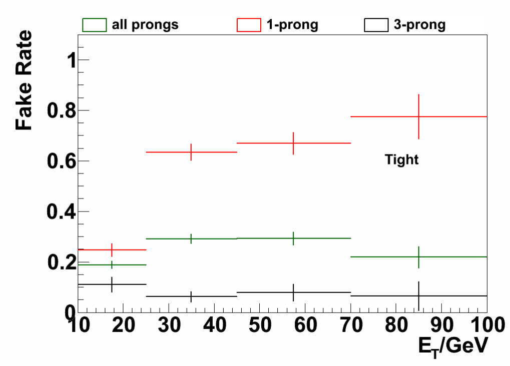
The identification fake rate was parametrized in four bins of E_T, which are the same bins the cut-based variables were optimized for the tau identification. The bins are $10-25$ GeV, $25-45$ GeV, $45-70$ GeV and $70-100$ GeV. The identification fake rate is shown in Figure 5.9 for fake tau candidates with 1-prong, 3-prongs and all-prongs and for the thresholds loose, medium and tight. The identification fake rate is lower for the tight threshold than for the loose threshold. In the case of the medium (tight) threshold, 1-prong fake tau candidates are passing the calo-based tau identification up to four (eight) times more than 3-prong fake tau candidates. In case of the loose threshold the deviations of the identification fake rate for the different prongs are smaller.



(a)



(b)



(c)

Figure 5.9: Distribution of the identification fake rate for the calorimeter-based approach with the three thresholds loose (a), medium (b) and tight (c). The fake rate is parametrized in bins of E_T and is calculated using jets from $Z \to ee + jet$ events.

The results of the validation process are shown in Figure 5.10 and 5.11. In row 1-3 the estimated and expected number of fake tau candidates passing the tau identification is plotted for the three thresholds loose, medium and tight. The y-axis are log-scaled. The differences between estimation and expectation can be seen more easily calculating the ratio (fourth row) presented in equation (5.9). This ratio has large deviations from being close to 1. Therefore a idendentification fake rate parametrized only in bins of E_T can not estimate the number of fake tau candidates passing the calorimeter-based tau identification caused by QCD dijet events. It is possible that one variable or especially E_T do not provide sufficent information to find common properties between jets in $Z \to ee + jet$ events and QCD dijet events. Therefore a parametrization with three variables is studied.

3-dimensional: R_{EM}, W_{strip}, f_{iso}

The identification fake rate was parametrized in three calorimeter variables: the electromagnetic radius R_{EM}, the isolation in the calorimeter f_{iso} and the strip width in the η layer, W_{strip}, all described in Sec. 4.2. Each variable was divided in two bins. The goodness of the fake rate parametrization is strongly depended on the bin limits. Therefore the bin limits are derived in an automatic process, which varies the middle bin limit, calculates the fake rate, estimates the number of fake taus and comparing this with the expected number of fake taus by calculating the ratio of both. The quality is checked by an chi-square approximation on the ratio. This ratio should be close to one over a broad range of E_T. The bin limits, which performs at best, are listed in Table 5.5. The variable distributions for fake tau candidates passing the calorimeter-based tau identification for the $Z \to ee + jet$ sample and QCD dijets is shown for all prongs (Figure 5.12), 1-prong (Figure 5.13) and 3-prongs (Figure 5.14). The values of the identification fake rate for each variable and E_T bin can be found in Appendix C.

		R_{EM}			f_{iso}			W_{strip}		
threshold	prong	low	mid	up	low	mid	up	low	mid	up
loose	1	0	0.04	0.3	-0.1	0.45	1	-0.005	0.005	0.06
	3	0	0.16	0.3	-0.1	0.25	1	-0.005	0.0	0.06
	all	0	0.16	0.3	-0.1	0.25	1	-0.005	0.01	0.06
medium	1	0	0.05	0.15	-0.1	0.0	1	-0.005	0.005	0.06
	3	0	0.07	0.15	-0.1	0.2	1	-0.005	0.015	0.06
	all	0	0.06	0.15	-0.1	0.15	1	-0.005	0.005	0.06
tight	1	0	0.06	0.1	-0.1	0.15	1	-0.005	0.01	0.06
	3	0	0.05	0.1	-0.1	0.3	1	-0.005	0.01	0.06
	all	0	0.05	0.1	-0.1	0.3	1	-0.005	0.025	0.06

Table 5.5: Bin limits of the three variables R_{EM}, f_{iso} and W_{strip}, which are used to parametrize the identification fake rate. Each variable is divided in two bins. The first bin ranges from "low" to "mid" and the second bin ranges from "mid" to "up".

The estimated and expected number of fake tau candidates passing th calo-based tau identification and the ratio of both, separately again for loose, medium and tight are shown for 1-prong (5.16), 3-prongs(5.17) and all-prongs (5.15) fake tau candidates. The parametrization works fine for the loose threshold. Regarding only fake tau candidates with three prongs, the agreement between estimated and expected fake tau candidates passing the calo-based loose tau identification is within 2%. For 1-prong fake taus, the deviation is especially for the second and third E_T bin large, being up to 35%. Using all fake taus without looking at the number of prongs, the deviation is about 6%, but mostly influenced by 3-prong fake taus. Looking at the medium threshold, the parametrizations works well for the first and third E_T bin, giving a deviation of 4% for 3-prong fake taus and up to 16% for 1-prong fake taus. This leads to an deviation of about 8% for all-prong fake taus. The second and fourth E_T bins have large deviations for 3-prong (up to 26%) and 1-prong fake taus (up to 50%), but decrease to about 12% if not considering the numbers of prongs. The deviations for the tight thresholds increase rapidly for higher E_T bins. While underestimating

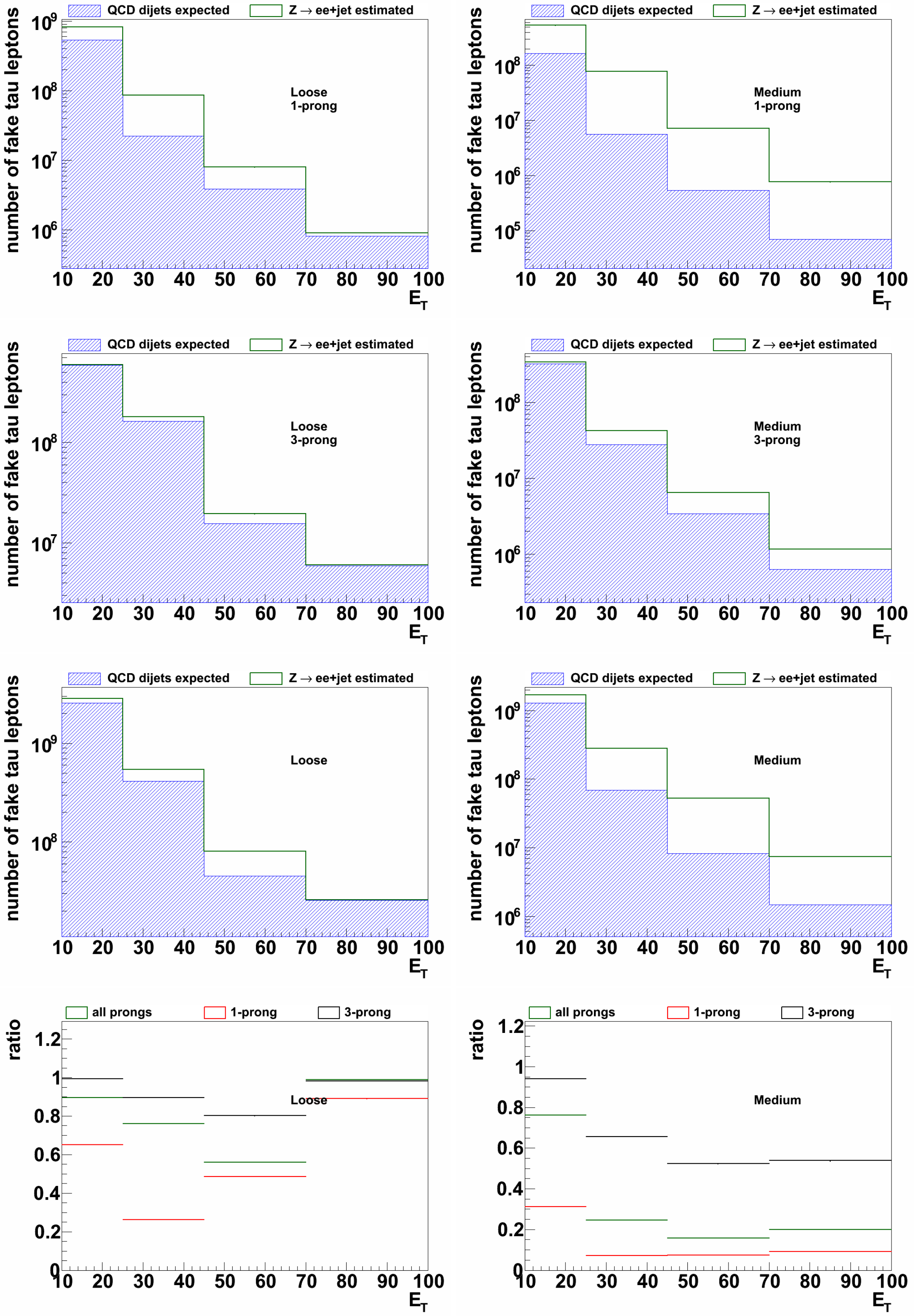


Figure 5.10: The estimated and expected numbers of fake tau candidates passing the calorimter-based tau identification are plotted for 1-prong (first row), 3-prong (second row) and all-prong (third row) for the thresholds loose (left column) and medium (right column). The y-axis is log scaled. The identification fake rate was parametrized in bins of E_{T}. In the fourth row, the ratios between estimated and expected fake taus are plotted.

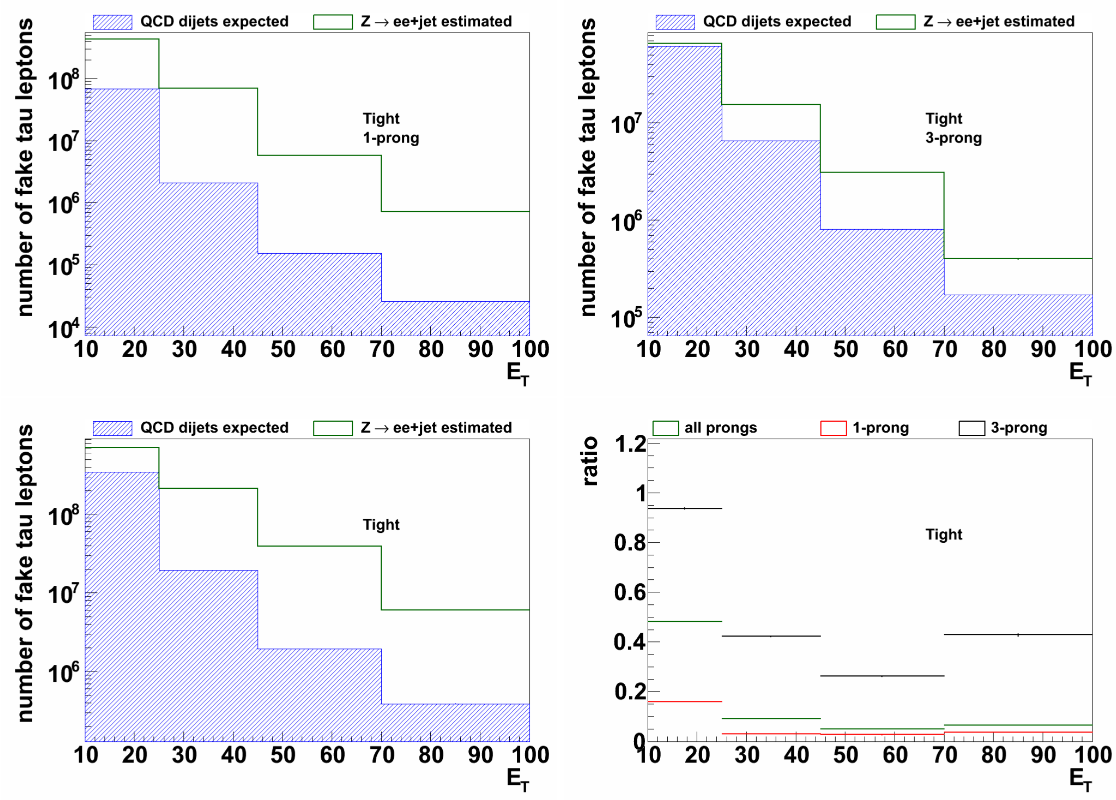


Figure 5.11: The estimated and expected numbers of fake tau candidates passing the calorimter-based tau identification are plotted for 1-prong (first row), 3-prong (second row) and all-prong (third row) for the tight threshold. The y-axis is log scaled. The identification fake rate was parametrized in bins of E_{T}. In the fourth row, the ratios between estimated and expected fake taus are plotted.

the numbers of fake tau candidates passing the tau identificatio of about 8% in the first E_T bin for 3-prongs, it is overestimated of about 22% for the fourth E_T bin. For 1-prong fake taus the deviation reach 80% in the fourth bin. Regarding to all-prong fake taus, the deviation is 8% in the first and second E_T bin, 14% in the third E_T bin and 32% in the last bin. Mostly the expected number of misidentified taus is overestimated. The uncertanties on the estimated number of fake taus is in order of magnitude of 0.1%, while the uncertainties on the expected number of fake taus is in the order of magnitude of 0.01%. The errors are dominated by the high statistic of the QCD dijet events expected in data.

The number of misidentified tau candidates in $Z \to ee + jet$ events is not sufficient to have statistical uncertainties lower than 10%. At least 100 misidentified tau lepton candidates should pass the tau identification to have sufficent statistics for each variable and bin the fake rate is parametrized on. After the cuts described in Sec. 5.2.2 there are about 1000 misidentified tau leptons that passed the loose calorimeter-based tau identification in $10\,\mathrm{pb}^{-1}$ of data $Z \to ee + jet$ events. Having 3 variables, each with two bins, results in eight bins, which must be filled. The identification fake rate is determined separately in four E_T ranges, giving 32 different identification fake rates. Thus, for the loose threshold three times more events are necessary. This calculation only regards to the loose threshold and all-prong fake taus. For medium the efficiency is about 60% and for tight of about 30%. Therefore assuming a flat distribution over all variables and bins ten times more events as for $10\,\mathrm{pb}^{-1}$ are neseccary for a theretical statistical uncertainty of 10% of the fake rate in all bins. Because the variable distributions are not flat as can be found in Figures 5.12, 5.13 and 5.14, even more data is required.

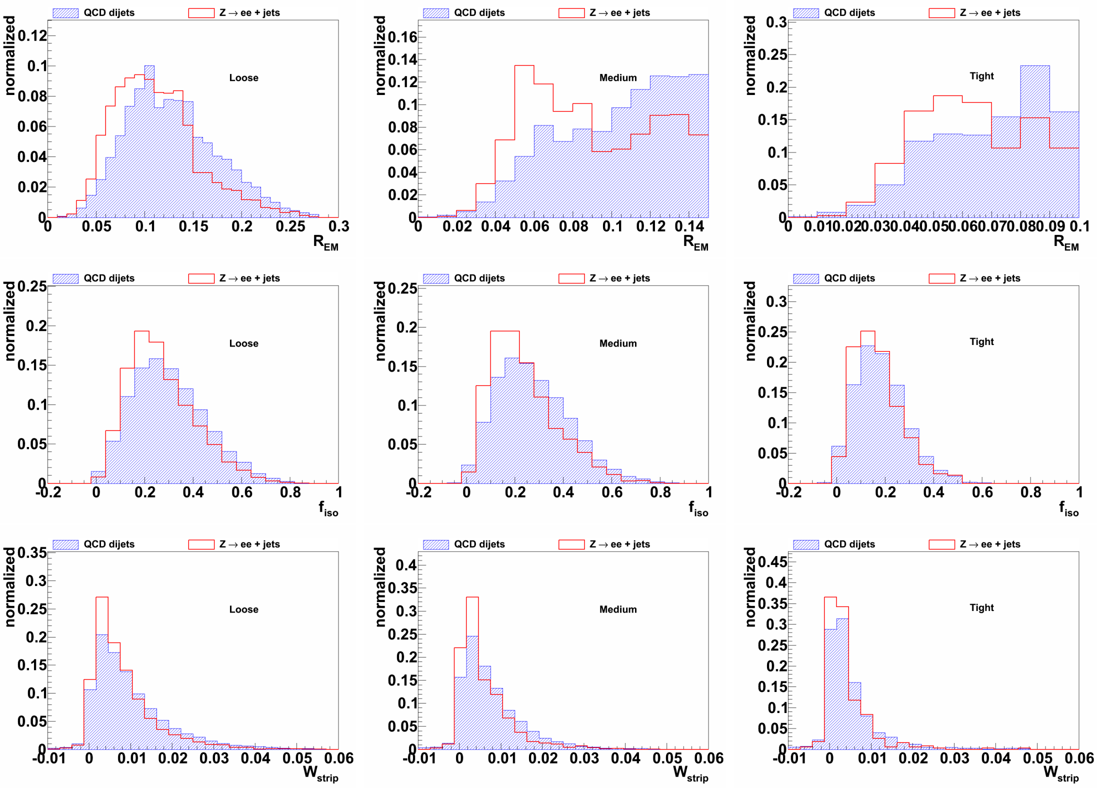


Figure 5.12: All prongs: Distribution of the three variables, R_EM, f_iso and W_strip, which are used to parametrize the identification fake rate, separately for the thresholds loose (left), medium (middle) and tight (right).

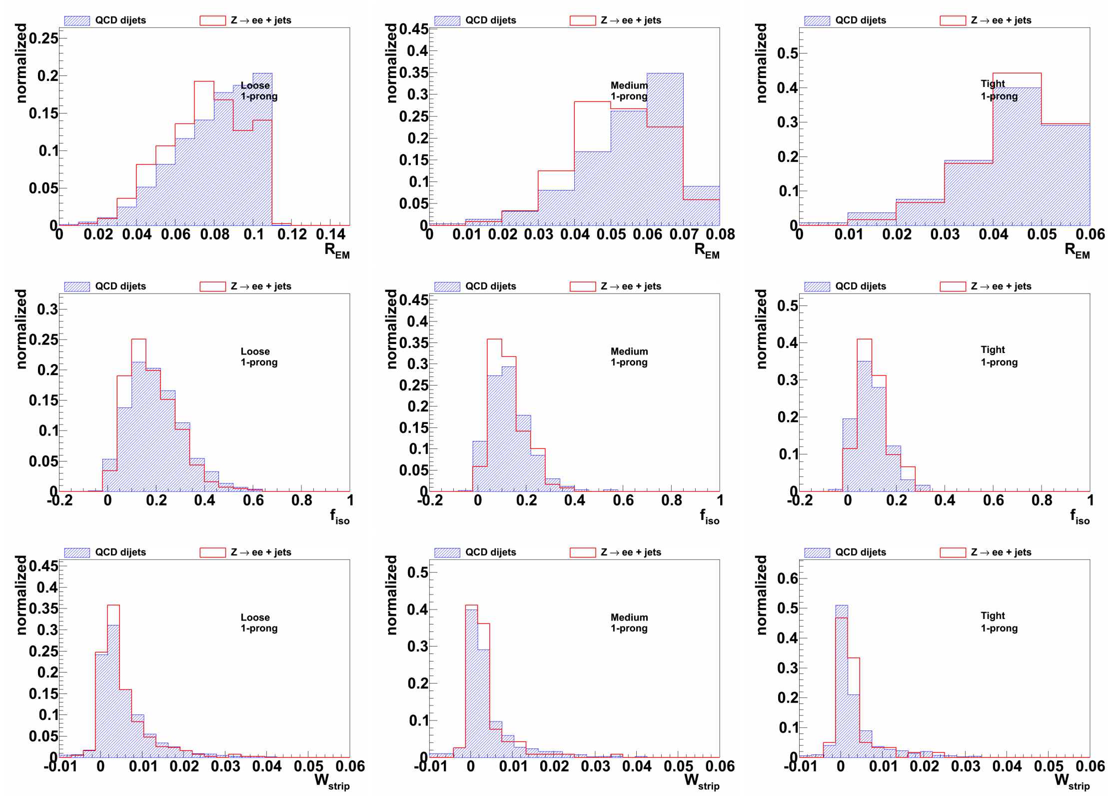


Figure 5.13: 1-prong: Distribution of the three variables, R_{EM}, f_{iso} and W_{strip}, which are used to parametrize the identification fake rate. The distributions are for 1-prong fake taus, separately for the thresholds loose (left), medium (middle) and tight (right).

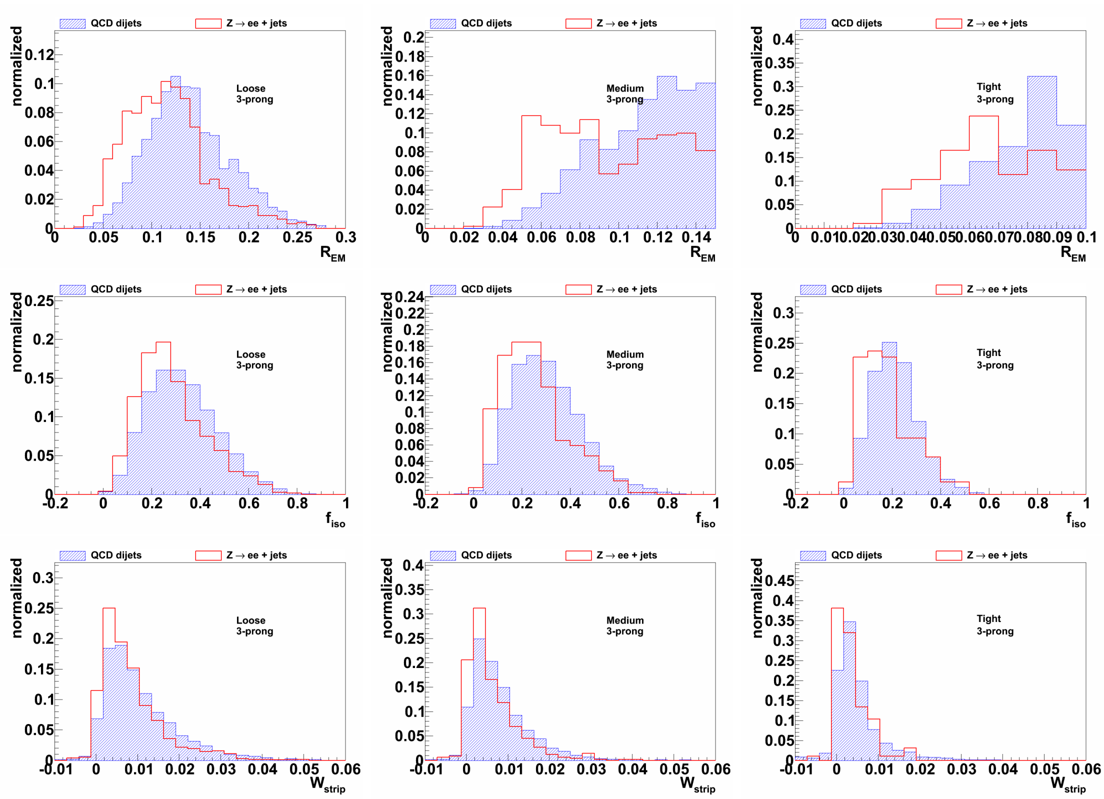


Figure 5.14: 3-prong: Distribution of the three variables, R_{EM}, f_{iso} and W_{strip}, which are used to parametrize the identification fake rate. The distributions are for 3-prong fake taus, separately for the thresholds loose (left), medium (middle) and tight (right).

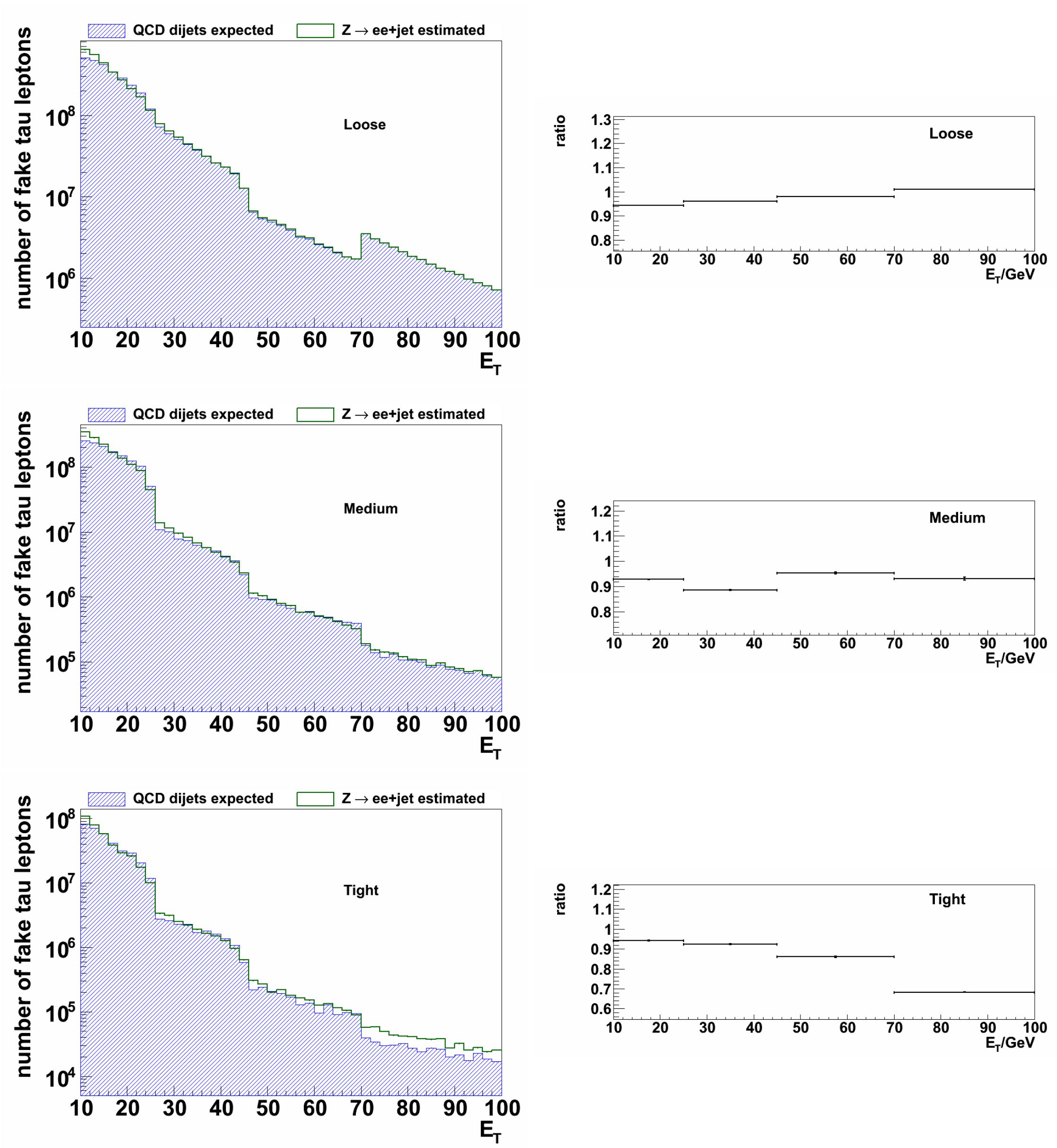


Figure 5.15: Left: Distribution expected (blue, hatched) and estimated (green, line) fake tau candidates (**all-prongs**) passing the calorimeter-based tau identification with the threshold loose (first row), medium (second row) and tight (third row).
Right: The ratio of expected fake tau candidates to estimated fake tau candidates.

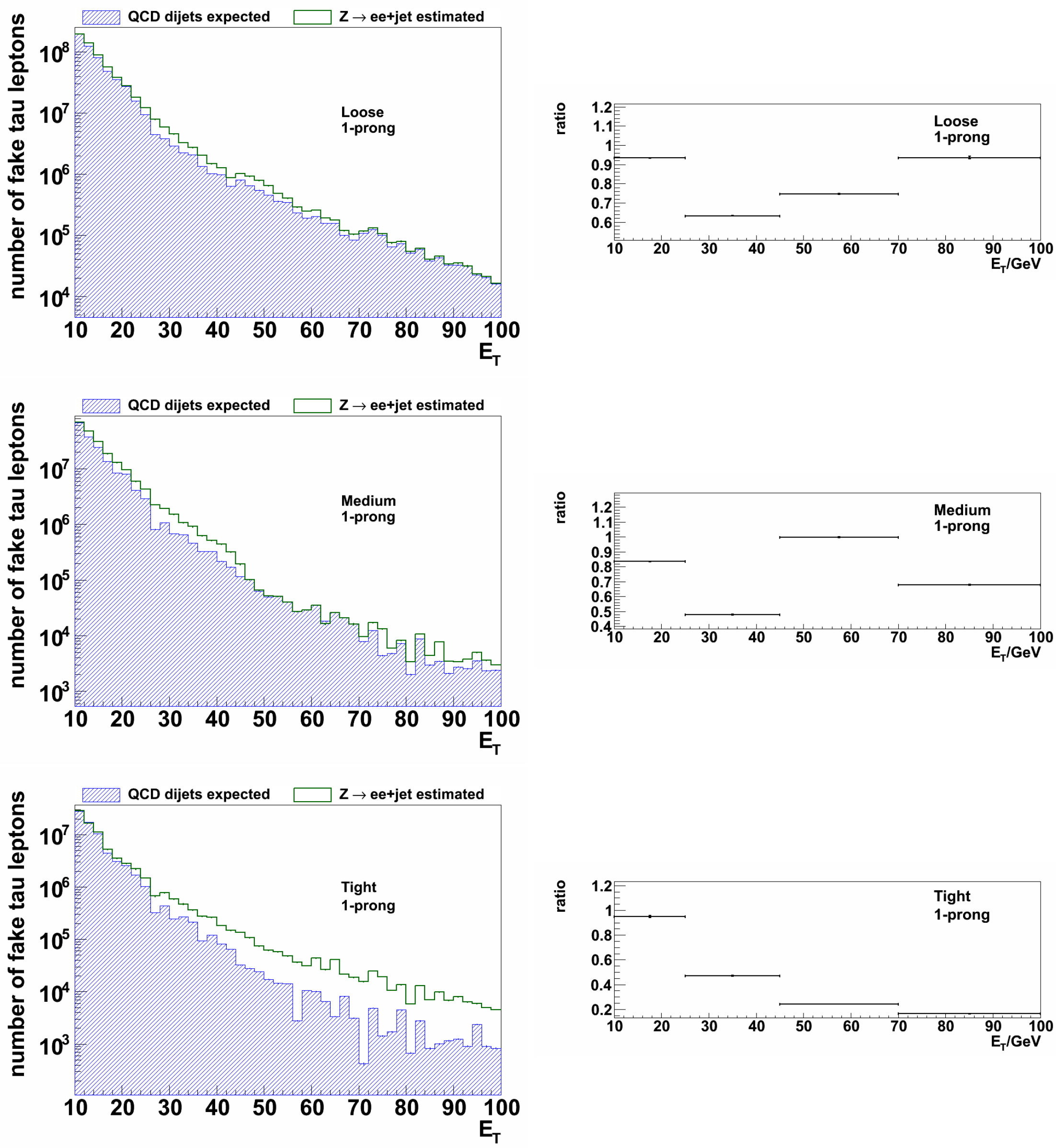


Figure 5.16: Left: Distribution expected (blue, hatched) and estimated (green, line) fake tau candidates (**1-prong**) passing the calorimeter-based tau identification with the threshold loose (first row), medium (second row) and tight (third row).
Right: The ratio of expected fake tau candidates to estimated fake tau candidates.

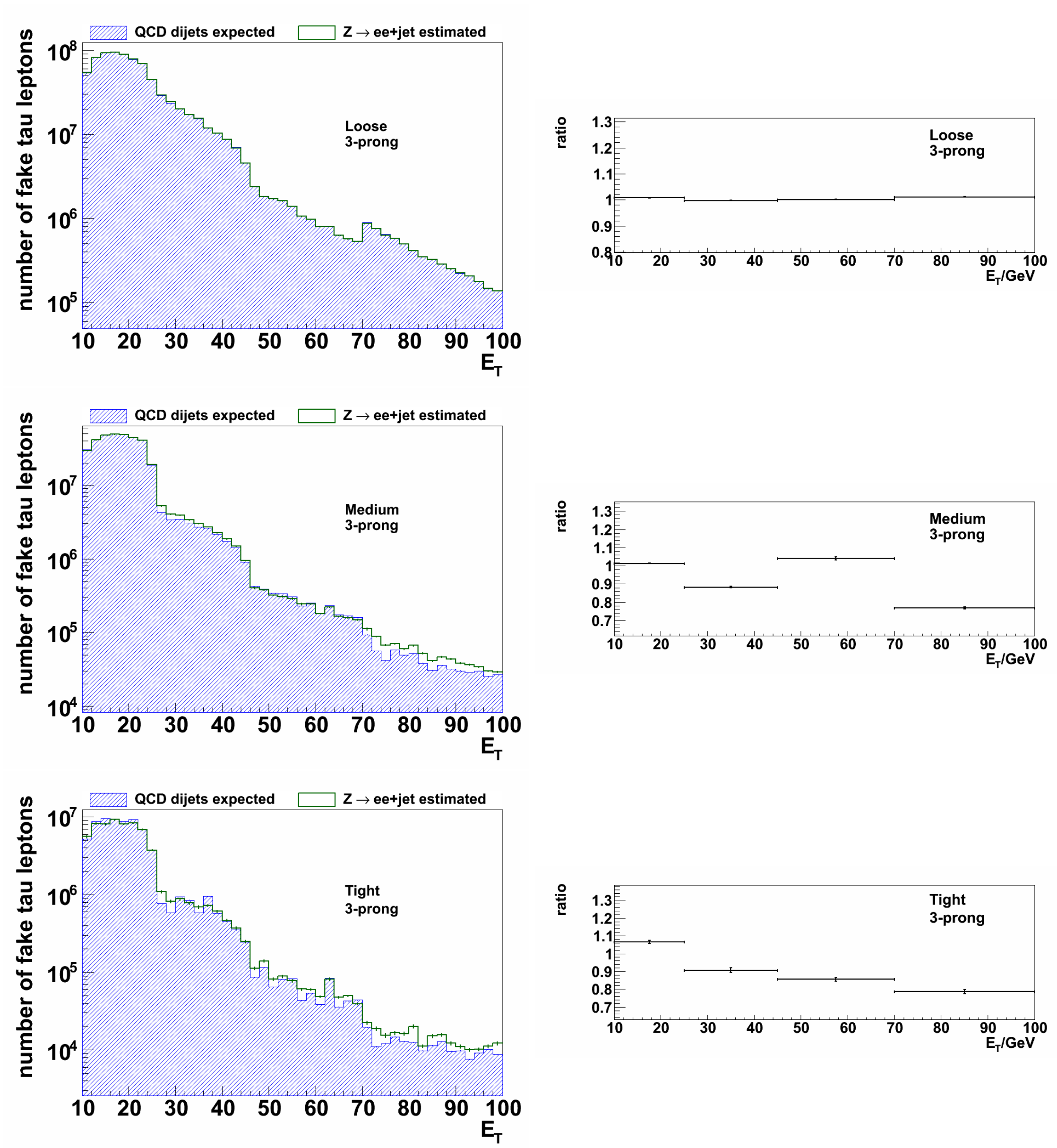
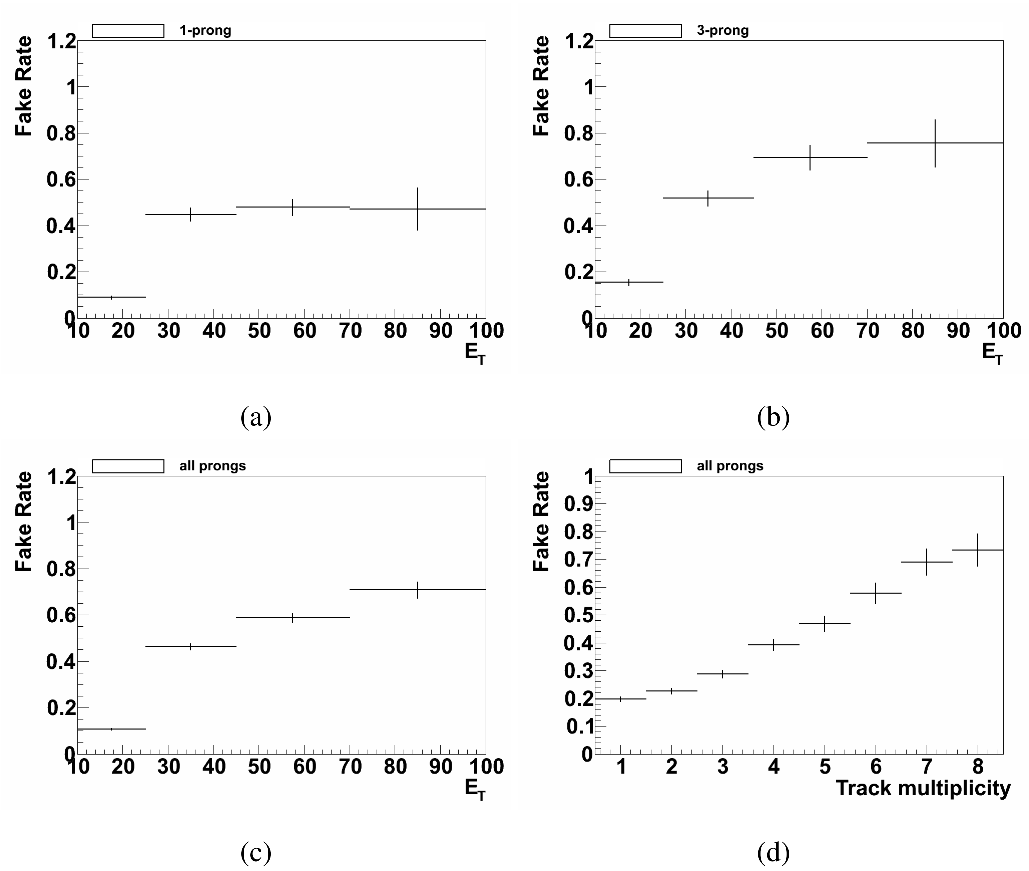


Figure 5.17: Left: Distribution expected (blue, hatched) and estimated (green, line) fake tau candidates (**3-prongs**) passing the calorimeter-based tau identification with the threshold loose (first row), medium (second row) and tight (third row).
Right: The ratio of expected fake tau candidates to estimated fake tau candidates.

5.6 Reconstruction Fake Rate

The reconstruction fake rate as defined in Eq. 5.2 is parametrized first in bins of E_T and is shown in Figure 5.18 a-c for 1-prong, 3-prong and all-prong jets. Hadronic jets with 3-prongs have a higher possibility to be reconstructed as a tau candidate than jets with one charged track. The reconstruction fake rate increase with higher jet E_T. The estimated number of fake tau candidates and the expected number of fake tau candidates is shown in Figure 5.19 a-c. The ratio of expected over estimated numbers of fake taus can be found in Figure 5.19d. The deviation from a ratio of 1 is up to 20% for the third and fourth E_T bin, but decrease to about of 60% for the first bin of E_T. Like for the identification fake rate a parametrization only in bins of E_T can not estimate the number of fake tau canidates.

Therefore following the reconstruction fake rate is parametrized in bins of the track multiplicity and can be found in Figure 5.18d. With increasing number of tracks, the reconstruction fake rate also increase. Jets with one charged track (three charged tracks) are reconstructed as tau candidate in 20% (30%) of the cases. The distributions of the expected and estimated number of jets passing the tau reconstruction and the ratio of both are shown in Figure 5.20. Concerning the large deviations from a ratio of one, the parametrization in bins of the track multiplicity works not sufficent. At next step the same parametrization is done but separately for the different E_T bins. It is found, that the deviations become smaller. The best result is achieved for the third (45–70 GeV) and fourth (70–100 GeV) E_T bin with deviations about of 10%.



(a) (b) (c) (d)

Figure 5.18: Reconstruction fake rate parametrized in bins of E_T for 1-prong (a), 3-prongs (b) and all-prongs (c) and in bins of the track multiplicity (d).

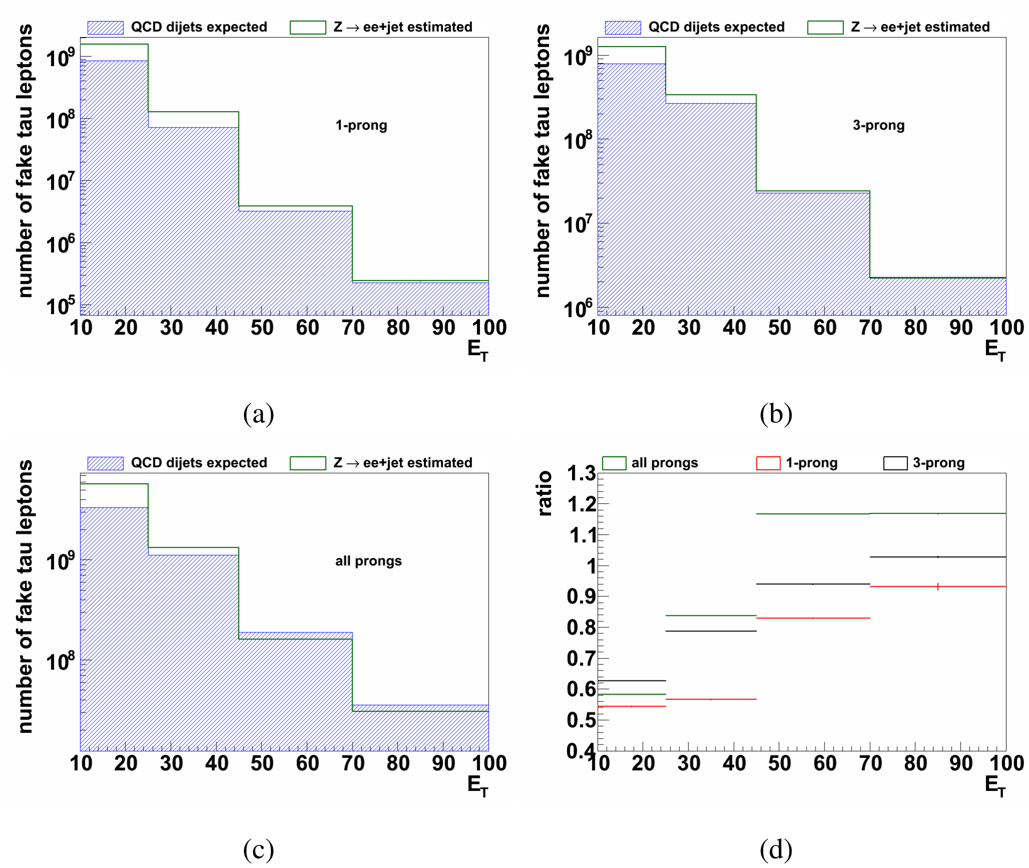


Figure 5.19: Distribution of the estimated (green, line) and expected (blue, hatched) number of QCD jets reconstructed as tau candidate for 1-prong (a), 3-prongs (b) and all-prongs (c). The ratio estimated number over expected number is shown in (d). The reconstruction fake rate is parametrized in bins of E_{T}.

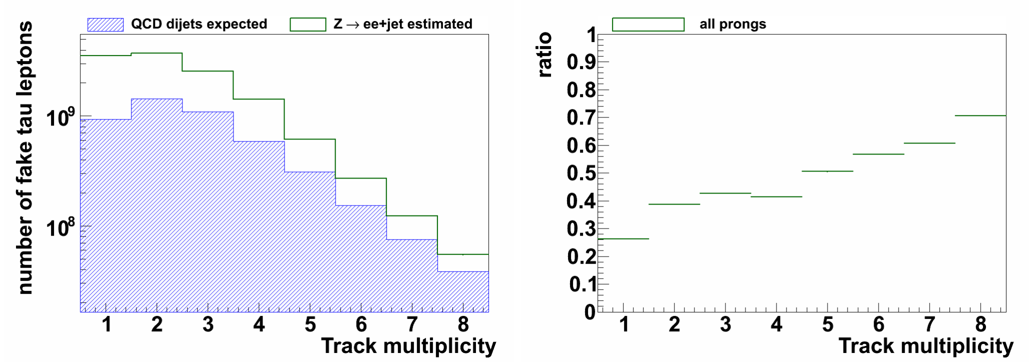
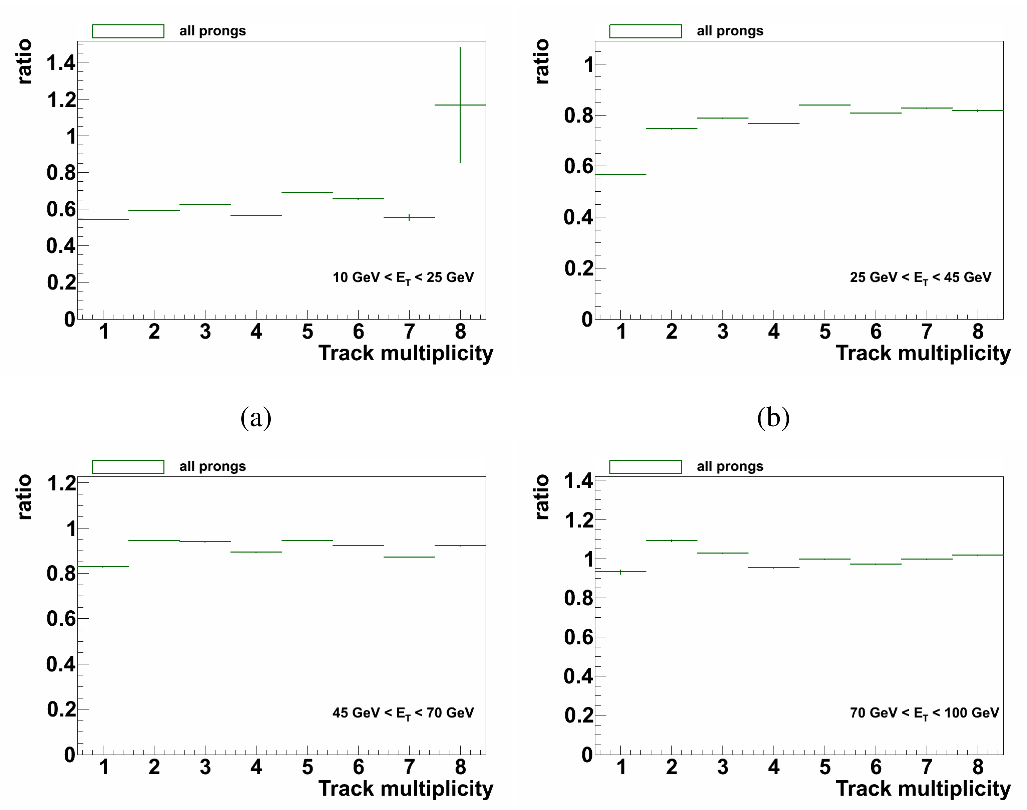


Figure 5.20: Left: Distribution of the estimated (green, line) and expected (blue, hatched) number of QCD jets reconstructed as tau candidate. The reconstruction fake rate is parametrized in bins of the track multiplicity.
Right: The ratio estimated number over expected number.



(a) (b)

(c) (d)

Figure 5.21: The ratio of estimated number of QCD jets reconstructed as tau candidate over the expected number is shown. The reconstruction fake rate is parametrized in bins of the track multiplicity but separately in the different E_T bins.

Chapter 6

Summary and Outlook

The number of misidentified tau leptons caused by QCD jets, which are reconstructed as tau candidates and identified as tau leptons, can be determined using the additional jet in $Z \rightarrow ee + jet$ events. The identification fake rate is parametrized in the three calorimeter variables R_{EM}, W_{strip}, f_{iso}, which are used at the tau calorimeter-based identification process. The number of fake tau candidates, which passed the tau identification can be estimated at a level of 6% for the loose threshold if including all prongs. In case of the medium threshold the estimation is at a level of 10%, while for the tight threshold the deviation is up to 30%. A parametrization in bins of E_{T} performs not sufficiently. The numbers of QCD jets, which are reconstructed as tau candidate can be estimated by using a fake rate as a function of the track multiplicity separately for different E_{T} ranges. Here the estimation can be done at a level of 10%. The statistic at the fake rate determination is not sufficient for $10\,\mathrm{pb}^{-1}$of data caused by the low cross section of $Z \rightarrow ee + jet$ events.

This study is influenced by the differences between the $Z \rightarrow ee + jet$ events and the QCD dijet events. Further studies have to determine, whether the results of the presented method are depended on Monte Carlo methods used to generate the event samples. Furthermore the origins of the jets in $Z \rightarrow ee + jet$ events and QCD dijets events can be study in more detail, to find possible quantities the misidentification rate can be parametrized sufficient. Also other combination of variables or using more than three can be studied.

Appendix A

Tau Identification Cuts

Cuts for the calorimeter-based safe tau identification method as implemented in the ATHENA tauRec package [28]:

	tight		medium		loose	
$E_T$range in GeV	1 prong	>1 prong	1 prong	>1 prong	1 prong	>1 prong
10–25	0.055	0.096	0.072	0.15	0.11	0.28
25–45	0.048	0.068	0.058	0.088	0.087	0.15
45–70	0.037	0.053	0.050	0.071	0.10	0.25
70–100	0.045	0.048	0.045	0.061	0.15	0.23
> 100	0.034	0.036	0.034	0.061	0.50	0.071

Table A.1: R_{em}

	tight		medium		loose	
$E_T$range in GeV	1 prong	>1 prong	1 prong	>1 prong	1 prong	>1 prong
10–25	0.32	0.62	0.65	0.90	0.75	0.96
25–45	0.20	0.42	0.24	0.68	0.89	0.79
45–70	0.26	0.25	0.26	0.25	0.31	0.25
70–100	0.058	0.18	0.14	0.20	0.70	0.50
> 100	0.049	0.18	0.60	0.20	0.60	0.47

Table A.2: f_{iso}

	tight		medium		loose	
$E_T$range in GeV	1 prong	>1 prong	1 prong	>1 prong	1 prong	>1 prong
10–25	0.033	0.048	0.045	0.056	0.047	0.057
25–45	0.0030	0.044	0.043	0.044	0.048	0.045
45–70	0.035	0.026	0.035	0.031	0.050	0.031
70–100	0.015	0.035	0.036	0.035	0.040	0.036
> 100	0.00069	0.030	0.030	0.037	0.030	0.038

Table A.3: W_{strip}

Appendix B

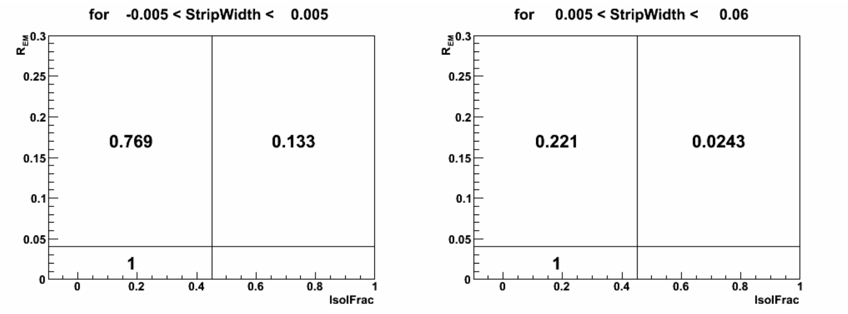
Datasets

```
mc08.105009.J0_pythia_jetjet.merge.AOD.e344_s479_r635_t53
mc08.105010.J1_pythia_jetjet.merge.AOD.e344_s479_r635_t53
mc08.105011.J2_pythia_jetjet.merge.AOD.e344_s479_r635_t53
mc08.105012.J3_pythia_jetjet.merge.AOD.e344_s479_r635_t53
mc08.105013.J4_pythia_jetjet.merge.AOD.e344_s479_r635_t53
mc08.105013.J5_pythia_jetjet.merge.AOD.e344_s479_r635_t53
mc08.106050.PythiaZee_1Lepton.recon.AOD.e347_s462_r635
```

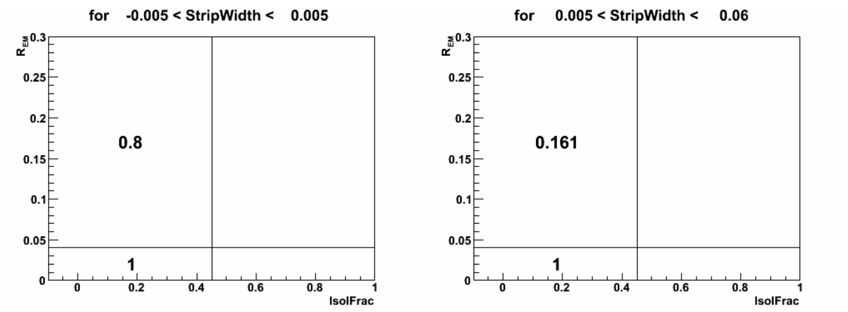
Appendix C

Identification Fake Rate

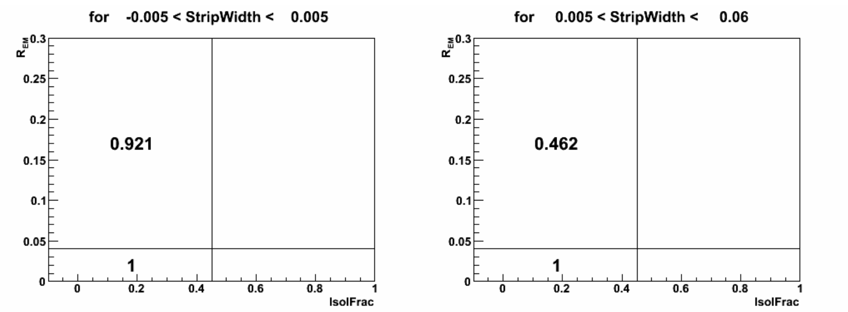
The bins are $10-25$ GeV (E_T bin 1), $25-45$ GeV (E_T bin 2), $45-70$ GeV (E_T bin 3)and $70-100$ GeV (E_T bin 4). In bins where no value is printed the identification fake rate is not defined.



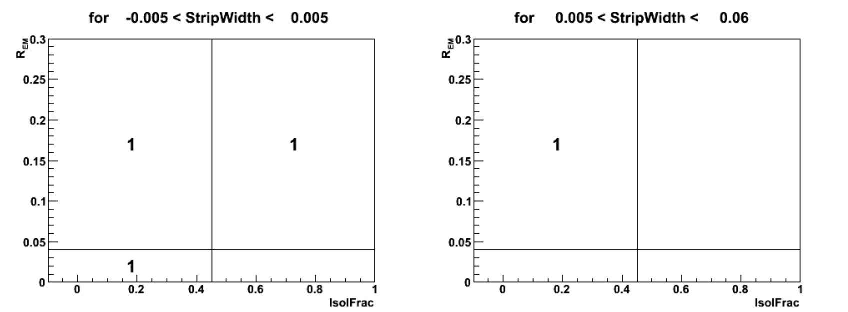
E_T bin 1



E_T bin 2

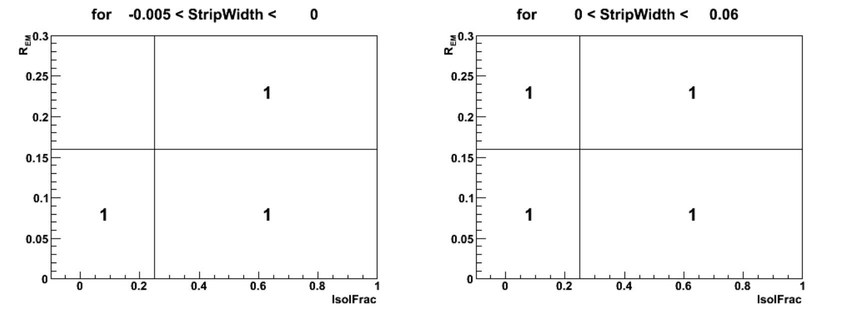


E_T bin 3

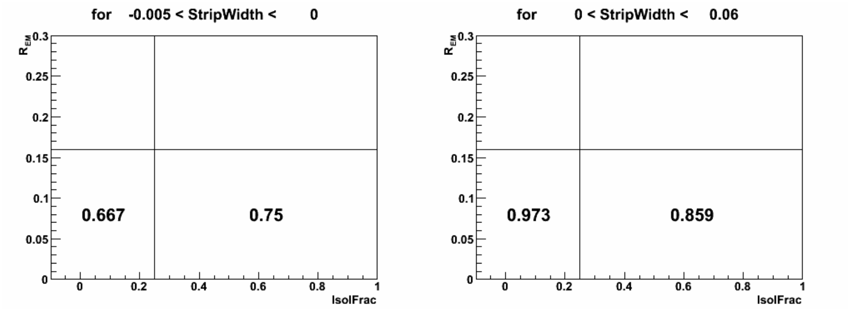


E_T bin 4

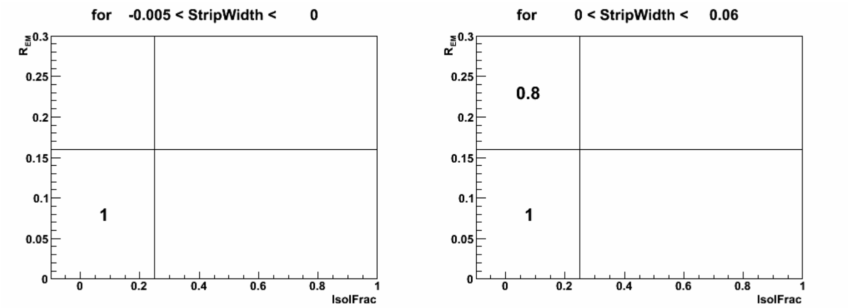
Figure C.1: calorimeter-based loose identification fake rate (1-prong)



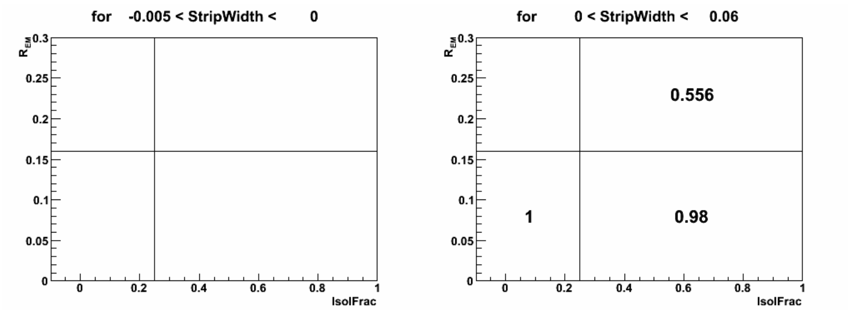
E_T bin 1



E_T bin 2

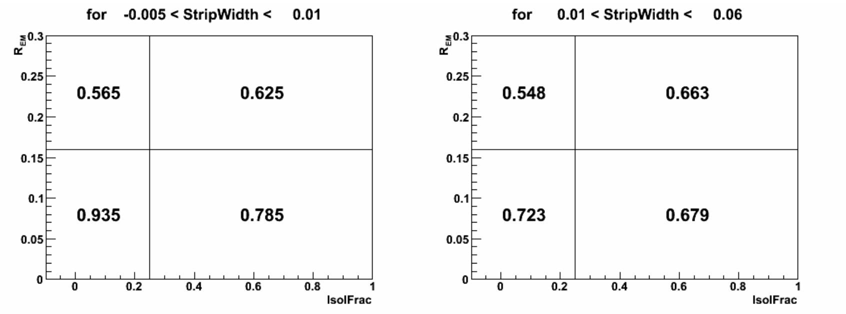


E_T bin 3

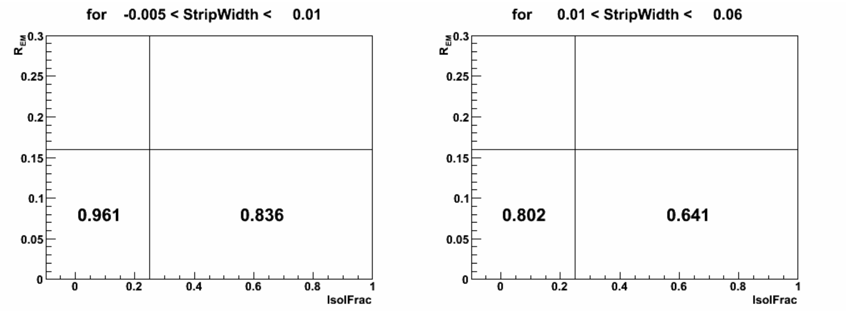


E_T bin 4

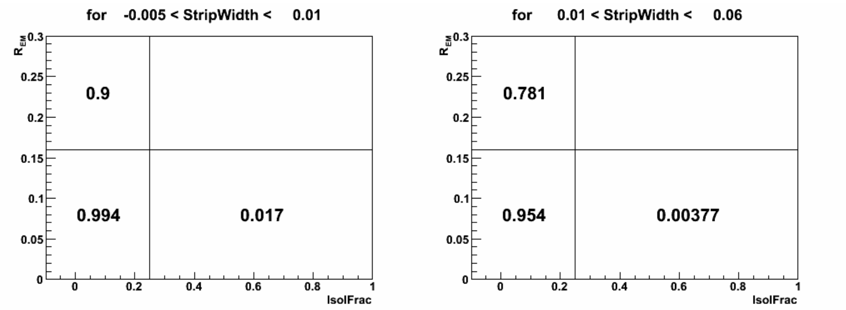
Figure C.2: calorimeter-based loose identification fake rate (3-prongs)



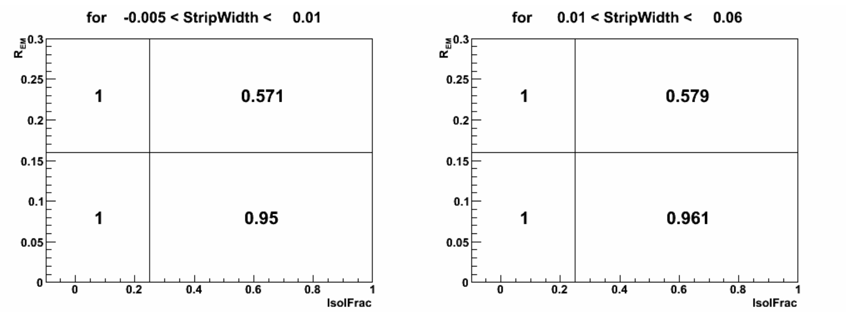
E_{T} bin 1



E_{T} bin 2

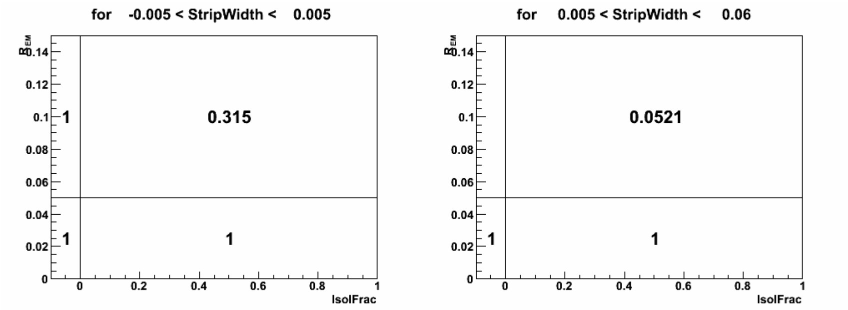


E_{T} bin 3

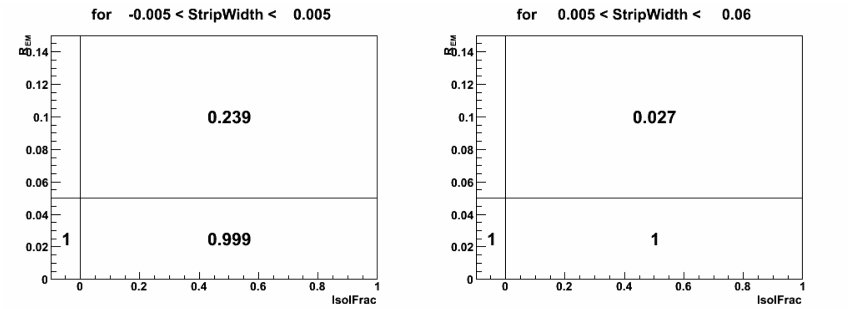


E_{T} bin 4

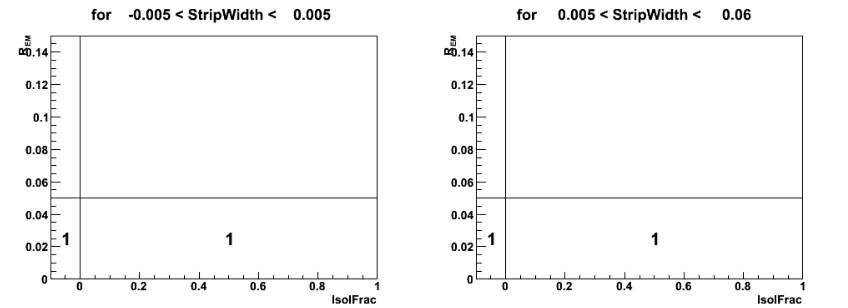
Figure C.3: calorimeter-based loose identification fake rate (all-prongs)



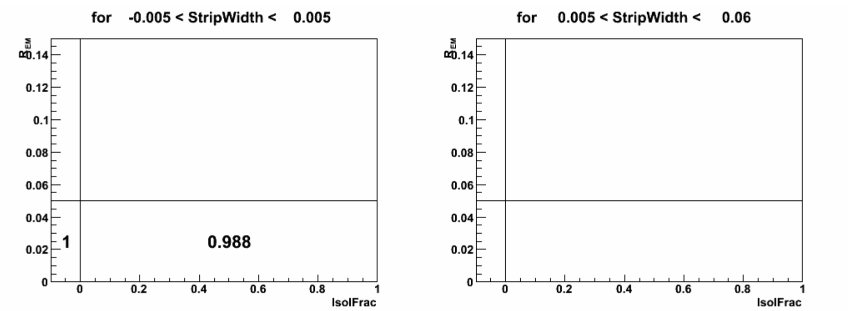
E_T bin 1



E_T bin 2

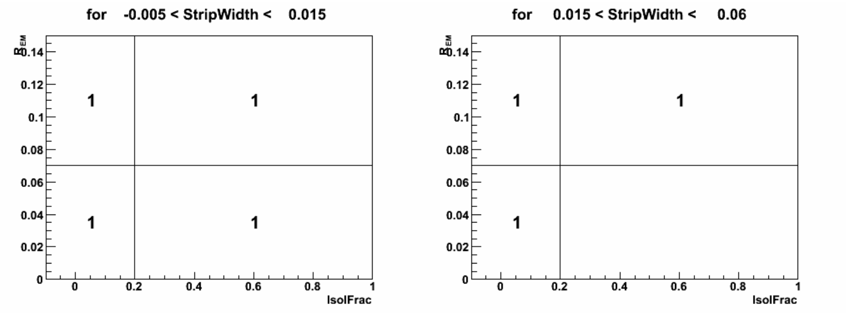


E_T bin 3

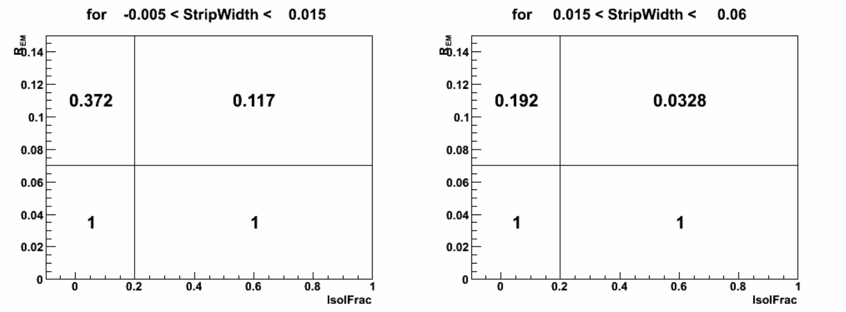


E_T bin 4

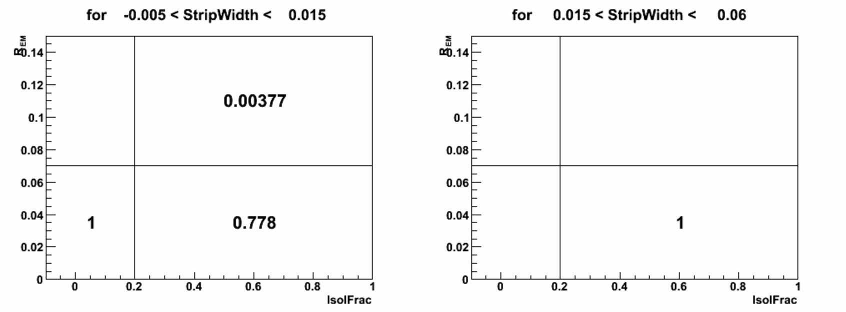
Figure C.4: calorimeter-based medium identification fake rate (1-prong)



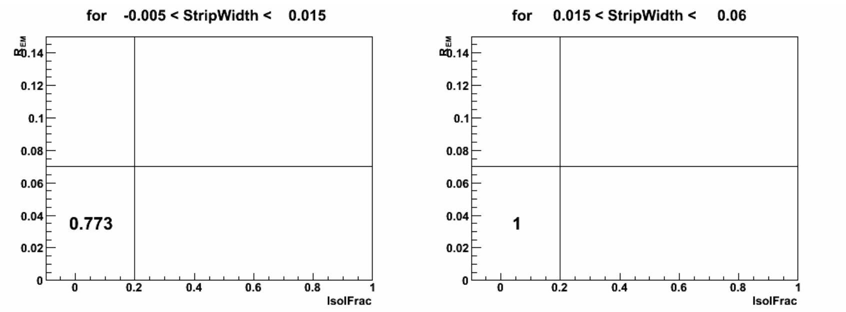
E_{T} bin 1



E_{T} bin 2

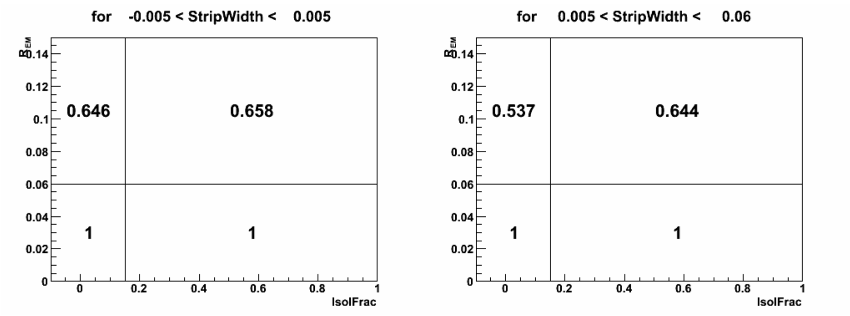


E_{T} bin 3

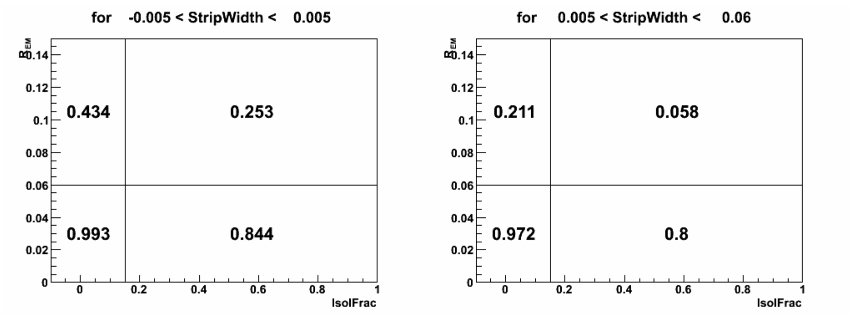


E_{T} bin 4

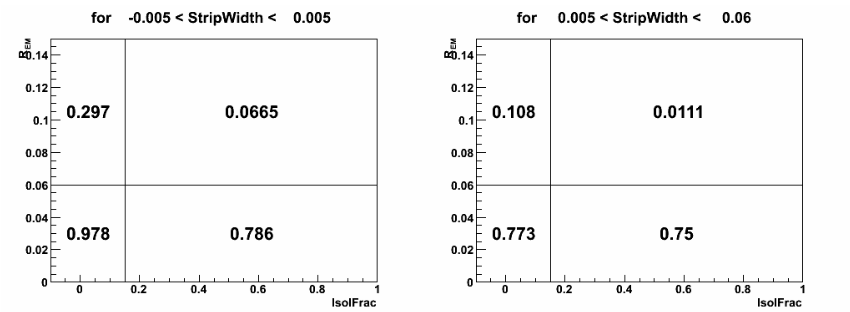
Figure C.5: calorimeter-based medium identification fake rate (3-prongs)



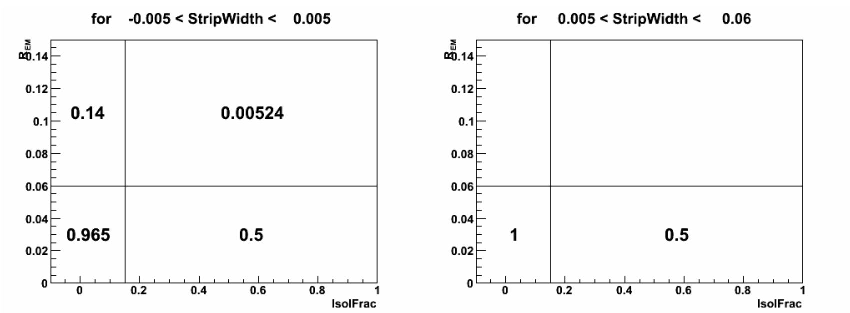
E_T bin 1



E_T bin 2

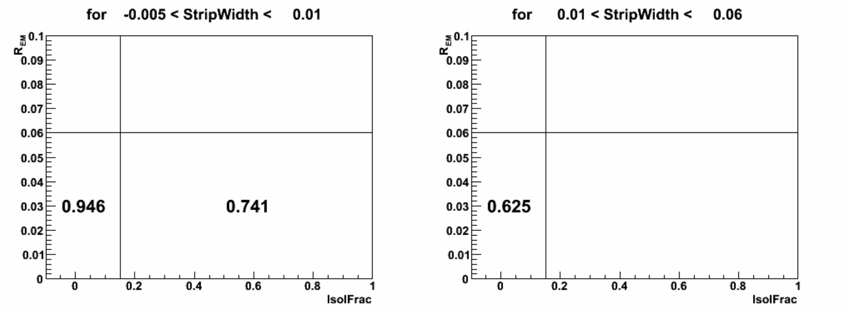


E_T bin 3

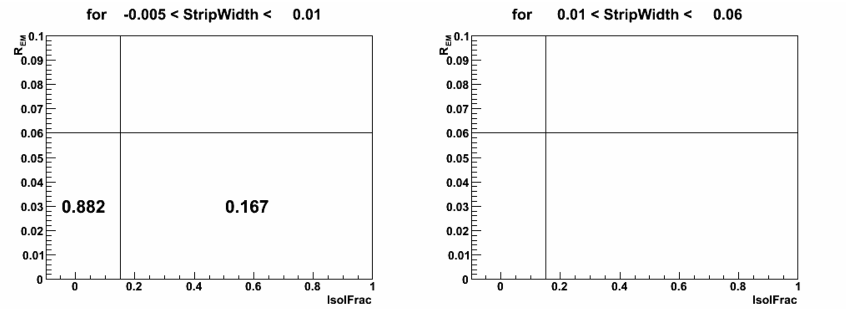


E_T bin 4

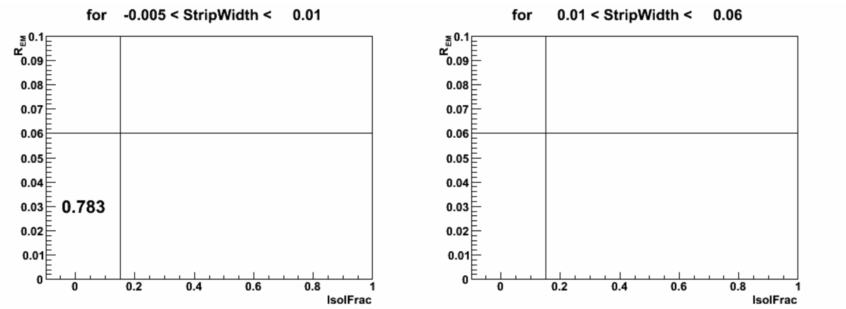
Figure C.6: calorimeter-based medium identification fake rate (all-prongs)



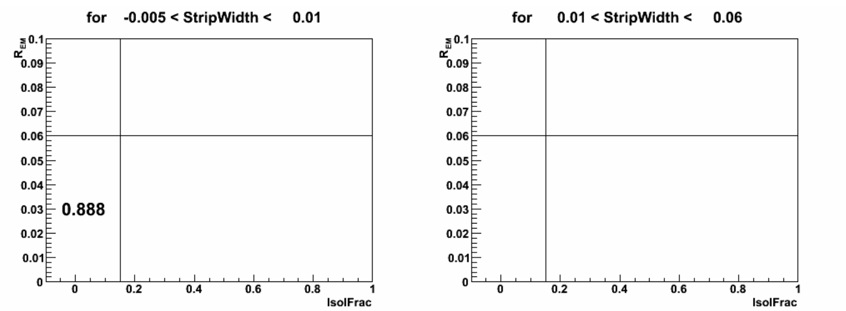
E_T bin 1



E_T bin 2

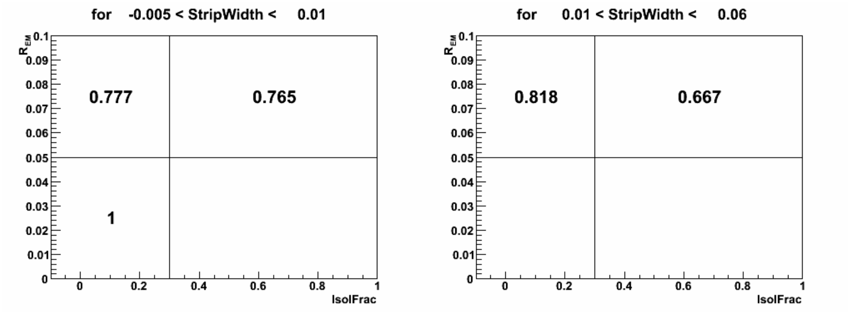


E_T bin 3

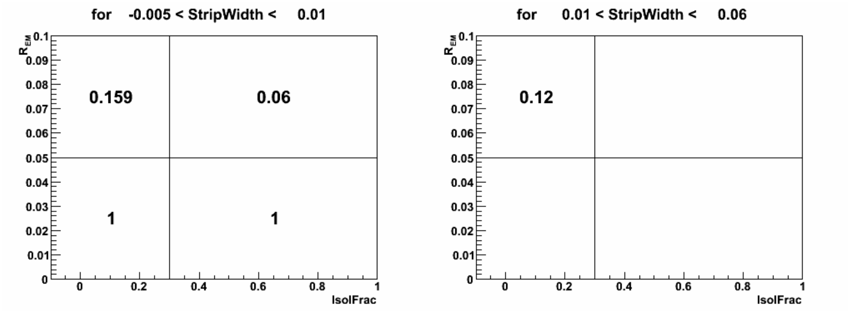


E_T bin 4

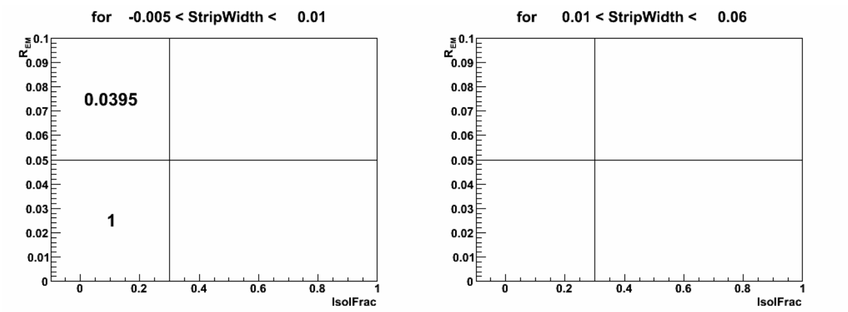
Figure C.7: calorimeter-based tight identification fake rate (1-prong)



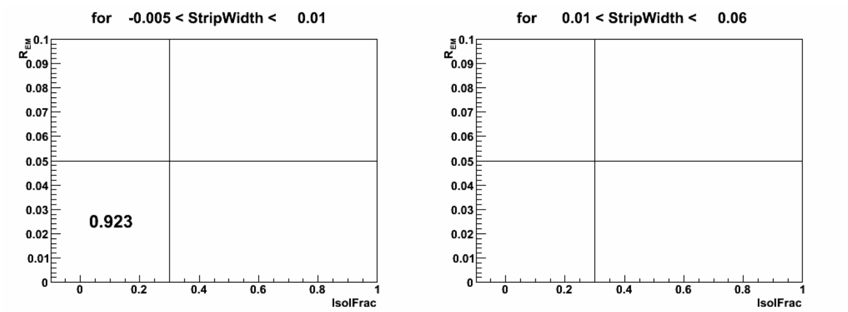
E_T bin 1



E_T bin 2

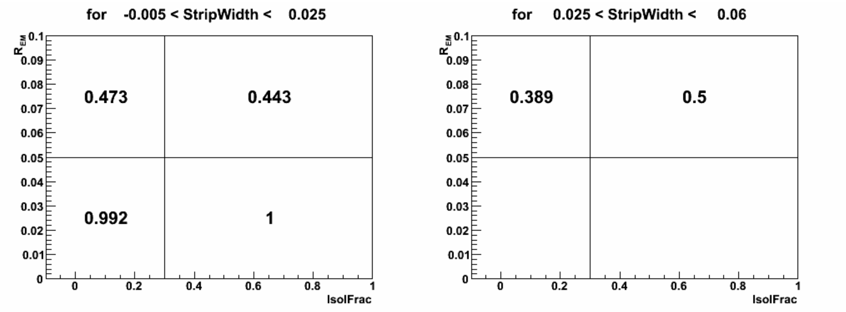


E_T bin 3

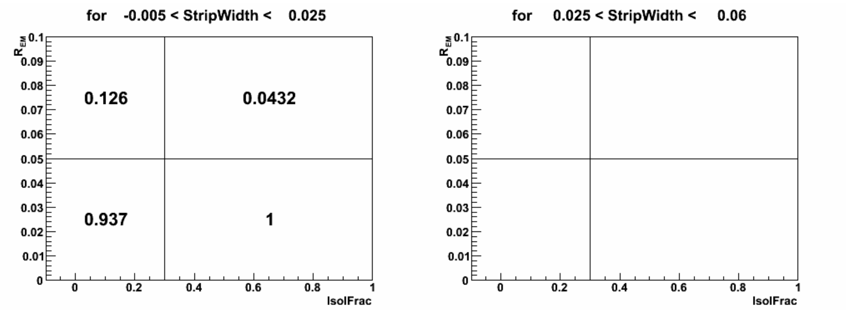


E_T bin 4

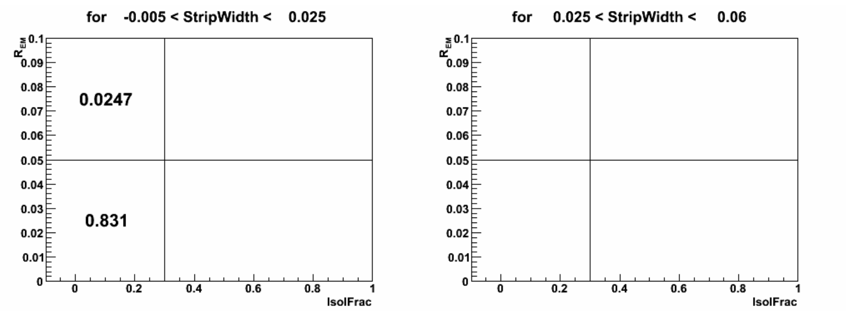
Figure C.8: calorimeter-based tight identification fake rate (3-prongs)



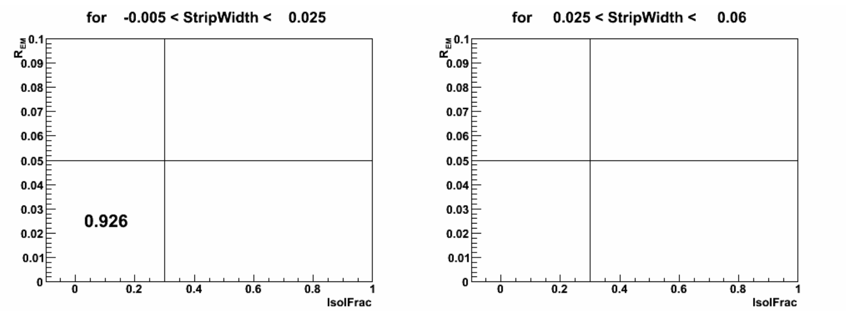
E_T bin 1



E_T bin 2



E_T bin 3



E_T bin 4

Figure C.9: calorimeter-based tight identification fake rate (all-prongs)

Bibliography

[1] D. Griffiths, *Introduction to Elementary Particles*. Wiley-VCH, 2nd ed., 1987.

[2] C. A. et al. (Particle Data Group), *Review of Particle Physics*, Physics Letters B667, 1 (2008) . URL: `http://pdg.lbl.gov`.

[3] I. J. R. Aitchison and A. J. G. Hey, *Gauge Theories in Particle Physics: A Practical Introduction: Volume 2, Non-Abelian Gauge Theories: Qcd and the Electoweak Theory*. Institute of Physics Publishing, 3rd edition ed., 2004.

[4] P. W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. **145** (May, 1966) 1156–1163.

[5] E. A. Paschos, *Electroweak theory*. Cambridge, UK: University Press, 2007.

[6] LEP Working Group for Higgs boson searches Collaboration, R. Barate et al., *Search for the standard model Higgs boson at LEP*, Phys. Lett. **B565** (2003) 61–75, `arXiv:hep-ex/0306033`.

[7] CDF Collaboration and D0 Collaboration Collaboration, *Combined CDF and DZero Upper Limits on Standard Model Higgs-Boson Production with up to 4.2 fb-1 of Data*, `arXiv:0903.4001 [hep-ex]`.

[8] ATLAS Collaboration, *Expected performance of the ATLAS experiment: detector, trigger and physics*, Tech. Rep. arXiv:0901.0512. CERN-OPEN-2008-020, Geneva, 2009.

[9] Atlas Collaboration, E. Richter-Was, *Prospect for the Higgs searches with the ATLAS detector*, Acta Phys. Polon. **B40** (2009) 1909–1930, `arXiv:0903.4198 [hep-ex]`.

[10] (ed.) Evans and (ed.) Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[11] ATLAS Collaboration, G. Aad et al., *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[12] CMS Collaboration, R. Adolphi et al., *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004.

[13] ALICE Collaboration, K. Aamodt et al., *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.

[14] LHCb Collaboration, A. A. Alves et al., *The LHCb Detector at the LHC*, JINST **3** (2008) S08005.

[15] LHCf Collaboration, O. Adriani et al., *The LHCf detector at the CERN LHC*, JINST **3** (2008) S08006.

[16] TOTEM Collaboration, G. Anelli et al., *The TOTEM experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08007.

[17] M. Lamont, *LHC: status and commissioning plans*, arXiv:0906.0347v1 (2009) .

[18] R. Assmann, *LHC Status*, in *ATLAS-D Conference, Bonn*. 2009.

[19] J. M. Campbell, J. W. Huston, and W. J. Stirling, *Hard Interactions of Quarks and Gluons: A Primer for LHC Physics*, Rept. Prog. Phys. **70** (2007) 89, `arXiv:hep-ph/0611148`.

[20] ATLAS Collaboration, *ATLAS detector and physics performance: Technical Design Report, 1*, tech. rep., Geneva, 1999.

[21] T. Sjostrand, S. Mrenna, and P. Skands, *PYTHIA 6.4 Physics and Manual*, JHEP **05** (2006) 026, `arXiv:hep-ph/0603175`.

[22] A. Christov, Y. Coadou, Z. Czyczula, M. Janus, L. Janyst, A. Kaczmarska, A. Kalinowski, J. Komaragiri, S. Lai, W. Mader, N. Meyer, E. Richter-Was, C. Ruwiedel, and M. Wolter, *Performance of the tau reconstruction and identification algorithm with release 14.2.20 and mc08 data*, Tech. Rep. ATL-COM-PHYS-2009-229, CERN, Geneva, Apr, 2009.

[23] P. Bechtle, B. Gosdzik, and S. Lai, *Cut based identification of hadronic tau decays*, Tech. Rep. ATL-COM-PHYS-2009-583, CERN, Geneva, Oct, 2009.

[24] M. Uhlenbrock and M. Janus, *Combined Talk about Tau Fake Studies*, in *ATLAS-D Conference, Bonn*. 2009.

[25] S. D. Ellis, J. Huston, K. Hatakeyama, P. Loch, and M. Tonnesmann, *Jets in hadron-hadron collisions*, Prog. Part. Nucl. Phys. **60** (2008) 484–551, `arXiv:0712.2447 [hep-ph]`.

[26] M. Warsinsky, *Studies of b-associated production and muonic decays of neutral Higgs bosons at the ATLAS experiment within the Minimal Supersymmetric Standard Model*. PhD thesis, TU Dresden, Germany, 2008.

[27] `http://home.fnal.gov/~paterno/images/effic.pdf`.

[28] `http://alxr.usatlas.bnl.gov/lxr-stb4/source/atlas/Reconstruction/tauRec/python/TauCommonSetIsTau.py?v=release_15_2_0`.

Danksagung

Diese Arbeit hätte ich ohne die Unterstützung von vielen Menschen nicht durchführen können. Als erstes möchte ich mich bei meinen Betreuern Jun.-Prof. Arno Straessner und Dr. Wolfgang Mader bedanken, die mir die Möglichkeit gaben, diese Arbeit bei Ihnen anzufertigen und immer mit vielen Ratschlägen zur Seite standen. Ebenso gilt mein Dank Prof. Michael Kobel der als Instituts- und ATLAS Gruppenleiter für die guten Rahmenbedingungen der Teilchenphysik an der TU Dresden sorgte. Für die vielen wichtigen Hinweise während der Diplomarbeit, der unzähligen interessanten wissenschaftlichen und nicht so wissenschaftlichen Diskussionen im Büro und in der Mensa möchte ich auch meinen Kollegen danken. Insbesondere Dr. Deepak Kar, Frank Seifert, Michael Stoebe und Kathrin Leonhardt für die zügigen und wertvollen Korrekturen. Ohne diese wäre die Arbeit nicht in dem schon sehr knapp bemessenen Zeitrahmen fertig geworden. Auch möchte ich all jenen Danke sagen, die meine Zeit am IKTP begleitet haben sowie das Studium in Dresden angenehm machten.

Der größte Dank aber gilt meinen Eltern und meinem Bruder, die mich das gesamte Studium nicht nur finanziell, sondern vor allem auch moralisch unterstützt haben. Sie standen mir bei allen Problemen immer mit großer Liebe zu Seite und haben trotz der Überlänge meines Studiums nie die Geduld und den Glauben an mich verloren.

Selbständigkeitserklärung

Ich versichere hiermit, dass ich die vorliegende Arbeit selbständig und ohne unzulässige Hilfe Dritter angefertigt habe. Es wurden keine anderen als die angegeben Quellen und Hilfsmittel benutzt. Die aus fremden Quellen übernommenen Resultate sind als solche kenntlich gemacht.

Dresden, 15.12.2009

Felix Friedrich