Universität Hamburg

Reconstruction of Photon Conversions in τ Lepton Decays in the ATLAS Experiment

Diplomarbeit

im Studiengang Physik

vorgelegt von:
Michael Böhler

Oktober 2008

Institut für Experimentalphysik
MIN - Fakultät
Universität Hamburg

Gutachter der Diplomarbeit	:	Dr. P. Bechtle Prof. Dr. P. Schleper
Vorsitzender des Prüfungsausschusses	:	Prof. Dr. W. Hansen
Departmentleiter	:	Prof. Dr. J. Bartels

Abstract

The ATLAS experiment is one of the experiments at the Large Hadron Collider (LHC) which is designed for the search of new elementary particles. To discover the Higgs boson or precisely measure SUSY scenarios, τ lepton final states are very powerful decay channels. Therefore the τ lepton decay modes have to be identified correctly. Due to interactions between photons from hadronic decay products of the τ lepton and detector material electron-positron pairs (photon conversions) may be produced. These lead to additional charged tracks changing the reconstruced τ lepton track multiplicity.

To avoid such missidentifications, this thesis introduces an explicit photon conversion identification in the very dense τ lepton decay environment. Existing tools had to be modified and a new electron identification method has been developed especially for this task. As a first result, the corrected τ lepton track multiplicity is presented.

Zusammenfassung

Das ATLAS Experiment ist eines der Experimente am Large Hadron Collider (LHC), das für die Suche nach neuen Elementarteilchen entwickelt wurde. Um das Higgs Boson oder SUSY Szenarien präzise zu vermessen ist der Nachweis von Zerfallskanälen mit τ Leptonen im Endzustand sehr wichitg. Dafür müssen die τ Zerfallskanäle allerdings korrekt nachgewiesen werden. Durch Wechselwirkungnen zwischen Photon, die im Laufe des hadronischen Zerfalls von τ Leptonen entstehen, und Detektormaterial können jedoch Elektron-Positron Paare (Photon Konversionen) erzeugt werden. Diese führen zu zusätzlichen geladenen Spuren, die die Anzahl der τ Spuren verändern.

Um solche Fehlidentifikationen zu vermeiden führt diese Diplomarbeit eine explizite Photon Konversions Erkennung innerhalb des sehr dichten τ Zerfallskegel ein. Bereits bestehende Hilfsprogramme mussten angepasst und eine neue Elektronenidentifikationsmethode, eigens für dieses Problem, eingeführt werden. Als erstes Ergebnis wird die korrigierte Anzahl der τ Spuren präsentiert.

Contents

Chapter 1

Introduction

The goal of elementary particle physics is to describe the fundamental constituents of nature and the observed forces between them. The current understanding of the theory of particle physics is specified by the Standard Model (SM) of the particle physics which is one of the most successful theories in the history of science. It has been tested with very high presicion in multiple experiments over a wide energy range. Almost all its predictions have been verified. The only missing particle, the Higgs boson, could not be discoverd yet. The so-called Higgs mechanism provides the explanation for the existence of the mass of elementary particles. Even the discovery of the SM Higgs will not solve all open questions of particle physics. In spite of the great success of the predictions of the SM, there are still theoretical concepts which cannot be included in the SM. For instance the existence of dark matter and dark energy in our universe, the hierarchy problem, and a theory of gravity, only to mention a few of them, cannot be described by the SM. However any new theory has to include the SM as low-energy limit. The most promising extension of the SM is Supersymmetry (SUSY), a new symmetry connecting fermions (matter particles) with bosons (force carriers). The next generation of particle accelerators, the Large Hadron Collider (LHC), which is in its commissioning phase right now, will reach the high energy and luminosity needed to investigate these new phenomena and discover or exclude the SM Higgs boson.

The LHC is built at CERN[1], near Geneva, across the French-Swiss boarder. The first collsions are expected in the spring of 2009. The LHC will extend the frontiers of particle physics. Bunches of up to 10^{11} protons will collide every 25 ns to provide a center of mass energy of 14 TeV at a design luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$. The high interaction rates and radiation doses, as well as the needed precision measurements, yield very high requirements on the detectors. Two multi purpose detectors, ATLAS (A Toroidal LHC ApparatuS) and CMS (Compact Muon Solenoid) are installed at the LHC.

The data provided by the detectors has to be analysed via specific algorithms to filter and to reconstruct the events the experiments are searching for. The fundametal tasks of the algorithms are to reconstruct and identify objects like electrons, muons and photons. The reconstruction and identification of τ leptons is much more complicated, due ti its decay. The τ lepton is the most massive lepton and its branching fraction of decays into hadrons is approximatly 65% [1]. The highest cross section at a hadron collider is the cross section of QCD jet production. Thus the differentiation of τ leptons

[1] European Organization for Nuclear Research

from QCD background jets is the biggest challenge of the τ lepton reconstruction due to the similar signatures of QCD jets and τ leptons decaying into hadrons. Nevertheless, it is important to have a powerful τ reconstruction algorithm because τ leptons play a decisive role in the search for the SM Higgs boson, heavy SUSY Higgs bosons, and SUSY scenarios. If the SM Higgs boson is in the mass range below 135 GeV, the τ lepton final states are very promissing channels for discovering the Higgs boson. Also many SUSY scenarios will produce multi-τ final states. In case of a τ pair production, informations on the spin correlation of the decaying particle can be determined. Therefore the τ decay modes have to be reconstructed correctly. Apart from the QCD background, additional compications have to be taken into account. If additionally to the charged pion a neutral pion (π^0) is produced during the τ decay, the π^0 decays with a brachning fraction of 98% into two photons. If at least one of the two photons interacts with the detector material and converts into an electron-positron pair, two additional charged tracks are observed within the decay cone of the τ lepton. This may lead to the missidentification of the decay mode. The goal of this thesis is to develop an explicit photon conversion identification in the environment of the τ decay cone to improve the τ lepton reconstruction at the ATLAS experiment.

The thesis is organised as follows: In Chap. 2 a very brief overview of the Standard Model of particle physics is given. Then an introduction of Supersymmetry with respect to τ final states follows. The chapter concludes with a more detail discussion about τ lepton decays, photon conversions, and Z^0 production at pp-colliders. Chapter 3 explains the experimental setup. At first the LHC is described. Then after a general introduction of particle identification the ATLAS detector is discussed in more detail. Chapter 4 describes the ATLAS Event Data Model. Both needed reconstruction algorithms, *TauRec* for τ reconstruction and the photon conversion reconstruction algorithm, are explained in Chap. 5. New software tools had to be developed especially for this study. The basic construction of the newly implemented tools are decribed in Chap. 6. Chapter 7 describes the improvment of the reconstruction and identification of photon conversions inside the τ decay cone. After achieving sufficient results in the photon conversion identification, the τ track multiplicity after an explicit conversion veto is discussed. Finally all results are summarised and an outlook of the ongoing work is given in the last chapter.

Chapter 2

Theoretical Overview

This chapter will give a short overview of the current understanding of the Standard Model (SM) of elementary particle physics. After introducing the particles and interactions of the SM it will briefly discuss the origin of mass and the Higgs mechanismn. In a section about Supersymmetry (SUSY) a possible extension of the SM will be presented with a special emphasis on the role of the τ reconstruction for SUSY discovery. The last section of this chapter covers the theoretical understanding of the main topic of this thesis, namely the τ-decay, photon conversions and the description of the Z boson production at pp coliders.

2.1 Introduction to the Standard Model

The Standard Model contains the electromagnetic, weak and strong interactions of elementary particle physics. Each interaction can be described by a quantum field theory based on local gauge invariance. An equation is defined as gauge invariant if it is not changed by a phase transformation. It is called local gauge invariant if the phase transformation is space- and time-dependent.

The wave equation of elementary particles with spin 1/2 without any interactions can be calculated with the Dirac-Equation:

$$(i\gamma_\mu \partial^\mu - m)\,\varphi = 0 \tag{2.1}$$

with the so-called *Gamma* matrices γ_μ, ∂^μ is the covariant derivative, the mass m of the particle and the particle wave function φ.

Describing the forces by gauge theories for each gauge field a gauge boson has to be introduced. Then the forces of each interaction can be described by the exchange of such gauge bosons (c.p. Tab. 2.1).

All known matter is built of a few elemetary particles. As known from chemistry all matter consists of atoms. These atoms again consist of eletrons, protons and neutrons. The electrons do not show any substructure and depend to the elementary particles, to the so-called leptons. Protons and neutrons have a substructure. They are composed of quarks. Altogether six leptons and six quarks are known. Both kind of particles have spin $\frac{1}{2}$ and are fermions. They are arranged in three generations. To each fermion exists an anti-fermion with inconverse quantum numbers, color in case of quarks and third component of weak isospin (cp. Tab. 2.3).

Table 2.1: *Standard Model gauge bosons and their properties.*

Interaction	couples to	Boson	Mass [GeV/c^2]	Electric Charge	Spin
electromag.	electric charge	photon (γ)	0	0	1
weak	weak charge	$W^{\pm}$, Z^0	80.4, 91.2	± 1, 0	1
strong	color	8 gluons (g)	0	0	1

Table 2.2: *Leptons and Quarks the particles of Standard Model with their properties.*

Familiy	Leptons	el. charge	Mass [MeV]	Quarks	el. charge	Mass [MeV]
1	ν_e	0	$< 2.2 \cdot 10^{-3}$	u	2/3	1.5 to 3.3
	e	-1	0.511	d	$-1/3$	3.5 to 6.0
2	ν_μ	0	< 0.170	c	2/3	1270
	μ	-1	105.7	s	$-1/3$	104
3	ν_τ	0	< 15.5	t	2/3	171200
	τ	-1	1776.8	b	$-1/3$	4200

2.1.1 Electroweak Interaction

The electromagnetic and weak interaction can be unified in the electroweak theory [2]. Due to the fact that weak interaction violates parity, the left-handed components are arranged in isospin doublets and the right-handed in singulets. In order to conserve the local gauge invariance a tripplet W_1^μ, W_2^μ, W_3^μ of vector fields has to be introduced for $SU(2)_L$ and a singulet B^μ for $U(1)_Y$. $SU(2)_L$ describes the transformation of the left-handed multiplets of the weak isospin I. The $U(1)_Y$ theory covers the weak hypercharge Y.

The covariant derivative in the electroweak theory has to be changed to solve the Dirac equation (c.p. to Equ. 2.1):

$$D^\mu = \partial^\mu + igT \cdot W^\mu + i\frac{g'}{2} Y B^\mu \tag{2.2}$$

with T = $\frac{\tau}{2}$ (τ Pauli-Matricies), Y = -1 for left-handed leptons and T = 0, Y = -2 for right-handed charged leptons.

The gauge bosons of the weak interaction are two charged $W^\pm$ bosons and one neutral Z^0. With respect to the coupling to the left-handed leptons the fields of the $W^\pm$ are mixings of the two fields $W_{1,2}^\mu$

$$W^{(\pm)\mu} = \frac{1}{\sqrt{2}}(W_1^\mu \pm W_2^\mu) \tag{2.3}$$

The other two fields W_3^μ and B^μ mix to Z^μ the field of the neutral Z^0 boson and to A^μ the photon field:

$$\begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix} = \begin{pmatrix} \cos\theta_W & \sin\theta_W \\ -\sin\theta_W & \cos\theta_W \end{pmatrix} \begin{pmatrix} B^\mu \\ W_3^\mu \end{pmatrix} \tag{2.4}$$

Table 2.3: *Standard Model fermions and their properties.*

Fermion	generation 1 2 3	elec. charge	color charge	weak lefth.	isospin righth.	spin
leptons	ν_e ν_μ ν_τ	0		$\frac{1}{2}$	-	$\frac{1}{2}$
	e μ τ	-1		$\frac{1}{2}$	0	$\frac{1}{2}$
quarks	u c t	$+\frac{2}{3}$	r,g,b	$\frac{1}{2}$	0	$\frac{1}{2}$
	d s b	$-\frac{1}{3}$	r,g,b	$\frac{1}{2}$	0	$\frac{1}{2}$

with θ_W the weak mixing angle (or Weinberg angle) which is defined by the following connection:

$$\cos\theta_W = \frac{g}{\sqrt{g^2+g'^2}}, \quad \sin\theta_W = \frac{g'}{\sqrt{g^2+g'^2}} \tag{2.5}$$

The Weinberg angle [3] is a free parameter of the SM and has to be determined experimentally. The bosons of the weak interaction have charge and can couple to each other. Their range is limited by their huge mass, it is $\approx 10^{-3}$ fm. Due to the fact that photons do not have any charge and mass their range is infinite.

As shown in Tab. 2.3 the quarks have weak isospin and hypercharge. However, their mass eigenstates are not identical with those of the weak interaction. By definition the weak eigenstates of the down type quarks are mixings and can be calculated with the Cabbibo-Kobayashi-Maskawa (CKM) matrix [4]:

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \tag{2.6}$$

The CKM Matrix is unitary and can be determined by tree mixing angles and the CP-violating phase. As an extension to the SM the observed neutrino oscillations can be described by a mixing of the neutrino mass eigenstates to their weak eigenstates. Such a mixing of neutrinos is comparable to the quark mixing. To determine the weak eigenstates of neutrinos the Maki-Nakagawa-Sakata (MNS) matrix is used.

Nevertheless the described gauge theories above only work with the assumption of massless neutrinos. Thus neutrinos occur only in lefthanded states.

2.1.2 Strong Interaction

The strong interaction is described by Quantum Chromo Dynamics (QCD) [5] which is based on a non-abelian gauge theory with a $SU(3)_C$ symmertry, where C denotes colour. To describe the quark model the mentioned $SU(3)_C$ symmertry has to be assumed. Therefore the derivative ∂^μ in the Dirac equation (cp. Equ. 2.1) has to be replaced by D^μ to assure the gauge invariance of the equation.

$$D^\mu = \partial^\mu + \frac{ig_s}{2}(\lambda_1 G_1^\mu + ... + \lambda_8 G_8^\mu) \equiv \partial^\mu + \frac{ig_s}{2}\lambda_j G_j^\mu \tag{2.7}$$

Then the gauge fields of the strong interaction can be described by gauge bosons. The gauge bosons of QCD are 8 massless, coloured gluons g_μ^i with i = 1,..,8 which couple to the colour charge. Due to the fact that gluons are coloured they interact with each other. Therefore the interaction range is very short ≤ 1 fm.

2.1.3 A Combined Symmetry of the Standard Model

Thus the Standard Model can be summarised in the symmetry group:

$$SU(3)_C \times SU(2)_L \times U(1)_Y. \tag{2.8}$$

with the gauge bosons:

- electronmagnetic interaction: massless photon γ
- weak interaction: massive bosons $W^{\pm}$ and Z^0
- strong interaction: 8 massless gluons g^i_μ

The description of the SM in form of the $SU(3)_C \times SU(2)_L \times U(1)_Y$ summetry group leads to the following problem: For an exact symmetry the gauge invariance is only preserved for massless gauge bosons. Additionally the SU(2) transformations of left-handed fermion doublets will only be gauge invariant if the particles in one doublet have the same masses. The neutrino is defined as massless in the SM so the electron would have to be massless too.

The only way to keep the definition of the SM described before with massive gauge bosons, leptons and quarks is to assume a not exact but broken symmetry. The Higgs mechanism is able to introduce this symmetry breaking [6].

2.1.4 Higgs Mechanism

The Higgs mechanism introduces the mass of the particles as an interaction with a background field with a non-vanishing vacuum expectation value (v.e.v.). The most simple Higgs structure which ensures a non-vanishing vacuum expectation value is a weak isospin doublet of complex fields (four states).

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \quad (I = \frac{1}{2}, Y = 1) \tag{2.9}$$

The following potential can be defined:

$$V(\phi) = \frac{\mu}{2}\phi^+\phi + \frac{\lambda}{4}(\phi^+\phi)^2, \quad \text{with } \lambda > 0 \text{ and } \mu^2 < 0 \tag{2.10}$$

and a v.e.v. of $v = \sqrt{\frac{-\mu^2}{\lambda}}$. The potential is shown in Fig. 2.1. The Higgs field has to be a doublet because it should give a mass to the gauge bosons of $\mathrm{SU(2)}_L$ and $\mathrm{U(1)}_Y$. Each of the four massless vector bosons has two degrees of freedom (polarisation) before getting its mass, so eight degrees of freedom (dof) are needed. The Higgs doublet has two complex (4 real Higgs) fields. This makes 12 degrees of freedom described by the Higgs doublet. After the interaction of the bosons with the Higgs fields the $\mathrm{W}^{\pm}$ and Z^0 are massive, three times three degrees of freedom are absorbed. The massless photon needs two degrees of freedom for the polarisation. To sum up: nine dof for the massive bosons and two for the photon makes 11. Thus there is one degree of freedom left for a neutral and spinless Higgs particle.

With the minimal Higgs mechanism the predicted masses of the gauge bosons are:

$$M_A = 0, \quad M_W^{\pm} = \frac{gv}{2}, \quad M_Z^0 = \frac{\sqrt{g^2 + g'^2}v}{2} \tag{2.11}$$

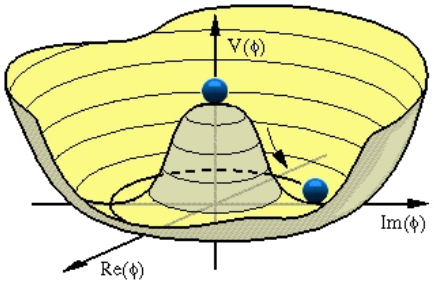

Figure 2.1: *It is shown hhe Higgs Potential. The rotational symmetry is broken spontaneously.*

With respect to Equ. 2.5 the relation of M_Z^0 and $M_W^{\pm}$ is: $M_Z^0 = \frac{M_W}{\cos\theta_W}$. This relation has been measured very precisely by the LEP experiments at CERN [7]. The mass predicted by theory of the Higgs boson is $m_H = 2\lambda v^2$. Up to now, the Higgs particle is the only particle of the SM which has not been experimentally discovered yet. The upper and lower bounds of the Higgs mass are shown in Fig. 2.2.

The lower limit of the Higgs mass is $m_H > 114.4\,\text{GeV}$ [8, 9] with a 95% CL. This is the result of direct Higgs searches at LEP-2. The upper limit of the Higgs mass is an indirect precision measurement using LEP and SLD, CDF, and D0 data [8]. The limit is $m_H < 154\,\text{GeV}$ and $m_H < 185\,\text{GeV}$ respectively, including the LEP-2 direct search limit of 114 GeV.

2.1.5 Shortcomings of the Standard Model

Although the Standard Model is a very successful theory which describes and predicts a lot of the observabels of the elementary particle physics, so far there are a few phenomena and theoretical aspects not included or explained by the SM.

1. There is no explanation of the ratio of matter/ anti-mattter in the universe [10] if neglecting an imbalance as initial condition. Even the CP-violation in the CKM mechanism, which is a possible candidate, cannot offer a solution because it is orders of magnitude too weak. Up to now no process for baryon number violation is known.

2. Measurements from astro physics predict that only 4% of the amount of energy in the universe is stored in matter, which is described by the SM. The other 96% are described by the Cosmological Standard Model, which predicts the consistence of the universe as 23 % dark matter, 73% dark energy and the rest as the visible matter. The weakly interacting neutrinos which are the only candidates in the SM for such dark matter are by several orders of magnitude to light to yield the amout of deterimned dark matter. So there is no explanation for 96 % of the energy in our universe by the SM [11].

3. Hierachy problem: If the SM should be valid up to the Planck scale $\Lambda_P = 10^{19}\,\text{GeV}$ the Higgs mass $M^2_{H_{bare}}$ must be tuned: $M_H^2 = M^2_{H_{bare}} + \delta M_H^2$. If the Higgs mass H_M is in electroweak scale the tuning δM_H^2 must be in order of $10^{19}\,\text{GeV}$. This

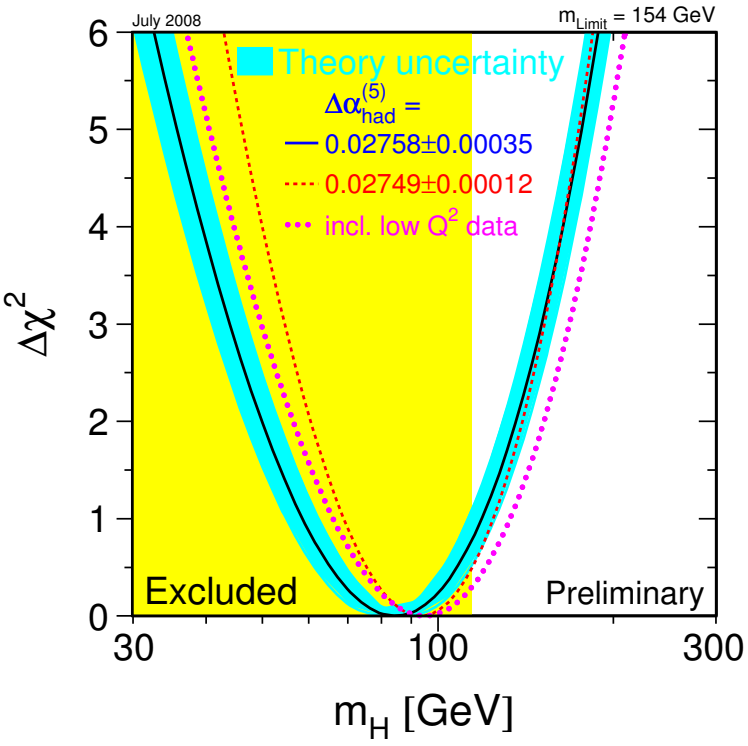

Figure 2.2: *Experimental limits of the Higgs mass [9]. The plot shows $\Delta\chi^2$ curve derived from precision electroweak measurements, performed at LEP and by SLD, CDF, and D0, as a function of the Higgs-boson mass, assuming the Standard Model. The preferred Higgs mass, the minimum of the curve, is at 84 GeV, with an experimental uncertainty of +34 and -26 GeV. The black line shows the value neglecting theoretical uncertainties. Taking these uncertainties into account the blue band has to be added. The Higgs mass is lower than about 154 GeV with an one-side 95 % confidence level including both the experimental and the theoretical uncertainty. The Higgs mass limit increases to 185 GeV when including the LEP-2 direct search limit of 114 GeV shown in yellow.*

kind of finetuning is not very elegant. But the SM does not give another possible explanation.

4. Trying to unify the gauge theories of the SM the Grand Unified Theories (GUT) has been defined. If such an unification exits the three coupling constants α_1, α_2 and α_3 should be descibed by one coupling at the GUT scale. Extrapolating therefore the three coupling constants to the GUT scale, they do not converge at one point (c.p. Fig. 2.3).

5. The SM is not able to include a theory of gravity.

As a possible extension of the SM the supersymmetry model described in the next section may solve some of the mentioned shortcomings of the SM.

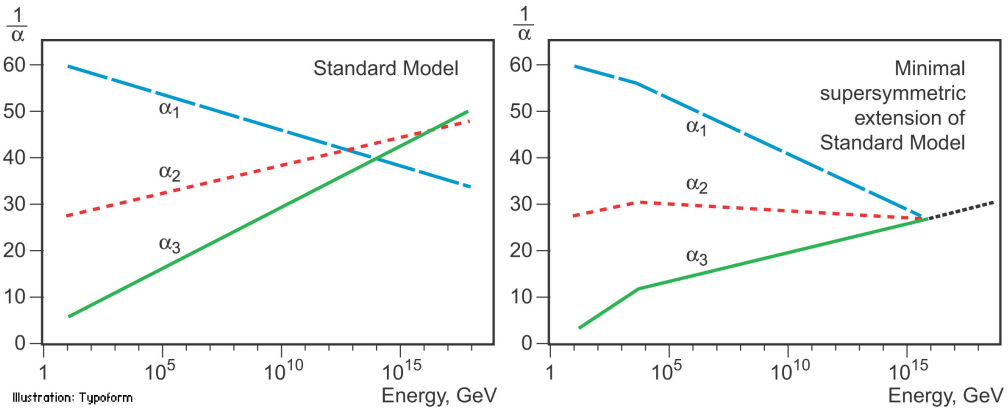

Figure 2.3: *Running inverse coupling constants in the Standard Model and the minimal supersymmetric model (MSSM). Without the additional loop contributions of the supersymmetric particles the coupling constants would not converge in one point for the SM. The unification of the coupling constants is only possible at the GUT scale in theories beyond the SM (e.g. MSSM) [12].*

2.2 Short Motivation for Theories beyond the Standard Model

Due to the success of the predictions of the SM a possible extension has to include the SM as a low-energy limit. The most promising extension is Supersymmetry (SUSY) [13]. SUSY introduces for each fermion a supersymmetric boson partner called *bino* and for each boson a supersymmetric fermion called *sfermion*. Thus for each particle with spin $\frac{1}{2}$ SUSY degrees of freedom with integer spin, the same quantum numbers and mass are added (c.p. Tab. 2.4) and spin $\frac{1}{2}$ SUSY particles for SM particles with integer spin. This is realized by the operator Q_α, which transforms a fermion into a boson and vice versa.

$$Q_\alpha |F\rangle = |B\rangle\,, \quad Q_\alpha |B\rangle = |F\rangle \tag{2.12}$$

By introducing these additional particles the divergencies responsible for the hierachy problem are canceled by the additional loop corrections of the additional fermions and bosons. Another advantage of this symmetry is that due to these additional loop contributions the running coupling constants α_i converge in one point (c.p. Fig 2.3), which is essential to define a grand unified theory.

A lot of shortcomings of the SM seems to be solved with this theory. The problem is that until now no SUSY partner of a SM particle has been discovered. The only possibility that this theory is realized in nature is a not exact but a broken symmetry. This means it exists a scale Λ_{SUSY} at which SUSY is broken. The scale Λ_{SUSY} should not be several orders of magnitude larger than the weak scale otherwise there will be a new hierarchy problem between the weak and SUSY scale. Also the unification of the coupling constants would not be possible at one point if Λ_{SUSY} reaches too large values.

Table 2.4: *Overview of SM particles with spin and their SUSY partners.*

Spin	SM particle	Superpartner	Spin
1/2	Leptons $(e, \nu_e, ...)$	Sleptons $(\tilde{e}, \tilde{\nu_e})$	0
1	Quarks $(u, d, ..)$	Squarks $(\tilde{u}, \tilde{d}, ..)$	1/2
0	Gluons	Gluinos	1/2
	W	Wino	
	Z^0	Zino	
	Photon (γ)	Photino $(\tilde{\gamma})$	
2	Graviton	Gravitino	3/2

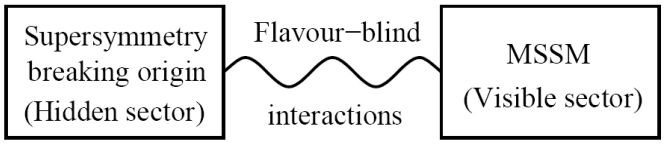

Figure 2.4: *The SUSY is broken in a hidden sector and mediated via a flavour-blind interaction to the visible sector. This interaction is described by a messenger field e.g a gauge field of gravity (mSUGRA) or new physics.*

SUSY Breaking Mechanism

The breaking mechanism can be described as it is shown in Fig. 2.4. SUSY is broken in a hidden sector and the breaking is mediated by a so-called messenger. This messenger brings the breaking by a flavour-blind interaction to the visible sector, the low-energy world.

A very promising candidate as messenger field is the gravity. It is called gravity mediated SUSY breaking scenario short mSUGRA. For a SUSY scale smaller than 1 TeV the scale of the hidden sector can be estimated with:

$$\Lambda_{SUSY} \approx \frac{\Lambda_{Hidden}}{\Lambda_{Planck}} \tag{2.13}$$

The scale of Λ_{Hidden} can be estimated in order of 10^{11} GeV. Due to specifc SUSY breaking mechanism the number of free parameters can be reduced.

Table 2.5: *Table of the five free parameters of the mSUGRA scenario.*

$\tan\beta$	ratio of v.e.v. of the two Higgs doublets in SUSY
$m_{\frac{1}{2}}$	mass scale of gauginos of SU(3), SU(2) and U(1)
m_0	mass scale of sfermions and Higgsino
A	common trilinear coupling between H and $\tilde{f}\tilde{f}$
sign(μ)	sign of the Higgsino mixing parameter

In some SUSY models (e.g. mSUGRA) several assumptions have been made to bring down the number of the additional free SUSY parameters from about 105 to just a few (5). To achieve that amongst others an additional quantum number the R-party is

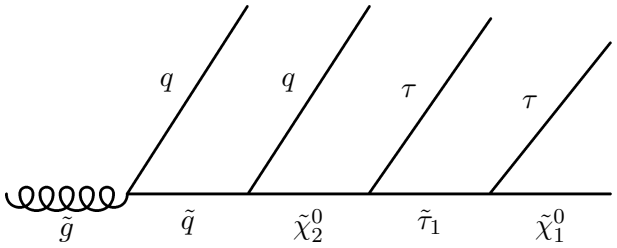

Figure 2.5: *SUSY decay chain with τ final states. SUSY processes with the shown decay chain will produce many jets and τ leptons. To discover such processes a sensitive τ reconstruction is needed. If both τ leptons can be reconstructed an invariant mass can be determined. This invariant mass distribution has a characteristic endpoint, which is determined by the mass of the two neutralinos ($\tilde{\chi}_2^0$, $\tilde{\chi}_1^0$) and the stau ($\tilde{\tau}_1$).*

introduced:

$$R = (-1)^{3(B-L)+S}, \tag{2.14}$$

with B the baryon number, L the lepton number and S the spin. In several SUSY models it is requiered that the R-parity is conserved at each vertex. Then SUSY particles can only be produced in pairs. Additionally a SUSY particle can never decay in a SM particle without producing a new sparticle. This means that the lightest SUSY particle will be stable and a good candidate for cold dark matter.

Apart from mSUGRA there are other SUSY breaking mechanisms like gauge mediated SUSY breaking (GMSB) and anomaly mediated SUSY breaking (AMSB). The main phenomenological difference of these models to mSUGRA is that in these models the lightest SUSY particle is the gravitino.

All these breaking mechanisms have in common that, if the R-parity is conserved the next-to-lightest particle (NLSP) determines the special signature of the decay. This is due to quantum number preservation at the decay vertices and leads to increased production of the SM partner of the NLSP. According to the five free parameters several scenarios with different NLSPs can be realized. Many of them have a stau as NLSP. Figure 2.2 shows an example of one possible SUSY decay chain with staus as NLSP. If such a SUSY model is realised in nature, many τ leptons will be produced in SUSY processes. To discover one of these models a very precise τ reconstruction is necessary. If both τs of one decay chain can be reconstructed their invariant mass can be determined. Via that invariant mass diffenrent SUSY scenarios can be distinguished. In case of a τ pair production, informations on the spin correlation can be determined [14]. The τ polarisation can also be used to distinguish between MSSM and Extra Dimension scenarios [15]. To achieve that the τ decay mode has to be correctly reconstructed.

2.3 Specific Theoretical Aspects for this Thesis

To give a comprehensive theoretical introduction in all SM topics covered by the subject of this work a short description of the τ lepton decay, the pair creation process of converted photons and the Z boson production at pp colliders have to be added. The latter aspect

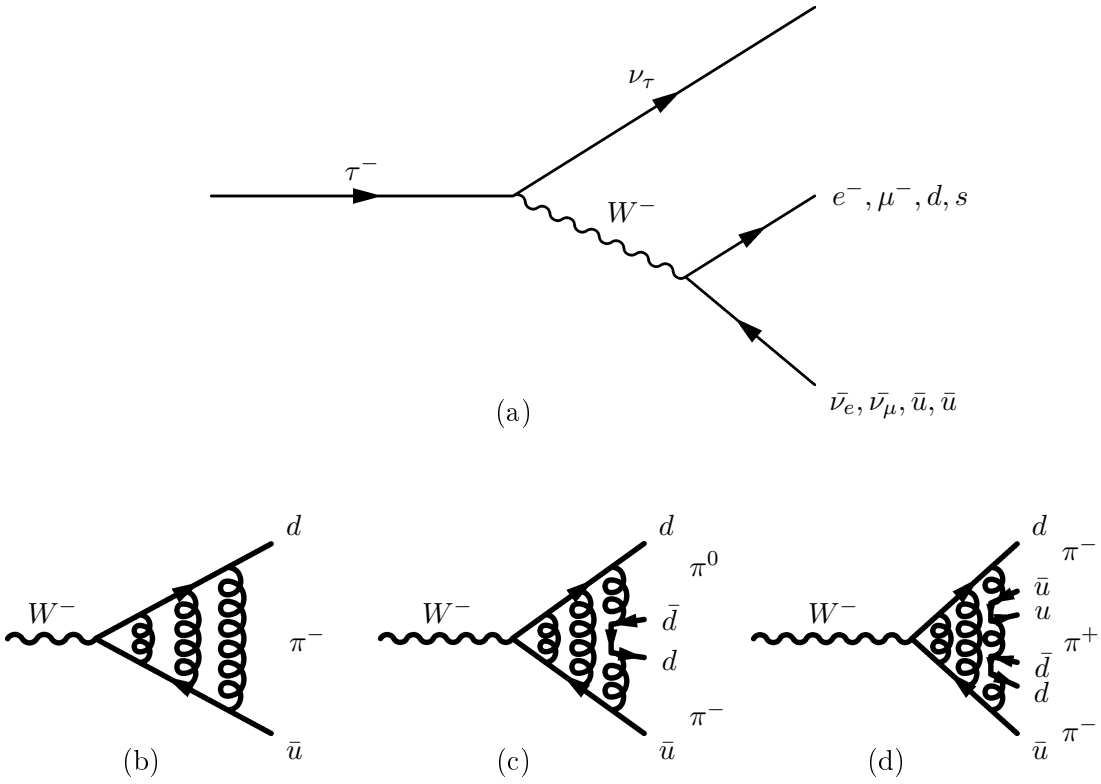

Figure 2.6: *Feynman diagrams of leptonic and hadronic τ decays.* ***(a)*** *Feynman diagram of possible τ decays with respect to the quatum number conservation.* ***(b)*** *Hadronic decay of the W^- boson.* ***(c)*** *Due to colour connection an additional π^0 may be produced.* ***(d)*** *In case of sufficient "stored" energy a W boson may decay into three charged pions.*

has to be explained in more detail because the $Z \rightarrow \tau\tau$ decay is applied as reference process and used to optimise the τ reconstruction as described in Chap. 7.

2.3.1 Decay of the τ Lepton

The τ lepton (m_τ = 1.78 GeV) is the most massive lepton. Due to the large mass difference to the other charged leptons (m_e = 0.51 MeV and m_μ = 105.6 MeV) it can decay into both of them whereas the muon only can decay into an electron. The large τ mass also yields hadronic τ decays meaning a decay into at least one charged pion (m_τ = 139.6 MeV) (cp. Tab. 2.6).

Due to lepton flavour conservation the τ^- decays into one ν_τ and one virtual W^-. This virtual W^- produced via a τ decay may than decay either into an e^--$\bar{\nu_e}$ pair, a μ^--$\bar{\nu_\mu}$ pair or into a quark pair (cp. Fig 2.6 (a)).

Neutrinos are weakly interacting particles and cannot be detected directly. Thus the total energy of a τ lepton is hard to reconstruct. This can be seen very clearly by comparing the transverse momenta of τ leptons and the transverse momenta of the decay products of hadronic τ decays, which is plotted in Fig. 2.7. More kinematical differences of τs and their hadronical decay products, e.g. the invariant Mass of $Z^0 \rightarrow \tau\tau$ events, are shown in the appendix Fig. A.1 and Fig. A.2).

Table 2.6: *τ decay branching ratios, based on 10^8 simlulated $Z \rightarrow \tau\tau$ events [16]. In about 40 % of all τ decay channels neutral pions (π^0) are produced.*

Decay modes	TAUOLA-CLEO
$\tau \rightarrow e\nu_e\nu_\tau$	17.8 %
$\tau \rightarrow \mu\nu_\mu\nu_\tau$	17.4 %
$\tau \rightarrow \pi^\pm\nu_\tau$	11.1 %
$\tau \rightarrow \pi^0\pi^\pm\nu_\tau$	25.4 %
$\tau \rightarrow \pi^0\pi^0\pi^\pm\nu_\tau$	9.19 %
$\tau \rightarrow \pi^0\pi^0\pi^0\pi^\pm\nu_\tau$	1.08 %
$\tau \rightarrow \pi^\pm\pi^\pm\pi^\pm\nu_\tau$	8.98 %
$\tau \rightarrow \pi^0\pi^\pm\pi^\pm\pi^\pm\nu_\tau$	4.30 %
$\tau \rightarrow \pi^0\pi^0\pi^\pm\pi^\pm\pi^\pm\nu_\tau$	0.50 %
$\tau \rightarrow \pi^0\pi^0\pi^0\pi^\pm\pi^\pm\pi^\pm\nu_\tau$	0.11 %
other modes with K	3.76 %
others	0.13 %

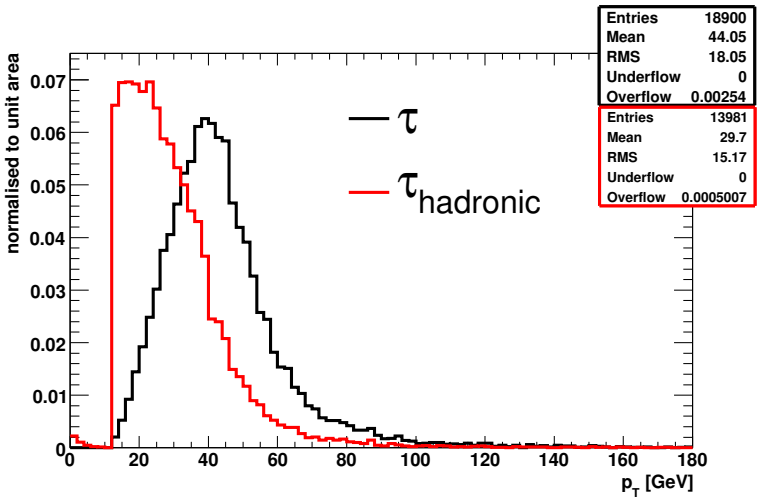

Figure 2.7: *Comparison of the transverse momenta of all τs and of the visible fraction of the hadronic τ decay products.*

To explain the Feynman diagram in Fig. 2.6 lower row in more detail, the hadronic W boson decay has to be discussed. If the W^- decays hadronically a d-$\bar{u}$ quark pair is created. The colour connection of the quarks leads to a meson production. In case of τ lepton decays either pions or kaons can be produced (e.g. in Fig. 2.6 a π^- is shown). Is enough energy "stored" in the quark pair, additional quarks-antiquark-pairs can be produced. The Feynman diagrams show the examples of a τ decaying into one π^- (b), into one charged and one neutral pion (c) and into three charged pions (d).

The branching fraction of producing a charged kaons via a τ leptons decay is much lower than the brachning fraction of leptonic τ decays or hadronic τ decays into pions (cp. Tab. 2.6). The reason is the difference in the coupling of the W boson to d-quark

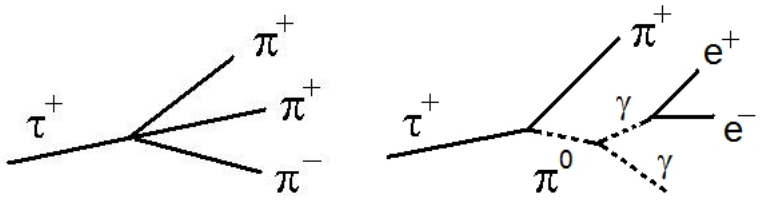

Figure 2.8: *3-prong τ decays and 1-prong τ decays with a photon conversion produce three charged particles. Left plot: τ lepton decays into three charged π's (3-prong) Right plot:τ lepton decays into one π^+ (1-prong) and one π^0, which decays into two γ's, one of them makes a e^+/e^- pair creation.*

($\propto \cos^2\theta_C = 0.9512$) and to s-quark ($\propto \sin^2\theta_C = 0.0488$).

If the τ lepton decays hadronically, two main signatures are differentiated: the so-called single-prong and the three-prong decays. Single-prong (1-prong) means that the decay products are one charged pion and one ν_τ ($\tau \rightarrow \pi^\pm\nu_\tau$) or one charged pion, one ν_τ and n neutral pion(s)($\tau \rightarrow n\pi^0\pi^\pm\nu_\tau$). About 23.4 % of this decay mode are decays without neutral pions and 76.6 % with neutrals.

For the three-prong (3-prong) decays, $\tau^\pm \rightarrow 3\pi^\pm\nu_\tau$ modes contribute 64.6 % and three-prong decays with neutrals ($\tau^\pm \rightarrow n\pi^0 3\pi^\pm\nu_\tau$) contribute 25.6 %.

Thus over all in about 40 % of all τ decay channels additionally to the one or three charged pions neutral pions (π^0) are produced. These π^0s decay with a branching fraction of 98.8 % into two photons [1]. If one of the photons makes a pair creation (this is decribed in Sec. 2.3.2) it is hard to distinguish a 3-prong decay from a 1-prong decay with pair creation, due to the additional charged tracks of the e^+-e^- pair (cp. Fig. 2.8). The aim of this thesis is to improve the existing τ reconstruction algorithm considering this problem.

Another important property of the τ lepton is the life-time, which leads to a decay length of $c\tau = 87.11\,\mu$m. Thus allows to reconstruct a secondary vertex for 3-prong decays. Spin-effects in τ lepton decays are important to determine the polarization of a decaying resonance like gauge bosons, Higgs- or SUSY particles. As already mentioned in Sec. 2.2, therefore the decay mode has to be correctly reconstructed.

2.3.2 Photon Conversions

Photon Conversions are e^+/e^- pairs which have been produced via the interaction of a photon with matter [18]. The leading order Feynman diagrams are shown in Fig. 2.9. The cross section of photons with energies above 1 GeV is completley dominated by pair productions (c.p. App. Fig. A.3). Interactions like photoelectric effects, Rayleigh or Compton scattering, which dominate at low photon energies can be neglected above 1 GeV (c.p. Fig. 2.10).

The radiation length (X_0) is defined as $\frac{7}{9}$ of the mean free path for pair production. If a photon passes an element heavier than helium, the radiation length can be approximated by [1]:

$$X_0 = \frac{716.4 g cm^{-2} A}{Z\,(Z+1)\ln\left(287/\sqrt{Z}\right)} \tag{2.15}$$

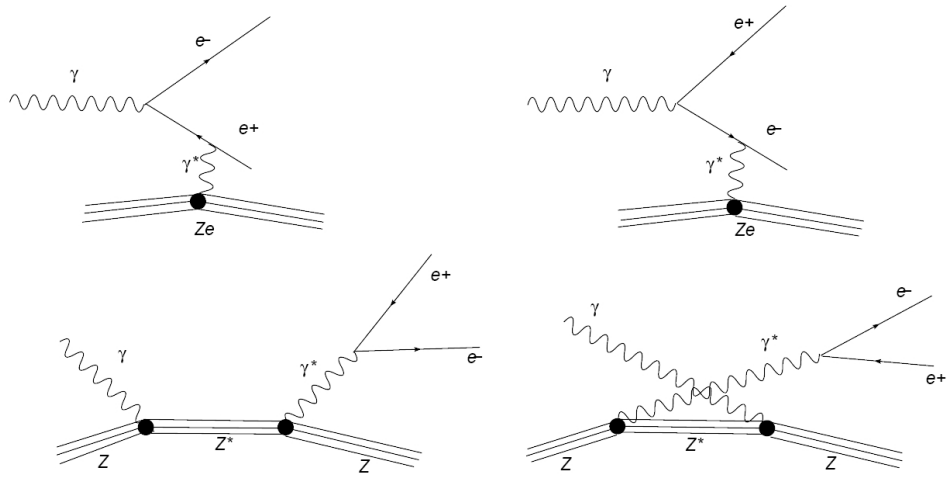

Figure 2.9: *Feynman diagrams of photon conversions in first order [17].*

with the atom mass A given in g/mol and the atomic number Z of the transversed material. In Fig. 2.10 one can see the increasing domination of photon conversions with higher photon energies. Therefore the differential cross section can be calculated [19]:

$$\frac{\partial\sigma}{dx} = \frac{A}{X_0 N_A}\left[1 - \frac{4}{3}x\,(1-x)\right] \tag{2.16}$$

with the fraction of the energy transferred from the incoming photon (with the energy E_{photon}) to the electron (positron) $x = E_{electron}/E_{photon}$ and the Avogadro number $N_A = 6.022 \times 10^{23}$.

Integration of Eq. 2.16 leads to the following total photon conversion cross section:

$$\sigma = \frac{7}{9}\frac{A}{X_0 N_A} \tag{2.17}$$

The photon conversion cross section is symmetric between x and 1-x (c.p. Fig. A.4 in the appendix). But this does not lead to a symmetrically shared momentum of the photon to the two produced leptons. A photon conversion is an interaction between a photon and an atom and not a photon decay. There are even photon conversions which are highly asymmetric in their momentum. This may be a serious problem, if it concerns a low energy photon conversion. Then the momentum of the electron (positron) may be below the reconstructable threshold of the experiment. The positron (electron) will be difficult to distinguish from prompt positrons (electrons).

2.3.3 Z Boson Production at pp Collider

Before it is possible to discover the Higgs boson or physics beyond the standard model the detector of a new particle physics experiment has to be understood very well. Therefore it is absolutly necessary to have some processes which are measured with high prescision and which are theoretically well understood.

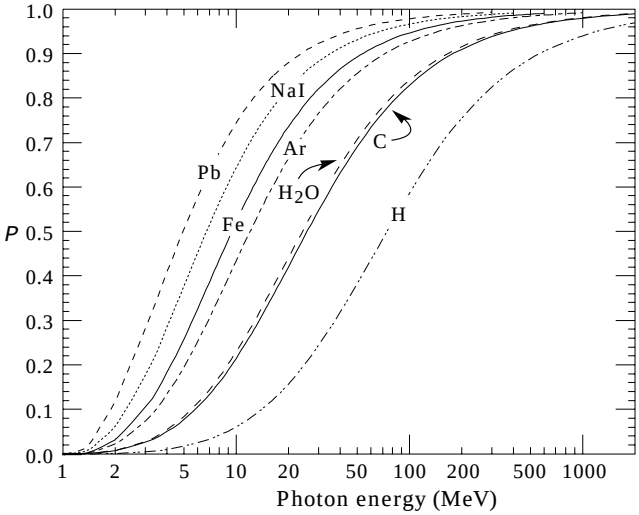

Figure 2.10: *Probability P that a interacting photon will produce a* e^+ e^- *pair. The interesting energy region is above* 1 GeV*, there the probability for each material is nearly 100 %. The probability of a photon with the attentuation length* λ *to convert in an absorber with the thickness t is* $P[1 - \exp(-t/\lambda)]$.

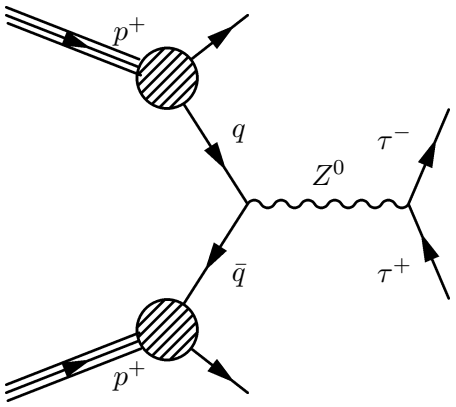

Figure 2.11: *Feynman diagram of* $Z^0 \rightarrow \tau^+\tau^-$ *decay.*

The production of heavy SM gauge bosons like the Z or the $W^{\pm}$ boson can be used to calibrate the detector, test the reconstruction algorithms, monitor the luminosity and study the trigger efficiency. Especially to study and improve τ reconstruction algorithms the $Z \rightarrow \tau\tau$ process with two final τ leptons is particularly important.

Properties like the total width Γ_Z, the mass m_Z and the partial decay width Γ_i of the Z boson have been measured accurately at CERN of the LEP experiments and at the SLC at SLAC. Some values of the Z boson properties are listed in Tab. 2.7.

At a Proton-Proton collider the Z boson is produced via a Drell-Yan process (c.p. Feynman diagram Fig. 2.3.3). A quark anti-quark anhiliation leads to a Z boson which decays into lepton anti-lepton or quark anti-quark pairs.

Protons consist of quarks and gluons. Due to the fact that the quark and anti-quark annihilate in the Drell-Yan process, it is not the full proton momentum which will be

Table 2.7: *Z^0 decay branching ratios [1].*

Decay modes	Fraction Γ_i/Γ
$e^+\ e^-$	3.363 ± 0.004 %
$\mu^+\ \mu^-$	3.366 ± 0.007 %
$\tau^+\ \tau^-$	3.370 ± 0.008 %
invisible	20.00± 0.06 %
hadrons	69.91 ± 0.06 %
m_Z	91.876± 0.0021 GeV
Γ_Z	2.4952± 0.0041 GeV

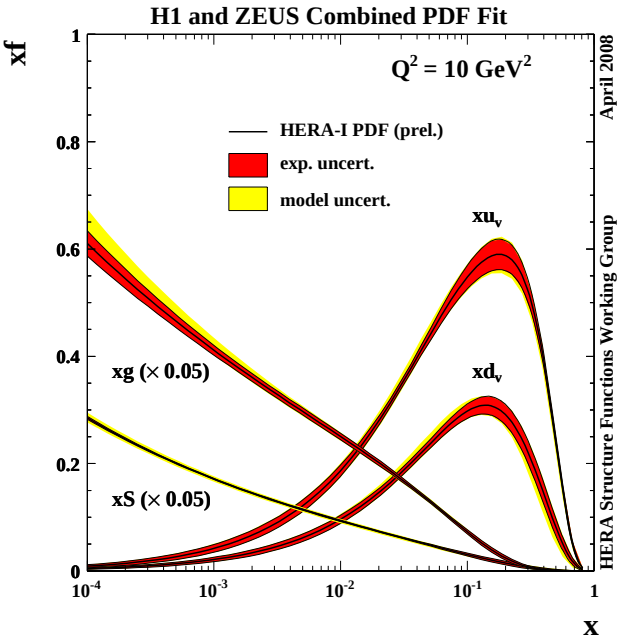

Figure 2.12: *Proton parton density function measured at HERA*

transferred. The effective center of mass energy can be calculated via:

$$\hat{s} = x_q x_{\tilde{q}} s \tag{2.18}$$

Here $\hat{s}$ is the reduced center of mass energy and s the center of mass energy of the protons. At the LHC with a total center of mass energy of $\sqrt{s} = 14\,\text{TeV}$ the momentum fraction of the partons can be estimated to:

$$\frac{m_{Z^0}^2}{s} \leq x_q x_{\tilde{q}} \approx 4 \cdot 10^{-5} \tag{2.19}$$

The HERA experiment has measured the proton parton density function (PDF) very precisely [20]. In Fig. 2.12 the prelimenary results are shown. The very low momentum fraction of the partons estimated by Eq. 2.19 means that the Z bosons produced at the LHC are mostly produced via see quarks.

Chapter 3

The ATLAS Experiment at LHC

This chapter presents the world largest experiment: the Large Hadron Collider (LHC) at CERN[1] near Geneva with its six different detectors. The LHC will provide first collisions in the spring of 2009. After giving a short overview of the identification of several elementary particles the ATLAS detector is described in more detail. Therefore each subdetector, the inner detector, the calorimeter system, the muon system and, the trigger system is explained in a seperate section.

3.1 The LHC Collider

The LHC [22] is a proton proton collider. The protons are accelerated up to a center of mass energy of $\sqrt{s} = 14\,\text{TeV}$ and brought to collisions in four different interaction points. As shown in Fig. 3.1, the four large experiments are placed at one of these collision points. ATLAS is installed at *point 1*, ALICE at *point 2*, CMS at *point 5* and LHCb at *point 8*.

LHC has been installed in the former tunnel of the LEP [7] experiment. The tunnel is 27 km long and between 50 m and 175 m below the earth's surface. The dipoles and quadrupoles which bundle and deviate the "bunches" of protons (10^{11} protons) have magnetic fields of 8.33 T. To achieve such strong magnetic fields the magnets are superconductors and have to be cooled to a temperature of 1.7 k. This is done with super-fluid helium.

Not only the center of mass energy is important to discover new physics, it is also important to have a high event rate. Figure 3.2 shows the expected cross sections and event rates of the most important processes at the LHC. For example the cross sections of Higgs bosons with an assumed mass of $m_H = 150\,\text{GeV}$ or $m_H = 500\,\text{GeV}$ are orders of magnitude lower than the cross sections of their background events. To detect a significant number of Higgs events many proton collisions have to be produced. To provide a sufficient number of collisions within an acceptable time scale a very high bunch crossing frequency is required. Every 25 ns two bunches collide. Due to the fact that each bunch has up to 10^{11} protons more than one collision can take place simultaneously and overlap, this is called "pile-up".

The event rate is characterised by the luminosity which is defined as:

$$L = \frac{N_P^2 N_B c}{4\pi\sigma_x\sigma_y U} \tag{3.1}$$

[1] European Organization for Nuclear Research

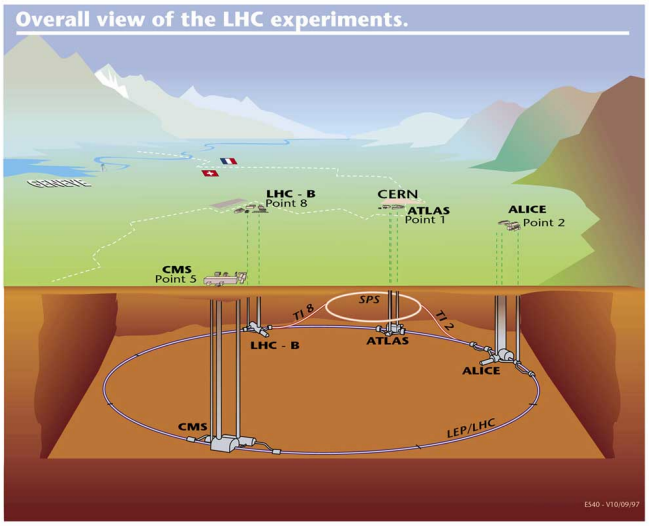

Figure 3.1: *The Large Hadron Collider and four of its experiments at CERN site near Geneva. Under earth's surface the experiments ATLAS, ALICE, CMS and LHCb are shown. The SPS ring (Super Proton Synchrotron) directly located under the CERN complex is one of the three pre accelerators needed to accelerate the protons to an energy of 450 GeV before injecting into the large LHC ring [21].*

Table 3.1: *Overview of the six experiments at LHC.*

ATLAS:	A Toroidal LHC Apparatus Is built as multi purpose detector to discover the Higgs boson and find physics beyond the standard model.
CMS:	Compact Muon Solenoid Same aim as ATLAS: ATLAS and CMS are built to cross check their results.
LHCb:	LHC B-physics experiment Is dedicated to B-physics and will investigate CP-violation.
ALICE:	A Large Ion Collider Experiment Investigates quark-gluon plasma produced via heavy ion collisions.
TOTEM:	TOTEM is dedicated to the measurement of the total cross section, elastic scattering and diffractive processes.
LHCf:	Large Hadron Collider forward experiment LHCf uses forward particles created inside the LHC as a source to simulate cosmic rays in laboratory conditions.

whereas N_P is the number of protons, N_B the number of bunches, U the circumference, c the speed of light and σ_x and σ_y are the beam width in x- and y-direction.

The luminosity at LHC will be raised stepwise from an initial luminosity ($L = 10^{31}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$) at the very first begining to a "low" luminosity phase ($L = 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$) to the "high" luminosity phase ($L = 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$).

Instead of refering to the total number of events for a special process often the integrated luminosity $\int L \cdot \mathrm{d}t$ is used. One year of data taking with low (high) luminosity corresponds an integrated luminosity of $10\,\mathrm{fb}^{-1}$ ($100\,\mathrm{fb}^{-1}$).

3.2 Short Introduction of Particle Identification

All multi purpose detectors are built by the same schema which is determined by the properties of the different elementary particles and the available methods to identify them. The two most important properties to specify a particle are the electric charge and the mass. The charge is determined via the deviation of the particle by a magnetic field. The mass cannot be measured directly. Instead other properties as the energy (E), the energy deposition ($\mathrm{d}E/\mathrm{d}x$), the trajectory through the detector, and the production vertex are measured. Here the criteria to detect the most important particles are listed:

Electron: High-energy electrons lose energy in matter by bremsstrahlung. This bremsstrahung (γ) can produce e^+/e^- pairs again and so on. Thus the electron energy can be determined by collecting the so-called electromagnetic shower in an electromagnetic calorimeter.

Photon: As described in Sec. 2.3.2 high-energy photons passing through matter create photon conversions. Thus they produce electromagnetic showers as electrons. They can be differentiated to e^- and e^+ due to their lack of electric charge (no tracks are left in the tracker, at least as long no conversion occurs).

Muon: Muons, like most relativistic particles, have mean energy loss rates close to the minimum of the Bethe-Bloch curve, they are said to be minimum ionising particles [1]. Thus almost no energy depositions in the calorimeters can be observed. The identification of muons is based on the absorption of all other charged particles before reaching the muon spectrometer. Thus only muons are registed in the outermost detector layers.

Tau: As described in Sec. 2.3.1 τ leptons will decay nearby the interaction point due to their mean life of $2.9 \cdot 10^{-4}\,\mathrm{ns}$. Therefore only the decay products can be detected, muons, electrons or several numbers of pions. Out of them the τ leptons have to be reconstructed via a reconstruction algorithm (see in Chap. 5).

Proton: Protons and charged hadrons (e.g. pions and kaons) deposit their energy in hadronic calorimeters. Due to nuclear interactions hadronic showers are induced. Via shower shapes and the amount of deposited energy the hadronic particles can be distinguished.

Neutron: The hadronic showers produced by neutrons look very similar as the proton showers. However they are very easy to distinguish due to the fact, that the electric uncharged neutron does not leave any tracks in the tracker.

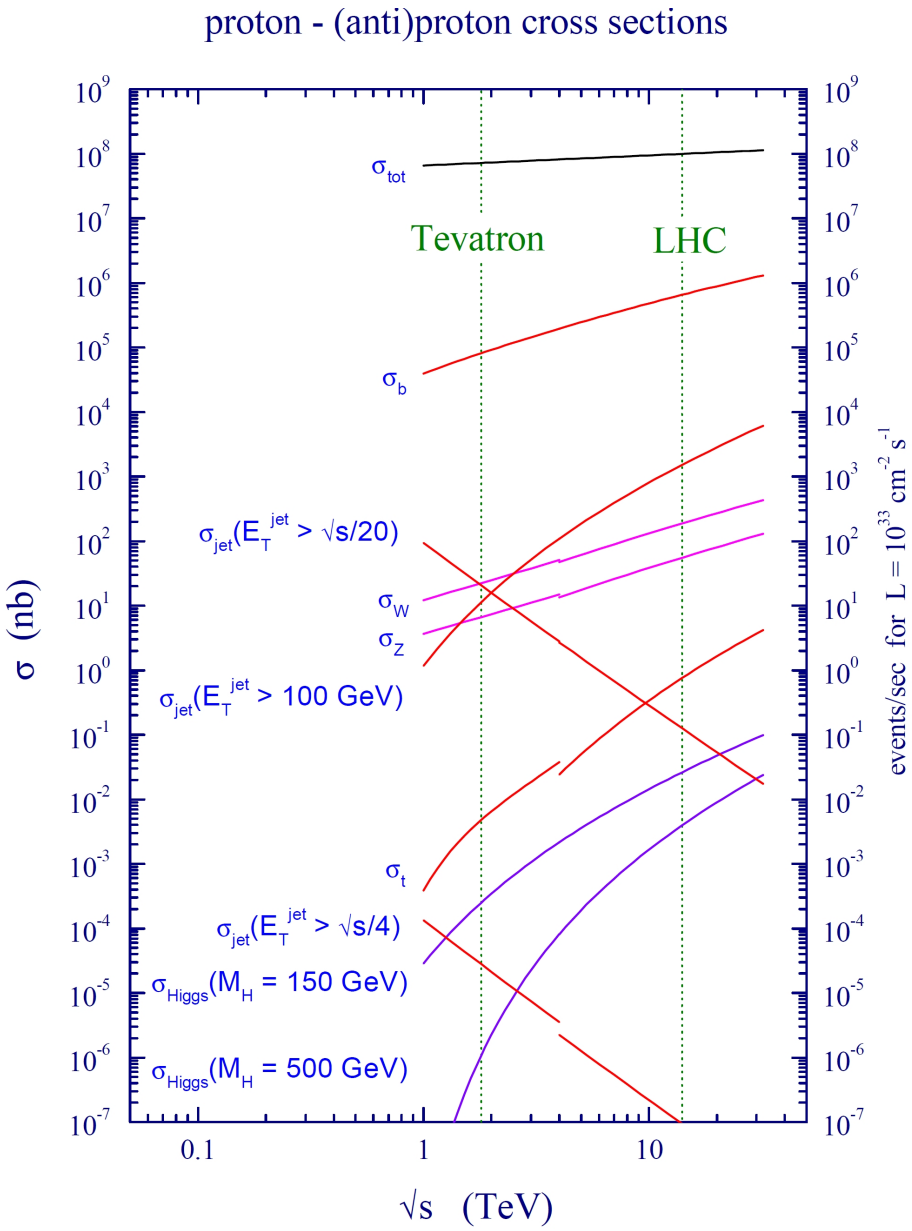

Figure 3.2: *The proton-proton cross section and the event rates at the TeVatron and LHC [23]. The cross section in nb and the event rate per sec. for a luminosity of* $10^{33}\,\mathrm{cm}^{-2}\mathrm{sec}^{-1}$ *are plotted over the center of mass energy in TeV. At* 14 TeV *the operation point of the LHC mainly b-quarks will be produced. The cross section of the Higgs boson production, the LHC is searching for, is orders of magnitude lower than the other cross sections shown in this figure.*

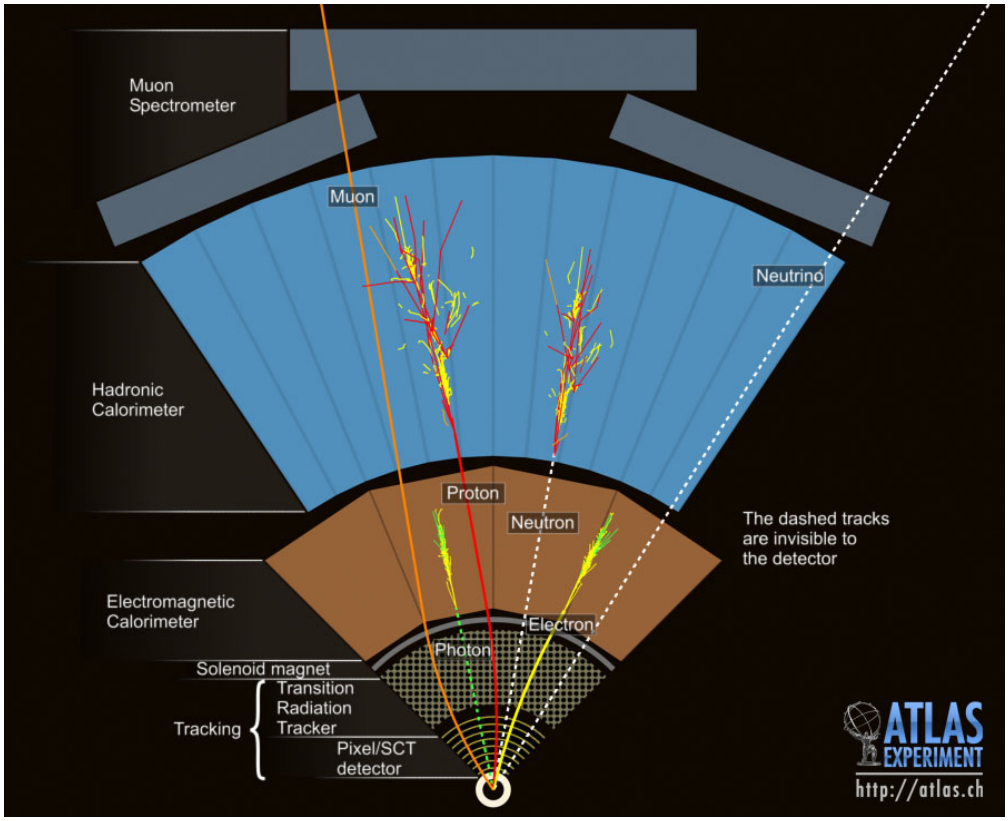

Figure 3.3: *Track overview of different particles [24]:* ***Electrons*** *leave tracks behind in the Inner Detector and deposite all their energy in the electromagnetic calorimeter.* ***Photons*** *behaves in the similar way but leaves no track.* ***Neutrinos*** *are weakly interacting, usually they are not detected.* ***Muons*** *pass most of the detector. They leave tracks in the tracker and in the muon spectrometers.* ***Protons*** *leave tracks and deposite all their energy in the hadronic calorimeter.* ***Neutrons*** *behave quite similar but without leaving tracks behind.*

Neutrino: Neutrinos (ν) are very difficult to detect. They are weakly interacting, thus a huge amount of absorber material would be needed. It is not possible to detect νs directly, the only possiblity to account for νs is to detect the electron of a $\bar{\nu_e} + p \rightarrow n + e^-$ interaction. Often huge Cerenkov detectors are used for neutrino searches. Multi purpose detectors like ATLAS are not able to identify νs directly. If the energy depositions of all other particles are produced in one event are measured very well, the amount of missing energy of neutrinos can be determined.

3.3 The ATLAS Detector

The ATLAS [25] detector is about 46 m long, more than 25 m high and weighs about 7000 t. It is built as a multi purpose detector to discover the Higgs boson and physics beyond the standard model (e.g. SUSY) but also to measure precisely SM observables such as the Z^0, the $W^{\pm}$ and the top quark mass. Hence the main challenge is to exploit the high event rate of the LHC and to resist the large radiation due to the exposure to

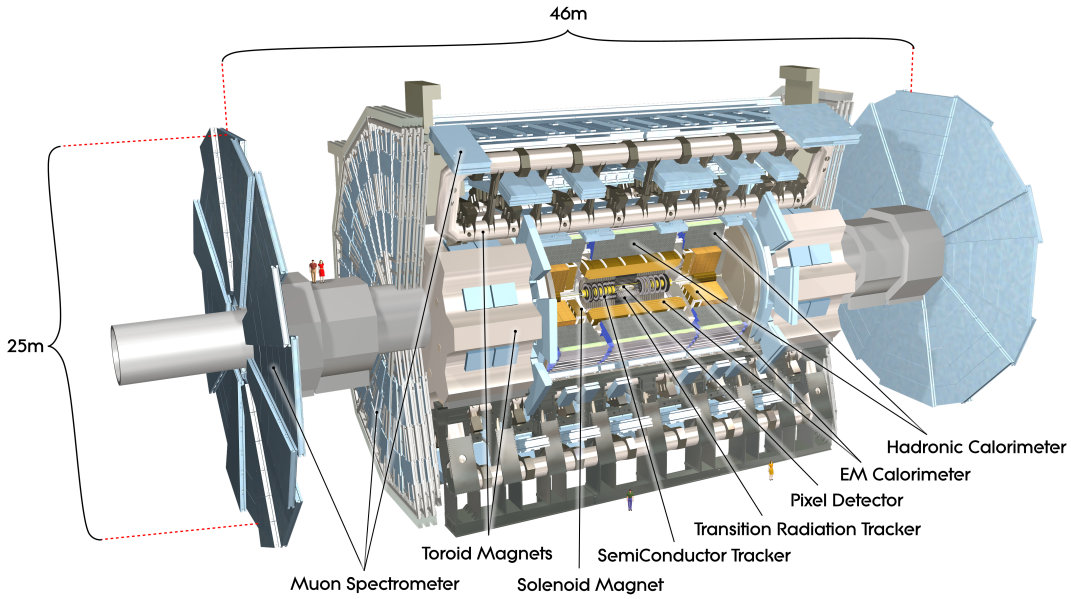

Figure 3.4: *The ATLAS detector in a schematical overview [25].*

an intense flux of high energetic particles.

In Fig. 3.4 the subdetectors are shown. Starting from the interaction point, the beam pipe is surrounded by the inner detector, which is covered by the solenoid providing a 2 T magnetic field. The next layers are the electromagnetic calorimeters encased by the hadronic calorimeters. Around all that the barrel toroid splitted in eight segments and two end-cap toroids produces a magnetic field of 0.5 T and 1 T in the central and the end-cap region, respectively. This magnetic field is needed to deviate the muons which are detected in the everything surrounding muon system.

The ATLAS Coordinate System

The ATLAS coordinate system is defined by the x-axis which is the axis pointing from the interaction point in the beampipe to the center of the LHC ring. Perpendicular to the x-axis pointing upwards is the y-axis defined and the z-axis is pointing in direction of the anticlockwise beam direction.

The azimuthal angle ϕ is defined as the angle around the beam pipe (z-axis). The polar angle θ is measured from the beam axis and defines the pseudo rapidity:

$$\eta = -\ln\left(\tan\frac{\theta}{2}\right). \tag{3.2}$$

The distance between to objects in the $\eta - \phi-$plane is defined as:

$$\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}. \tag{3.3}$$

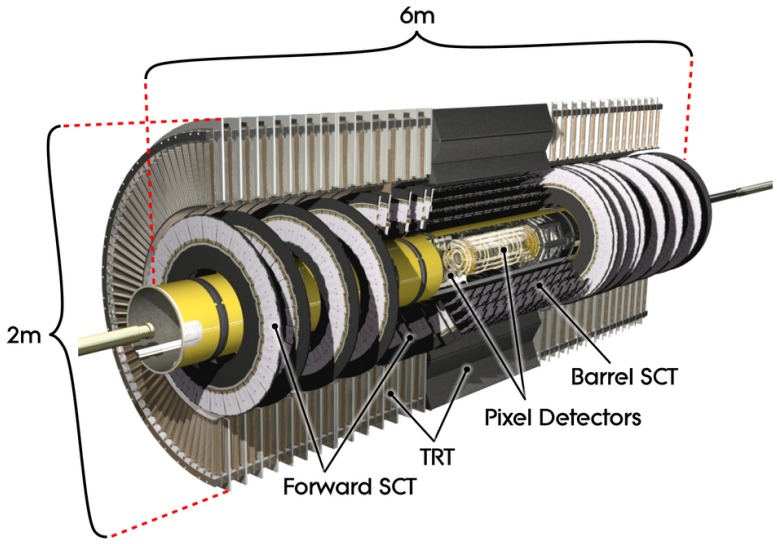

Figure 3.5: *The ATLAS Inner Detector (ID). In the middle is the very highly granulated pixel detector for vertexing and precise tracking near the interaction point. Around that the Silicon Microstrip Detector, which covers a much larger area, is placed. Most volume of the ID is filled with the Transition Radiation Tracker [25].*

3.3.1 Inner Detector

The Inner Detector (ID) is in total 6 m long with a radius of 2 m [26]. It covers a region of $|\eta| < 2.5$ and consistes of three subdetector systems: the silicon pixel detector, the semiconductor tracker (SCT) and the transition radiation tracker (TRT). Surrounded by the solonoid with a magnetic field of 2 T it is responsible to measure momentum, primary and secondary vertices, the sign of the electric charge and to distinguish between electrons and pions in the TRT.

The two most inner subdetectors are silicon detectors which detect passing charged particles by providing electron-hole pairs. The transition radiation tracker works like a drift chamber. These three systems make sure that enough hits are measured to reconstruct the track and vertices very precisely. The two silicon based sytems provide up to 11 hits in the barrel region and the TRT system up to 36. The more layers available the more precisely the tracks and vertices can be determined. The limitation is that each detector system means additional material in the flight path of the particles. As described in Sec. 2.3.2 this means that the possibility of photon conversions increases. A schematical view of the ID is shown in Fig. 3.5 and the subdetectors are described below.

Pixel Detector

The pixel detector [27] is designed to achieve a very high granularity to measure the charged tracks, the impact parameters and secondary vertices of short-lived particles like τ leptons and b-mesons with the highest presicion. Due to the closeness to the beam the

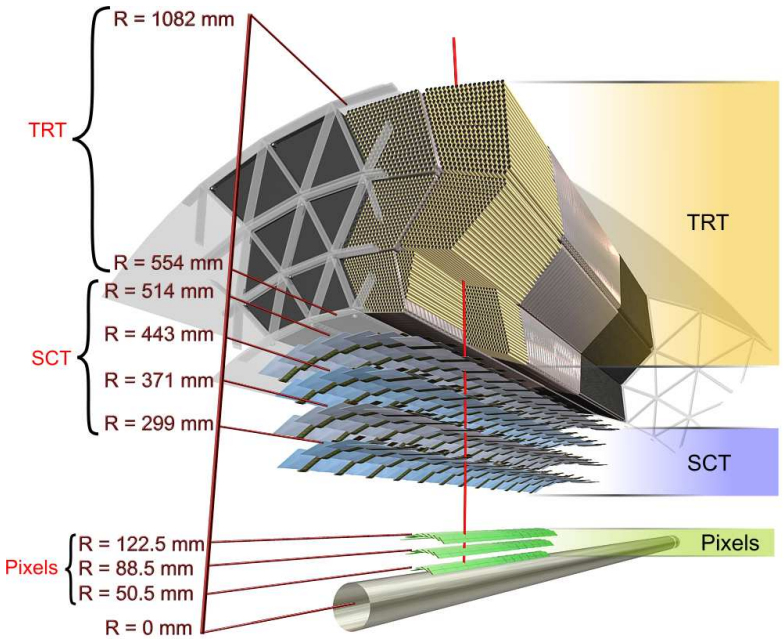

Figure 3.6: *Barrel region of the ATLAS Inner Detector. Shown are the three subsystems Pixel, SCT and TRT of the ID. The precise positions of the several layers are registered [25].*

pixel detector is exposed to the most radiation, e.g. the predicted ionisation dose of the innermost layer of the barrel pixel detector is 160 kGy/y.

The pixel detector provides three layers in the barrel region at an avarage radii of 5 cm, 9 cm and 12 cm and four disks at the $|z|$-position of 47.3 cm, 63.5 cm, 77.8 cm and 107.2 cm in the end-cap region. The layers and disks consist of rectangular cells (pixels) of silicon with a size of $50\,\mu\text{m} \times 300\,\mu\text{m}$ which provides a spatial resolution of $r\phi \times z = 12\,\mu\text{m} \times 60\,\mu\text{m}$.

Silicon Microstrip Detector

Overall the semiconducter tracker (SCT) covers an area of $61\,\text{m}^2$. Therefore the granularty could not be as high as the one of the pixel detector. The SCT consists of silicon strips with a width of $80\,\mu\text{m}$ and a length of 12 cm. To improve the resolution in the length of such strips, two of such detectors are glued toghether back-to-back at a 40 mrad angle.

The SCT system similar to the pixel detector is separated in a barrel and an end-cap region. The barrel region consists of four layers (c.p. Fig. 3.6) and end-cap of nine axial wheels. The SCT is designed to provide four times two precision measurements for each track, providing a spatial resolution of $r\phi \times z = 17\,\mu\text{m} \times 580\,\mu\text{m}$.

Transition Radiation Tracker

The transition radiation tracker (TRT) is placed in the barrel region between the radii 55 cm and 108 cm and consists in each of the end-cap regions of 18 wheels.

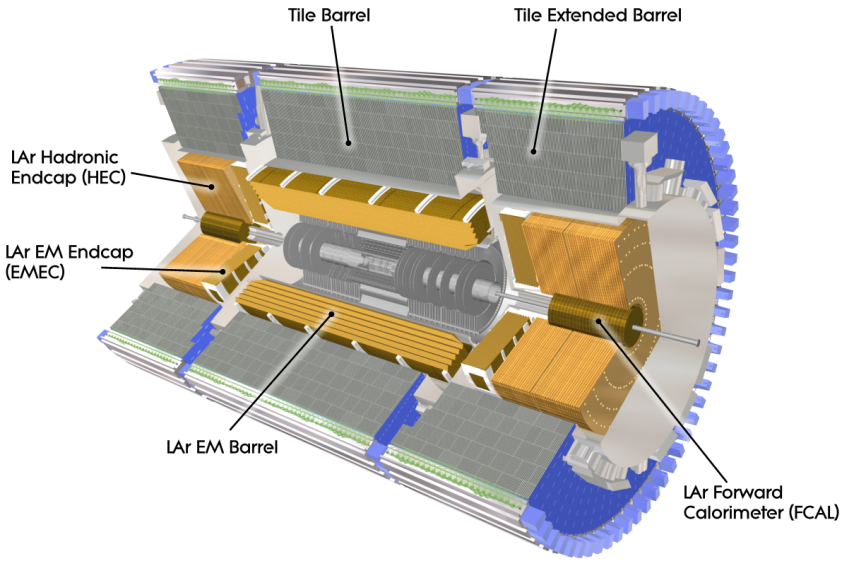

Figure 3.7: *The ATLAS Calorimeter system. It contains of the electromagnetic calorimeter (ECAL) and the hadronic calorimeter (HCAL). The two moduls of the ECAL are the LAr EM Barrel and the LAr EM Endcap. The HCAL consists of the Tile Barrel and the Extended Tile Barrel, the Hadronic Endcap and the Forward Calorimeter near the beam axis [25].*

The TRT is based on straw detectors with a diameter of 4 mm and a length of 37 cm to 150 cm filled with a nonflammable gas mixture of $Xe(70\%)$ $CO_2(27\%)$ and $O_2(3\%)$. The anode of the drift tubes are gold-plated tungsten wires and the cathodes are the aluminum cased tubes.

Two types of informations are provided by the straw tubes, the distance of the closest approach of the track to the wire and the discrimination between tracking and transition radiation hits.

Ultra relativistic charged particles, e.g. electrons, produce transition radiation when they cross the surface between two media. By detecting such transition radiation electrons can be differentiated from e.g. pions. Therefore the drift tubes are interleaved with polypropylene fibres serving as transition radiation material. The discrimination of tracking and transition radiation hits is achieved by two thresholds: low-threshold 200 eV for tracking hits and high-threshold 5 keV for transition radiation hits. In Chap. 7.3 it will be described in detail how charged tracks are identified for improving τ-reconstruction.

The distance of closest approach of the track to the wire is calculated from the time, it takes for the ionisation electron cluster created along the track to drift to the central wire, which is measured. The spatial resolution per straw tube is 170 μm, which is, compared to the resolution of the pixel detectors, rather low. But due to the many, typically 36, measurements for each track the relatively bad spatial resolution can be compensated.

3.3.2 Calorimeter System

As shown in Fig. 3.7 the calorimeter system encloses the ID and the solenoid. It is responsible for the energy measuremnt and to provide a 4-vector of neutral particles and

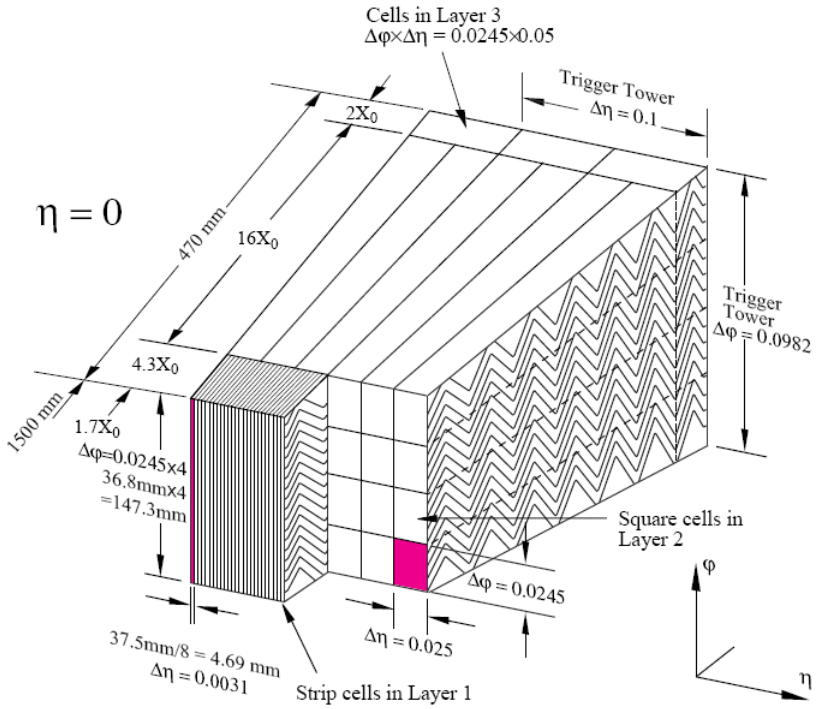

Figure 3.8: *The ATLAS Electromagnetic Calorimeter is built in three differnet layers. The first layer is very high granulated in η to separate γ and π^0 but not in ϕ. Hence the second layer is squared to get an η and ϕ resolution. The third layer the so-called tail catcher registers an occurring leakage of calorimeter [25].*

jets with energies from 10 GeV to 1 TeV [28]. To determine the missing transverse energy ($E_{T,miss}$) the full coverage of the solid angle is absolutely neccessary.

The electromagnetic and hadronic showers described in Sec. 3.2 are detected in fluid argon and plastic scintilator plates. Besides the energy measurements the shower shapes are used to identify the miscellaneous particles. The shower of a hadronically decayed τ lepton e.g. can be differentiated from a QCD-jets shower by its narrow shape. To measure such shapes the calorimeters have to be highly granulated.

The calorimeter system consists of the electromagnetic calorimeter (ECAL) covering $|\eta| < 2.5$ and the hadronic calorimeter (HCAL) whose components cover $|\eta| < 1.7$ (hadronic barrel), $1.5 < |\eta| < 3.2$ (hadronic end-cap calorimeters) and $3.1 < |\eta| < 4.9$ (forward calorimeters).

Electromagnetic Calorimeter

The Electromagnetic calorimeter (ECAL) consists of 1.5 mm to 2.2 mm thick layers of lead or stainless steal as absorber material and 4 mm thick layers with liquid-argon (LAr). The LAr is used because of its radiation resistance and long-term stability. As shown in Fig. 3.8 the several layers are arranged in an accordion geometry to provide a complete symmetry in ϕ. To correct energy losses due to the ID, solenoid and the cooling system a presampler is upstreamed the ECAL.

Around $\eta \approx 1.5$ in the region between the barrel and end-cap, the so-called "crack-

region", the ECAL performance is reduced. In the central region ($|\eta| < 2.5$) the ECAL is segmented into three samplings (c.p. Fig. 3.8):

1. Sampling: The first sampling has a very high granularity in η ($\Delta\eta \times \Delta\phi = 0.0031 \times 0.098$) and a thickness of 4.3 X_0. Whereas X_0 is the radiation length, the distance in which an electron loses $1/e$ of its energy. The high resolution in η ensures a separation of γ and π^0.

2. Sampling: In the second sampling most of the electromagnetic energy is deposited. It is segmented in squared towers of $\Delta\eta \times \Delta\phi = 0.0245 \times 0.0245$ and is 16 X_0 thick.

3. Sampling: The third sampling records if the shower leaks out of the ECAL. The thickness is 2 X_0 and the granularity is $\Delta\eta \times \Delta\phi = 0.05 \times 0.0245$. Especially to distinguish electrons from pions this sampling is very important. It records if there is a high fraction of the shower leaking out. If that is the case, a pion has probably activated the electromagnetic shower.

More details about using informations from electromagnetic calorimeter to identify electrons is described in Chap. 7.4.

Hadronic Calorimeter

The hadronic calorimeter (HCAL) is designed to measure jet energies and their directions. These jets are hadronisations from quarks and gluons and from hadronic τ lepton decays. The missing transverse energy is determined by the HCAL, the ECAL and the muon spectrometer. The HCAL is placed around the ECAL and consists of three parts:

Hadronic Tile Calorimeter: This part is made up of $4 - 5$ mm thick lead layers as absorber and 3 mm thick plastic scintillator plates. Two sides of the scintillators are read out by wavelenght shifting fibres into two seperate photomultiplier tubes. In η, readout cells are built by grouping fibres into the photomultipliers as pseudo-projective towards the interaction region.

Hadronic Endcaps (HEC): Due to the high radiation to which this part is exposed the absorber material is copper plates and the active material is liquid-argon as in the ECAL.

Forward Calorimeter (FCAL): To measure the energy as close as possible to the beam axis the FCAL ranges up to 1° to the beam axis. This leads to an extreme radiation level. The absorber consisting of tungsten is circumfluent of the active material which is liquid-argon, too.

3.3.3 Muon System

Muons do not interact via the strong force and due to the large mass compared to electrons, they do not make any bremsstrahlung. Muons are minimal ionising particles. Thus they do not deposit all their energy in the calorimeters. To guarantee that only muons are detected by the muon chambers an additional magnetic field is installed outside of the calorimeter systems.

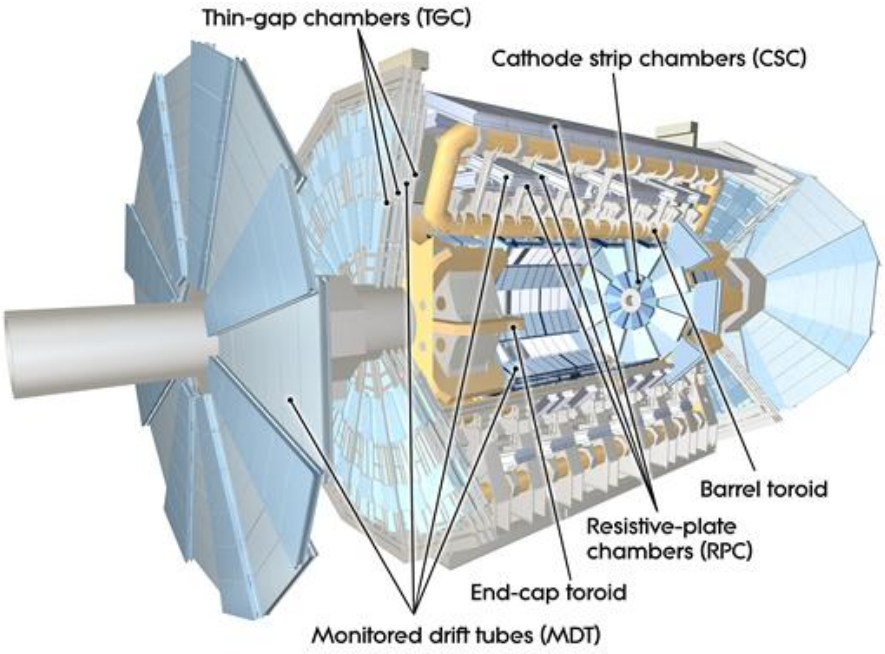

Figure 3.9: *The ATLAS Muon system divided in a detection and a trigger system. Both are seperated into barrel and end-cap region. The detection system are the Monitored Drift Tubes and the Cathode Strip Chambers and the trigger system are the Resistive Plate Chambers and the Thin Gap Chambers. The end-cap toroids and the barrel toroid produce a* 1 T *and a* 0.5 T *magnetic field to deviate the muons for better identification [25].*

In a region with $|\eta| < 1$ eight superconducting toroid magnets are symmetrically installed providing a magnetic field of 0.5 T (c.p. Fig. 3.9 in yellow). Two smaller end-cap magnets cover the region $1.4 < |\eta| < 2.7$ and provide a magnetic field of 1 T. The muon spectrometers are able to measure the muon momentum with higher precision than the ID. Due to the larger distance from the interaction point the curvature can be measured more precisly.

Muons are detected via Monitored Drift Tubes (MDTs) in the barrel region and via Cathode Strip Chambers (CSCs) in the end-cap region [29]. Additionally to that a muon trigger system is installed. This is needed because the readout of the muon precision measurement detectors is too slow to use it as trigger. It consists of Resistive Plate Chambers (RPCs) and Thin Gap Chambers (TGCs) in barrel and end-cap region respectively.

Precision Measurment Detectors

MDTs consist of aluminium tubes with a diameter of 30 mm and cover $|\eta| < 2$. The anode is a tungsten-rhenium wire. The tubes are filled with a gas mixture of 93 % Ar and 7 % CO_2 with a pressure of 3 bar. Muons are detected by ionisation. Each tube has a spatial resolution of 80 μm. Overall there are three layers of MDTs installed at radii of 5 m, 7 m and 10 m.

CSCs are installed in the forward region ($2< |\eta| <2.7$), where a large particle rate is expected. To differentiate signal from background the CSCs have a high granularity, which is realised by a smaller distance from anode to cathode and another gas mixture (30% Ar and 50 % CO_2 and 20% CF_4). The CSCs are multiwire proportional chambers which means that two cathodes are able to measure two coordinates. The anode-cathode spacing is equal to the anode wire pitch 2.54 mm. This leads to a patial resolution of better than 60 μm.

Trigger Detectors

RPCs are gas detectors consisting of parallel bakelite plates without wires. An electric field between the plates allows avalanches from ionising tracks. The signal is read out via capacity coupling to metallic strips, which are mounted on the outer surface of the resitive plates. It provides a resolutions of 1 cm with a digital readout in 1 ns.

TGCs are based on the same technology as CSCs. Due to a small wire distance a short drift time can be achieved thus a good timing resolution is feasible.

The muon trigger chambers provide three informations: bunch-crossing identification, well defined p_T threshold and measurement of the muon coordinate orthogonal to the direction determined by the precision-tracking chambers.

3.3.4 ATLAS Trigger

Due to the very high bunch crossing rate of 25 ns data of 1 Petabyte per second is produced. It is not possible to read out, store or analyse this amount of data. Therefore the ATLAS Trigger system [31] selects the "interesting" events. The fundamental principle of the event selection is as follows:

1. For the first decisions only few informations of all events are available.
2. Then less events have to be analysed, hence all informations within defined regions of interest are used for a more detailed selection.
3. In a final step the events selected above can be analysed with the full event informations.

These steps of the event selection are realized in the ATLAS trigger system in different trigger levels. The schematic overview of the ATLAS trigger system is shown in Fig. 3.10.

Level 1 trigger [30]: Level 1 is hardware based because of the high event rate 40 MHz. The muon trigger system described above and the calorimeter informations (not with full granularity) are used to select events with high transverse momentum. To reduces the event rate to 25 kHz the L1 trigger has to make the decision in 2.5 μs. Therefore one or more **R**egions **o**f **I**nterest (RoI) are defined. RoIs are regions within the detector where processes with interesting features have been selected. RoI data includes informations about type, identified feature and passed criteria e.g. thresholds.

Level 2 trigger: After that the Level 2 trigger analyses the RoIs defined by Level 1. To select the events in more detail Level 2 uses the full calorimeter granularity and the Inner Detector informations in the RoIs (approximately 2 % of the total event

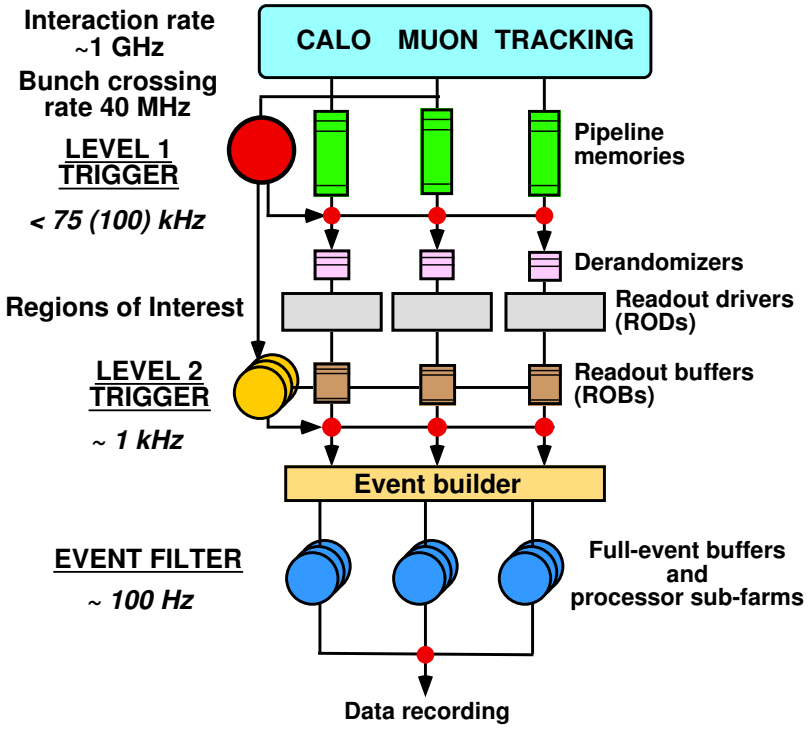

Figure 3.10: *The three steps of the ATLAS Trigger System: the* ***Level 1*** *trigger, the* ***Level 2*** *trigger and the* ***Event Filter****. It reduces the total data rate from* $1\,^{PB}/_{s}$ *to* $100\,^{MB}/_{s}$ *[30].*

data). Using tighter p_T cuts reduces the event rate to approximately 3.5 kHz with an event processing time of 40 ms.

The Event Filter (EF): The EF is able to reconstruct the full event. Its selections are implemented using full offline analysis procedures. So e.g. track fitting, vertex reconstruction and bremsstrahlung recovery for electrons can be calculated. The EF has to reduce the event rate to 100 Hz with an average processing time of 4 s. Additionally the EF is able to monitor the lower trigger levels and detector performance with the full event informations.

Thus the data rate has been reduced in total from $1\,^{PB}/_{s}$ to $100\,^{MB}/_{s}$.

Chapter 4

Data Simulation and the ATLAS Event Data Model

Before such a huge experiment as the ATLAS detector can be run studies have to be made to find the optimal setup. This is usually done by simulations. Once the experiment is running, simulation remains necessary to understand its output. First the physical model has to be simulated by an event generator. Then the detector interactions and the detector output have to be studied. Out of this data reconstruction algorithms for the several particles can be developed. For all these steps several tools are implemented and combined in a framework for the ATLAS experiment named ATHENA [32] (a schematical overview of the full chain from generating an event to the analysis is given in Fig. 4.1).

This chapter will first give an introduction in *PYTHIA* as an example of a typical Monte Carlo (MC) event generator. Then it describes the *GEANT4* detector simulation tool. The following section briefly summarises the data reconstruction and particle identification. In the last section an overview of the ATLAS Event Data Model (EDM) is given.

4.1 Event Generation with PYTHIA

The *PYTHIA* [34] program generates high-energy physics events of colliding particles. Out of two incoming particles their interaction and the outgoing particles are derived. *PYTHIA* works with pp, $p\bar{p}$, e^+e^- or $\mu^+\mu^-$ as incoming particles. To describe the full event the event generator has to simulate several physics aspects. These aspects have to follow the evolution of such an event (c.p. Fig. 4.2).

Hard Processes and Parton Distribution

As already mentioned in Chap. 2.3.3 protons contain partons. Thus the parton distribution has to be included when simulating the pp collision. Therefore Fig. 4.2 a) shows the hard subprocess producing a W^+ boson. This can be calculated through matrix elements. The W^+ boson decay is affected by to the hard subprocess.

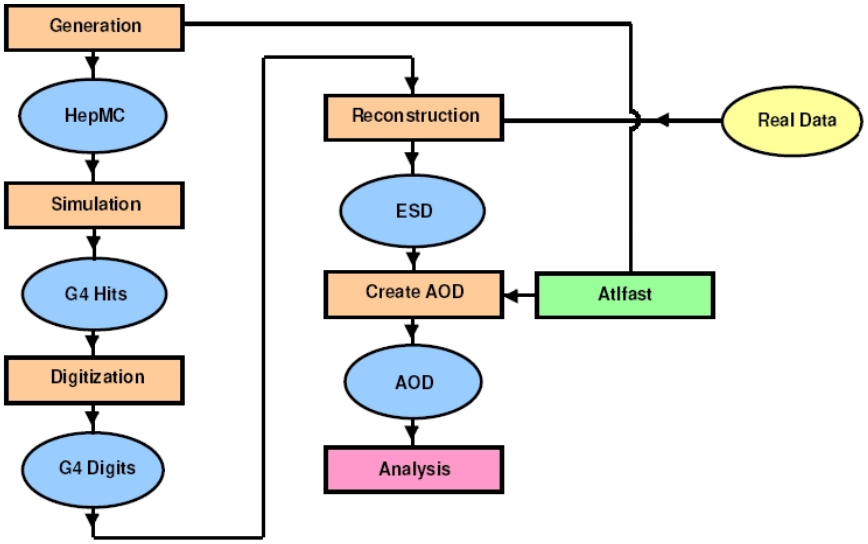

Figure 4.1: *Chain of simulating and reconstructing data in ATHENA the ATLAS framework [32]; (blue ellipse) the data formats and (red rectangle) the processes. The left chain describes the event and detector simulation the right chain the reconstruction and data preparation (this part has to be passed by real data too (yellow ellipse)). To simulate rare events (e.g. SUSY events) high statistics are required. Therefore a fast simulation is available (ATLFAST green rectangle). It produces in one step the generation of the events, a simplified detector simulation and reconstruction. It is not as exact as the full simulation but much faster [33].*

Initial- and Final Radiation

In all processes with charged or colour charged particles initial- or final state photon or gluon radiation will proceed (c.p. Fig. 4.2 b) in green and c) in blue). Due to large corrections to the overall topology of the event these radiations have to be included into the simulation.

The corrections can be determined by two approaches. The matrix element method is the more exact method. It calculates the Feynman diagrams order by order. But it is much more time consuming due to the fact that calculations in higher order become more and more complex. This approach becomes less relevant for the full structure of the event for higher energies. Then the second approach, parton-shower, which only approximates the full matrix elements by simplifying the kinematics, is used. Therefore an arbitrary number of branchings of one parton into two or more are used to describe multijet events.

Beam Remnants and Multiple Interaction

Due to the proton structure only partons of the protons are involved in the hard process. They will take a fraction of the total beam energy. The rest is left in the proton rests, the beam remnants. If for instance a u quark is involved in the hard process the beam remnant is an ud di-quark with a antitriplet colour charge. Thus the beam remnant is colour connected to the hard interaction (c.p. Fig. 4.2 e)).

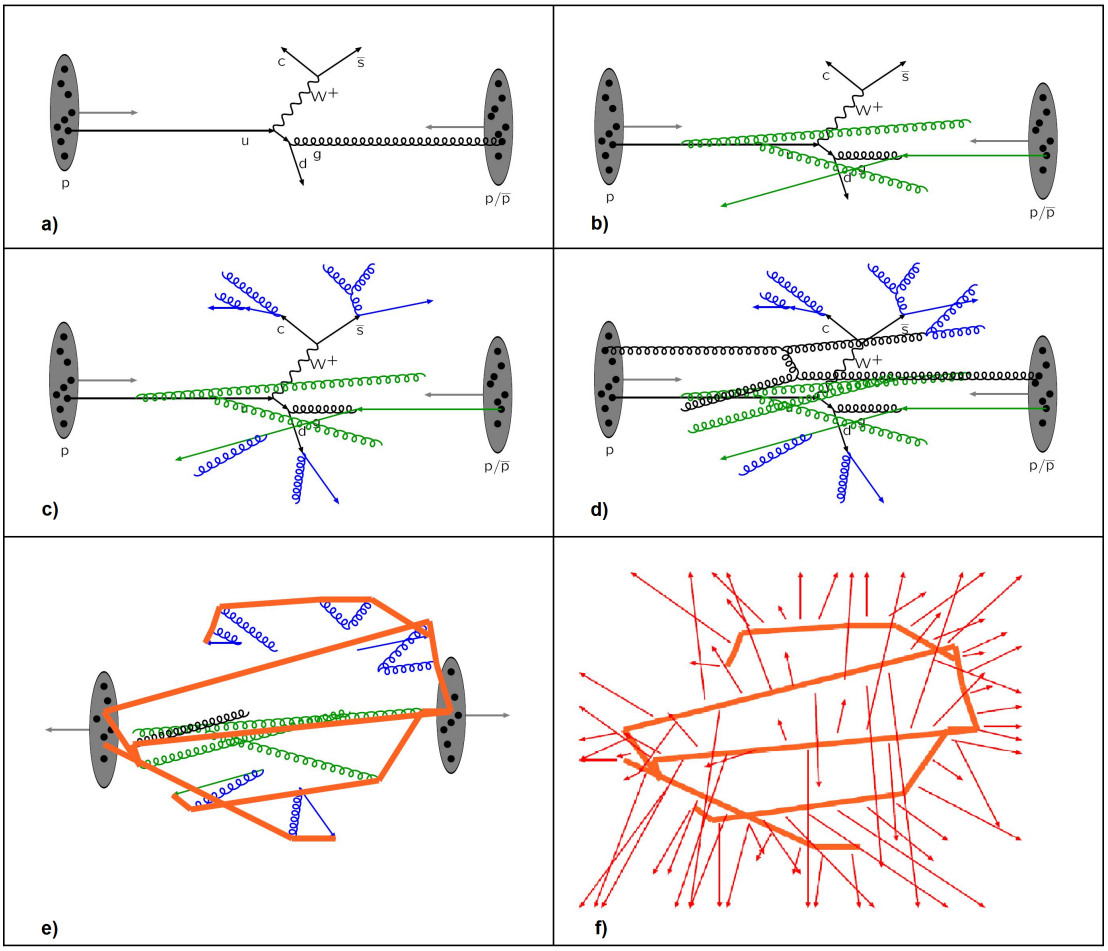

Figure 4.2: *Several steps of MC event simulation by PYTHIA [35]:* ***a)*** *The hard subprocess of a pp collision can be described by matrix elements. The produced intermediate W-boson decays via electroweak physics.* ***b)*** *Initial state radiation can be described as spacelike parton shower (here gluon radiations in green).* ***c)*** *Final state radiation can be described as timelike parton shower (in blue).* ***d)*** *Protons contain multiple partons. If more than one parton of a proton interact via a hard process it is called multiple interactions (shown in black + initial- (green) and final (blue) radiation)* ***e)*** *After these processes all coloured particles and beam remnants are connected by colour confinement strings (red).* ***f)*** *Primary hadrons are produced from the string fragments. Unstable hadrons decay further on.*

Also associated with the substructure of the proton are multiple interactions. Due to the multitude of partons there is a probability that more than one parton will have a hard interaction. These underlying events are called multiple interactions (c.p. Fig. 4.2 d) gluon interactions in black).

Hadronisation and Decay of unstable Particles

coloured partons like gluons and quarks are colour connected. If they diverge the colour confinement is responsible for the hadronisation. This means that the energy "stored"

in the colour connection transforms the coloured partons into colourless hadrons (this is shown in Fig. 4.2 f)). In *PYTHIA* the hadronisation is implemented via the *Lund* string fragmentation framework.

Many hadrons are unstable and decay further. At the end of this step a realistic event is simulated. To detect all the generated final state particles a particle detector is needed. Therefore the interaction of these particles with the several detector materials have to be simulated too. This can be done with *GEANT4* and is described in the next section.

4.2 Detector Simulation with GEANT 4

The **GE**ometry **AN**d **T**racking (GEANT4) program [36] is a system of detector description and simulation tools. Originally it was designed for High Energy Physics experiments. By now it is also used in medical and biological sciences, radio-protection and astronautics. It simulates the passage of elementary particles through matter and their interactions.

In particle experiments as ATLAS, GEANT4 is used for design studies, detector optimisation and to develop and test reconstruction tools. Once real data is available, GEANT4 is used to interpret the experimental results.

Before a full simulated event can be reconstruced the generated particles of the MC event generator (e.g. PYTHIA) have to cross the detector. This passage through the detector is simulated by GEANT4. Therefore the exact geometrical volumes of the several subdetectors (which have been described in Chap. 3.3) and the physical effects of the interactions of the simulated particles with the several materials and magnetic fields have to be taken into account. Particles like e^-/e^+ pairs or bremsstrahlungs photons which are produced by photon/electron matter interactions as described in Chap. 2.3.2 are generated by GEANT4. Thus all the photon conversions which are studied in this thesis are simulated in this step, whereas the τ decays are simulated in the MC generator.

Then GEANT4 records the trajectories of the particles passing through the detector. This is done via simulating energy losses of the particles due to material interactions. The positions where such interactions taking place and the amount of energy are so-called hits, which are stored as simulated data. Via a graphical representation it is able to visualise the detector and the tracks of the particles.

The last step of event simulation is to convert the detector hits into so-called "digits". This means that the simulated signals, e.g. track hits or calorimeter clusters have to be converted into digital signals looking like the output of the readout elements of the detectors. During this transformation detector resolutions and other information losses due to hardware inefficiencies have to be taken into account. Then the simulated event can be reconstructed as a real detected event (last data format in the left column Fig. 4.1).

4.3 Particle Reconstruction and Identification

Before identifying several particles produced in one event, out of the detector informations - the "digits"- reconstructed objects have to be built. This is necessary in both cases: For simulated data produced via an MC generator and GEANT4, as well as for real data recorded during data taking.

Therefore two kinds of informations are available. Firstly all the hits in the Inner Detector and the muon spectrometer used to reconstruct tracks. Secondly the energy depositions in the electromagnetic and the hadronic calorimeter systems.

To reconstruct the tracjectory of a particle a track fitter combines space points to a track. The algorithm starts with the inner most subdetector and tries to combine several hits to one track. Afterwards if it is possible a so-called global Inner Detector track through the three subsystems is built. The hits registed by the muon spectrometer are also fitted to a track. If it is possible to match a track found by the Inner Detector and the muon system a global detector track is built.

The response of the calorimeter cells are evaluated via clustering algorithms. Therefore several quite different algorithms has been developed. In general a clustering algorithms builds clusters out of cells with significant energy deposition above noise thresholds lying nearby. By matching tracks found via the Inner Detector to such clusters either jets, in case of hadronic clusters, or electrons, in case of electromagnetic clusters, are created. Both τ leptons and photon conversions studied in this thesis are very complex and much more difficult to reconstruct and to identify. Therefore the reconstruction and identification of these objects are described in more detail in Chap. 5.

Having identified all particles of an event it can be decided if the event is interesting or if it can be removed. Due to the large amount of produced data a very detailed event selection has to be done. To register the removed events and to slim out the stored events several types of data sets are produced. They are described in the next chapter.

4.4 The ATLAS Event Data Model

After the data selection by the trigger system the ATLAS detector will produce approximately 3 PB of raw data per year. It is unfeasible to distribute such a vast amount of data to the world wide collaboration. Therefore several datasets with different amount of stored informations are produced. Begining with the raw data the several datasets are becoming smaller and smaller and the content becomes more and more specific for different analyses [37].

Raw Data Objects

The data coming out of the Trigger system is the so-called *byte-stream*. The byte-stream informations are converted in C++ objects. These are the *Raw Data Objects (RDO)*.

For simulated data truth objects called *Simulation Data Objects* (SDO) can be stored at this level too. In the SDOs information about all the particles and/or noise that contributed to the existence of the signal produced in the given sensor are stored. These SDOs are used to cross check the reconstruction algorithm during the implementation. More about these truth object is described in Chap. 6.

Event Summary Data

Produced from a raw data object the detailed informations of the detector reconstruction are stored in *Event Summary Data (ESD)*. To allow a fast tuning of the reconstruction and calibration the ESDs contain the sufficient informations for particle identification, track re-fitting and jet-calibration. The ESD size is about 500 kB per event.

Analysis Object Data

With a target size of 100 kB per event the *Analysis Object Data* (AOD) is a summary of the reconstructed event which contains all sufficient informations for a commom analysis. It can be produced out of ESDs. It was planned to make several tailor-made streams of AODs for the several needs of the physics community. This is now shifted to the next step to the derived physics datasets.

Derived Physics Datasets

The *Derived Physics Datasets* (DPD) are classified in three data sets. The D^1PD (Primary), D^2PD (Secondary) and D^3PD (Tertiary). To bring down the amount of data the interesting events and objects have to be selected. This is realized by three methods: Skimming which means removing uninteresting events; Thinning removes unnecessary objects and Slimming removes unneeded details from objects.

Primary DPD: It is foreseen to have in $\mathcal{O}(10)$ D^1PDs. They are not very specific to an analysis but more tailor-made as the AODs. The D^1PDs are additionally subdivided into so-called *performance DPDs* and *physics DPDs*. To simplifiy this classification the *performance DPDs* are more ESD like, to re-reconstruct them for detector calibration and performance studies while the *physics DPDs* are more tailor-made meant for physics analyses. These *physics DPDs* are smaller than the AODs and thus can be delivered much faster to the differnet users. The data structure of D^1PDs is still the same as the AOD structure. Thus D^1PD can be re-reconstruced if the calibration has to be changed.

Secondary DPD: The D^2PDs are more specific than the D^1PDs. For the production the standard ATHENA framework or an analysis framework like EventView[1] can be used.

Tertiary DPD: Whereas the above described datasets are written in the POOL format the D^3PD does not need to be POOL-based. It is typically the output of an analysis like a ROOT[2] ntuple. D^3PDs can be produced out of each step of the data processing. The D^3PDs production out of an ESD is described in more detail in Chap. 6 because that is the production method used to develop the specific tools needed for this thesis. To summarize this section the full ATLAS analysis flow is shown in Fig. 4.3. From left to right the flow of the data out of the detector to the final histograms and the publications is illustrated.

[1] EventView is an analysis framework which factorise the complex process of physics analysis into well defined modules that represent a single task or a grouped operation.[38]

[2] Root is a Object Oriented framework for large scale data analysis [39]

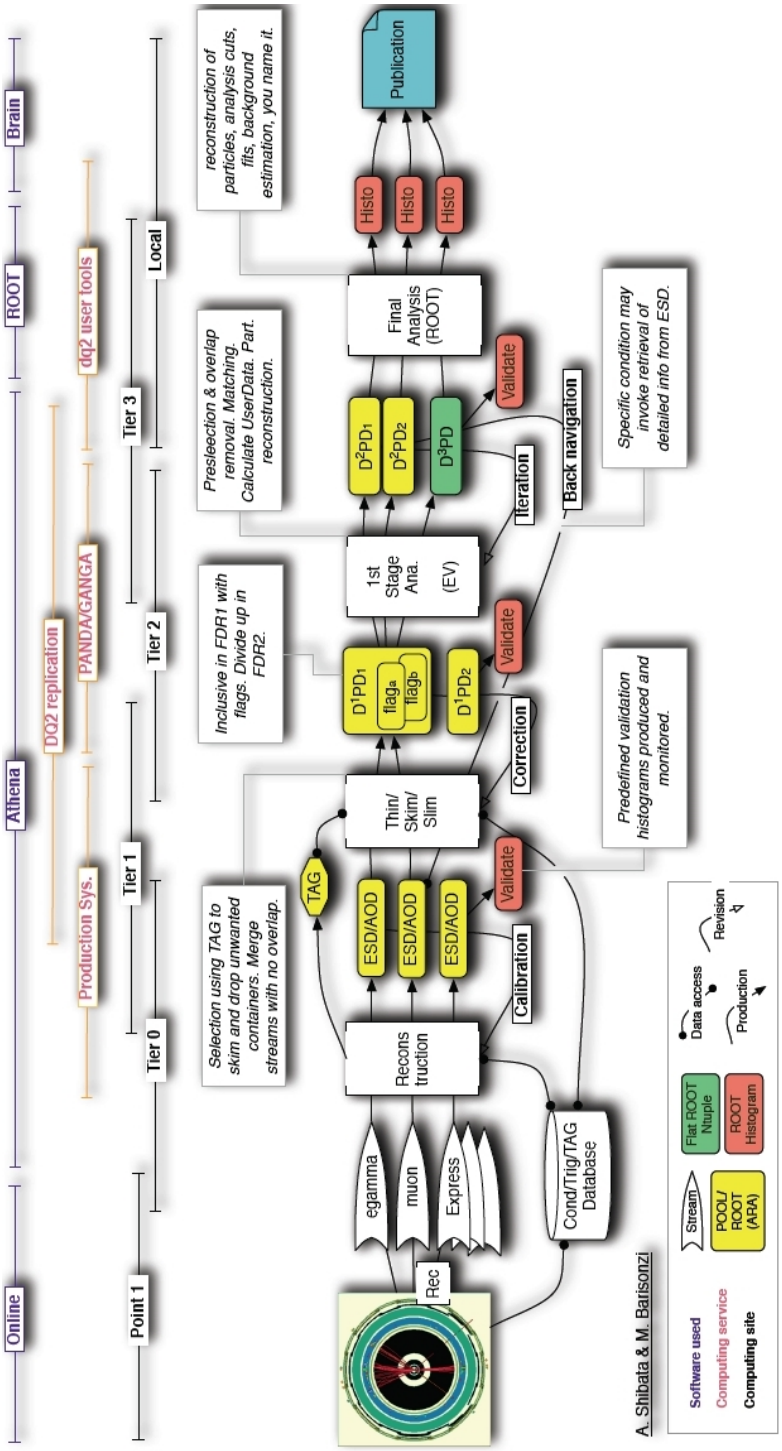

Figure 4.3: *Full ATLAS analysis flow model [40]. Out of the detector (on the left) several streams (e.g. egamma) "bring" the data to the reconstruction. There the ESDs and AODs are produced. After skimming thining and slimming the D^1PDs are available. The 1. stage analysis (e.g. done by EventView (EV)) produce the more specific D^2PDs and D^3PDs. Based on D^2PDs or D^3PDs a final (e.g. ROOT based) analysis can be done. The results (histograms) then can be plotted in the publications. The lines above that schema represent the used frameworks (blue), the delivery systems (red) and the storage locations (black).*

Chapter 5

The τ Lepton and Conversion Reconstruction Algorithms

Several complex reconstruction algorithms are needed to extract the complete information of each ATLAS event. This chapter will only concentrate on the specific algorithms for τ lepton and photon conversion reconstruction. Reconstructing photon conversions in the very dense τ evironment is a very complex and difficult task. The existing ATHENA offline reconstruction [32] contains tools for photon conversion reconstruction and tools for τ reconstruction. Both algorithms are optimized for different purposes. This chapter will give a short summary of the existing tools, their purposes and some shortcomings by combining the algorithms with the existing definitions. First the τ reconstruction algorithm is explained to give an impression of the number of tracks within a so-called τ cone. Afterwards the photon conversion algorithm is described in more detail, because the major focus of this thesis is on the optimisation of this tool for the very special τ environment.

5.1 Reconstruction and Identification of τ Leptons

Since it would be difficult to distinguish electrons and muons from leptonic τ decays and from primary interactions, it is only possible to reconstruct specifically hadronic τ lepton decay modes. This reduces the reconstructable fraction of the τ leptons by 35,2 %. The background of the reconstruction of hadronic τ decays at a hadron collider is strongly dominated by QCD jets.

The basis to distinguish hadronic τ decays from QCD jets is the lower track multiplicity and the more narrow decay cone. Therefore characteristics of the track system and the shapes of the calorimeter showers have to be taken into account. Thus the τ reconstruction algorithm is a so-called higher level algorithm because it uses objects which have been already reconstructed by other algorithms. The track reconstruction is done by the Inner Detector and the dedicated algorithms. The topological clustering of the energy deposition in the electromagnetic and hadronic calorimeters has been completed by the cluster reconstruction.

In the ATLAS offline reconstruction software two complementary algorithms for τ reconstruction have been implemented at present [16].

Calorimeter based (*TauRec*): Building the τ candidates is based on the clusters in

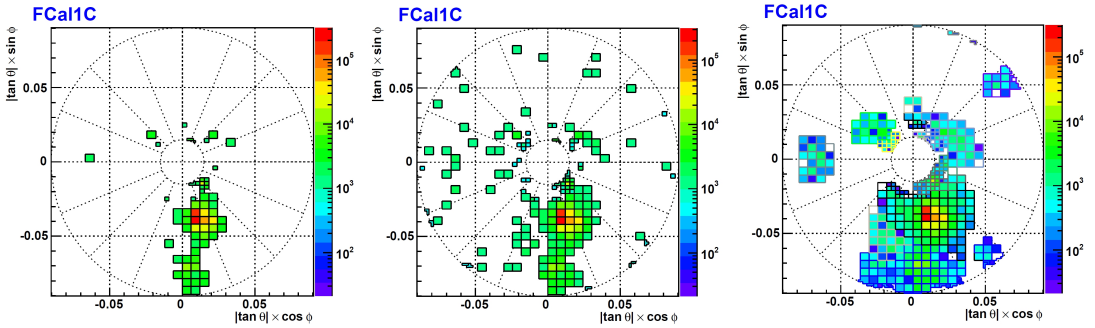

Figure 5.1: *Example of Topological Clustering in the forward calorimeter. All plots are 2 dimensional projections with respect of the FCal position to the interaction point.* ***Left:*** *For building a cluster cells with a measured energy larger than 4σ above the noise threshold are used.* ***Middle:*** *Then cells nearby with an energy deposition larger than 2σ above the noise threshold are added.* ***Right:*** *As last step the fired cells closed by are added to the cluster. [41]*

the hadronic and electromagnetic calorimeters. For identification then the track and the calorimeter informations are used.

Tracking based (*Tau1P3P*): This approach is based on seeds which are built out of "good" tracks. These seeds require low track multiplicity in the core region of the reconstructed object. Then the energy of the candidate is calculated via an energy-flow algorithm.

For the future both approaches will be merged to one algorithm to get the advantages of both methods. For this thesis the *TauRec* algorithm is used only. Therefore it will be described in more detail in the next section.

The *TauRec* Reconstruction Algorithm

Building a τ object is split in two steps. First the τ candidates have to be built. This is done in the **reconstruction** phase. Then the τ candidates have to be identified as τ. This is done during the **identification** which is based on a likelihood discrimination to separate τ jets from QCD jets.

5.1.1 Reconstruction

The seed for a τ candidate is a so-called *Cone4TopoJet*. Therefore the topological clustering algorithm [42] adds a variable number of nearby calorimeter cells with a significant energy deposition above the noise threshold to a cluster.

The *Topological Clustering* algorithm works in three steps. Each step is shown in Fig. 5.1 from left to right. First the algorithm searches for the cell with the most significant deposition (4 σ above the noise), which is not yet associated to a cluster. Then it adds all cells nearby which are above a neighbourhood threshold (2 σ above the noise) to the cluster. In case a cell already belongs to another cluster they are merged. The last step is an additional iteration of neighbouring cells. If they have an energy depostion

they are also added to the cluster. These neighbouring relations are applied in all spatial directions and different calorimeter layers. Such a cluster is called a *TopoCluster*.

After this step typically for each particle one cluster has been built. Then a seeded cone algorithm, in case of the τ reconstruction a cone of $\Delta R = 0.4$, is used to combine these *TopoClusters* to a *TopoJet*.

Then each *Cone4TopoJet* with a transverse energy $E_T > 10\,\text{GeV}$ and an absolute pseudorapidity $|\eta| < 2.5$ is considered as τ candidate. All tracks reconstructed by the Inner Detector are associated to the τ candidate if they pass the track selection criteria below.

Track Selection Criteria:

$\Delta R <$ **0.3:** Each associated track has to lie in a cone smaller than $\Delta R < 0.3$ around the TopoJet center.

$p_T >$ **1.0 GeV :** The transverse momentum of each track has to be larger than 1.0 GeV. This cut reduces tracks of minimum bias events.

$d0 <$ **1.5 mm :** The pseudo impact parameter stands for the smallest distance of the track from the primary vertex. This cut supresses secondary vertices of particles with a longer lifetime (decay length) as the τ lepton.

$\frac{\chi^2}{ndf} <$ **3.5:** The χ^2/ndf value shows the goodness of the track fit of the Inner Detector algorithm. The cut on this variable guarantees a needed track quality for the associated tracks.

Number Si Hits (Pixel + SCT) $\geq$ 6: At least six of the silicon subdetector layers (Pixel and Micro Strip Detector) have to have measured a hit (cp. Sec. 3.3.1). If the track for example only is detected in the barrel region up to eleven hits are possibe (cp. Fig. 3.6 three Pixel layers plus four double Micro Strip layers).

Number Pixel + B-Layer Hits $\geq$ 1: This cut requires at least one hit in the inner most silicon detector. As described in the next section many conversions take place in a more outer region of the detector. The π^0 and two photons in which the π^0 decays before making a photon conversion do not leave any track, any Pixel or B-Layer hit. Thus this cut is very powerful to supress tracks from photon conversions, but it also rejects pion tracks (cp. fig 7.1 where the track ID before and after applied track selection criteria are shown). Nevertheless an explicit photon conversion track reconstruction algorithm will improve the vetoing of such tracks.

The track selection criteria with a special focus on how they effect the reconstruction of electrons and positrons within a τ candidate is shown in Fig. 5.2. After building the τ candidates a preselection for the identification is made: Only candidates with a track multiplicity between 1 and 3 are kept. In Fig. 5.3 the track multiplicity of τ candidates before the likelihood cut is shown.

5.1.2 Identification

After building the τ candidates the full calorimeter and track informations are available. Out of the eight variables described below likelihood values binned in several p_T bins in

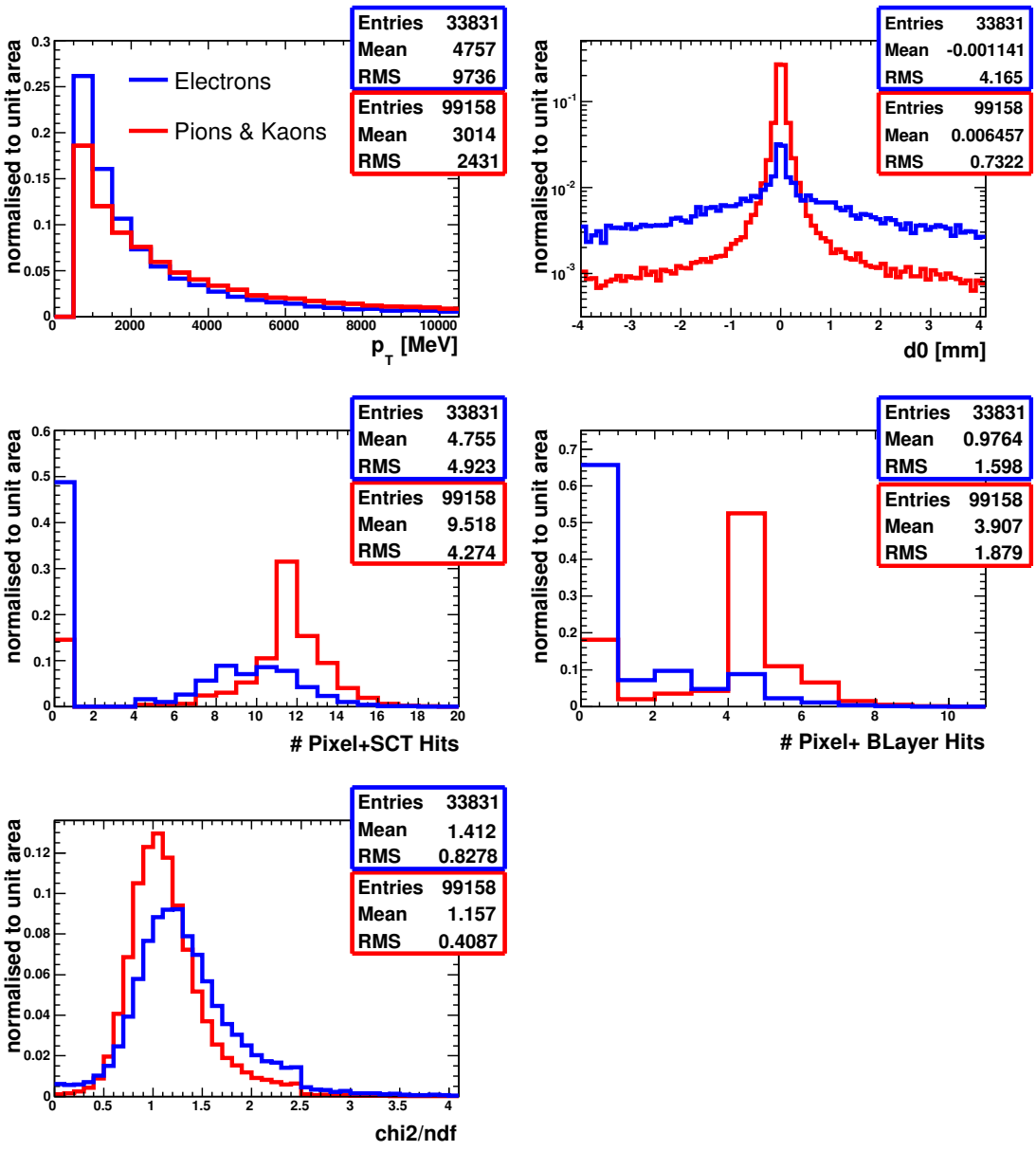

Figure 5.2: *TauRec track selection criteria for all tracks in a cone around the TopoCluster center $\Delta R < 0.3$. Here only the electrons and positrons are differentiated from pions and kaons because here it should be shown how the track selection criteria effect the electron/positron reconstruction in τ candidates.*

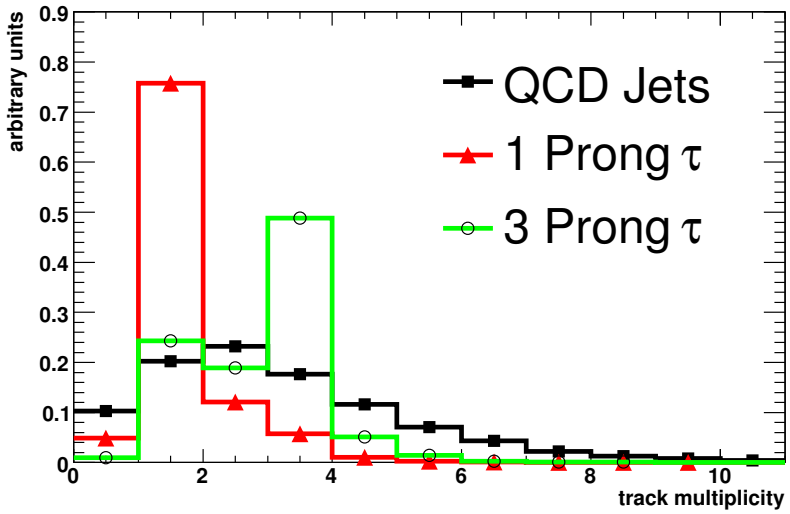

Figure 5.3: *Track multiplicity of τ candidates matched to 1 and 3 prong τ decay modes and in black matched to QCD background. Preselection of likelihood discrimination is the limitation of the track multiplicity $1 \leq N_{Tr} \leq 3$. This selection cuts the tails of the one and three prong decays to larger track multiplicities off. If the additional tracks are photon conversion tracks an explicit photon conversion veto may improve the reconstruction algorithm.*

GeV (10, 17, 32, 45, 70, 100, 150, 220, 500 and higher) are calculated. It determines how likely a τ candidate comes from a hadronic τ decay or from a QCD Jet.

EM-Radius: To make use of the more narrow shower profile of the τ decay the EM-Radius is calculated.

$$R_{em} = \frac{\sum_{i=1}^{n} E_{Ti}\sqrt{(\eta_i - \eta_{cluster})^2 + (\phi_i - \phi_{cluster})^2}}{\sum_{i=1}^{n} E_{Ti}} \tag{5.1}$$

i runs over the electromagnetic calorimeter cells in the cluster of $\Delta R < 0.4$, n denotes their number and E_{Ti} is the transverse energy in the cell i. This variable is quite powerful in the low p_T region where the QCD jets are much broader than the τ jets, but for higher E_T it becomes more and more inefficient. This is due to the higher boost and the fact that the high boosted QCD jets are getting more narrow.

Isolation in the calorimeter: The isolation fraction criteria is also based on the narrowness of the τ jets. The well colimated τ decays are supposed to have a higher fraction of their energy in a small cone. Therefore the fraction of the energy in a region $0.1 < \Delta R < 0.2$ and the total energy is calculated.

$$\Delta E_T^{12} = \frac{\sum_{j=1}^{n'} E_{Tj}}{\sum_{i=1}^{n} E_{Ti}} \tag{5.2}$$

Where j runs over all electromagnetic calorimeter cells in a cone of $0.1 < \Delta R < 0.2$ around the τ axis, n' denotes their numbers, i runs over all electromagnetic

calorimeter cells in a cone of $\Delta R < 0.4$ of the center of the object with the number n. E_{Tj} and E_{Ti} denote the deposited transverse energies in the cells j and i, respectively. E_T^{12} is also strongly dependend on E_T and less efficient for higher boosted candidates.

Number of assiciated tracks: This is the number of tracks lying in a cone of $\Delta R < 0.3$ around the TopoCenter which have passed the track selection criteria. This variable can be faked by additional e^+ and e^- tracks of photon conversions. Thus all studies on the impact of photon conversions on the τ reconstruction have to be done before the likelihood discrimination.

τ Charge: The charge of a τ is defined as the sum of the charges of the decay products. Track pairs from photon conversions should not fake the total charge because such a pair is in addtion neutral. But if just one of the conversion tracks can be reconstruced due to reconstruction inefficiencies the τ charge is influenced, too.

Number of Hits in η-strip layer: To differentiate the τ leptons from QCD jets the fine η strips (cp. Fig. 3.8) in the electromagnetic calorimeter are used, too. In a cone of $\Delta R < 0.4$ around the cluster center all energy depositions larger then 200 MeV are counted as hits. In case of low energy candidates a significant fraction of τs with nearly no hits are counted in the η layer. In comparable E_T regions QCD jets deposit energy and thereby leave hits. Even for high E_T jets a discrimination between QCD and τ jets is possible, because the QCD jets make more hits than τ jets. Photon conversion tracks can fake this variable, too. Electrons mainly deposit their energy in the first and second electromangetic calorimeter samplings.

Transverse energy width in the η-strip layer:

$$\Delta\eta = \sqrt{\frac{\sum_{i=1}^{n} E_{Ti}(\eta_i - \eta_{cluster})^2}{\sum_{i=1}^{n} E_{Ti}}} \tag{5.3}$$

Where E_{Ti} is the transverse energy in the η-layer, i runs over all cluster cells in $\Delta R < 0.4$. This variable is comparable to the EM-Radius, powerful in low E_T region and more and more inefficient for higher boosted jets.

Lifetime signed pseudo impact parameter significance: The so-called pseudo impact parameter is a 2-dimensional impact parameter of the smallest distance of the track from the beam axis. Out of this value the lifetime signed pseudo impact parameter significance is calculated.

$$\sigma_{IP} = d_0/\sigma_{d_0} \times sign(\sin(\phi_{cl} - \phi_{tr})) \tag{5.4}$$

Where d_0 is the pseudo impact parameter, σ_{d_0} the error, and ϕ_{cl} and ϕ_{tr} the coordinates at the calorimeter for the cluster and the point of closest approach for the track. It is defined that if the decay happens in flight direction it has positive sign. For particles with a lifetime like τ leptons positive values are expected. Due to the increasing decay length with higher boosts the identification power increases with E_T.

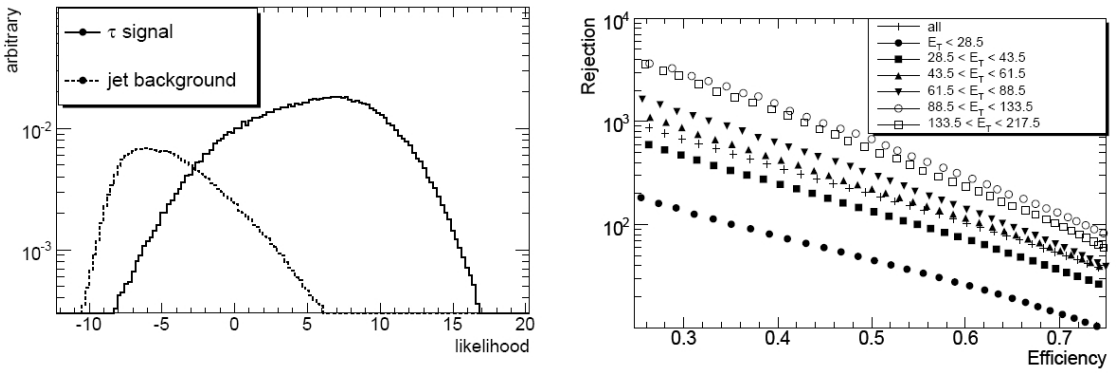

Figure 5.4: ***Left:*** *The Likelihood (LLH) distribution of QCD jets and τ jets calculated of the above described variables. The LLH is applied after the preselection of the track multiplicity (1≤ N_Tr ≤3).* ***Right:*** *Rejection of QCD background over the reconstruction efficiency of τ jets for different p_T bins. Results taken from [16]*

E_T over p_T of the first track: Due to the fact that for τ jets a concentration of the energy in the leading track is expected E_T over p_T of the leading track will lead to a discrimination value (in contrast to the uniform energy distribution in QCD jets). This variable is peaking around value of one for τs, whereas for 1 prong decays without neutral pions a value of one is expected. 1 prong decays with π^0s or 3 prong decays result in values above one. This variable does not very strongly depend on E_T for τs but QCD jets getting more and more signal like for higher E_T. Therefore it has a good discrimination for low and medium E_T.

The distributions of all likelihood input variables described above are plotted in Fig. A.5 in the appendix. For the three discrete values N_{Tr}, $N_{\eta-hits}$ and τ_{charge} the ratios are directly taken from the histograms, for the continuous variables R_{em}, ΔE_T^{12}, $\Delta\eta$, $d^{d_0}/_{\sigma_{d_0}}$ and $^{E_T}/_{p_{T,1}}$ fits of appropriate functions in several E_T bins have been performed for the distribution.

A likelihood value determined via a multivariate analysis with the above discribed variables leads to a distribution shown in Fig. 5.4 (left). It shows a good discrimination power between τ jets and QCD background. For the standard τ definition a cut on LLH> 4 is chosen. The QCD-jet rejection over the τ reconstruction efficiency for different p_T regions is shown in Fig. 5.4 (right).

As described in Chap. 2.3.1 the decay modes with neutral π^0s have to be studied for photon conversions. Therefore Fig. 5.5 shows in the upper row the reconstruction efficiencies for the different τ decay modes with and without neutral pions. The left plot shows the efficiencies for τ candidates after reconstruction and the right plot for τs after reconstruction and identification via the LLH cut. The efficiency is defined as (e.g. 1 prong decay):

$$\text{eff.}_{1Prong} = \frac{reco_{all} \overset{matched}{\rightarrow} truth_{1Prong}}{truth_{1Prong}}. \tag{5.5}$$

The efficiency for the reconstruction of τ candidates is between 90% and 100 %. The losses are due to inefficiencies in the track reconstruction for very low p_T tracks for example. But most of the τs are picked up as candidates.

After reconstruction and identification both 1 prong decays with neutral pions and 3 prong decays with neutral pions are significantly inefficient in comparison to decay modes without neutral pions. Here the already mentioned problem of the additional charged tracks from the photon conversions and the preselection of the track multiplicity during the identification via the LLH leads to additional inefficiencies.

The lower plots (Fig. 5.5) with the purtiy before (left) and after identification (right) show the large fraction of 1 prong τ decays with π^0s. This is theoretically expected: 76.6 % of 1 prong decays contain π^0s (cp. Tab. 2.6). The black distribution is the purity of truthmatched τs. The shown "purity" for the several τ decay modes is defined as the τ purity times the branching fraction of the individual mode. To calculate the plotted purity the following formula is used (e.g. for 1 prong τ decay):

$$\text{pur.}_{1Prong} = \frac{reco_{all} \overset{matched}{\rightarrow} truth_{1Prong}}{reco_{all}}. \tag{5.6}$$

Such a definition has to be used, because the reconstructed τ candidates/τs after likelihood do not have the decay mode information a priori.

The QCD background, the main impurity above the black distribution in the lower plots, could be reduced from $\approx$ 60 % in the left plot to $\approx$ 20 % in the right plot by the LLH cut. The other entries are the purity fractions from the total τ purity. Additional efficiency and purity plots as comparsion of τ candidates and τs after the LLH identification are shown in the appendix Chap. A.2.1.

To study the effects of an explicit photon conversion track reconstruction algorithm the track multiplicity of the τ candidates has to be corrected. After such a correction the LLH variable has to be recalculated. This thesis concentrates on the photon conversion track reconstruction in the very dense τ environment. To give an impression of a possible improvement, the corrected track multiplicity is shown in Chap. 7.7.

5.2 Reconstruction and Identification of Photon Conversions

Before describing the photon conversion reconstruction and identification it has to be specified how the experimental setup of the ATLAS detector biases the production of photon conversions. As described in Chap. 2.3.2 photon conversions are photon-material interactions. Thus the probability of a photon converting into an electron-positron pair is proportional to the amount and density of material in the trajectory of the photon. For the ATLAS Inner Detector the radiation length as a function of the pseudorapidity is shown in Fig. 5.6.

Thus the probability of a conversion depends on the pseudorapidity and the distance from the beam axis. This is shown in Fig. 5.7. In the most central region, $|\eta| < 0.5$, the probability is lowest. There no services are installed and the amount of tracker material is greatly reduced. Overall about 40 % of all photons will convert before reaching the surface of the electromagnetic calorimeter [25]. As described in Sec. 3.3.1 the Transition Radiation Tracker covers only a region of $|\eta| < 2.1$. The lack of TRT detector elements at higher pseudorapidities results in a limitation of the photon conversion reconstruction to the pseudorapidity of $|\eta| < 2.1$.

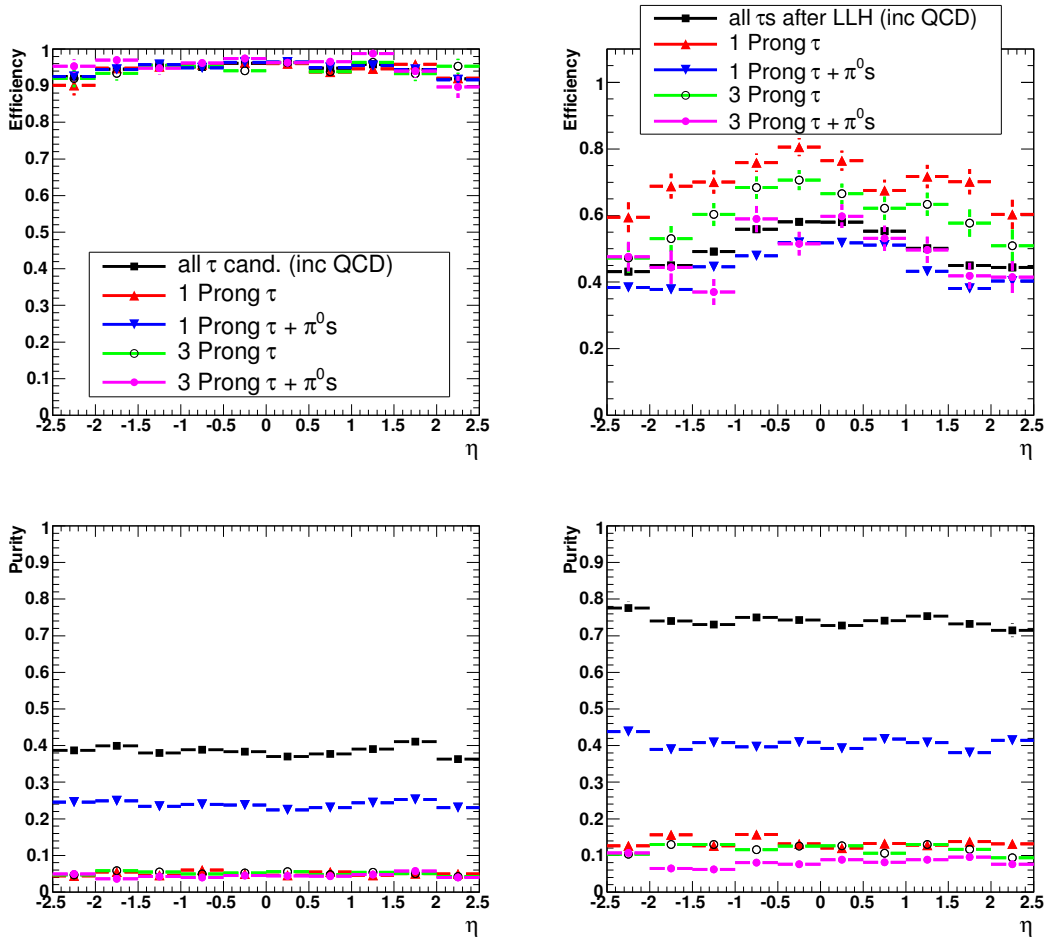

Figure 5.5: *Efficiency and purity in η of τ candidates in the left column and τs after likelihood identification in the right column, for several decay modes.*

Based on the event (Pythia) and detector (GEANT4) simulation of about 10 000 $Z \rightarrow \tau\tau$ events the photon conversions shown in Fig. 5.8 occur in the Inner Detector. Each dot represents a simulated vertex of a photon conversion. It has to be mentioned that not every conversion comes directly from a τ decay, additional prompt photons and photons from QCD jets lead to photon conversions too. Then it has to be taken into account that only tracks which can be reconstructed may impact the τ track multiplicity. Finally, only tracks with a transverse momentum larger than $p_T > 500\,\text{MeV}$, a pseudo-rapidity smaller $|\eta| < 2.1$ and a distance from the beam axis smaller than $R < 800\,\text{mm}$ can be reconstructed by the *InDetConversionFinder*, the conversion reconstruction tool. With that constraints the equivalent plot has changed to Fig. A.8 shown in the appenix Sec. A.2.2.

Like the *TauRec* algorithm, the photon conversions are reconstructed in two steps. First the so-called *VxCandidates* are built. This step is the photon conversion reconstruction and is described in the Sec. 5.2.1. Section 5.2.2 covers the identification of photon conversions, which is done in a second phase. Due to the optimisation of these tools to a much higher region of the p_T-spectrum than the one of photon conversions

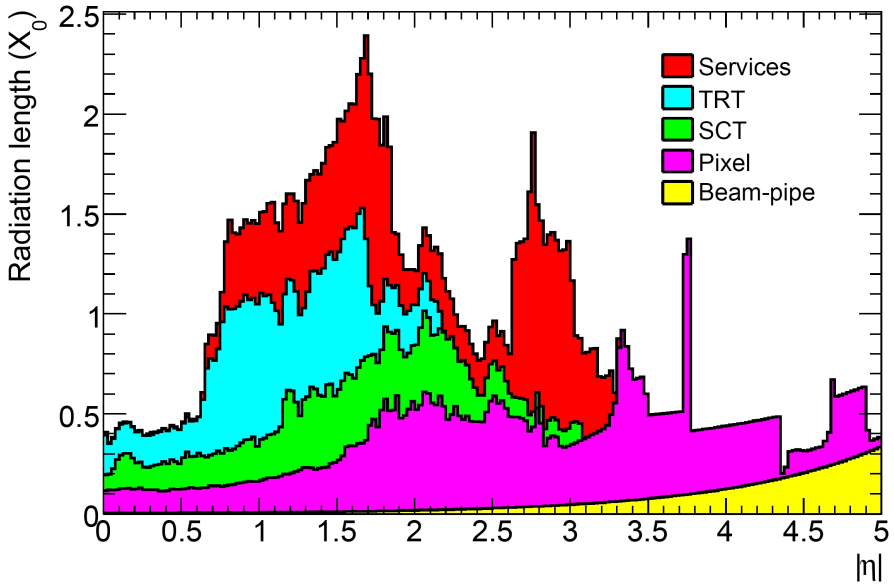

Figure 5.6: *The reason of photon conversions are several interactions with the detector material. Plotted are the radiation length of the different detector materials as a function of the pseudorapidity η. This plot is taken from [17].*

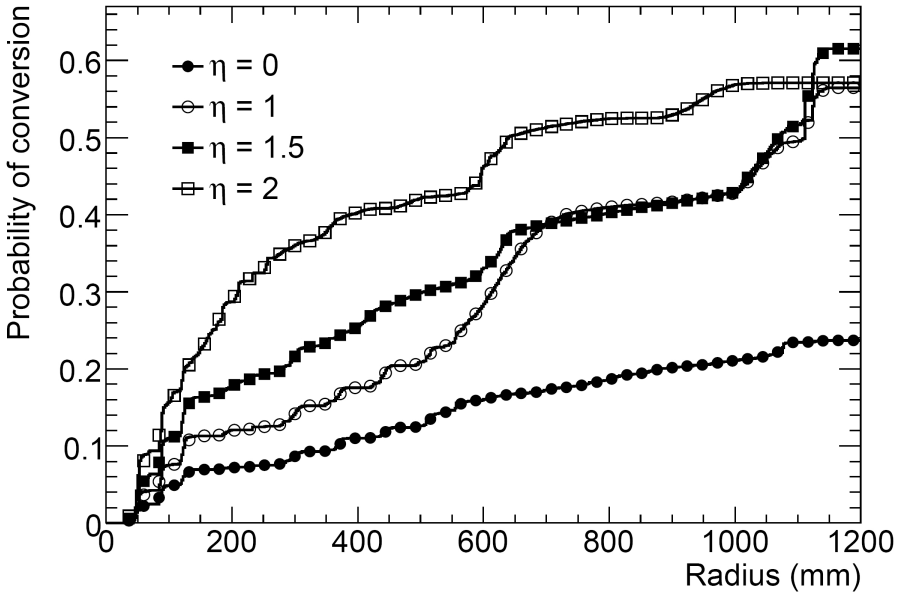

Figure 5.7: *The probability that a photon converts inside the ATLAS detector as a function of the radius R (distance from beam axis) for several pseudorapidities η. Plot taken from [17].*

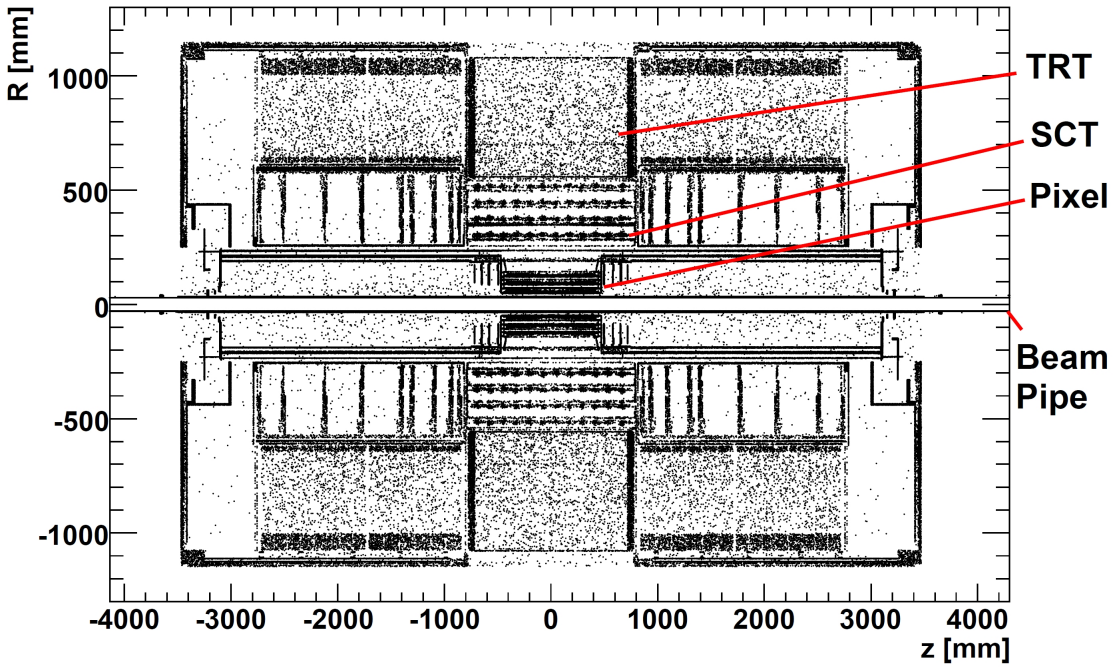

Figure 5.8: *Vertices of truth photon conversions in the ATLAS Inner Detector. Each black dot is the vertex of a simulated photon conversion (by PYTHIA and GEANT4 cp. Chap. 4). The beam pipe, several layers and discs of the silicon detectors and the structure of the barrel and end-cap TRT sections are clearly visible (cp. Fig. 3.5).*

from τ decays, the default definition of the photon conversion reconstruction cannot be used to study photon conversions in τ candidates. More about the inefficiency of the photon conversion reconstruction tools at low p_T and within very dense environments is explained in Sec. 5.2.3.

5.2.1 Building VxCandidates

The seeds of photon conversions (*VxCandidates*) are described in more detail, because these *VxCandidates* - before identification - pick up about 55% of electron and positron tracks in the τ candidates. The conversion reconstruction algorithm is implemented within the ATHENA framework [32] as part of the Inner Detector reconstruction software. It is one of the last algorihms in the post-processing phase. The reconstructed tracks are used to build positive/negative track pairs. After fitting a secondary vertex and adding so-called single track conversions, which are conversions with only one reconstructed track, the conversion seeds are stored as a vertex candidate collection named *VxCandidates*. With a matching of the VxCandidates to electromagnetic clusters, the photon conversion identification is done.

Tracking

The usual tracking method is a so-called Inside-Out track reconstruction algorithm. It reconstructs the tracks from the interaction region to the outer layers of the Inner De-

tector. Due to the fact that many photon conversions occur in a more outer region of the detector, they do not make any hits in the more inner region. Thus only using an inside-out tracking would not reconstruct many conversion tracks. To avoid such inefficiencies the tracking for photon conversions consists of three track selections: Inside-Out, Outside-In and standalone TRT tracks.

Inside-Out Track Reconstruction: For the seed building of the track reconstruction space points inside the Pixel and SCT subdetectors are used. Beginning with three space point combinations a geometrical tool searches for additional hits. Then a combinatorial Kalman-filter/smoothing formalism is used to add successive hits to the track. Not all seeds become tracks. E.g. if tracks share to many hits or do not pass the track selection criteria they will be discarded [25].

After this stage each "good" track candidate is assigned a TRT extension. This means a track candidate is extrapolated into the TRT, where transition radiation hits are searched. Have such hits been found they are used to refit the whole track. A "global" Inner Detector track has been built [43]. If a TRT extension has been found and the track quality of the refitted track is worse than before, the track fit without the extension is stored. Thus three categories of tracks are preserved:

1. Tracks without TRT extension (no extensions have been found)
2. Tracks with TRT extension - extension is used for final track fit
3. Tracks with TRT extension - extension is not used for final track fit (the quality of the track fit with the TRT extension is worse than before)

The reason for TRT extensions which are not useful for the track fitting are mostly outliers occuring due to large material interactions e.g. bremsstrahlung.

Outside-In Track Reconstruction: The starting point of Outside-In track reconstruction (back tracking) is the TRT. A histogramming technique formes initial track segments. The TRT is divided in 12 pseudorapidity slices to improve the longitudinal accuracy. Before a final Kalman-filter smoother procedure determines the track parameter of a segment a cut on a minimum number of straw hits has to be applied [43].

Then the reconstructed TRT segments are extended into the Si subdetectors. There space-point seeds are searched. At least two space point hits in the outer most three SCT layers are required. With the extension in the SCT subdetector the track can be refitted. Comparable to the Inside-Out reconstruction a "global" Inner Detector track is built and stored in a dedicated track collection. To minimize double counting the back tracking procedure excludes the TRT-straw hits and the Si detector space points which have been already assigned to inside-out tracks [17].

Figure 5.9 shows the improvement of the reconstruction efficiency for photon conversions of single photons with a p_T of 20 GeV using the back tracking algorithm as addition to the Inside-Out reconstruction. The efficiency mainly increases in a region between 300 mm and 450 mm. This is due to the minimum of at least two of the outermost three reqiured SCT hits. The second outermost SCT layer is at a radius of 443 mm (cp. Fig. 3.6). To reconstruct conversion vertices at larger radii tracks without any pixel or SCT hits have to be taken into account, too. This is done with the standalone TRT tracks.

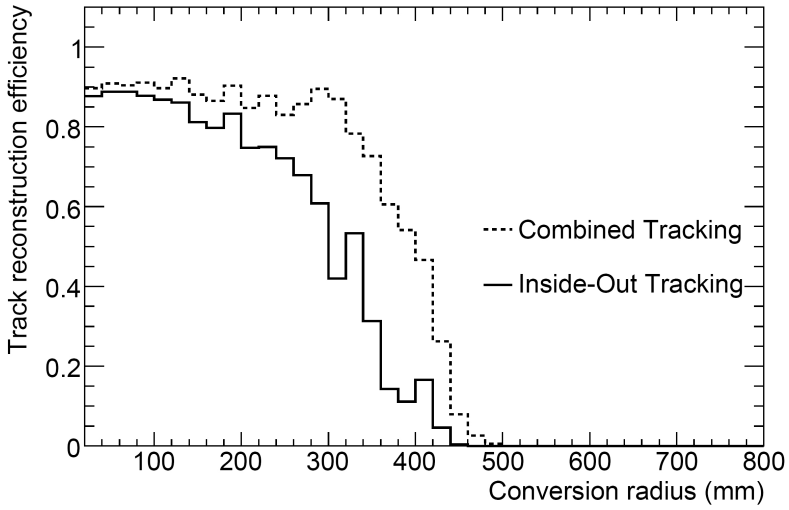

Figure 5.9: *Track reconstruction efficiency for conversions from 20 GeV p_T photons as function of the conversion radius. Solid line tracks reconstructed with Inside-Out tracking and dashed reconstructed with combined (Inside-Out and Outside-In) tracking are shown [17].*

Standalone TRT tracks: TRT segments that have not been assigned to one of the above described tracking methods are then used as basis of one more track collection, the standalone TRT tracks. The TRT segments are transformed into tracks. Therefore based on the segment local parameters the track parameters assigned to the surface of the first straw hits are computed. Then the perigee parameters are calculated, but no track refitting is done. To supress a final ambiguity all tracks which share too many straw hits are rejected. Then the three track types of Inside-out, Outside-In and the standalone TRT tracking are merged to one track collection. The merged track collection is then the input for the next steps, building track pairs and fitting the secondary vertices.

The track reconstruction efficiency for photon conversions of single photons with p_T of 20 GeV as function of the conversion radius R is shown in Fig. 5.10. With the addition of standalone TRT tracks to the merged track collection, the photon conversion track reconstruction up to a conversion radius of 800 mm is possible. The decrease of the Si track reconstruction with higher radii can be explained with the measure of the inner two subdetectors of the Inner Detector (up to $R = 514$ mm). The TRT covers the radius from 554 mm to 1082 mm (cp. Fig. 3.6). Due to the constraint that a sufficient number of straw hits are needed to solve any track ambiguities in the TRT module no conversion track can be reconstructed for larger radii than 800 mm. An increase of the track reconstruction of early conversions with the TRT tracks is due to incurred losses, (e.g. bremsstrahlung) only a part of the track can be reconstruced.

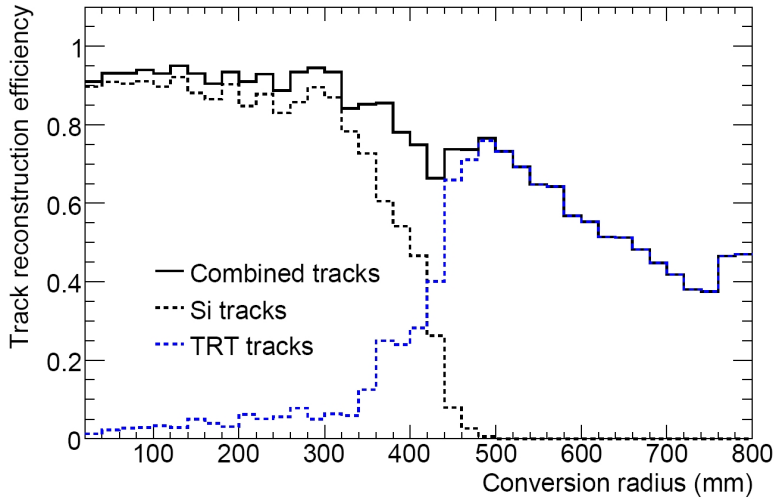

Figure 5.10: *Track reconstruction efficiency for conversions from 20 GeV p_T photons as function of the conversion radius. Si tracks means tracks from combined inside out and outside in tracking. Combined means additionally standalone TRT tracks are used too. [17].*

Building Track Pairs

Using the reconstructed track collection described above positive/negative track pairs are built. Only a fraction of these track pairs come from true photon conversions. To reduce the amount of CPU time by identifying too many wrong track pairs, a few quality cuts are already made during the reconstruction. The cut flow tables of all cut variables are listed in the tables A.1,A.2 and A.3 in the appendix therefore the abbreviations are quoted in brackets.

Track Selection: Cuts on the perigee (*Impact d*0), the longitudinal track impact parameter (*Impact z*0) and on the minimum transverse momentum are applied. Cutting on the ratio of high threshold TRT hits over the total number of TRT hits (*TR ratio*) preselects electron like tracks. All the mentioned cuts have been optimised for $H \rightarrow \gamma\gamma$ events with $m_H = 120\,\text{GeV}$.

Track Pair Selection: The track pairs are of three different types:

1. Pairs based on tracks with Si hits
2. Pairs based on one Si track and one standalone TRT track
3. Pairs based on two standalone TRT tracks

To reduce combinatorial background a series of cuts are applied on these track pairs independent of the track pair types decribed above. Based on the non existent photon mass the difference in the polar angles (*Polar Angle*) of the two daughter tracks should be small. The distance of the first hits (*FirstHitsdR*) of the tracks should be reasonable close. As well the distance of minimum approach (*MinDistance*) between the two tracks is checked. If the track pairs fail these selection they are rejected.

To enhance the vertex fitting a reasonable initial estimate of the vertex position is needed. Therefore the track-helix projection on the $R-\phi$ plane can be derived with the track perigee parameters. In case of a homogenous magnetic field, as it is the case in ATLAS, the track-helix projection is circular. Then the vertex position estimation is either the point of the intersection of the two circles or if they do not intersect the point of minimum approach of the circles. If the circles do not intersect and the minimum approach is not under a specific cut value, the track pair is rejected. If the two circles have more than one intersection point the distance of the two intersection points along the z-axis (dXY) is calculated and above a cut-off the track pairs are discarded. The arc lenght of the $R-\phi$ plane projection ($MinArcLength$, $MaxArcLength$), the distance from the track origin and the actual point of intersection ($VertexR$) are additional cut-off criteria. If after all these criteria there are still two intersection points, the one with the smallest distance transversed along the z-axis (dz) is selected. By applying the above discribed cuts a combinatorial background rejection of at least two orders of magnitude can be achieved (cp. cut flow table Tab. A.2). The initially estimated vertex falls within a few millimeters of the actual conversion vertex. All deviations are dominated by inaccuracies of standalone TRT tracks.

Vertex fitting

For the track perigee which is assigned during the track reconstruction, the track is assumed to come from the primery vertex. For conversions which can take place in more outer regions this is a rather poor assignment. Therefore with the help of the initial estimated secondary vertex the tracks are refitted. The new track perigee is calculated as an extrapolation from the first hit of each track in the pair to the estimated secondary vertex. Therefore the material encountered on the trajectory has to be taken into account. Then the newly computed track perigee parameters are used to calculate the vertex fit. The vertex fit is based on a fast-Kalman filtering method and uses the full 3-D informations of the refitted conversion tracks. After iterations of the vertex fitting the new vertex position with an according error matrix and a χ^2 value of the vertex fit are calculated.

A vertex candidate is accepted, if some post selection cuts on the χ^2 of the vertex fit ($Fit\ \chi^2$), the reconstructed invariant photon mass ($InvariantMass$) and the reconstructed photon transverse momentum ($Photon\ p_T$) are passed. The track pair selection and the secondary vertex fitting with its cuts lead to a combinatorial background rejection by almost three orders of magnitude.

The resolution of the vertex position for each Cartesian coordinate is shown in Fig. 5.11. The values plotted in these histograms are calculated as the differences of the values of the vetrex position of the reconstruced and the truthmatched coordinates divided by the truth coordinate.

Figure 5.12 shows the reconstruction efficiency for single photons with p_T of 20 GeV as a function of the conversion radius. The efficiency significantly drops at conversion radii larger the $R > 400$ mm for both track pairs and conversions after secondary vertex reconstruction. This is due to the inefficiency of reconstructing both tracks in the track pair. To enhance the conversion reconstruction due to such losses so-called single track conversions are defined.

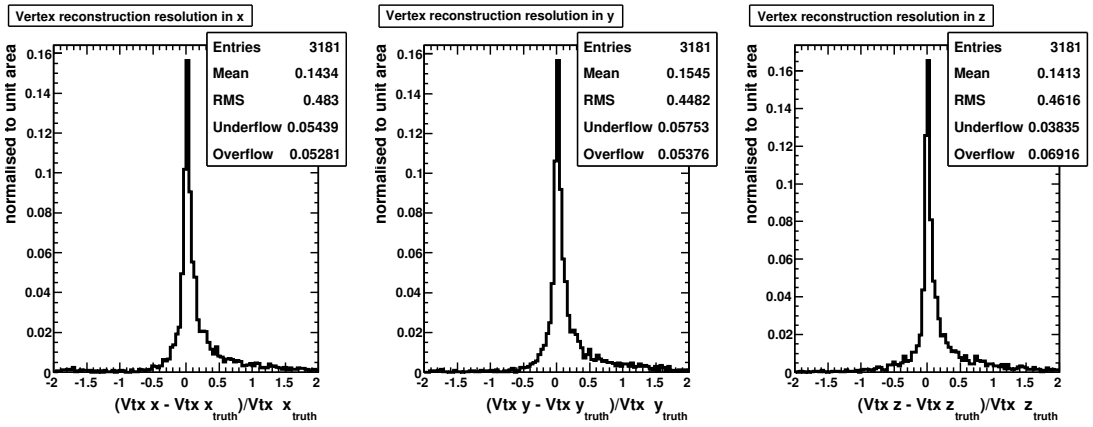

Figure 5.11: *The resolutions of reconstructed conversion vertices are shown in the x, y and z direction from left to right. Therefore in each direction the difference of the reconstructed and the truth photon conversion vertex position is calculated and divided by the value of the truth position.*

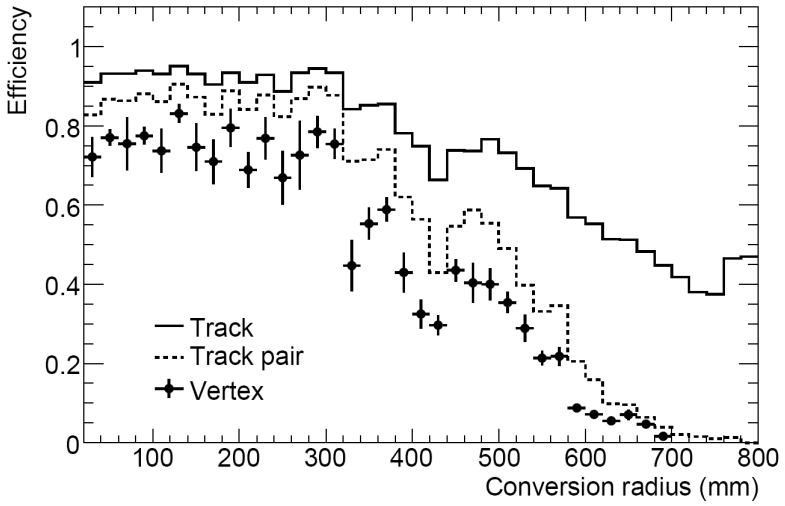

Figure 5.12: *Track and vertex reconstruction efficiency for conversions from 20 GeV p_T photons as function of the conversion radius.* ***Track*** *is the efficiency of the track reconstruction alone,* ***Track pair*** *is the efficiency after the track pair building and* ***Vertex*** *is the vertex reconstruction efficiency [17].*

Single Track Conversions

Several reasons may lead to a lack of reconstructing both tracks of a conversion. For instance, asymmetrically decaying photons where only one track has a high enough transverse momentum to be reconstructed, or conversions which take place so late in the Inner Detector that the tracks do not traverse a long enough distance that both tracks can be distinguished, are rejected by above described conversion selection criteria.

After the vertex fitting the photon conversion candidate pairs which have passed all cuts are marked as "assigned". The candidate pairs which have not been assigned are

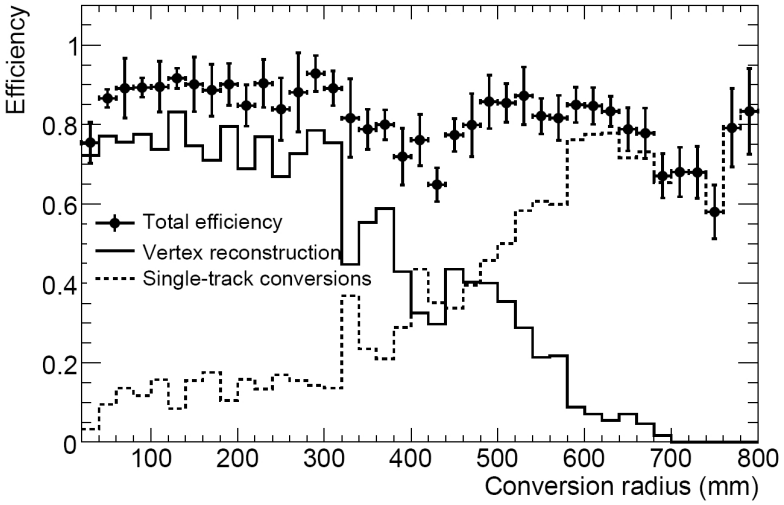

Figure 5.13: *Single track conversion and vertex reconstruction efficiency for conversions from 20 GeV p_T photons as a function of the conversion radius. After adding the single track conversions to the efficiency distribution an enhancement especially for larger conversion radii can be observed [17].*

examined once more. They are selected as single track conversions if their first hits lie beyond the pixel vertexing layer and they are determined as electron like by their ratio of high threshold TRT hits over the total number of TRT hits.

The position of the secondary vertex is then the first detector layer where the track has been measured. Thus the inaccuracy for single track conversion vertices depends on the thickness of the detector layer they first have been measured. Due to the higher straw density the discrepancy of the vertex position is much smaller for TRT tracks than for Si tracks.

The single track conversions are stored, as the conversions passing the secondary vertexing fits, in the *VxCandidate* collection. The only difference is that single track conversions have only one associated track. The reconstruction efficiency for *VxCandidates* with and without single track conversions is shown for single photons with p_T of 20 GeV in Fig. 5.13. Single track conversions become more and more dominant for larger conversion radii. As described above this is due to a less good track pair reconstruction for larger radii.

5.2.2 Photon Conversion Identification

The above described *VxCandidate* reconstruction is done by the Inner Detector reconstruction algorithm. At this stage no calorimeter information is available. To decide if a *VxCandidate* really comes from a photon conversion the reconstructed conversions (*VxCandidates*) have to be matched with electromagnetic clusters.

Therefore for each electromagnetic cluster loosely matching tracks are searched. The tracks have to lie in a window with an angluar distance of $\Delta\eta < 0.05$ and $\Delta\phi < 0.10$ to the center of the cluster and have a E/p value smaller than 10. If the track fullfil these criteria the track is extrapolated to the surface of the electromagnetic calorimeter. If the extrapolated track lies within $\Delta\eta < 0.05$ and $\Delta\phi < 0.05$ to the cluster center, the cluster

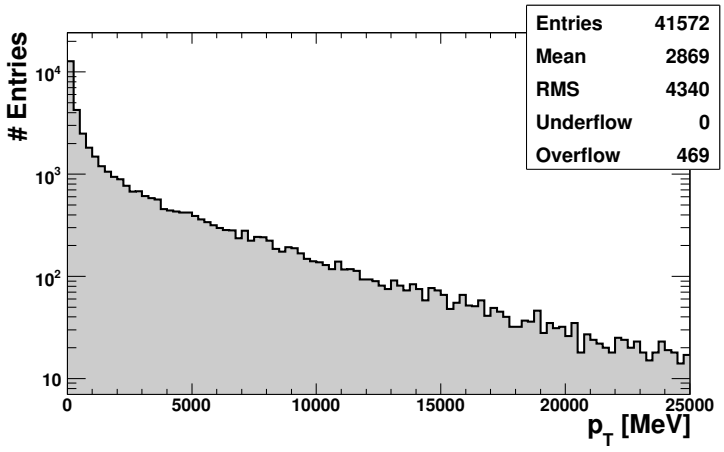

Figure 5.14: *p_T distribution of truth photons in τ decay cone of $Z \rightarrow \tau\tau$ events. Most photons have a much lower transverse momentum as the photons from $H \rightarrow \gamma\gamma$ for which the algorithm has been optimised. Thus the lower efficiency for low p_T photon conversions shown in Fig. 5.15 is the main limitation for the reconstruction of the conversions coming from τ decay.*

is labeled as matched to a track and it is flagged as electron-like. If the cluster cannot be matched to a track it is flagged as photon-like [17].

If a cluster is flagged as electron-like it has to be re-examined. If a *VxCandidate* can be matched as "best-match" to the cluster, the tracks of the *VxCandidates* are assigned as tracks from this photon conversion. As a last cut it has to be checked if both tracks have either zero or both more than zero B-layer hits. This condition has to be done to avoid flagging electrons which undergo bremsstrahlung and producing afterwards a photon conversion as wrong paired photon conversion.

5.2.3 Shortcomings of the Photon Conversion Reconstruction Algorithm in the τ Environment

Several problems occur if the default photon conversion reconstruction algorithm above described is used within the τ environment. Firstly the photons produced in τ decays basically have a transverse momentum lower than 2 GeV. The reconstruction efficiency for low p_T photon conversions decreases in comparison to the above used examples with a photon transverse momentum of 20 GeV (cp. Fig. 5.15). The decreasing of the reconstruction efficiency from $p_T = 5\,\text{GeV}$ to $p_T = 2\,\text{GeV}$ is due to losses of bremsstrahlung. The photon conversion efficiency is limited by the minimum transverse momentum of $p_T > 500\,\text{MeV}$ for the track reconstruction. To guarantee a track quality the cut of minumim track p_T cannot be loosened.

The comparison of this decrease in the photon conversion efficiency for low p_T single photons to the simulated transverse momentum distribution of photons from τ decays shows the importance of an improvement for low p_T photon conversions reconstruction (cp. Fig. 5.14).

Additionally the very dense τ environment, meaning the fact that the seeds of τ leptons are jets which may contain of several tracks, make a cluster matching problematical.

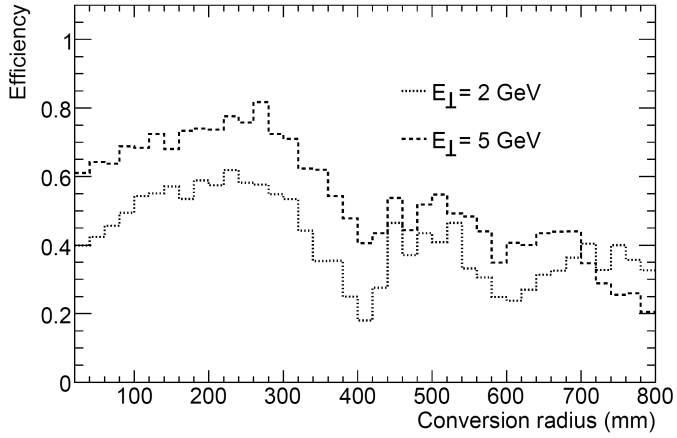

Figure 5.15: *Vertex reconstruction efficiency for low p_T conversions from 2 GeV (5 GeV) p_T photons as function of the conversion radius. Even in a much "cleaner" than the very busy τ environment the efficiency of low p_T conversions drops [17].*

For the very clean photon conversions of the $H \rightarrow \gamma\gamma$ processes the method works very well. But if the conversion comes from a τ decay the tracks of the photon conversion lie within the τ decay cone described in Sec. 5.1. The electromagnetic clusters produced by charged pions cannot be easily distinguished from clusters of the electron or the positron. An additional p_T cut at $p_T > 2\,\text{GeV}$ in the photon conversion identification (sec. 5.2.2) totally kills the identified photon conversions for this study. Figure 5.16 shows a comparison of an older release (*ATHENA* rel. 13.0.40) without a p_T cut at 2 GeV(left) and the a newer release which is the basis for this thesis (rel. 14.2.0 right). This problem is discussed in more detail in Chap. 7.2.

5.3 Conclusion

After explaining the method of operation of the τ and the photon conversion reconstruction algorithm it has to be mentioned, that both the standard definition of an identified τ and the standard definition of an identified photon conversion cannot be used to reconstruct photon conversions in τ decays. The first problem is postponed: The result of this study will not be the τ efficiency after the LLH identification. Instead this study will concentrate on the achieved track multiplicity after vetoing the conversion tracks, as a first result.

A solution to the second problem, not using the standard definition of a photon conversion, needs a lot of development. Thus another method to clean the *VxCandidates* from combinatorics and wrong particles has to be implemented. Therefore two independent methods have developed in context of this thesis.

In order to use the result from the τ specific photon conversion identfication algorithm to improve the τ track multiplicity, a link between the reconstructed particles of both algorithms is needed. For quick and easy changes during the developement it would be very helpful to have all needed variables in ROOT readable ntuples. On that level no connections between τ and conversion variables are available. Special tools providing such informations for ROOT ntuples has to be implemented.

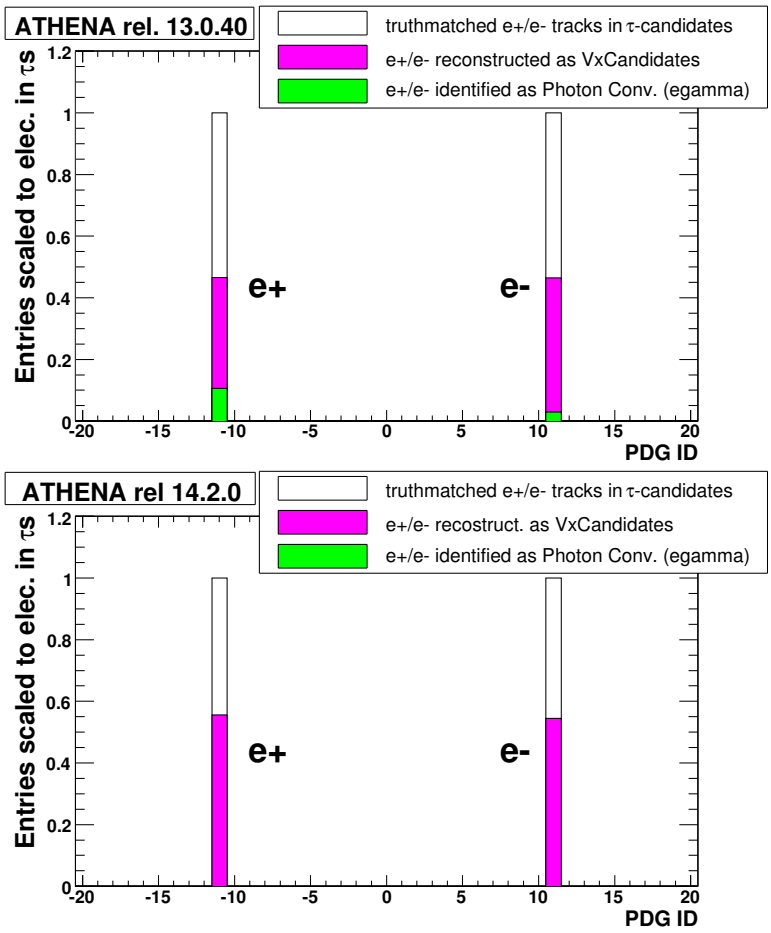

Figure 5.16: ***Upper plot:*** *histogram of the truthmached track ID of tracks reconstructed as track of a τ candidate (white), reconstruted as track of a τ candidate and as VxCandidate (magenta) and reconstructed as track of a τ candidate and identified by standard egamma tool as "good" photon conversion. Used sample $Z \rightarrow \tau\tau$ reconstructed with ATHENA rel. 13.0.40.* ***Lower plot:*** *Due to an additional p_T cut ($p_T > 2$ GeV) in ATHENA rel. 14.2.0 no identified photon conversions (green) can be found in a τ candidate anymore.*

A short summary about the newly implemented tools and other improvements for the data analysis are described in Chap. 6.2.3.

Chapter 6

Development of Specific Software Tools

The ATLAS Event Data Model (EDM) described in Chap. 4.4 consists of several steps with different data formats with a specific amount of stored event informations. To provide these data formats miscellaneous tools have been implemented. This chapter concentrates on the tools needed to produce D^3PDs.

It was decided to use the D^3PD data format for the study of this thesis to have an easy and fast access to the simulated and reconstruced data. Due to the fact that the tools providing D^3PDs are in the developing phase right now, the tools to store variables of photon conversions in the D^3PDs have not been implemented at the begining of this thesis. Therefore the tools explained in this chapter had to be enhanced. First, a very brief summary of the already existing tools is given. The most important tools developed for, and in the context of this thesis are explained in more detail.

6.1 Introduction to Existing Tools for Data Preparation

Within the ATHENA framework smaller frameworks to provide ROOT readable ntuples (D^3PDs) already exist. The specific tools providing the informations needed for this thesis only have to be added to these frameworks. Before describing the newly developed and updated tools the frameworks have to be introduced very briefly.

TauDPDMaker

To produce the tailor made DPDs several physics groups developed specific *DPDMakers*. The one developed by the *Tau performance group* is named *TauDPDMaker* [44]. This does not mean that only τ specific informations are stored in such DPDs. It is developed as a general tool with the attention to have all informations to make a complete analysis with τ leptons, which means that e.g. electrons, muons and jets have to be stored in the *TauDPDMaker* output, too.

The input of the *TauPDPMaker* can be data directly from a MC generator, RAW data, ESDs, AODs or D^1PD from *TauDPDMaker*. The abbreviations of the several data formats are explained in Sec. 4.4. The *TauDPDMaker* produces primary, secondary and/or tertiary DPDs. This is achieved technically in three steps: pre-processing, D^1PD-

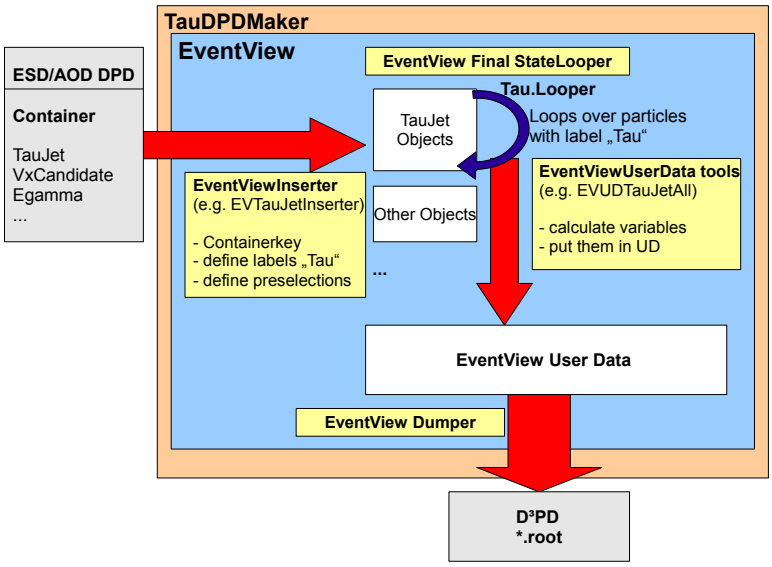

Figure 6.1: *Schematical Overview of the EventView usage of TauDPDMaker producing D^3PD. Mainly these four EventView tools are used: the EVInserter, EVToolLooper, EVUserData and the EVDumper.*

making and D^3PD-making. The pre-processing is optional and the D^1PD and D^3PD making is totally independent. Thus they can be produced simultaneously or separately.

The TauDPDMaker is a package within the ATLAS ATHENA framework based on the programming languages *Python* and C++. To produce D^1PD and D^2PD the package uses other tools within the ATHENA framework to skim, slim and thin the DPDs (cp. Chap. 4.4). To produce the D^3PD which are simple ROOT ntuples the TauDPDMaker calls the analysis framework *EventView* (see next subsection).

As basic τ specific D^3PD the *TauPDPMaker* provides a D^3P named *Control Sample D^3PD*. It is a multipurpose ntuple providing all necessary informations needed for a complex τ analysis on dijet, W or Z^0 events. To study photon conversions in τ decays very specific D^3PDs have to be produced.

EventView

EventView is an object-oriented physics analysis framework [38]. It is at the end of the computing workflow shown in Fig. 4.3 and amongst others it is able to produce ROOT ntuples. These ntuples can be analysed or transformed into simple histograms with the analysis framework ROOT [39]. ROOT has been developed in the mid 1990's aimed at solving the data analysis challenges of high-energy physics (HEP). In the meanwhile ROOT has become a standard high energy physics framework and provides a large selection of HEP specific utilities such as histogram building and fitting. But it cannot handle the high data flow provided by the LHC. Therefore the EventView framework preselects the data and transformes it into a ROOT readable format. Then the utilities of ROOT can be used to produce all the plots needed to visualise the miscellaneous physical quantities.

EventView consist of a few core classes from which all other classes inherit. To

summarise the most important functionalities applied by TauDPDMaker a very simplified overview is listed here and shown in Fig. 6.1:

EVInserter Tool: The *EventViewInserters* are responsible to provide the particle objects stored in so-called containers by the ATHENA offline reconstruction for further analyses. Therefore the several objects are "inserted" into the an "EventView" and labeled (e.g. as "tau" cp. Fig. 6.1). For this study no further preselections are made and all particle objects from the input collection (e.g. an ESD) are inserted, labeled and then handled as *finale state objects*. In principle preselections, overlap removal and other data quality cuts can be applied with this tool.

EVUserData Tool: Quantities like sum of p_t of jets, the number of τ tracks or more detailed informations like the electromagnetic radius of a τ jet which are calculated during an analysis are computed with the aid of this tool. Therefore the needed values are read out from the particle objects and either stored directly or used to compute more complex variables. After the calculation the variables are stored including their values in the *UserData*.

EVToolLooper: For each "label" a specific object of the class *EventViewToolLooper* loops over all inserted particles and calls other classes like *EVUserData* to be executed upon these objects. Figure 6.1 shows as example how τ specific variables are calculated. A *Tau.Looper* from type *EventViewToolLooper* loops over all as "Tau" labeled objects and calls the *EVUDTauJetAll* tool. More than one tool can be run during a loop and optional truthmatching methods can be switched on or off.

EVDumper Tool: At the end dumper tools are used to output the UserData either directly on screen, as XML file or as ROOT ntuple. The ATHENA tool ATLANTIS [45] which produces event displays, as shown in Fig. A.10 in the appendix, needs such XML files as input format for example. The ROOT ntuples can be read into a standalone ROOT analysis.

Thus no cuts or other changes have been applied in *EventView*. It has been generally used to apply two functions. Firstly to convert the container structure from ESD or AOD level into ROOT readable ntuples. Secondly to calculate additional variables by calling ATHENA tools from the so-called *ToolService* which is very helpful to use already existing code and to avoid writing additional source code.

6.2 Special Tools needed to store Conversions and other useful data for this study in D^{3}PDs

To store all needed informations about photon conversions new *EventViewUserData* classes for truth photon conversions, reconstructed data of the *VxCandidate* collection and identified photon conversions had to be developed. Also the *EventViewUserData* class which provides all τ specific variables had to be updated. Additionally a new tool, able to link together indentical objects used by different algorithms (e.g. same tracks used by τ reconstruction and photon conversion reconstruction) has been developed. Due to the fact that this development was a very important step to get the final results

presented in Chap. 7 the recent developements will be outlined briefly in the following subsections.

6.2.1 The New *EVUDVertex* class

Based on a quite similar tool at a higher level of the ATHENA framework a new *UserData* tool to handle *VxCandidates* has been implemented. To share as much code as possible this tool has been written as general *vertex* tool. Thus the input container can be steered via a *python* file. Not only the conversion specific *VxCandidates*, but also primary vertices or other secondary vertex collections can be handled with this tool.

Due to technically special features of photon conversion objects (vertex objects in general) no sufficient inserter tool has been available. Therefore the EVUDVertex tool directly inserts the objects without using an inserter and provides the needed *UserData*. This bypass has to be done due to the inheritance structure of the inserter classes.

Specific conversion routines can be switched on or off via additional options. Here a brief overview of several functions of *EVUDVertex* class which can be separately used should be given:

Variables of VxCandidate: The data available of the tracks at each vertex like track p_T, η and so on are stored in the *UserData*. Also vertex informations like x, y and z coordinates of the vertex position or quality data like the χ^2 value of the vertex fit are stored. The most important variables provided by the *EVUDVertex* class are listed in Tab. A.4 in the appendix.

Conversion Track Truthmatching: Another very difficult task for photon conversions is track truthmatching. Composited particles like τ leptons can be truthmatched by a ΔR cone matching method. ΔR cone matching means that the distance between truth and reconstructed object are calculated with Eq. 3.3. If the calculated value is smaller than a specific matching value (e.g. < 0.1) the reconstructed particle is matched to the truth particle and vice versa. This method works fine for composited objects. Due the fact that tracks often lie in narrow cones, e.g. in τ decays or QCD jets, this method is not feasible for track truthmatching. Especially the tracks of photon conversions of highly boosted photons are very nearby and the geometrical variables show relative differences of 10^{-5}.

It is needed a much more sensitve truthmatching method for tracks. Therefore the so-called SDOs (cp. Sec. 4.4) are provided during the digitalisation of the simulated data. This means that for each simulted track additionally to the truth 4-vectors the truth space points are stored. These space points are the exact 3-dimensional positions, where the truth particles pass the detector. By comparing the truth with the reconstructed space points a very exact link between the truth and the reconstructed objects are available. Out of the space points both the truth and the reconstructed tracks are built. Accordingly for simulated data so-called *ElementLinks* between truth tracks and reconstructed tracks are provided. This track truthmatching method is shown in Fig. 6.2. The rows from top to bottom are the several phases of building both truth and reconstructed tracks. The *TrkTrack* (*TrackParticles*) are the reconstructed tracks at ESD (AOD) level, the *TrackTruth* and *ParticleTruth* are the equivalent steps of truth tracks. Via the *ElementLink* a

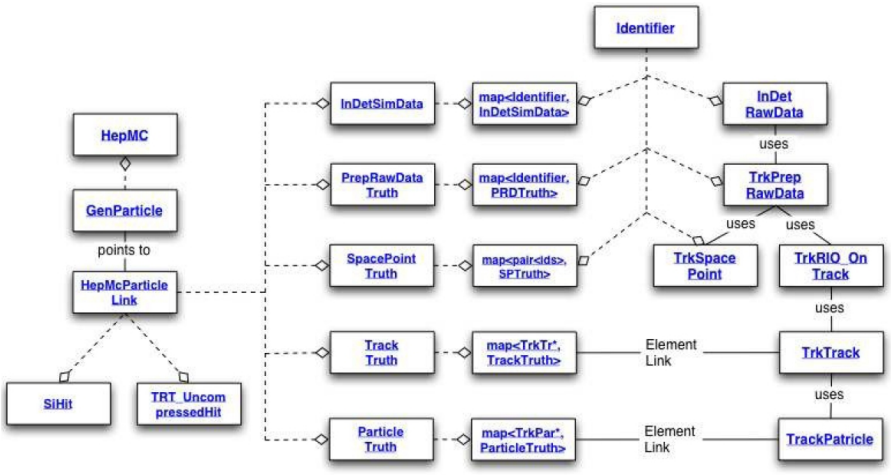

Figure 6.2: *Track truthmatching method via ElementLinks. During digitalisation of simulated data additionally to the Raw Data Objects (right column) the Simulation Data Objects (SDO) (second column from left) are provided too. The reconstructed and truth tracks are connected via so-called Element Links. From upper to lower row several stages of truth and reconstructed tracks are shown.*

track based truthmatch is possible. The *EVUDVertex* tool uses these *ElementLinks* for track truthmatching and stores the linked truth informations in *UserData*.

Special integrated Conversion Index Tool: As described in Sec. 5.2 the tracks used to build photon conversions are refitted with the additional constraints of the secondary vertex. Thus the unrefitted tracks have to be used to compare with tracks used to build τ candidates. Otherwise no agreement would be found. Easier as to compare the conversion tracks and the τ tracks is simply to store the index of the track in the track container, where all tracks of one event are stored. Due to the very complex structure of *VxCandidates* and the fact that the unfitted tracks are used for the truthmatching too, the track index for tracks from photon conversions is provided directly by the *EVUDVertex* tool. For other particle objects a more general index tool has been implemented (see Subsec. 6.2.3).

Track Extrapolation to the Electromagnetic Calorimeter: An additional function is able to call a tool from ATHENA framework *ToolService* which is able to extrapolate the track to the surface of the electromagnetic calorimeter with respect to the magnetic field. There the best matching cluster is searched and all associated informations like the energy deposition in the different electromagnetic calorimeter layers are stored in the *UserData*.

6.2.2 The New Classes *EVUDTruthegammaConv* and *EVUDegammaConv*

EVUDTruthegammaConv has been developed to store not only the truth informations of truthmatched *VxCandidate* objects but all available truth informations about photon conversions. These truth informations are necessary to cross check if the *EVUDVertex* tool works properly.

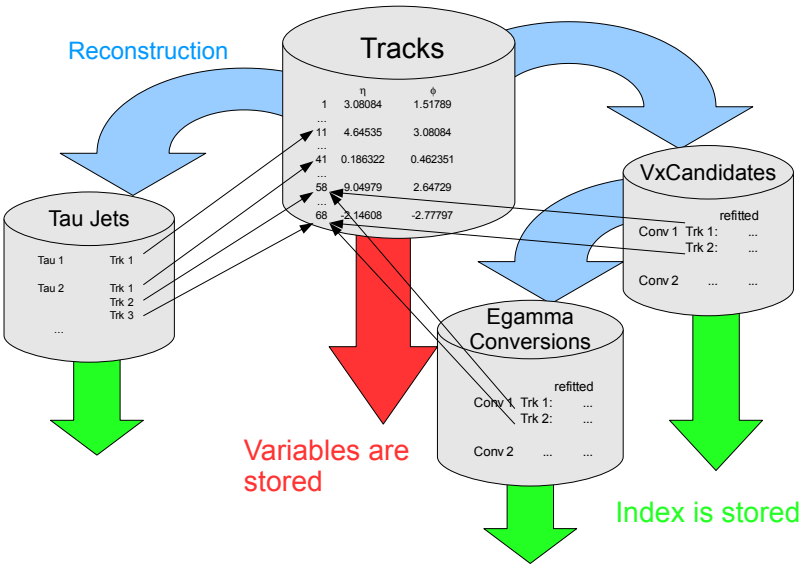

Figure 6.3: *Overview of the method of operation of the IndexTool. Properties of tracks like η, ϕ and p_T are stored once (red). Combined objects like τs or photon conversions consisting of tracks only store the specific track indices (green). This reduces the storage space of the ntuple and makes a comparison of used tracks of several objects much easier.*

EVUDegammaConv is another tool for cross checking *EVUDVertex* tool results. Conversions stored by the *EVUDegammaConv* tool in the *UserData* are the identified photon conversions (cp. Chap. 5.2.2).

6.2.3 The New *EVUDIndex* class

The *IndexTool* is, as already mentioned, developed to store the index of a track which has been used to built a particle object. Therefore a base class has been written from which more specific *IndexTools* like *TauIndexTool* inherit.

By using this tool several improvements can be achieved. As shown in Fig. 6.3 instead of storing the whole track informations twice, once for the tracks and once for a more specific object like τs, the properties of the tracks are stored once and other objects like τs only store one variable, an integer containing the index of the specific track. This technique is more or less a reproduction of the pointer structure used in the higher levels of the data preparation. Thereby the implementation takes care of using the *IndexTool* only in case of having stored all applied tracks in the *UserData* before.

More important for this study is of course the possibility to use the stored track indices in the ntuple to figure out which tracks have been used to built photon conversions and τ candidates. The dependecies shown in Fig. 6.3 as black arrows can be evaluated easily by comparing the value of two integer variables.

Using these newly developed and updated tools all needed informations for an extensive study of photon conversions can be stored in D^{3}PDs. These ROOT readable D^{3}PDs are of size about $45\,^{kB}/_{\text{evt}}$.

Chapter 7

Improvements of Photon Conversion Reconstruction in τ Environment

This chaper contains all results which have been developed during this thesis. As first step the content of τ jets have to be discussed. Thereby the important role of photon conversions in the τ track multiplicity is explained. In the second section photon conversions reconstructed by the default photon conversion reconstruction tool as described in Chap. 5.2 are tested for using in the τ environment. Due to the optimisation of this tool for other purposes the existing photon conversion identification is quite inefficient and has to replaced by a more compatible with τ leptons. As a first ansatz to replace the identification some studies with data from the Transition Radiation Tracker (TRT) are described in the beginning of Sec. 7.3. Then the results of using a more general tool based on the informations from the TRT are presented. Comparable with the standard conversion identification a method extrapolating the reconstructed track to the surface of the electromagnetic calorimeter has been worked out and the results are shown in Sec. 7.4. Due to the implementation of the conversion reconstruction much combinatorial background has been created during this phase. A special method has been tested to remove this background. This method and its results are described in in Sec. 7.5. After a comparison of the different innovations in conversion purifying the best method is choosen to veto the conversion tracks in the τ candidate track multiplicity. These results are presented in the last section.

7.1 Content of the Reconstructed τ Cone

Before explaining the methods of vetoing photon conversion tracks in τ candidates the identity of the tracks which have been picked up from τ reconstruction algorithm *TauRec* have to be analysed in more detail. In Fig. 7.1 an extract of the reconstructed tracks identified via track truthmatching are shown, in red with applied default Track Selection Criteria (TSC) explained in Sec. 5.1 and in black all tracks with $p_T > 500\,\mathrm{MeV}$ which have been found in a cone of $\Delta\,\mathrm{R} < 0.3$ around the direction of τ candidates. All particles found in the τ cone are listed in Tab. A.5 in appendix.

Due to the fact that the *TauRec* algorithm reconstructs hadronic τ lepton decays most of the electrons within the cone around a true τ have to originate from photon conversions. If comparing the red and the black histogram it is clearly visible that the

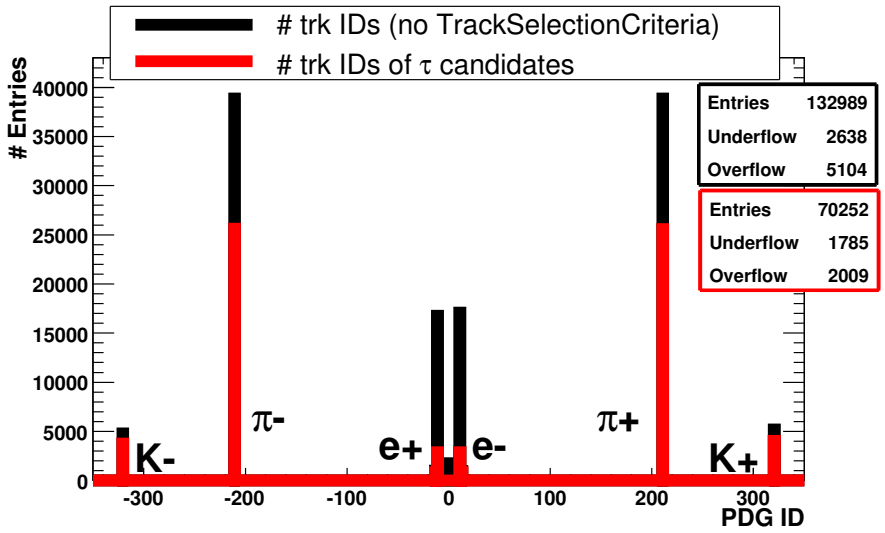

Figure 7.1: *Truthmatched tracks in τ candidates with and without applied Track Selection Criteria. In both shown histograms the entries are the number of reconstructed tracks of 10 k $Z \rightarrow \tau\tau$ events. The number of tracks passing the Track Selection Criteria is displayed in red whereas the black distribution shows the τ candidates without applying any TSC.*

TSC already supress a lot of electrons and positrons. Thus the cuts are very powerful to supress photon conversions.

For a more detail look into the origin of the electrons and positrons in τ candidates Fig. 7.2 shows in two histograms the main sources of electrons. The most important aspects are firstly that the fraction of electrons from conversions increase if the TSC are not applied which is also an evidence for supressing conversions by the TSC. Secondly the TSC are important to guarantee a minimal track quality. A large fraction of "other parents" in the right histogram are not matched tracks what is an indicator for corrupted tracks. Roughly two-thirds of electon tracks in τ candidates after applying TSC still come from photon conversions. This means that additional charged tracks affect the τ track multiplicity. If the additional tracks can be vetoed explicitly more τ candidates will pass the likelihood selection due to the track multiplicity variable of the likelihood discrimination.

7.2 Reconstruction of Conversions From τ Candidates with default Egamma Tools

After analysing the content of τ candidates it has to be studied how these conversion tracks can be identified during the reconstruction of real data. Developed with quite different purposes as described in Sec. 5.2 a photon conversion reconstruction tool already exists. In the Fig. A.10 and Fig. A.11 in the appendix two event displays are shown. The number of tracks in a $Z \rightarrow \tau\tau$ event is much higher compared to $H \rightarrow \gamma\gamma$ events. Apart from the different p_T regions of the photons in both processes this is one reason for the insufficient identification of the default conversion reconstruction tool in the dense

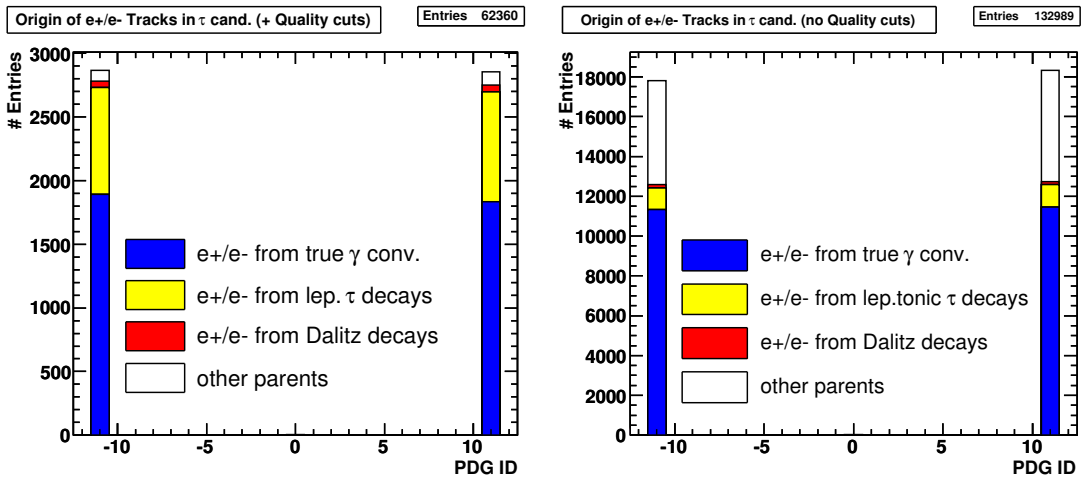

Figure 7.2: *Origin of electron tracks in τ candidates.* ***Left****: The TauRec default Track Selection Criteria have been applied.* ***Right:*** *No TSC have been applied, the only criteria is the tracks have to lie within a cone of $\Delta < 0.3$ around the τ candidate and has a larger transverse momentum than p_T >500 MeV.*

τ environment. Using that tool to reconstruct and identify the tracks which have been picked up as τ candidate tracks leads to the plots shown in Fig. 5.16.

The prototypes of the tools described in Chap. 6 have been developed in the ATHENA release 13.0.40. The corresponding plot is the left one (cp. Fig. 5.16). A slight fraction of reconstructed and identified photon conversions could be found in the τ candidates. In comparison to the fraction of the *VxCandidate* tracks found within the τ cone the identification reduces the reconstructed tracks by a factor of seven. Thus even in release 13.0.40, where a small identified fraction of conversions could be found, the identification algorithm has to be replaced by a more τ specific one. Even though the reconstruction phase building the *VxCandidates* is much more efficient in comparison to the identification it reconstructs only about 46.5 % of the electrons within the τ candidates. Compared with the fraction of truthmatched photon conversions which is 65% there is a lot of space for improvements. With this results it was decided to concentrat on the *VxCandidates*, try to find a purifying method and use them to veto the tracks.

Due to many improvements in ATHENA release 14.x.x (e.g. including single track conversions) the conversion reconstruction could be upgraded. Thus the fraction of reconstructed *VxCandidates* overlapping with the τ candidates could be improved to a fraction of 55 %. During the same upgrade cuts in the identification phase of converions have been added (e.g. a cut on $p_T > 2\,\mathrm{GeV}$) which supresses many low p_T conversions from τ decays. Nevertheless all plots excepting Fig. 5.16 the left plot are produced with ATHENA release 14.2.0.

However the overlapping *VxCandidates* cannot be used directly to veto tracks in τ candidates. During the reconstruction it has not been checked if all these tracks are only electrons. The content of the *VxCandidate* collection without any electron identification is listed in Tab. 7.1.

If this collection is used to veto τ tracks without any electron identification many charged pions would be rejected, too. Thus it will not bring any improvement. So the

Table 7.1: *Track identity of VxCandidate tracks. Total number and fraction of electrons, pions and kaon tracks with statistic errors are listed.*

Truth ID	Entries	Fraction	Combined
e^-	8654 ± 93.03	30.97 ± 0.33	
e^+	8437 ± 91.85	30.19 ± 0.33	61.16 ± 0.66
π^+	3830 ± 61.89	13.71 ± 0.22	
π^-	4154 ± 64.45	14.87 ± 0.23	28.57 ± 0.45
K^+	509 ± 22.56	1.82 ± 0.08	
K^-	530 ± 23.02	1.90 ± 0.08	3.72 ± 0.16
total:	27944		

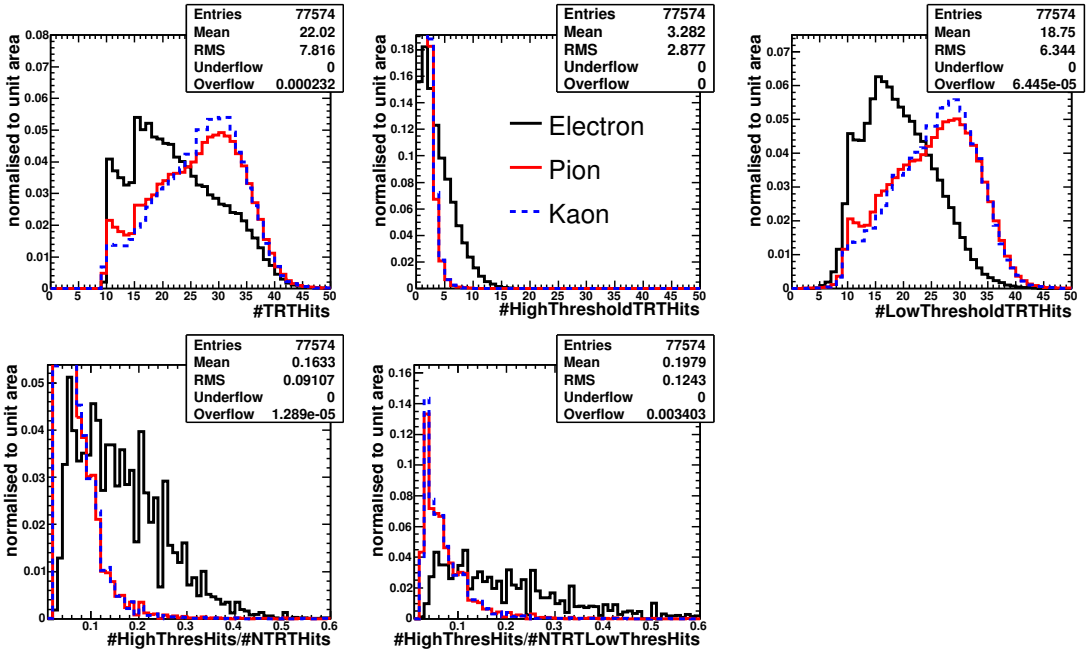

Figure 7.3: *Several TRT variables for electrons pions and kaons. From upper left to lower right: number of total TRT hits, number of TRT hits above high threshold, number of low threshold TRT hits, ratio of high threshold hits over the total number of TRT hits and ratio of high threshold hits over low threshold hits for each track. The two lower plots show a sufficient discrimination.*

VxCandidates first have to be cleaned from the charged pions before using to veto any tracks in the τ candidates.

7.3 Electron Identification with the TRT

As briefly mentioned in Chap. 3.3.1 the Transition Radiation Tracker is able to distinguish electrons and pions due to different TRT signatures. The highly relativistic electrons create transition radiation whereas not relativistic particles like pions induce so-called tracking hits. Two independet thresholds allow to differentiate these signals.

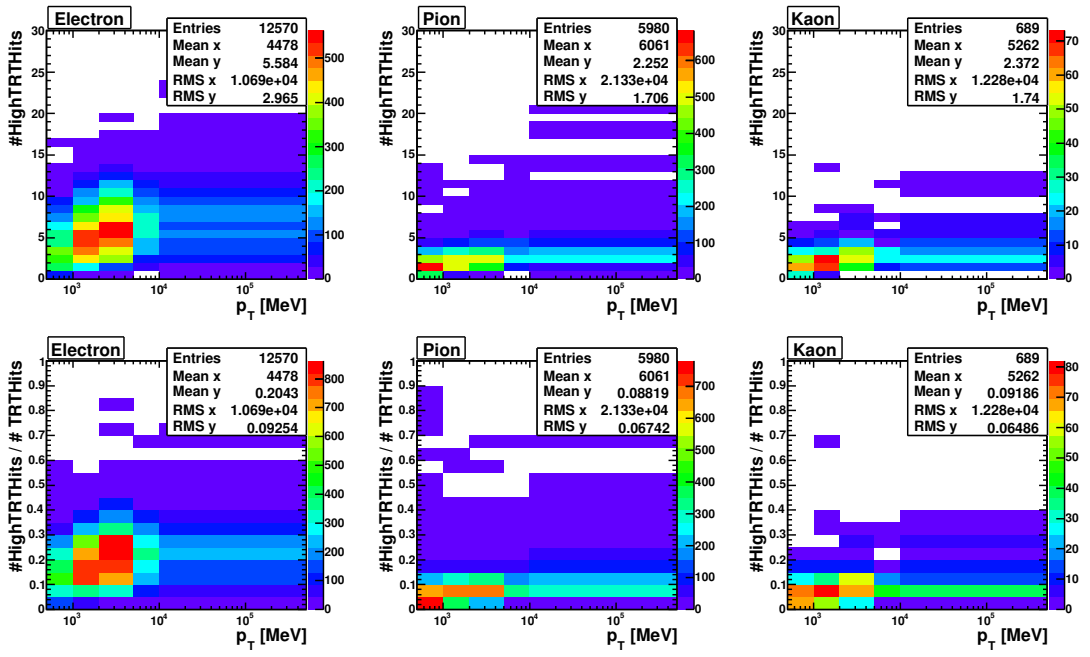

Figure 7.4: *The two mostly sufficent TRT variables plotted for electrons, pions and kaons over the p_T distribution of the tracks from VxCandidates. Both distributions show a clear peak in the electron plots which can be used to differentiate from pions and kaons.*

7.3.1 Discrimination of Electrons and Pions via Several TRT Observabes

Figure 7.3 shows several distributions for eletrons, pions and kaons. In the upper row from left to right the total number of TRT hits, the number of TRT hits above the high threshold (transition radiation) and the number of low threshold hits per tracks is exemplified. In the lower row the fraction of the number of high threshold hits over the total number of hits (left) and the fraction of the number of high threshold hits over low threshold hits (right) are plotted. The bins of all histograms are scaled to the total number of entries.

In both histograms in the lower row a clear peak is visible for low ratios of pions and kaons whereas the electrons have a larger tail. So one of these two distributions seems to be appropriate to distinguish electrons from background. To achieve a more detailed analysis of these TRT spectra the dependence of the transverse momentum of the track is studied and shown in the appendix A.4.3 for the same distributions in several p_T regions. Even the 2D plots in Fig. 7.4 point out that the informations from the TRT are quite useful for the pion cleaning. Both plots of electrons show a clear peak around 5 high threshold hits and $p_T = 2\,\text{GeV}$ (upper plot) and ratio of high over all TRT hits around 0.2 and $p_T = 2\,\text{GeV}$ (lower plot). In contrast to that, the peaks of the pions and the kaons in both cases are shifted to a lower number of high threshold hits (lower ratio of high threshold over total number of TRT hits) and lower p_T values.

Instead of directly using the TRT informations, the Inner Detector offline reconstruction algorithm offers a more elaborated method which amongst others is based on these values.

7.3.2 Inner Detector TRT Electron ID

A so-called *TRT ElectronPIDTool* provides an electron probability for each track. This probability is based on Inner Detector track informations only. The tool calculates three independent variables. Each is a probability of a track being an electron:

High Threshold Probability (p^e_{HT}): To calculate the probability a liklehood ratio is used:

$$p^e_{HT} = \frac{\prod_i p^e_{HT,i}}{\prod_i p^e_{HT,i} + \prod_i p^\pi_{HT,i}} \tag{7.1}$$

Where i loops over the hits of a track. This variable depends on the Lorenz γ factor (i.e. energy), the distance into the TRT (i.e. hit radius) and the TRT module side walls (i.e. ϕ_{track}). The difference of the high threshold hits of electrons, pions and kaons are shown Fig. 7.4 upper row.

Bremsstrahlung Probability: If an electron makes bremsstrahlung a photon is emitted. Due to such energy losses the curvature of the electron is changed. By defining a relative curvarture a discrimination between electron, muons and pions is possible.

Time over Threshold Probability: The discrimination value is based on the differences of the time over threshold variables of electrons and pions. The probability is calculated via a likelihood discrimination comparable to the high threshold probability.

The tool is still under construction and therefor only the first two probabilities are combined to a PID value. The combined electron probability (eProb) is calulated as follows:

$$ePrab = \frac{p^e_{HT} \cdot p^e_{brems}}{p^e_{HT} \cdot p^e_{brems} + (1 - p^e_{HT}) \cdot (1 - p^e_{brems})} \tag{7.2}$$

The electron probability is strongly dominated by the high threshold hits due to the good discrimination power of that variable. Using this tool to provide the *ePrab* value of all *VxCandidate* tracks the plot shown in Fig. 7.5 can be produced. The *ePrab* values of pions and kaons peak at low values smaller than $ePrab < 0.25$ and at $ePrab = 0.4$ whereas the electrons show a high peak for probabilities larger than $ePrab > 0.9$ and the same peak at $ePrab = 0.4$. The peak at $ePrab = 0.4$ of all distributions can be explained by the definition of a default value. If a tracks does not deposite any high threshold hits in the TRT the probability for that track being an electron is set to $ePrab = 0.4$.

Using this tool to distinguish the electrons from pions and kaons in the *VxCandidates* a good discrimination can be achieved. Cleaning the *VxCandidates* using preselection cuts, described later on, and a cut on the electron probability at $ePrab > 0.9$ leads to an electron purity of 94 %. Due to the high decrease of the efficiency of the reconstruced conversion after such a cleaning another method has been tested. Therefore the electromagnetic cluster informations have been taken into account.

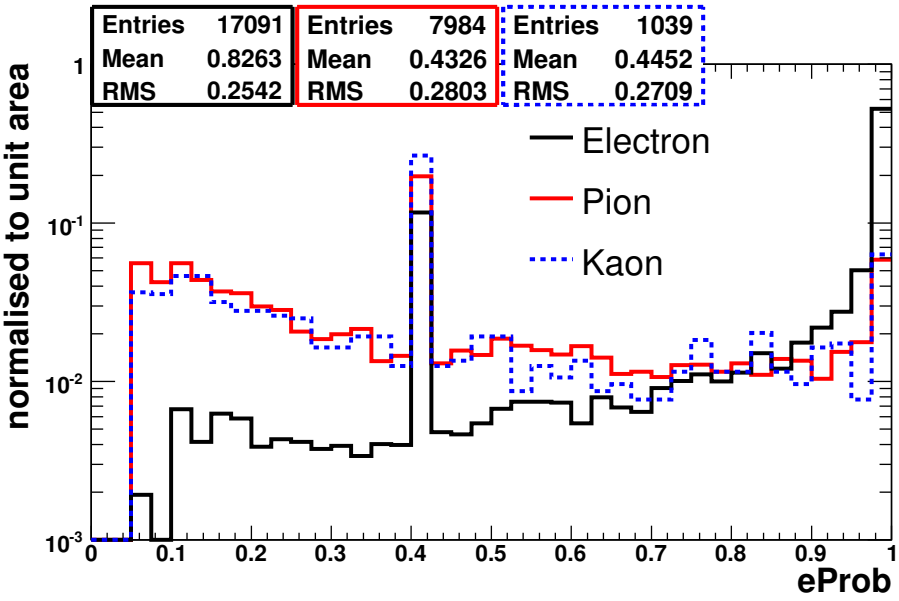

Figure 7.5: *Electron probability derived by the TRT electron PID tool*

7.4 Identification Using Electromagnetic Calorimeter Informations

Using the electromagnetic calorimeter observables means an extrapolation of the conversion tracks to the surface of the electromagnetic calorimeter (ECal) is necessary. This is done during the production of D^{3}PDs by the *EVUDVertex* tool described in Chap. 6.2.1. Then the cluster found nearby the extrapolated track is associated to the track. The variables described in the following section are stored to extract a discrimination variable.

7.4.1 Additional Variables from Electromagnetic Calorimeter

All histograms in Fig. 7.6 show the distribution of tracks matched to truth electrons which is the signal distribution and of tracks which have not been matched to electrons, the background. The first variable shows if the track is only found by the TRT module or not. This is important for geometrical reasons because the TRT barrel does not measure precise η angles. In case of such tracks one of four default η values are stored (cp. to the four visible spikes in Fig. A.9). The second and third histograms show the deviation in ϕ and η of the direction of the nearest calorimeter cluster and the extrapolated track. The following histograms are several combinations of the momentum P or the transverse momentum p_T of the track determined by the tracker and the energy depositions in the first, second and third layer of the ECal (cp. structure of the ECal in fig. 3.8). The geometrical setup of the three layers is designed to differentiate electron like showers from hadronic showers. Thus electrons should deposit most of their energy in the first and second layers whereas for hadronic particles like pions often a leakage into the hadronic calorimeter via the third ECal layer should be observable.

Analysing these histograms no variable for a direct discrimination has been found.

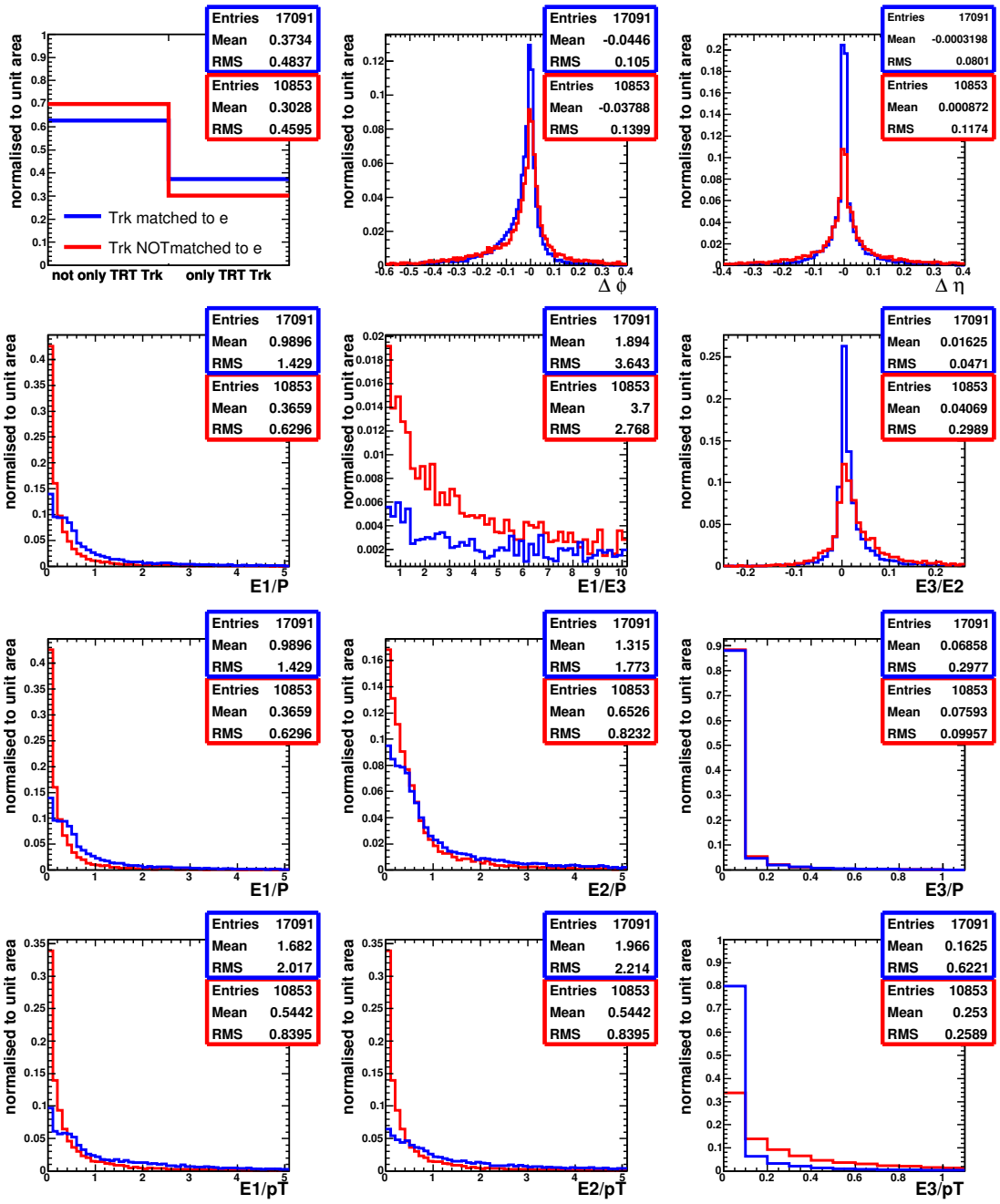

Figure 7.6: *Several variables from track extrapolation to the ECal*

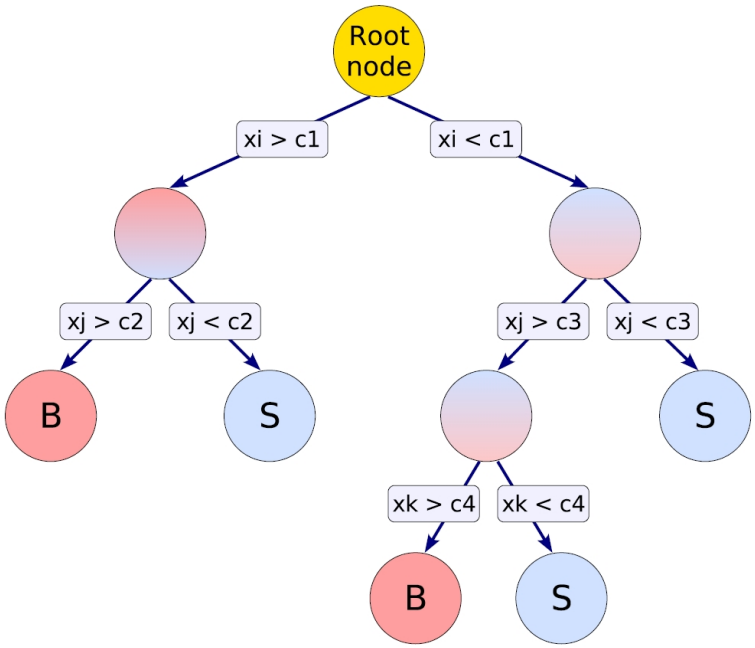

Figure 7.7: *Schema of a simple decision tree [46]. The phase space of a variable (x) is splited into smaller regions. After scanning the whole training sample the nodes are assigned signal or background like depending on the majority of the entries.*

To study if the combination of some or all of these variables will lead to a useful discrimination value a multivariate analysis by TMVA [46] has been performed.

7.4.2 Brief Introduction in TMVA techniques

To optimise the input variables and to find the best discrimination method several TMVA classifiers and optimisation methods have been tested. Here only the methods which have been used with the optimised setup should be explained very briefly.

Variable decorrelation

Due to the fact, that e.g. Boosted Decision Trees or multidimensional likelihood approches underperform due to correlations between the input observables, TMVA offers a method to decorrelate the linear correlations of such observables. Therefore the linear correlations of the training sample is measured and the square-root of the covariance matrix is computed. Then the linear decorrelation of the input variable is estimated by multiplying the square-root of the matrix by the initial variable ntuple. This has to be done seperately for the signal and the background sample. Only linearly and Gaussian distributed variables can be completely decorrelated. This method has been applied during the optimisation and it has attained better results than without decorrelation.

Boosted Decision trees (BDT)

A decision tree [47] is a binary yes/no (smaller or larger than) decision finder. As shown in Fig. 7.7 the phase space of one variable (x) is devided into smaller regions. Therefore

a defined sample (training sample), with known signal and background identity of each event, is split by such binary cuts. The sequence of repeating yes/no decisions builds a decision tree. A defined break condition stops the sequence. Then the final *leaf* nodes are evaluated if the entries are more signal or more background like. Due to the problem that such a procedure is very instable with respect to statistical fluctuations of the training sample the trees are *boosted*.

The boosting algorithm builds several trees with the same input variable but reweighted. This means that e.g. in the so-called *adaptive boost* algorithm events which are misclassified during the training of a tree are giving a higher weight in the training of the next tree. So a forst of trees is created which is less susceptible to statistical fluctuations.

Nevertheless an overtraining of a boosted decision tree has to be avoided. Overtraining may occur if too many decisions are executed on a variable with insufficient degrees of freedom. Such an overtraining can be found if the statistics of the number of events in one node is to low. Then a quite reasonable signal and background discrimination can be achieved for the training sample, but the result is not sample independent. This can be seen in Fig. 7.9 where the the Kolmogorov-Smirnov test shows the probability of the agreement between the training sample and test sample shape. The upper left plot shows the result of a strongly overtrained BDT. No agreement between training and test sample could be found. The upper right plot shows the optimised training. To avoid overtraining a *pruning* function [48] runs over the tree from the last to the first decision and removes statistically insignificant nodes. It has been found that first growing the tree to its maximum and then cut it back is more efficient than interrupting the node splitting at an earlier stage. This is due to the possibility that apparently insignificant splits can lead to quite sufficient splits further down the tree.

To summarise the performance of boosted decision trees two main aspects have to be mentioned. On the one hand it is a very easy discrimination method which becomes more complex and less transparent by reweighting (boosting) the desicions. Thus one has to take care not to overtrain a BDT which is very likey the case. But on the other hand BDTs are also very sensitive on very poorly discriminating input variables. This can be seen in the output of BDT response in Fig. 7.9 (upper right plot) compared to the result of a simple likelihood discrimination (lower right plot).

Predictive Learing via Rule Ensembles (RuleFit)

Another very promesing result is given by the *RuleFit* classifier [49]. It is based on an ensemble of so-called rules to creat a scoring function. The easiest way to get such an ensemble of rules is to extract it from a forest of desicion trees. Each node can be regarded as one rule. Then lineare combinations of the rules are created weighted with coefficients which are calculated via a regularised minimisation procedure. The linear combination of all rules defines a so-called score function. After fitting a *RuleFit* response is provided.

As the authors from TMVA categorise the *RuleFit* classifier as very untransparent the desicion was made to use the BDT instead of the slightly better performing *RuleFit* classifier (shown in Fig. 7.10).

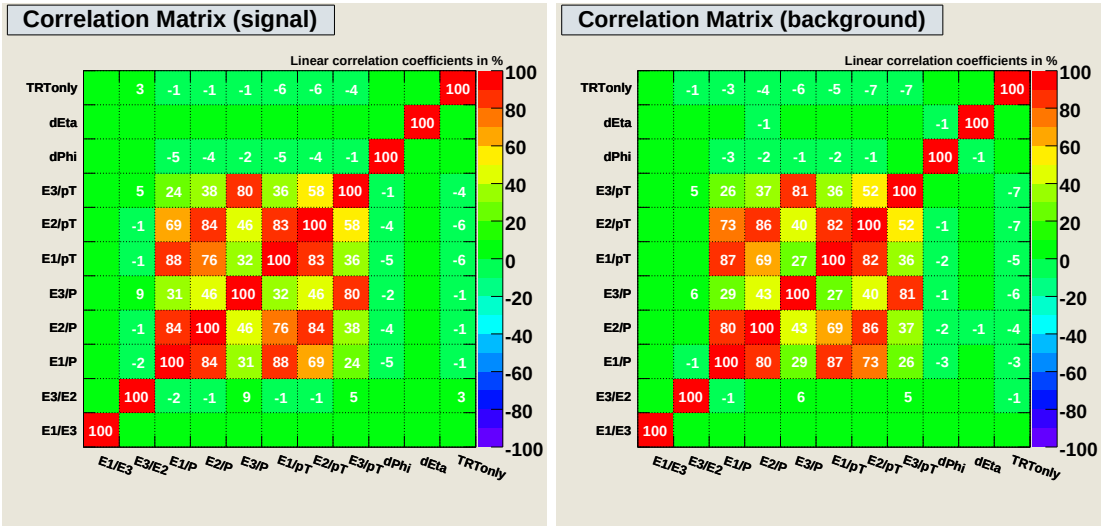

Figure 7.8: *Linear correlations between TMVA input variables for signal and background.* ***Left:*** *Signal means VxCandidate truthmatched to electrons.* ***Right:*** *Background in VxCandidate container, e.g. pions and kaons. The smaller the number (green) in one bin the less linearly correlated the variables are.*

7.4.3 Multivariate Analysis for Discrimination

First all observables shown in Fig. 7.6 has been used as TMVA input variables. Large linear correlations have been found between some of these observables (cp. Fig. 7.8). Testing several input variable combinations led to the result that using the variables of the first six histograms shown in Fig. 7.6 yields most uncorrelated and best discrimination.

Additional to several input variables different TMVA classifier has been tested. As example the overtraining check of TMVA for some classifiers are shown in Fig. 7.9. The training and input variable choice has been optimised for the boosted decision tree classifier with decorrelated input variables (BDTD) (upper right plot). The upper left plot should be given an example of a strong overtrained BDT classifier.

The results of the background rejection versus the signal efficiency for the different classifier are shown in Fig. 7.10. The BDT achieves the best result, but due to the mentioned overtraining another classifier has to be chosen. The Likelihood discrimination has the worst performance, which also can be seen in the distribution of the Likelihood response in Fig. 7.9 lower right plot. Despite of the slightly better performance of the RuleFit classifier the BDTD results have been selected to use for further analysis. The reason therefore is the more transparent and less complex implementation of the BDTD.

The plots for the best cut value and the depending efficiency and significance for the different classifiers are shown in the appendix Fig. A.17.

To get a more comparable result to the electron probability of the TRT electron PID tool (**eProb**) and in order to allow for a simple way of combining the informations, it was decided to calculate an electron probability out of the BDTD response distribution based on the informations from electromangnetic calorimeter (**ECal eProb**).

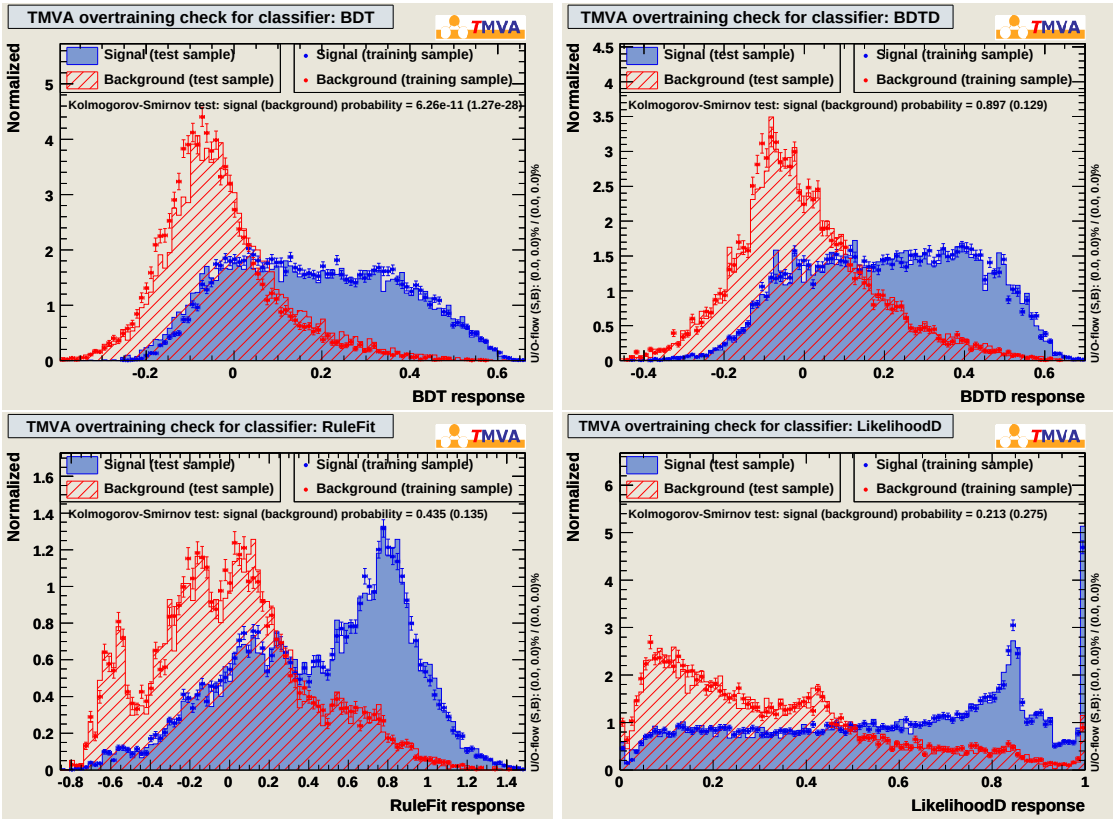

Figure 7.9: TMVA offers a so-called overtraining test. After training the classifiers with a so-called training sample with specified signal and background events, it has to be tested if the applied cuts also work with another sample. The overtraining test checks how likely the response shape of the training and the testing sample agree - the Kolmogorov-Smirnov probability. **Upper left:** This plot shows a strongly overtrained BDT. It can be seen clearly, that the shape of the training sample (signal blue and background red) plotted as dots with error bars does not agree with the result of the test sample, which is shown in the filled histograms. **Upper right:** The optimal trained BDT with decorraleted input variables (BDTD) shows a much better agreement of training and test sample. **Both lower plots** also fit very well, but the shape of the likelihood discrimination does not differ much of signal and background.

7.4.4 Building an Electron Probability with ECal observables

The ECal eProb value is calculated via the BDT output of TMVA. Therefore a "look-up" histogram with truthmatched data has been created. The electron probability is defined as the probability for a track with a certain BDT value belonging to the signal distribution. This can be calculated by dividing the number of entries for each bin of the BDT response of the signal by the number of entries of the signal and the background of the BDT value in each specific bin.

$$ECal\ eProb_{bin} = \frac{signal_{bin}}{(signal + background)_{bin}} \tag{7.3}$$

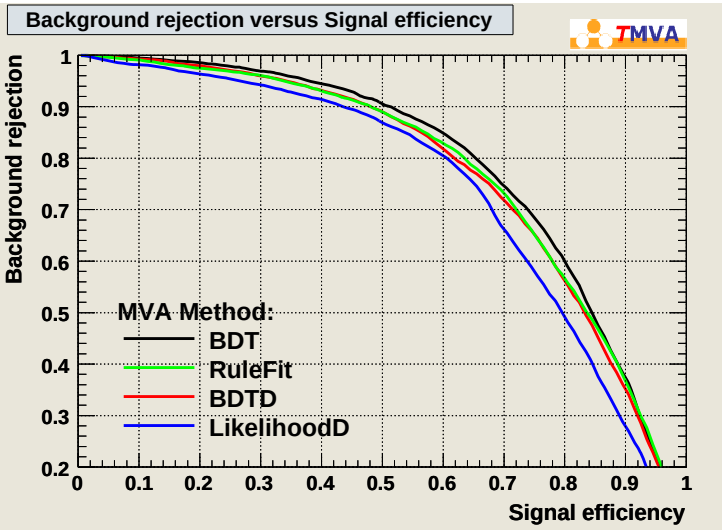

Figure 7.10: *Rejection over efficiency for the several TMVA classifiers. The best possible discrimination would reach the upper right corner (1,1) which means for 100% signal efficiency 100% of the background would be rejected. The distribition of BDT is as example of a strongly overtrained boosted decision tree, thus the other distributions have to be compared.*

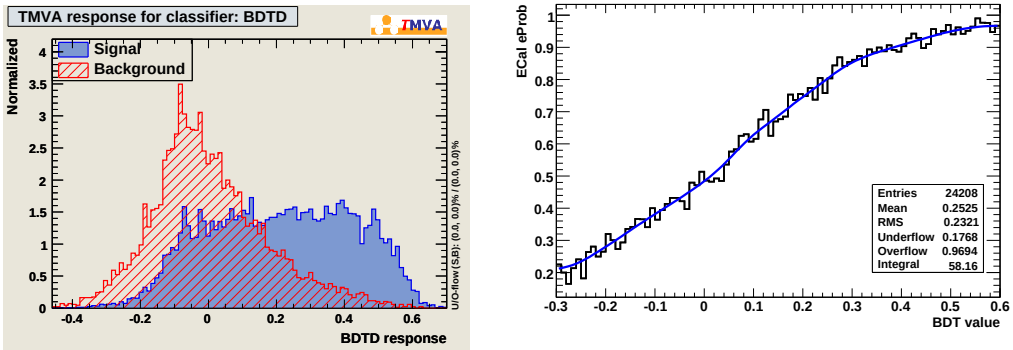

Figure 7.11: *For each bin of the BDTD response distribution the ratio of signal/(signal + background) is calculated and written into the so-called "look up" histogram. Via that histogram for each BDTD response the depending electron probability (**ECal eProb**) can be determined.*

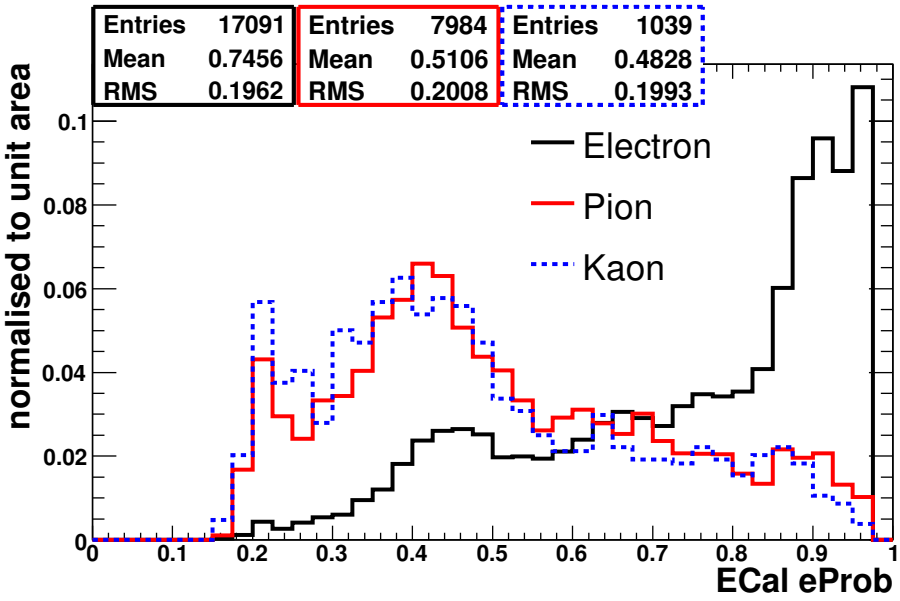

Figure 7.12: *Electron probability calculated of ECal observables only. Distributions of VxCandidate tracks truthmatched to electrons, pions and kaons.*

The right plot of Fig. 7.11 shows the "look-up" histogram. For each track a BDT value can be calculated via TMVA. With the distribution of the "look-up" histogram for each BDT value the electron probability can be determined. By using the "look-up" histogram without smoothing large fluctuations in the electron probability are visible. These fluctuations are due to the form of the "look-up" distribution. The distribution has been calculated bin by bin thus unphysical fluctuations occur. To avoid such fluctuations the "look-up" histogram has to be smoothed. Several steps of such a procedure are shown in Fig. A.18 in the appendix. The method shown in the lower middle histogram has been chosen as the best smoothing method.

Using the function shown in blue in the right histogram Fig. 7.11, the ECal eProb for the *VxCandidate* tracks can be calculated. To differentiate electrons, pions and kaons the tracks has been truthmatched and the several distribitions are plotted in Fig. 7.12.

To compare the results using the ECal eProb variable as discrimination value with the results of the eProb discrimination (cp. Sec. 7.3.1) the same preselection cuts have been applied. To achieve a comparable electron purity a cut on the ECal eProb value $>$ 0.85 has to be applied. Using the ECal eProb variable for identification leads to a much lower photon conversion efficiency than cutting on the eProb value.

Thus it has to be checked if a combined electron probability of Inner Detector observables (eProb) and the electromagnetic calorimeter values (ECal eProb) leads to an improvement.

A reason for the less good discrimination with the informations from the ECal only may be due to the very busy τ environment which means a lot of neaby tracks depositioning a lot of energy in the ECal distorting the shapes in the calorimeters. This has also been observed by the default photon conversion identification.

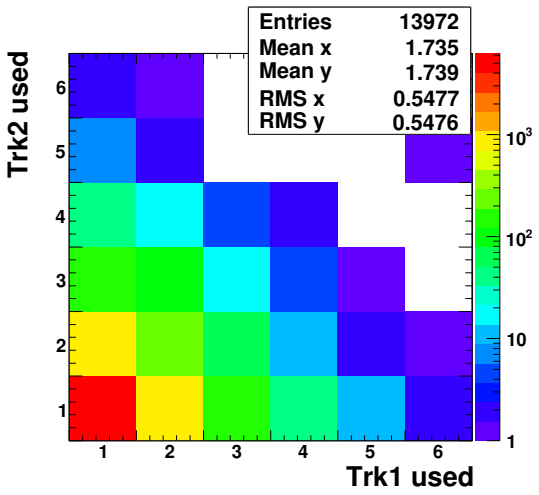

Figure 7.13: *Number of how often a tracks used to build VxCandidates*

Table 7.2: *Cut flow table of preselection cuts on VxCandidates*

Cuts	Eff.	Purity	el. Pur.
no cuts			
no cuts	34.4 %	22.8%	47.4 %
preselection cuts			
in τ cone $\Delta R < 0.3$	36.1 %	28.4 %	58.7 %
$\chi^2_{Prob} > 0.00001$	29.0 %	36.5 %	67.9 %

7.5 Combinatorics

Before comparing different combinations of **eProb** and **ECal eProb** it has to be mentioned that due to the method of reconstructing *VxCandidates* a large fraction of combinatorial background is provided by the reconstruction algorithm.

In Fig. 7.13 the quantity of how often one specific track has been used to built a *VxCandidate* is plotted. The x-axis shows how often track_1, the positive conversion track, has been used to build a *VxCandidate* and the y-axis how often the negative track (track_2) has been reused. A mean value of 1.7 for both tracks shows how much the *VxCandidates* are impurified by combinatorial background.

7.5.1 Applying Preselection Cuts

To reduce the combinatorial background two preselection cuts have been chosen. First to improve the τ reconstruction algorithm only photon conversions which may change the τ track multiplicity are taken into account. Thus only conversions in the τ cone with $\Delta R < 0.3$ around the direction of a τ candidate are included. The second preselection cut is based on the goodness of the secondary vertex fitting. A value of $\chi^2_{Prob} > 0.00001$ has been chosen to cut on the χ^2 probability which has been derived during the vertex reconstruction. Using these preselection cuts yields the efficiencies and purities listed in Tab. 7.2.

The reason of introducing a second purity the *electron purity* is the fact that if electon tracks from photon conversions should be vetoed it does not matter if the electron comes from a combinatorial background conversion or from a correctly reconstructed conversion. It is only important to make sure that it is definitly an electron. Thus the electron purity is the interesting value for this study.

Nevertheless applying the mentioned preselection cuts does not yield a sample of *VxCandidates* with a satisfied amount of reduced combinatorial background (cp. 7.14 magenta distribution). Thus a function has been implemented ensuring that one specific track is only used once to build a *VxCandidate*. This is described in the next section.

7.5.2 Selection of the "best" *VxCandidates*

An additional contraint is needed to decide which of the *VxCandidates* built with one specific track may be the most likely photon conversion and which are the combinatorial background. As an example two of the analysed variables, the distance from the interaction point and the invariant mass of the *VxCandidate* are shown in Fig. A.19 in the appendix. As the distribution of the invariant mass has a better discrimination power, this variable has been chosen to evaluate which *VxCandidate* should be kept. The function gathering out the most likely *VxCandidate* is implemented as follows:

1. It loops over all positive tracks of *VxCandidates*

2. If tracks are found, which are used to build more than one *VxCandidate*, the invariant mass of all those pairs are calculated. The invariant mass has to be calculated out of the tracks before the secondary vertex refit of the tracks because otherwise the invariant mass is set to zero during that fit and no discrimination is possible.

3. Then the pair with the smallest invariant mass is stored in a *new VxCandidate* collection. All *VxCandidates* building of only once used tracks are stored there too.

4. After looping over all positive *VxCandidate* tracks a loop over all negative tracks in the *new VxCandidate* collection has to be done.

5. Again, if there are *VxCandidates* found, built of the same negative tracks, the pair with the smallest invariant mass is kept, the other pair(s) is (are) removed from the new collection.

After that cleaning all left over *VxCandidates* consist of different tracks. To reduce combinatorial background before running this method a cut on one of the electron probability variables has to be done. This wil additionally reduce the computing time. In Fig. 7.14 the number of reusing track_1 and track_2 is shown seperatly. The black distribution shows the *VxCandidates* without any cuts, applying the preselection cuts leads to the magenta distribution. Cutting additionally on the electron probability eProb $> 90\%$ results in the grey histogram. If the combinatorial supressing function is run on top of all these cuts the distribution in blue is left.

This method has been tested as an additonal tool to achieve the best possible cleaned *VxCandidate* collection. While the method achieves a high purity (only once used tracks

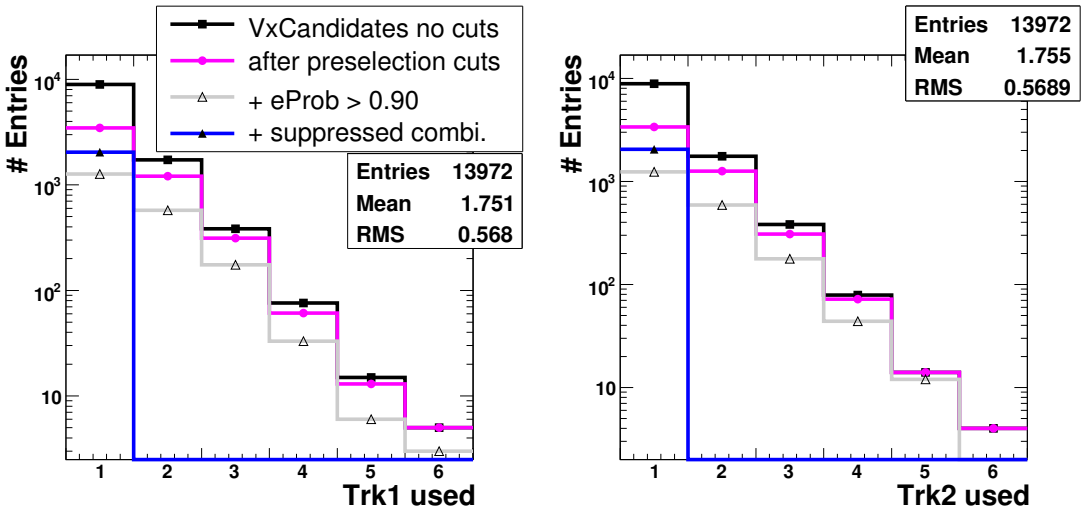

Figure 7.14: *Number of reused tracks to build VxCandidates. Both plots show how often one specific track has been used to built a VxCandidate. Trk1 are the tracks of positive charged particles and Trk2 are the negative charged tracks.*

are left in the blue distribution), the resulting efficiency after that kind of cleaning is too low to use it for further studies.

After presenting the two different implementations of electron probabilities and the combinatorial background cleaning method the performances have to be compared and the best method or the best combination of methods have to be chosen to identify the photon conversions in τ candidates.

7.6 Comparison of the Different Identification Methods

The definitions and the variables of the electron probabilities eProb and ECal eProb introduced in the sections above are linearly uncorrelated. Figure 7.15 shows 2-dimensional plots of the electron probabilities eProb versus the ECal eProb of truthmatched electron, pion and kaon distributions. The calculated linear correlation factors are 7.21 % for electrons, 8.95 % for pions and 11.59 % for kaons. The slightly higher correlation factor of the kaons may be due to smaller statistics.

Thus several combination of the two probabilities may lead to a improved performance. The probabilities and some possible combinations listed in Tab. 7.3 have been examined as discrimination variables, to identify the *VxCandidates*.

Table 7.3: *The listed variables have been analysed to chose the best discrimination variable for the electron identification of photon conversion originating in τ decays.*

1.	eProb
2.	ECal eProb
3.	eProb and ECal eProb
4.	eProb $\cdot$ (ECaleProb)
5.	$\text{eProb}_{\text{Trk}_1} \cdot (\text{ECal eProb})_{\text{Trk}_1} \cdot \text{eProb}_{\text{Trk}_2} \cdot (\text{ECal eProb})_{\text{Trk}_2}$

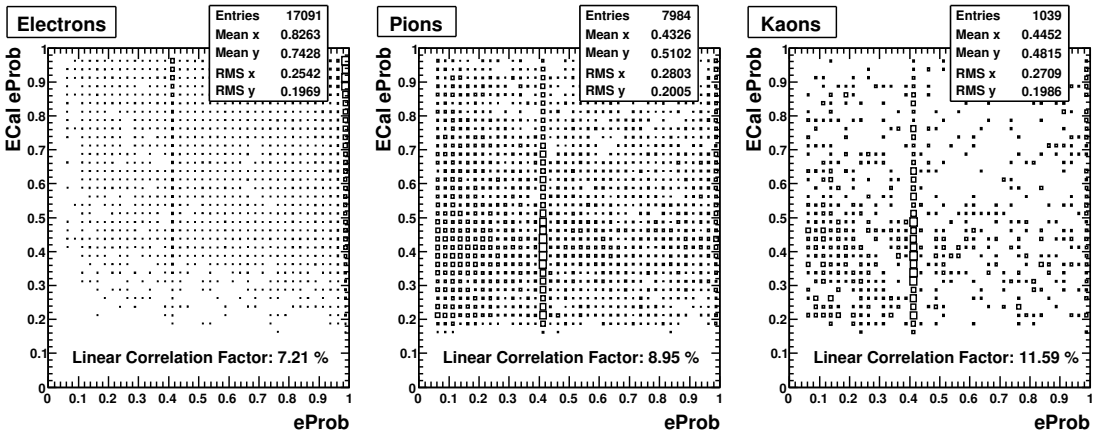

Figure 7.15: The 2-dimensional plots do not show correlation between the electron probabilities eProb and ECal eProb for electrons, pions and kaons.

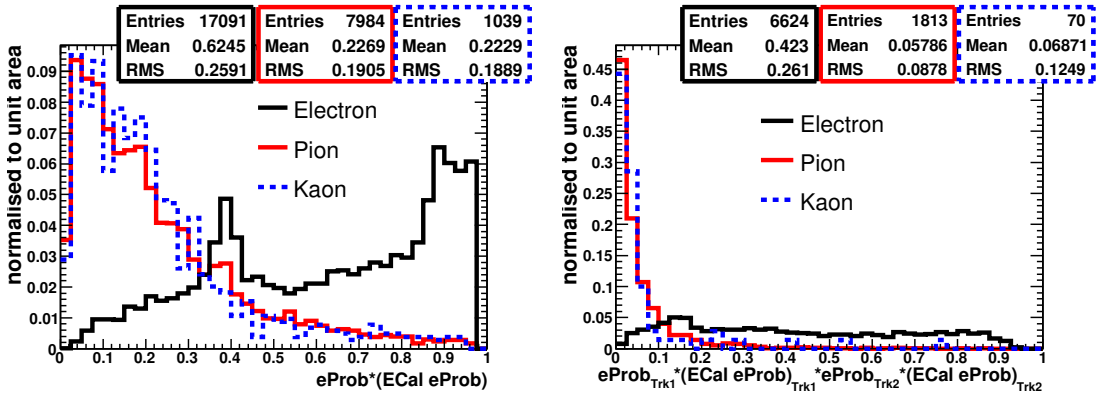

Figure 7.16: *Two possible combinations of the two electron probabilities: eProb and ECal eProb.* ***Left****: eProb and ECal eProb combined via the multiplication of the two probabilities of each track.* ***Right:*** *Here a combined electron probability value for the whole conversion, means for both tracks together is calculated.*

The important distributions, which have to be analysed, are the eProb distribution (fig. 7.5), the ECal eProb distribution (fig. 7.12), and the combinations of eProb and ECal eProb of one track (Fig. 7.16 left) and the combination for a whole conversion candidate, meaning the multiplication of both probabilities of both tracks (Fig. 7.16 right).

To compare the performance histograms of the electron purity versus the reconstruction efficiency of each probability and probability combination have been plotted. Therefore each discrimination variable has been scanned for a probability from 0 to 1 in 0.05 steps. The result of all variables listed in Tab. 7.3 are shown in Fig. 7.17. The exact values of the efficiencies and the electron purities are listed in more detail in Tab. A.6 in the appendix.

The distribution in green shows the performance of ECal eProb alone. This result gives the worst electron discrimination of all. This has to be probably attributed to the dense τ environment an the difficulty to discriminate a cluster created by an electron

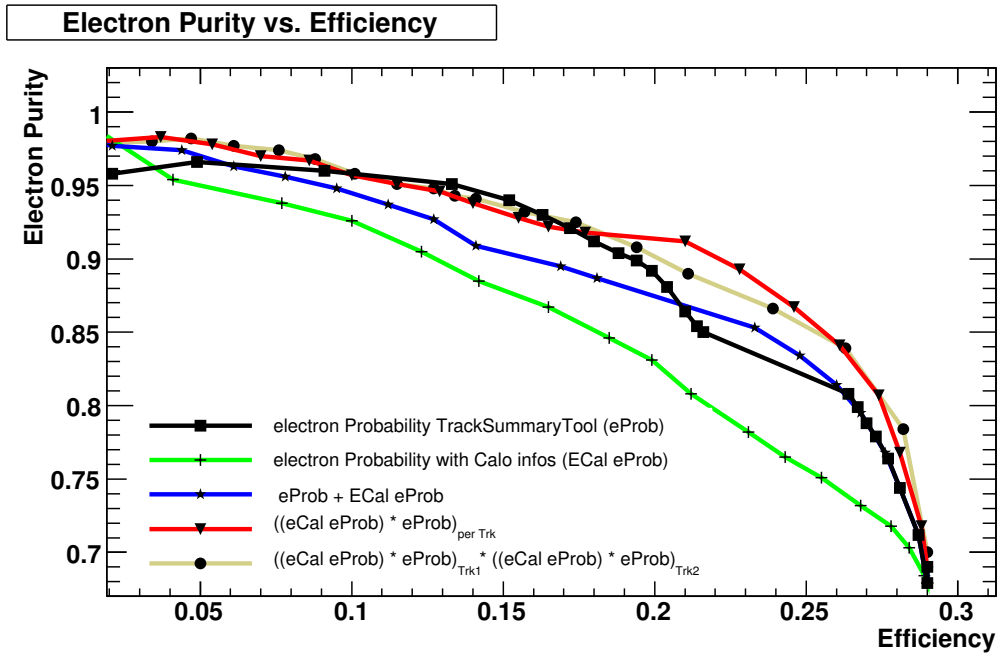

Figure 7.17: *Electron purity vs reconstruction efficiency of VxCandidates for several cleaning methods. Most robust method is the eProb of the Inner Detector which has been chosen to use as conversion veto in τ candidates.*

from the energy deposition of nearby passing neutral and or charged pions. Even the combination of the Ecal eProb and the eProb (in blue) by cutting on both variables on a specific value does only slightly improve the bad ECal eProb performance. Thus the multiplied combinations of the two probabilities (red and lighgrey) and the eProb alone (black) are considered. The eProb performance is worse than the combined ones in the higher efficiency region but it is comparable and slightly better in the high purity region.

As the cleaned *VxCandidates* are meant as vetoing tool, after the identification only a very pure electron collection should be remain. Thus electron purities higher than 90 - 95 % have to be achieved. In that region the three distributions show more or less the same performance.

To study the several performances in more detail additional plots showing the efficiency for comparable electron puritiy ($\approx$94%) and the electron puritiy for comparable efficiency ($\approx$15%) are pictured for the conversion radius R, the transverse momentum of the conversion tracks p_T and the pseudo rapidity η in the figures (7.18 and 7.19).

Firstly the distributions of the orignal *VxCandidates* without any cuts have to be discussed. Then the characteristics of the different combined electron probabilities are mentioned. The efficiency in R dereases with higher conversion radii. Two dips can be observed. The first one at radii of around 370 mm this is where the second SCT layer in the barrel region is installed. If a photon conversion does happen at larger radii the conversions have to be reconstructed via the Outside-In tracking, their performance is slightly worse than the Inside-Out tracking. The second and more dramatically derease is at radii larger than 650 mm there the reconstruction efficiency of the TRT standalone tracking decreases due to the measure of the TRT module. A slightly higher efficiency

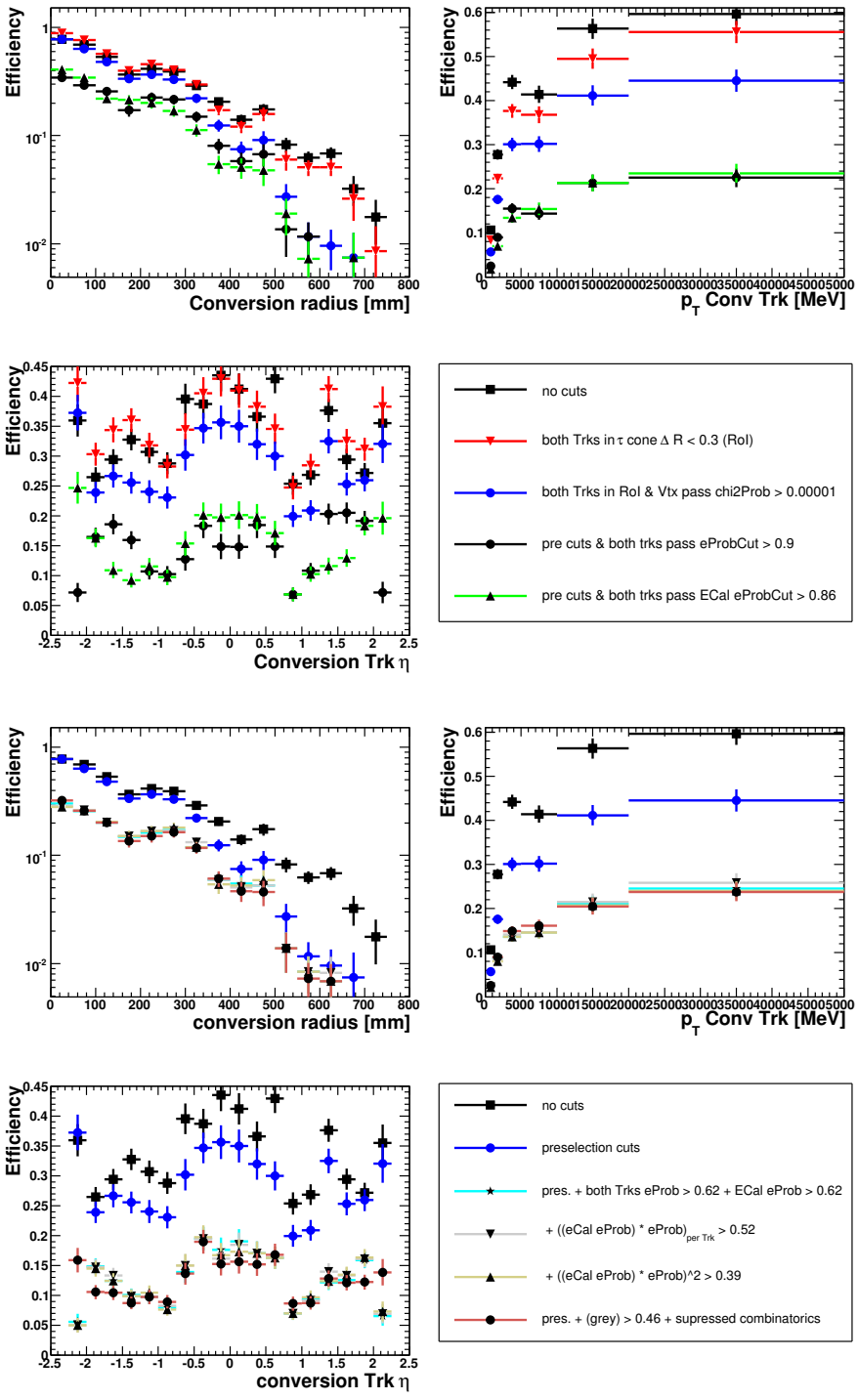

Figure 7.18: Efficiency of photon conversion reconstruction after *VxCandidate* identification with an electron purity of 94%. The upper three plots show the performances of the individual and the preselection cuts in R the conversion radius, p_T the conv. track transverse momentum and η the pseudorapidity of the tracks. The lower three plots show the same variables, but for several combined cuts.

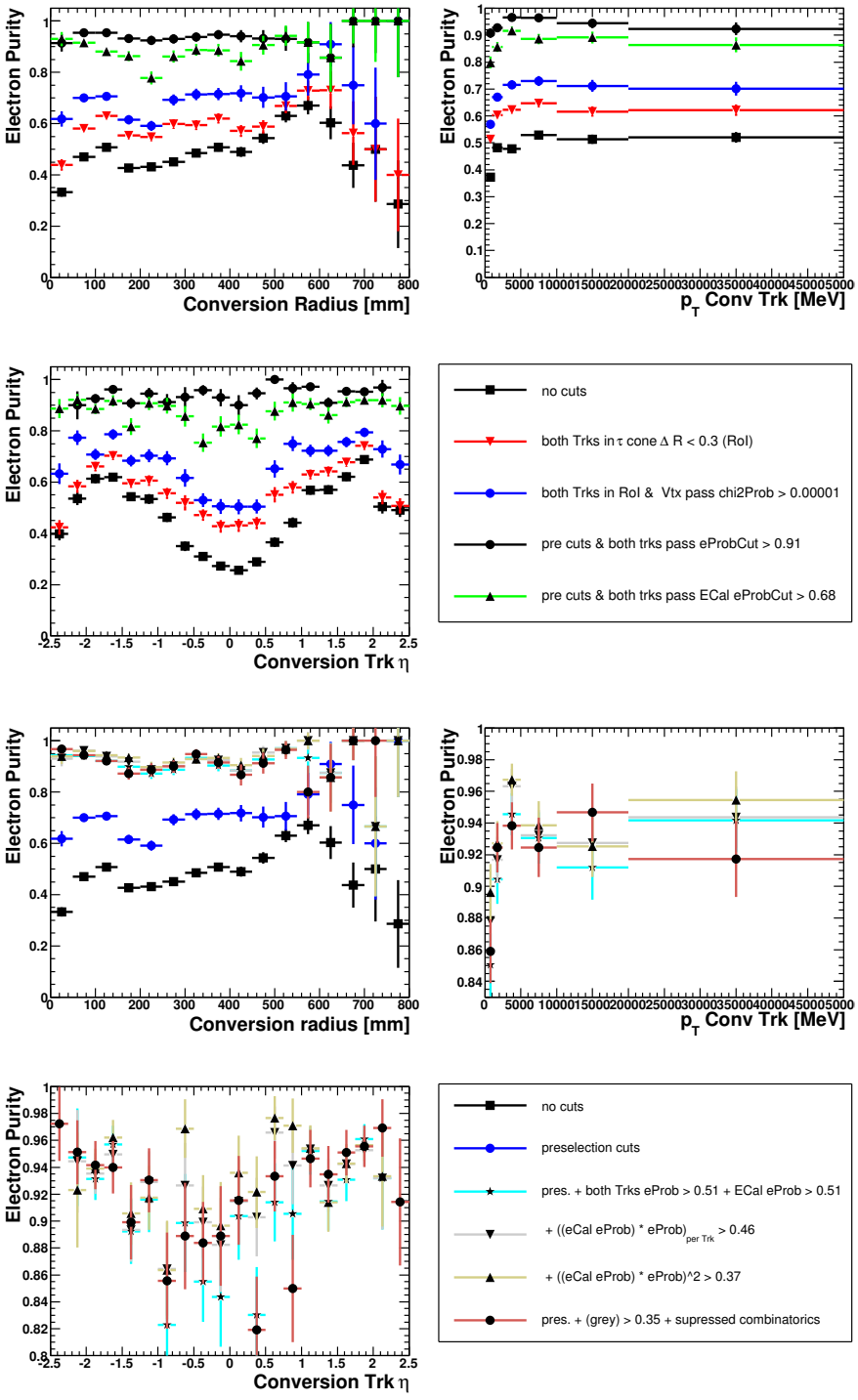

Figure 7.19: Electron purity of photon conversions after *VxCandidate* identification with an efficiency of 15%. The upper three plots show the performances of the individual and the preselection cuts in R the conversion radius, p_T the conv. track transverse momentum and η the pseudorapidity of the tracks. The lower three plots show the same variables, but for several combined cuts. The electron purity distribution in p_T and η of the combined probability cuts are in the same region, therefore only a detail of the plot is shown.

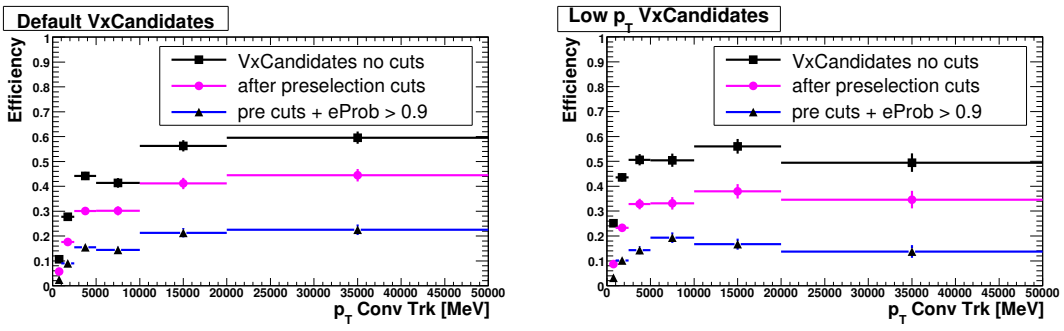

Figure 7.20: *Comparison of the default (left) and the low p_T VxCandidates (right). In both histograms distributions of the VxCandidates without cuts, after applying the preselection cuts and after identifying the photon conversions via a cut on the eProb value are plotted.*

after the cut on $\Delta R < 0.3$ between the conversion tracks and a τ candidate can be observed. After that cut only conversions lying in the τ cone of one τ candidate are analysed. The efficiency increases due to neglecting both reconstructed and truth photon conversions not lying in the τ candidate cones. Most efficient discrimination is based on the eProb cut alone.

The shape of the efficienies in p_T look the same for all cut variables. The efficiency is best for tracks with a transverse momentum larger than 20 GeV and drops for smaller transverse momenta. This is due to the optimisation of the *VxCandidates* to photon conversions from $H \rightarrow \gamma\gamma$ events. Some optimisations of so-called low p_T conversions have been made, but not yet used here. These low p_T *VxCandidates* are optimised for minimum bias events. Figure 7.20 compares the p_T distributions of the default *VxCandidates* and the low p_T *VxCandidates* without any cuts, after the preselection cuts and after the photon conversion identification with a cut on eProb > 0.9. The low p_T *VxCandidates* show much higher efficiencies for conversions with track momenta below 10 GeV, whereas the default *VxCandidates* are more efficient above 10 GeV. Thus it has to be cross checked if this may improve the conversion veto in the future.

The efficiency in η shows a plateau in the middle $|\eta| < 1$. There the efficiency is about 10 % higher than for larger pseudo rapidity values. This is due to the measure of the TRT barrel, which ends at $|\eta| = 1$. Apart from the preselection cuts the η distributions of all probabilities /combinations drop. They show more or less the same shape.

The distributions of the electron purities for efficiencies around 15 % support the results from the efficiency study. The only difference is that the results of the combined cuts only differ between 1.5 and 2.5 %. Thus the plots have to be zoomed such that those marginal diviations can be observed.

Analysing these plots lead to the result, that the combined electron probabilities are strongly dominated by the eProb value, whereas the result of the eProb discrimination alone is only slightly worse than the best combination.

Thus the result of this extensive study is to use initially only the eProb variable as discrimination value to identify the *VxCandidates*. This decision has been made, because the eProb variable is not based on such a complex multivariate analysis as the ECal eProb is and therefore it is more robust and suitable for early data taking. Apart from the

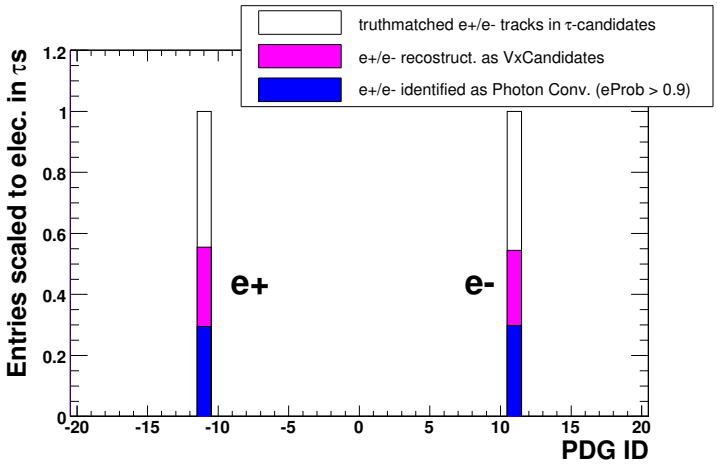

Figure 7.21: *Fraction of tracks reconstructed as τ candidate track and photon conversion track after conversion identification. The white histogram shows the truthmatched electron and positrons in τ candidates, whereas in magenta the fraction of overlapping with VxCandidates and in blue the fraction overlapping with identified conversion are added.*

already achieved purity the combination of the probabilities on which the eProb variable is based as decribed in Chap. 7.3.2 is still under construction. Some very preliminary results [50] promise further improvements.

By using the TRT electron PID tool based electron probability as discriminating variable only a cut on eProb > 0.9 leads to a conversion reconstruction efficiency of 15.2% and an electron purity of 94%. Based on this cut the purified *VxCandidates* can be used as photon conversion veto in τ candidate tracks, described in the following section.

7.7 Improved Track Multiplicity of τ-Candidates

After identifying photon conversions as described above, each remaining track which is also stored in the τ candidate track collection can be used to veto or to tag a τ track. The overlapping tracks are shown in Fig. 7.21, in white the truthmatched tracks of the τ candidate, in magenta the overlapping tracks of default *VxCandidate* and in blue the fraction of overlapping tracks after applying the preselection cuts and a cut on the eProb > 0.9 on the *VxCandidates*.

In the left column of Fig. 7.22 the corrected track multiplicity for τ candidates is shown without any applied track selection criteria. The right column shows the track multiplicity with the default *TauRec* track selection criteria.

The y-axis shows the number of τ candidates with a specific number of reconstructed tracks (x-axis). The different colours differentiate the truthmatched decay modes 1 prong, 3 prong and QCD background, whereas the solid and dashed lines are the different modes before and after conversion veto.

Comparing the upper row, both decay modes and the QCD background contain higher track multiplicities than after applied track selection criteria (TSC). The higher track multiplicity of the QCD background is also corrected by the photon conversion veto. On

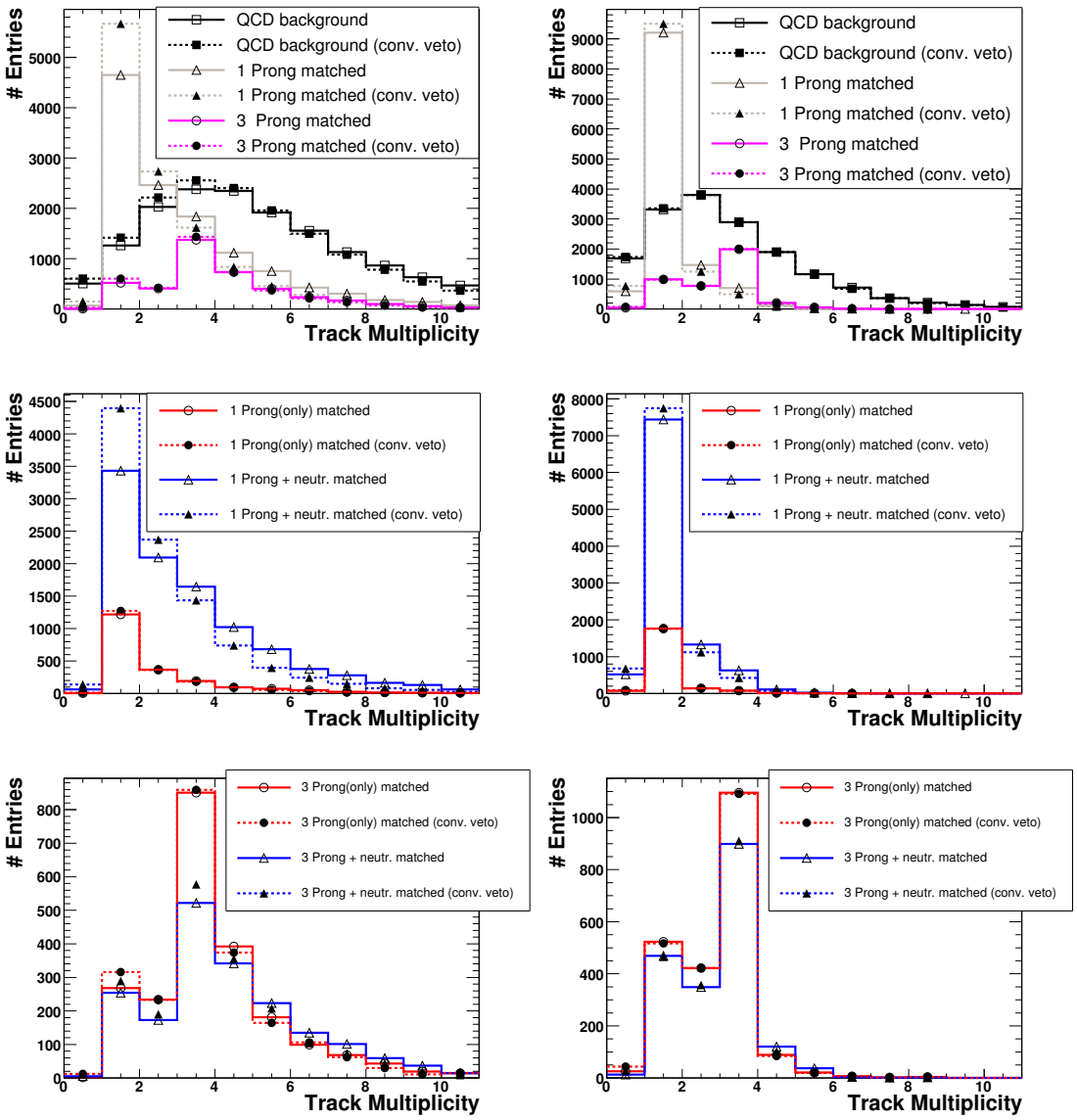

Figure 7.22: *Track multiplicity of τ tracks before (solid lines) and after (dashed lines) photon conversion veto.* ***Left column:*** *No track selection criteria are applied.* ***Right column:*** *Track multiplicity with the default TauRec TSC.* ***From top to bottom:*** *distributions of 1 prong, 3-prong and QCD jets, 1-prong subdivided into 1-prong with and without π^0s and 3-prong subdivided into 3-prong with and without π^0s.*

the one hand this is expected, because many photon conversions are created within a QCD jet, on the other hand this correction may change the QCD track multiplicity to lower values so that it looks more signal like. If applying the TSC the change of QCD tracks decreases to a marginal correction. Thus the photon conversion veto should not increase the number of fakes too much, if the TSC are considered. But before this cleaning method could be implemented in the *TauRec* reconstruction algorithm additional fake rate studies need to be done.

Table 7.4: *Track multiplicity of τ candidates truthmatched to 1 prong decays before and after the explicit photon conversion veto. The figures listed are the fraction of all reconstructed 1 prong decays.*

Track Multiplicity:	0	1	2	3	4	5	> 5	total
1 Prong only	0.59	14.53	1.16	0.63	0.11	0.05	0.01	17.08
1 Prong only veto	0.76	14.49	1.11	0.58	0.09	0.03	0.01	17.08
1 Prong with pi^0	4.29	61.27	10.96	5.16	0.91	0.19	0.15	82.92
1 Prong only veto	5.62	63.81	9.20	3.48	0.61	0.08	0.13	82.92

Table 7.5: *Track multiplicity of τ candidates truthmatched to 3 prong decays before and after the explicit conversion veto. The figures listed are the fraction of all reconstructed 3 prong decays.*

Track Multiplicity:	0	1	2	3	4	5	> 5	total
3 Prong only	0.64	12.81	10.34	26.84	2.18	0.51	0.32	53.64
3 Prong only veto	1.08	12.66	10.34	26.72	2.06	0.47	0.32	53.64
3 Prong with pi^0	0.32	11.49	8.55	22.02	2.94	0.91	0.15	46.36
3 Prong only veto	0.69	11.39	8.74	22.31	2.50	0.59	0.15	46.36

Apart from that, a good improvement for 1 prong decays in the distribution without TSC can be achieved. This is due to the higher fraction of truth photon conversion tracks in the τ candidates without TSC. Unfortunately the overall reconstruction performance is a lot worse if no TSC are applied. Which TSC can be loosen and which has to be applied also have to be studied in more detail. As first step the improvement of the track multiplicity with applied TSC should be discussed. The improvement is not as strong as without TSC but a slight enhancement can be observed. The total numbers of the upper two histograms and the fraction of the several decay modes are listed in Tab. A.9 and Tab. A.10 in the appendix.

Both track multiplicity distributions after conversion veto are dominated by the corrections in 1 prong decays (cp. the two upper plots in Fig. 7.22 with the plots in the middle). There only changes in case of additional neutral pions are observed. This is theoretical expected. Only if a τ decays into charged and neutral pions, photons from the decaying π^0 are predicted. The detected fractions of around 17% of 1 prong decays without neutral pions and around 83% with π^0s agree with the theoretical predictions mentioned in Chap. 2.3.1.

The two lower plots in Fig. 7.22 show the corrections for 3 prong decay modes with the corresponding Tab. 7.5. The enhancement is only marginal. But in both modes without and with applied track selection criteria a shift towards three recontructed tracks can be observed if additional neutral pions are produced. Nevertheless the result of mainly changing distributions with neutral pions adverts to the correctness of the cleaning method and the vetoing in general.

7.8 Conclusion

If the explicit photon conversion veto is run on top of the *TauRec* default tracks selection criteria, which have been partly applied to suppress conversion tracks, still an improvement can be achieved.

Thus the combination of applying the track selection criteria on the τ candidates and using the more robust and more conventional TRT electron PID tool of the Inner Detector to identify the photon conversions, which is the most pessimistic setup leads to the following enhancement:

29.7% of the truthmatched electrons are reconstructed both as τ candidate track and as track of identified photon conversions. After achieved an electron purity of 94% these overlapping tracks can be vetoed in the τ track multiplicity. Doing that the correctly reconstructed track multiplicity of 1 prong decays can be increased by a factor of 3.13% whereas the 3 prong decay reconstruction can only be advanced by a factor of 0.19%.

Chapter 8

Summary and Outlook

The reconstruction of τ leptons at a hadron collider such as the LHC is a difficult but very important task. Both the discovery of the Higgs boson and the precise study of SUSY need a very good τ reconstruction, as in both cases important τ lepton final states exist. Even information on the spin correlation in case of τ pair productions can be determined. Therefore the τ decay mode has to be reconstructed correctly. Interactions of photons with detector material lead to photon conversions. Thus there additional charged tracks in the τ lepton environment have to be taken into account. This thesis describes several developments to achieve an explicit photon conversion reconstruction and identification in the τ decay cone.

First the content of τ candidate tracks has been analysed. Even if the *TauRec* track selection criteria are applied, which have been partly developed to suppress photon conversion tracks, about 8.2% of all τ tracks are electrons and positrons. About 66% of these tracks originate from photon conversions. In order to identify these tracks correctly an explicite photon conversion identification is needed.

After testing the default photon conversion reconstruction and identification algorithm in the τ environment, which has been optimised for photon conversions of $H \rightarrow \gamma\gamma$ processes, it was clear that the existing identification method cannot be used for the given task. A first check revealed that the seeds of photon conversions (*VxCandidates*) provided by the default reconstruction algorithm are sufficiently reconstructed. Thus 55% of the $e^{\pm}$ tracks of τ candidates are reconstructed as *VxCandidates*. Due to the fact that nearly 30% of the *VxCandidates* are pion tracks an electron identification is necessary.

To achieve the highest possible electron purity of *VxCandidates*, two electron identification methods have been tested. The first one provides an electron probability for each track based on the information from Transition Radiation Tracker (TRT). Additionally a method providing an electron probability, based on the information from electromagnetic calorimeter (ECal), has been newly implemented. Both probabilities are sufficiently uncorrelated and can be applied seperately or combined. Due to the very complex structure of the ECal based electron probability, at first only the TRT based electron probability (eProb) is used to identify photon conversions. A cut on eProb $> 90\%$ purifies the *VxCandidates* to an electron purity of 94% with a photon conversion reconstruction efficiency of 15.2%. More important than the photon conversion reconstruction efficiency is the fraction of reconstructed and identified photon conversions overlapping with conversion tracks picked up by the τ reconstruction algorithm as τ track because these are

the tracks which lead to wrong τ track multiplicities. This fraction is 29.7%.

If the *VxCandidates* after electron identification are used as explicit photon conversion veto in the τ candidates an improvement of the τ track multiplicity can be achieved. This conversion veto can be applied either on top of the already existing *TauRec* track selection criteria or without such quality cuts.

If no track selection criteria (TSC) are applied the enhancement is significantly larger. This is due to the fact that the TSC are designed to supress photon conversion tracks. It has to be mentioned that running the *TauRec* algorithm without any track quality cuts is more an academic example of the maximal improvment than a realistic result. The improvement of the correctly reconstructed track multiplicity of 1 prong decays can be enhanced by a fraction of 10.2% in case of 3 prong decays by a factor of 3.8% whereas a shift of the QCD background to one track τ leptons is 1.7% and to three track τ leptons 2.86%. Even if the track selection criteria are applied a slight improvement of the track multiplicity can be achieved. The improvement of the correctly reconstructed track multiplicity of 1 prong decays can be increased by a factor of 3.7%, the 3 prong reconstruction can only be advanced by a factor of 0.8%, whereas even the QCD background is shifted by a factor of 0.6% to a track multiplicity of three and a factor of 0.5% towards one track. All mentioned fractions are the difference of the number of tracks with the correct track multiplicity matched to the τ decay mode divided by the total number of all reconstructed τ candidates.

Methods for an explicit photon conversion reconstruction in the very dense τ environment have been developed. First steps to correctly reconstruct the τ track multiplicity, correctly accounting for photon conversions have been presented. Additional studies on the specific low p_T *VxCandidates* developed for minimum bias events, shortly presented in Sec. 7.6, are needed to check if they can improve the shown results. Further studies of the effect of the conversion veto applied on QCD background with respect to track multiplicity changes are needed. After these checks, the study presented above can be implemented as an extension of the already existing τ reconstruction algorithm. Using the explicit conversion veto information, at least one of the currently used track selection criteria can be re-tuned to optimise the reconstruction of τ leptons.

List of Figures

List of Tables

Bibliography

[1] PDG. Review of Particle Physics. *Journal of Physics G*, 33, 2006.

[2] S. Glashow, *Nucl. Phys.* **22** (1961) 579, S. Weinberg, *Phys. Rev. Lett.* **19** (1967) 1264, A. Salam, in *Elementary Particle Theory*, ed. N. Svartholm, Stockholm, "Almquist and Wiksell" (1968), 367.

[3] S. Weinberg. A Model of Leptons. *Phys. Rev. Lett.*, 19:1264–1266, 1967.

[4] N. Cabibbo. Unitary Symmetry and Leptonic Decays. *Phys. Rev. Lett.*, 10:531–532, 1963.

[5] H. Fritzsch, Murray Gell-Mann, and H. Leutwyler. Advantages of the Color Octet Gluon Picture. *Phys. Lett.*, B47:365–368, 1973.

[6] P. W. Higgs. Broken Symmetries, Massless Particles and Gauge Gields. *Phys. Lett.*, 12:132–133, 1964.

[7] J Alcaraz et al. A Combination of Preliminary Electroweak Measurements and Constraints on the Standard Model, 2006. oai:cds.cern.ch:1016509. Technical Report hep-ex/0612034. ALEPH-2006-001 PHYSICS-2006-001. CERN-L3-310. CERN-PH-EP-2006-042. DELPHI-2006-014 PHYS-948. L3-Note-2833. LEPEWWG-2006-01. OPAL-PR-419, CERN, Geneva, Dec 2006. Preprint not submitted to publication.

[8] R. Barate et al. Search for the Standard Model Higgs Boson at LEP. *Phys. Lett.*, B565:61–75, 2003.

[9] The LEP Electroweak Working Group. The LEP Electroweak Working Group (LEP EWWG) Combines the Measurements of the four LEP Experiments ALEPH, DELPHI, L3 and OPAL. `http://lepewwg.web.cern.ch/LEPEWWG/`, date 27 Sept 2008.

[10] E. W. Kolb and Michael S. Turner. *The Early Universe.* Addison-Wesley, Redwood City, USA, 1990. Frontiers in physics, 69.

[11] C. L. Bennett et al. First Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Preliminary Maps and Basic Results. *Astrophys. J. Suppl.*, 148:1, 2003.

[12] Douglas L. Smith. Professor of Physics David Politzer is Caltech's newest Nobel Laureate. `http://eands.caltech.edu/articles/LXVII3/quark.html`, date 27 Sept 2008.

[13] S. P. Martin. A Supersymmetry Primer. 2006.

[14] Z. Was T. Pierzchala, E. Richter-Was et al. Spin Effects in Tau Lepton Pair Production at LHC. *Acta Phys Polon*, B32:1277–1296, 2001.

[15] Y. Coadou K. Assamagan. The Hadronic Tau Decay of Heavy Charged Higgs in ATLAS. *ATLAS Physics Note*, ATL-PHYS-2000-031, 2000.

[16] P. Bechtle et al. Reconstruction and Identification of Hadronic τ Decays with ATLAS. Technical Report ATL-COM-PHYS-2008-068, CERN, Geneva, May 2008.

[17] D. Zerwas D. Froidevaux et al. Electron/Photon CSC Chapter. Technical Report ATL-COM-PHYS-2008-110, CERN, Geneva, Aug 2008.

[18] H.A. Bethe and W. Heitler. *Proc. R. Soc. A*, 146:83, 1934.

[19] Y.S. Tsai. *Rev. Mod. Phys.*, 46:815, 1974.

[20] S. Chekanov et al. A ZEUS Next-to-Leading-Order QCD Analysis of Data on Deep Inelastic Scattering. *Phys. Rev.*, D67:012007, 2003.

[21] Collier Benedikt et al. *LHC Design Report V3.* CERN, Geneva, 2004.

[22] Collier Brüning, Olivier et al. *LHC Design Report V1.* CERN, Geneva, 2004.

[23] J. M. Campbell, J. W. Huston, and W. J. Stirling. Hard Interactions of Quarks and Gluons: a Primer for LHC Physics. 2006.

[24] The ATLAS Collaboration. The ATLAS Experiment. `http://atlas.ch/photos/events.html`, 08 Okt 2008.

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

[26] ATLAS Collaboration. *ATLAS Inner Detector: Technical Design Report, Volume I.* Technical Design Report ATLAS. CERN, Geneva, 1997.

[27] Norbert Wermes and G Hallewel. *ATLAS Pixel Detector: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1998.

[28] ATLAS Collaboration. *ATLAS Calorimeter Performance: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1996.

[29] ATLAS Collaboration. *ATLAS Muon Spectrometer: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1997. distribution.

[30] ATLAS Collaboration. *ATLAS Level-1 Trigger: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1998.

[31] Peter Jenni, Nessi, et al. *ATLAS High-Level Trigger, Data-Acquisition and Controls: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 2003.

[32] ATLAS Collaboration. *ATLAS Computing: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 2005. revised version submitted on 2005-06-20 16:33:46.

[33] C. Arnault et al. The Atlas Computing Workbook. 2007.

[34] Torbjorn Sjostrand, Stephen Mrenna, and Peter Skands. PYTHIA 6.4 Physics and Manual. *JHEP*, 05:026, 2006.

[35] Torbjorn Sjostrand. Monte Carlo Generators for the LHC (1/4). 2008.

[36] A. Dellacqua et al. GEANT-4: An Object Oriented Toolkit for Simulation in HEP. CERN-DRDC-94-29.

[37] K. Cranmer. The ATLAS Analysis Architecture. *Nucl. Phys. Proc. Suppl.*, 177-178:126–130, 2008.

[38] K Cranmer, A Farbin, and A Shibata. EventView - the Design Behind an Analysis Framework. Technical Report ATL-SOFT-PUB-2007-008. ATL-COM-SOFT-2007-012, CERN, Geneva, Sep 2007.

[39] R. Brun and F. Rademakers. ROOT: An Object Oriented Data Analysis Framework. *Nucl. Instrum. Meth.*, A389:81–86, 1997.

[40] M. Barisonzi. Top studies in FDR I and II. `https://indico.desy.de/getFile.py/access?contribId=6&resId=0&materialId=slides&confId=675`, 28 Jan 2008.

[41] S. Menke. ATLAS D-Meeting Calorimeter Cluster Treatment - Introduction. `http://indico.cern.ch/getFile.py/access?contribId=33&sessionId=1&resId=0&materialId=slides&confId=31799`, 2008.

[42] W Lampl, S Laplace, Lelas, et al. Calorimeter Clustering Algorithms: Description and Performance. Technical Report ATL-LARG-PUB-2008-002. ATL-COM-LARG-2008-003, CERN, Geneva, Apr 2008.

[43] T Cornelissen, M Elsing, I Gavrilenko, W Liebig, E Moyse, and A Salzburger. The new atlas track reconstruction (newt). *Journal of Physics: Conference Series*, 119(3):032014 (6pp), 2008.

[44] David Côté. TauDPDMaker. `https://twiki.cern.ch/twiki/bin/view/Atlas/TauDPDMaker`, date 27 Aug 2008 - 04:06:10.

[45] N. P. Konstantinidis, J. Drohan, Z. Maxa, et al. The Atlantis Event Visualisation Program for the ATLAS Experiment, 2005.

[46] Andreas Hocker et al. TMVA: Toolkit for Multivariate Data Analysis. *PoS*, ACAT:040, 2007.

[47] R Ohlsen L. Breiman, J.Friedman and C.Stone. Classification and Regression Trees. *Wadsworth*, 1984.

[48] Y.R. Quinlan. Symplifying Desicion Trees. *Int. J. Man-Machine Sudies*, 28:221, 1987.

[49] J.Frieman and B.E. Popescu. Predictive Learning via Rule Ensembles. *Technical Report, Statistics Department, Stanford University*, 2004.

[50] S Mehlhase and T C Petersen. A probability based approach to PID in the TRT Detector of ATLAS. Technical Report ATL-COM-INDET-2006-017, CERN, Geneva, Oct 2006.

Appendix A

Additional Plots

A.1 Theoretical Overview

A.1.1 Decay of the τ Lepton

Additional plots to chapter 2.3.1:

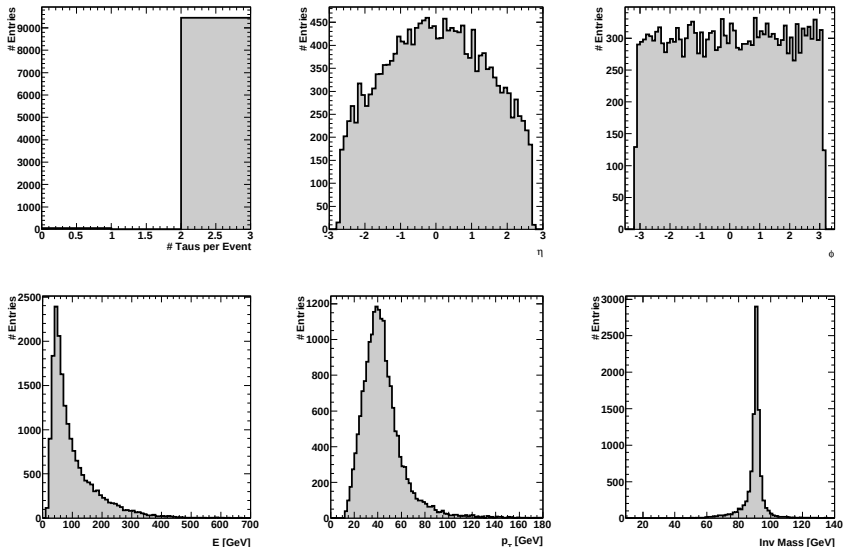

Figure A.1: *Kinematics of τ leptons from 10,000 simulated $Z^0 \rightarrow \tau\tau$ events. It is shown (1) the number of τ leptons per event, (2) η distribution of both leptons, (3) ϕ distribution, (4) energy of the τs, which can be larger than the Z mass, because the Z boson has not been produced in rest system, (5) transversal momentum (p_T) distribution, (6) invariant mass of the two τ shows a narrow peak, the Z boson mass ($m_Z^0 = 91.876\,\mathrm{GeV}$).*

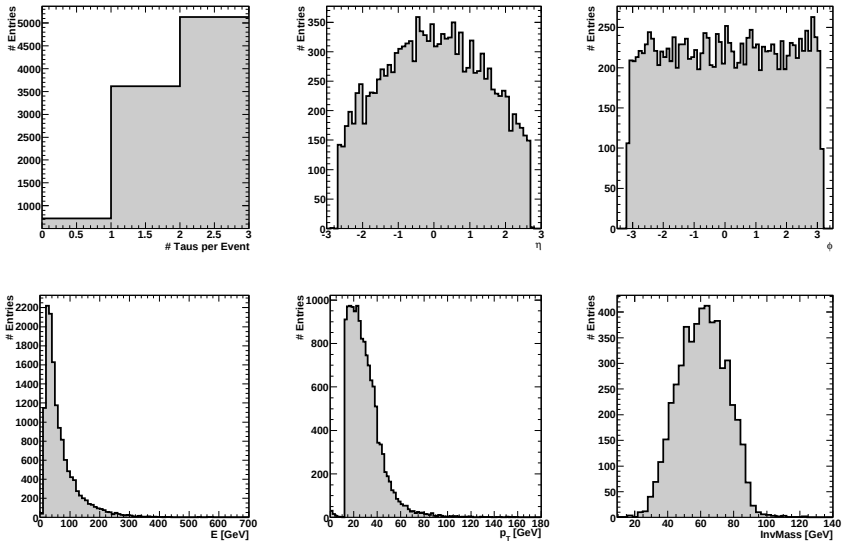

Figure A.2: *Kinematics of the τ leptons from simulated $Z^0 \rightarrow \tau\tau$ events. It is shown (1) the number of hadronically decayed τ leptons per event, (2) η distribution of the combined decay products of the τ, (3) ϕ distribution, (4) energy distribution of the combined τ decay products, which is smaller than the total energy distribution of the τs, (5) transversal momentum (p_T) distribution, (6) invariant mass of the hadronically decayed τs shows a much broader peak than the invariant mass the τ-leptons this is because of the loss of the energy borne away by the neutrinos.*

A.1.2 Photon Conversions

Additional plots to chapter 2.3.2:

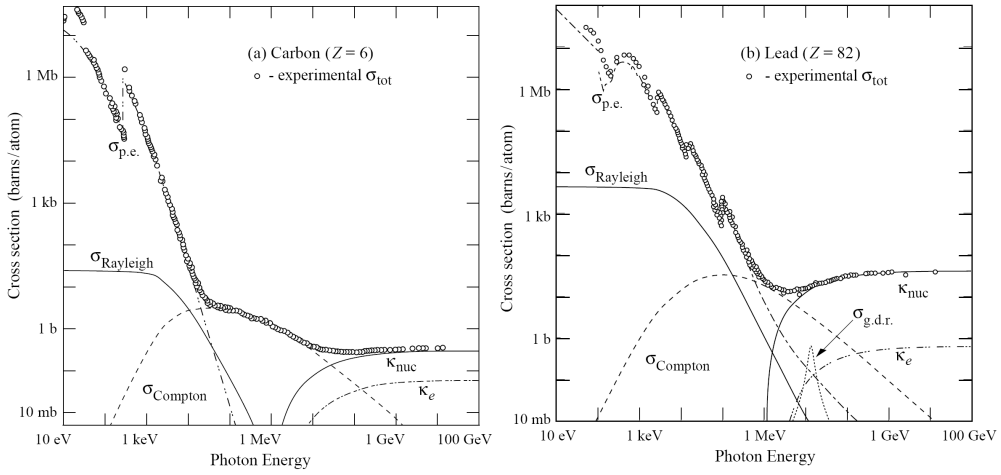

Figure A.3: *Shown are the total photon cross sections as function of the energy in a) carbon and b) lead for different processes: $\sigma_{p.e.}$ atomic photoelectric effect (photon absorption and electron emission); $\sigma_{Rayleigh}$ Rayleigh (coherent) scattering (atom neither ionized nor excited); $\sigma_{Compton}$ incoherent compton scattering; κ_{nuc} pair production in the nuclear field; κ_e pair production in the electromagnetic field; $\sigma_{g.d.r.}$ photonnuclear interactions (Giant Dipol Resonance).*

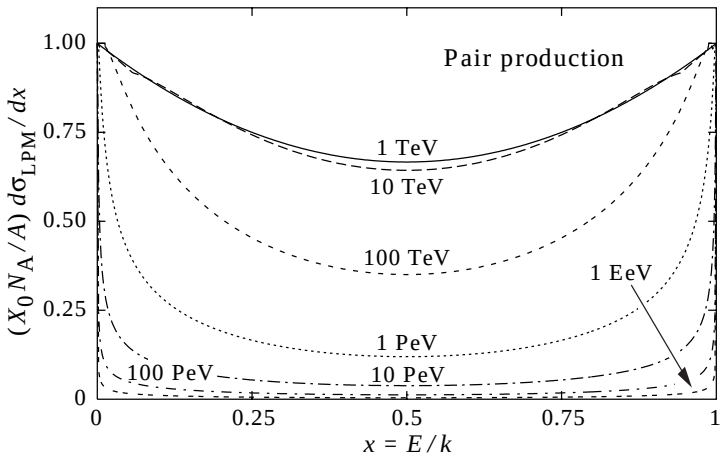

Figure A.4: *The normalized photon conversion cross section $d\sigma^{LPM}/dy$ versus fractional electron energy $x = E/k$.*

A.2 The Reconstruction Algorithms

A.2.1 The τ Reconstruction

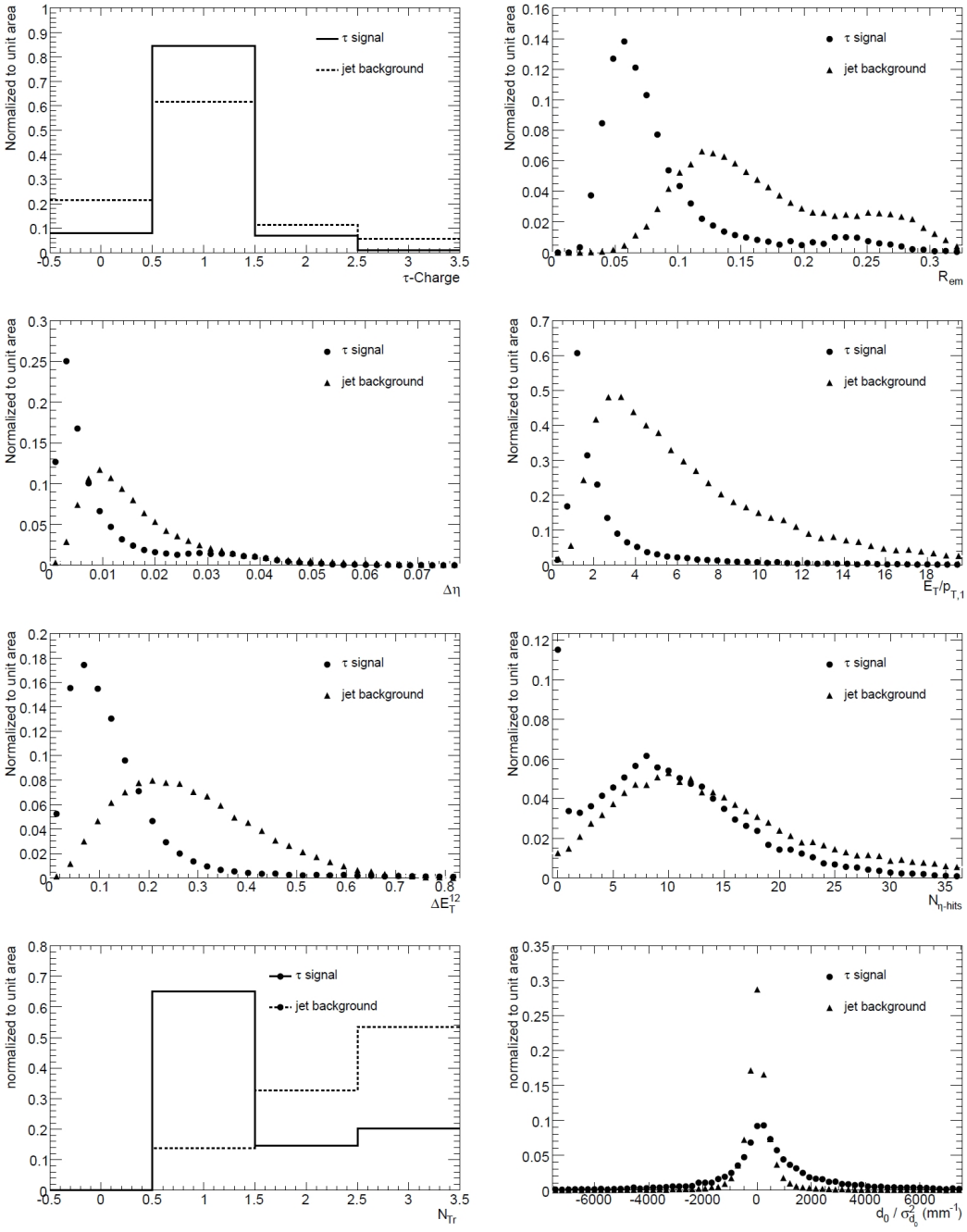

Figure A.5: *All likelihood input variables from TauCSC note [16].*

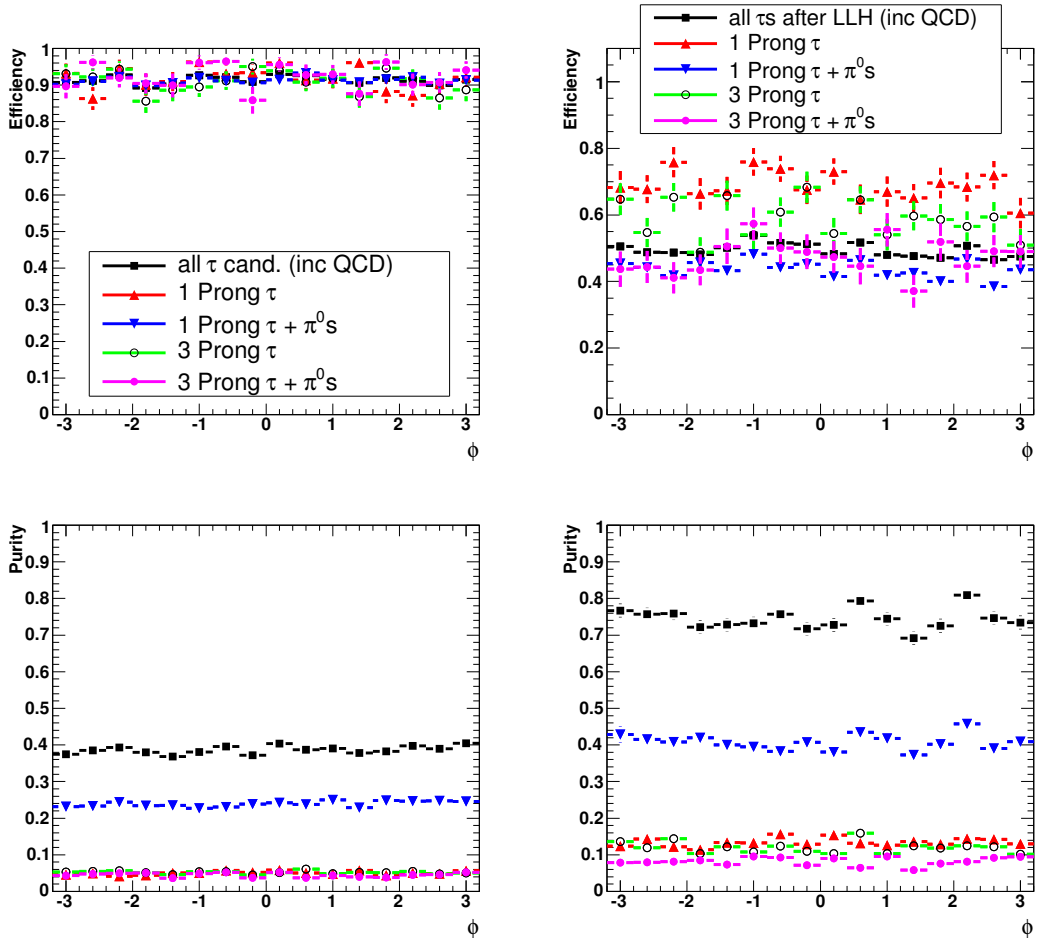

Figure A.6: *Efficiency and purity in ϕ of τ candidates in the left column and τs after likelihood cut in the right column, for several decay modes*

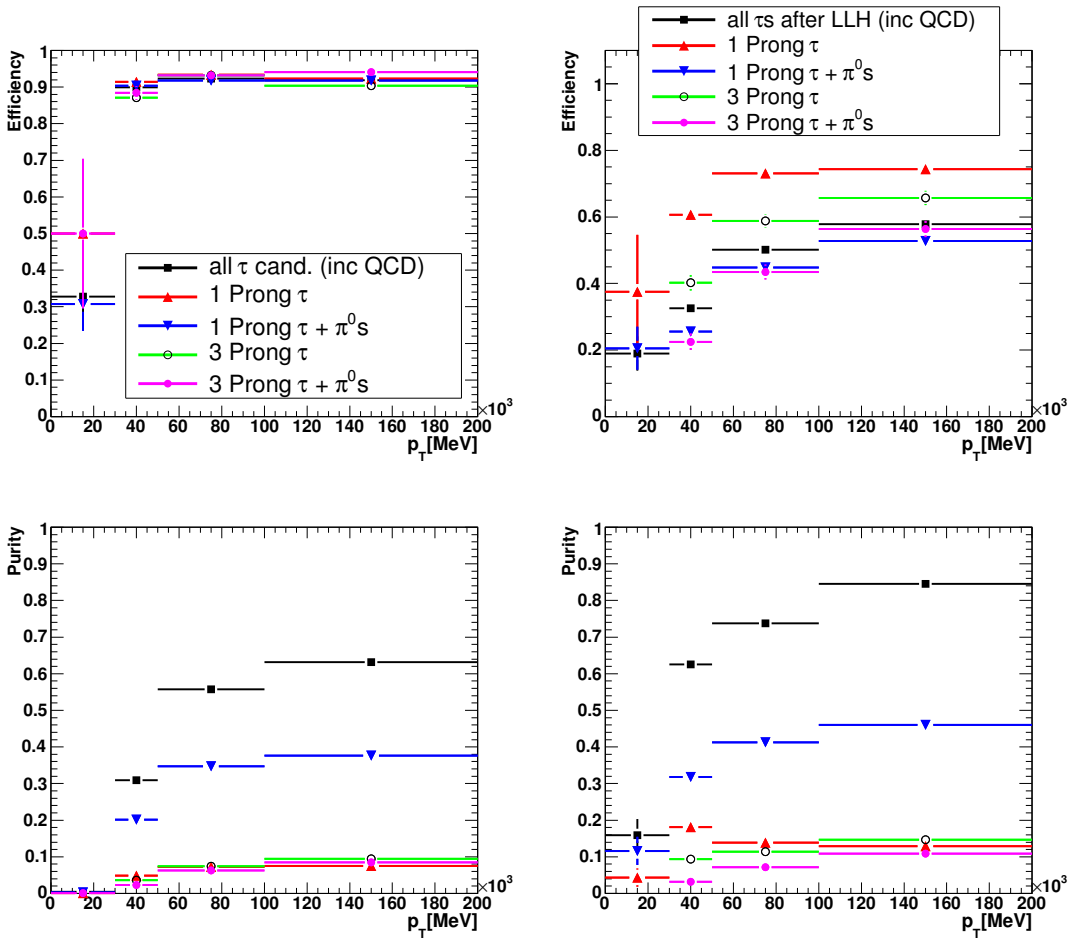

Figure A.7: *Efficiency and purity in p_T of τ candidates in the left column and τs after likelihood cut in the right column, for several decay modes*

A.2.2 Reconstruction and Identification of Photon Conversions

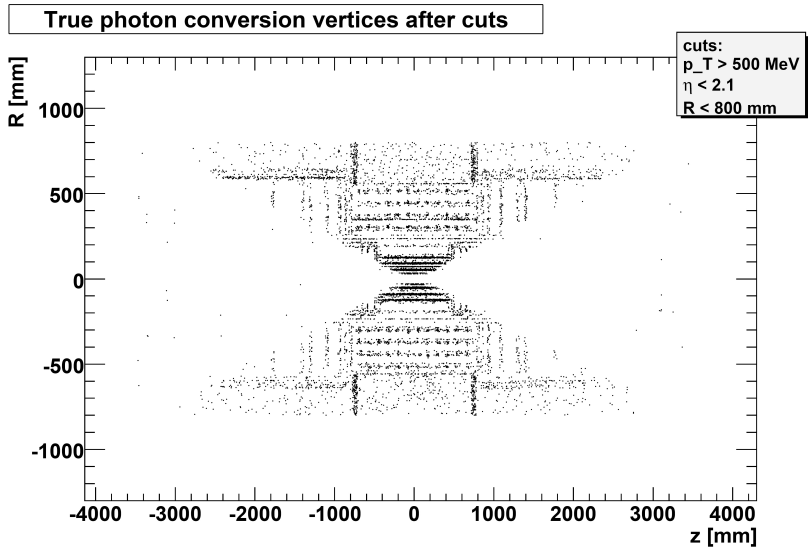

Figure A.8: *Truth vertices of photon conversions which are reconstructable. Due to a lack of TRT elements for a pseudorapidity $|\eta| > 2.1$, the photon conversion reconstruction algorithm has a cut of at $|\eta| = 2.1$. Only tracks with hits within the Inner Detector ($R < 800\,\mathrm{mm}$) and with a transverse momentum $p_T > 500\,\mathrm{MeV}$ are reconstructable. The shown vertcies have passed these cuts. These fraction of the total simulated truth photon conversions are used to calculate the efficinecy and purity of the photon conversion algorithm.*

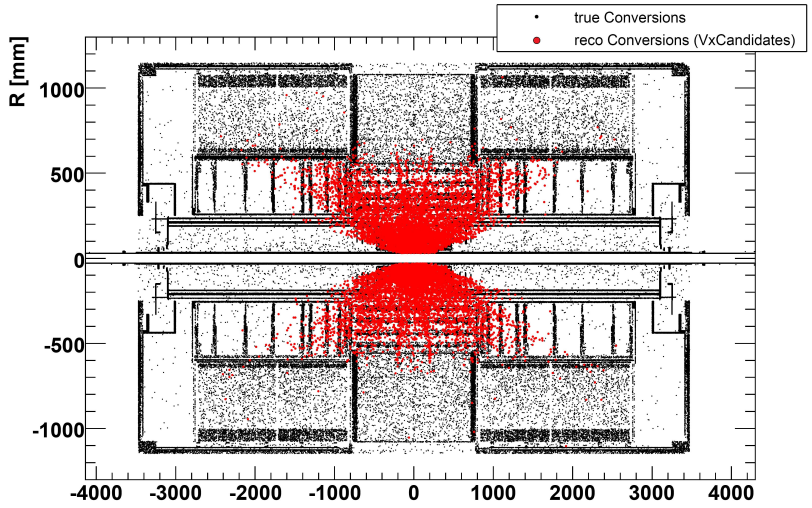

Figure A.9: *All vertices of truth photon conversions and the reconstructed vertices by the InDetConversionFinder tool, the seeds for photon conversions, are shown.*

In Chap. 5.2 the following cuts have been described. These quality cuts are applied during the reconstruction of photon conversions (*VxCandidates*). For each reconstruction step the cutflow table is listed below. All results are taken from [17]:

Track Selection

Table A.1: *Track selection cuts. Cumulative efficiency and rejection rates are presented.*

Cut	Efficiency	Rejection
No Cuts	0.7378	1.00
Impact d0	0.7334	1.16
Impact z0	0.7316	1.18
TR ratio	0.7119	2.12

Track Pair Selection

Table A.2: *List of cuts employed during the track pair selection for the three possible types of track pairs.The cumulative efficiency and rejection rate are presented. See text for the definition of the cut variables.*

Cut	Efficiency	Rejection
Polar Angle	0.7070	10.8
First Hit dR	0.7049	12.5
Min Distance	0.6994	16.3
dXY	0.6970	16.5
Vertex R	0.6959	16.6
Min Arc L	0.6935	40.3
Max Arc L	0.6890	111.6
dz	0.6870	111.9

Vertex Fitting

Table A.3: *Post-vertex fit selection cuts. Cumulative efficiency and rejection rates are presented.*

Cut	Efficiency	Rejection
Fit Convergence	0.6870	171.5
Fit χ^2	0.6710	288.9
Invariant Mass	0.6626	353.9
Photon p_T	0.6625	377.1

A.3 Developement of Specific Software Tools

EVUDVertex Tool

Table A.4 shows most of the variables which are stored in the D^3PDs by using the *TauDPDMaker* calling the *EVUDVertex* class. Naturally, the other newly implemented *EVUD* classes store addtional variables in the D^3PD, too. Due to the importance of photon conversions for this thesis, the variables provided to describe the conversions have been chosen as an example.

Table A.4: *List of informations of conversion variables provided by EVUDVertex class.*

Conversion Vertex Informations	**Truthmatched Informations**
Conv NumVertices	Conv isConv
Conv Vtx x	Conv is elec pair
Conv Vtx y	Conv Vtx truth barcode
Conv Vtx z	Conv Vtx xtruth
Conv Vtx sigx	Conv Vtx ytruth
Conv Vtx sigy	Conv Vtx ztruth
Conv Vtx sigz	Conv Trk truth PDGID
Conv Vtx chi2	**TRT PID Variables**
Conv Vtx pt	Conv Trk eProbComb
Conv Vtx ndf	Conv Trk eProbHT
Conv Vtx chi2prob	Conv Trk eProbToT
Conv Vtx numTracks	Conv Trk truth PDGID
Conversion Track Informations	**Variables for ECal Identification**
Conv Trk chi2	Conv Trk onlyTRT
Conv Trk d0	Conv Trk ECal is
Conv Trk sigd0	Conv Trk ECal dEta
Conv Trk z0	Conv Trk ECal dPhi
Conv Trk sigz0	Conv Trk ECal E1
Conv Trk phi	Conv Trk ECal E2
Conv Trk sigphi	Conv Trk ECal E3
Conv Trk theta	
Conv Trk sigtheta	
Conv Trk qOverP	
Conv Trk sigqOverP	
Conv Trk eta	
Conv Trk px	
Conv Trk py	
Conv Trk pz	
Conv Trk Index	

A.4 Improvements of Photon Conversion Reconstruction in τ Environment

A.4.1 Contant of reconstructed τ cone

Table A.5: *Total number and fraction of track IDs of τ candidates with and without Track Selection Criteria.*

	with TSC			no TSC		
truth ID	entries	fraction		entries	fraction	
π^+	25596	36.4%		38855	29.2%	
π^-	26635	36.5%	72.9%	38860	29.2%	58.4%
e^-	2877	4.1%		17074	12.8%	
e^+	2868	4.1%	8.2%	16757	12.6%	25.4%
K^+	4039	5.7%		5194	3.9%	
K^-	3772	5.4%	11.1%	4787	3.6%	7.5%
μ^-	833	1.2%		982	0.7%	
μ^+	834	1.2%	2.4%	989	0.7%	1.4%
not matched	4	$5 \cdot 10^{-5}$	-	1748	1.3 %	1.3%
p^+ remnant	3795	5.4%	5.4%	7742	5.8 %	5.8 %
total:	70252			132989		

A.4.2 Comparision of Event Displays of Events with Photon Conversions

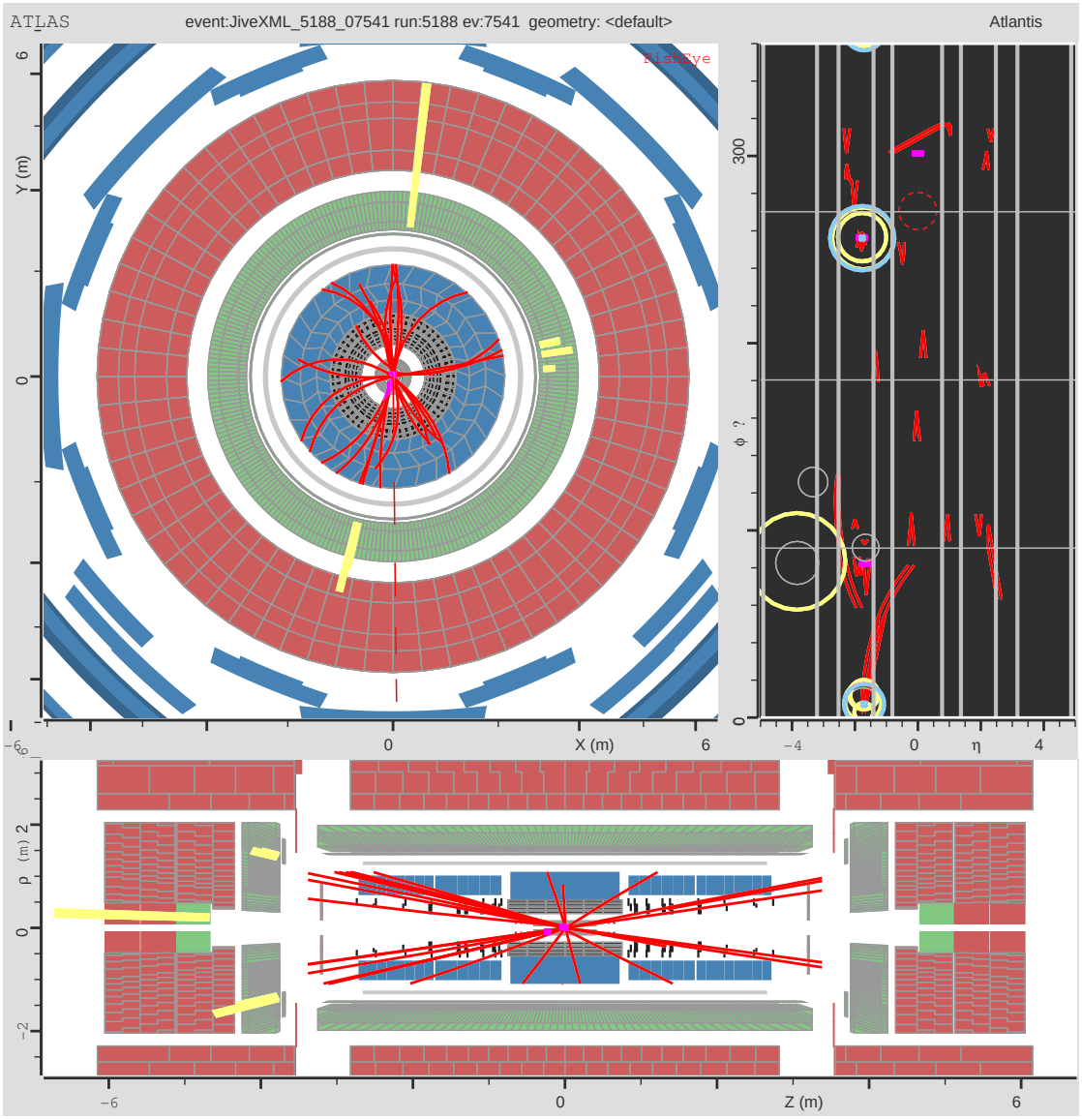

Figure A.10: *Event Display of photon conversion produced in a τ decay from a $Z^0 \rightarrow \tau\tau$ event produced by ATLANTIS. The upper left plot shows the X-Y plain, the upper right plot the η-ϕ plain and the lower plot the Z-ρ plain of the ATLAS detector. The τ jets are marked in yellow. The photon conversion takes place in the Pixel Detector marked in magenta. The τs deposit their energies in the electromagnetic calorimeter (green) and in the hadronic calorimeter (red). In blue the muon chambers are displayed. Much more tracks as in case of the $H \rightarrow \gamma\gamma$ event are left in the detector. Thus the correct identifications of the electron and positron tracks of the photon conversion are much more complex.*

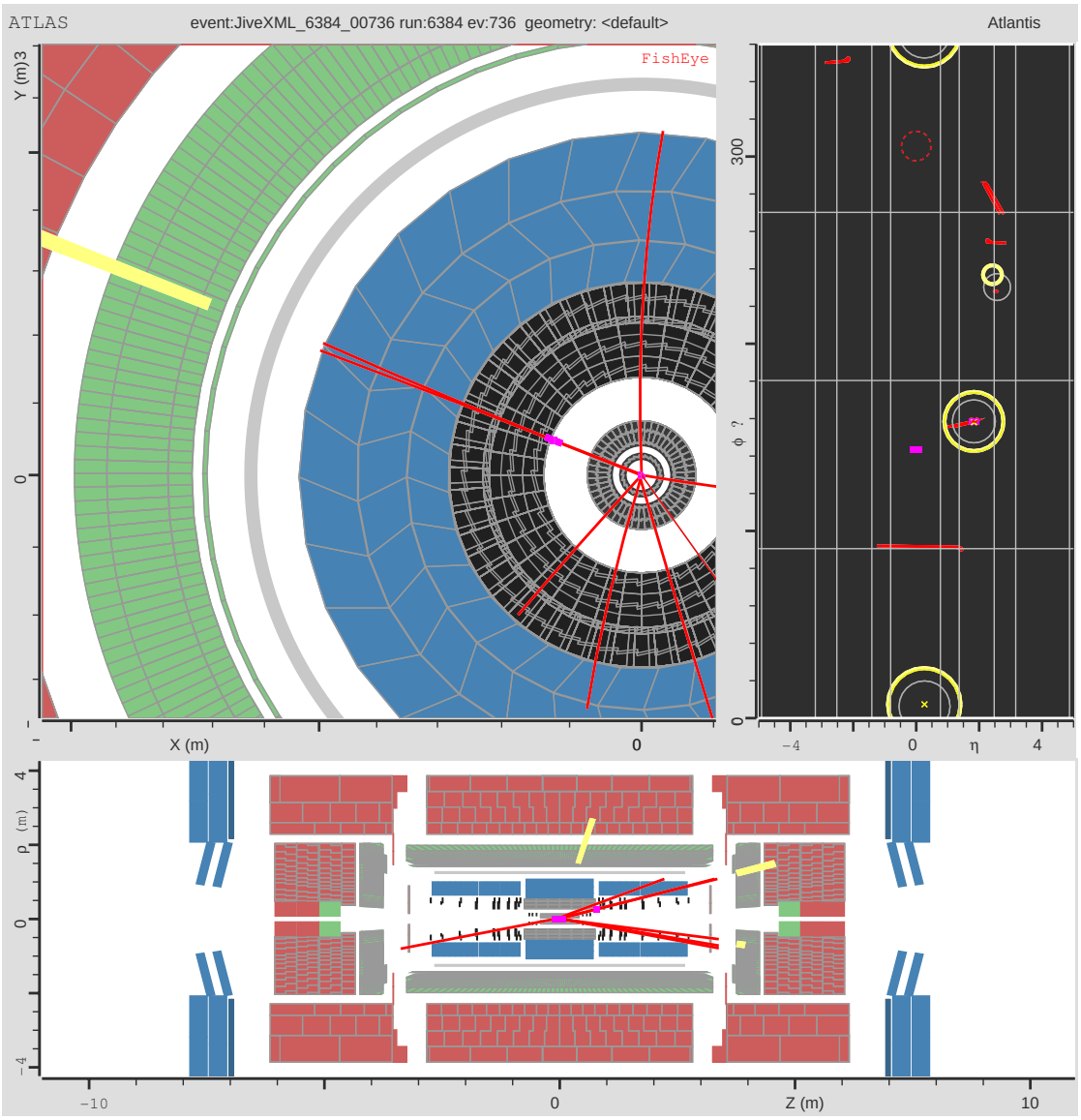

Figure A.11: *Event Display of a photon conversion in a $H \rightarrow \gamma\gamma$ event produced by ATLANTIS. The upper left plot shows the X-Y plain, the upper right plot the η-ϕ plain and the lower plot the Z-ρ plain of the ATLAS detector. The yellow circles in the η-ϕ plain assign reconstructed photon objects. The photon in the middle shows an additionally magenta dot which marks a vertex candidate. In this case it is the reconstructed vertex of a photon conversion. In the X-Y plain the photon conversion can be seen very clearly. In case of $H \rightarrow \gamma\gamma$ events the photons conversions can be isolated from other tracks much better than in the τ decay cone. It seems that at least one track of the photon conversion tracks comes from the interaction point (IP). This is due to the fact, that the event informations used to built these displays cannot use the specific track selections used for conversion tracks and therefore assumes that every track comes from IP.*

A.4.3 Electron Identification with TRT

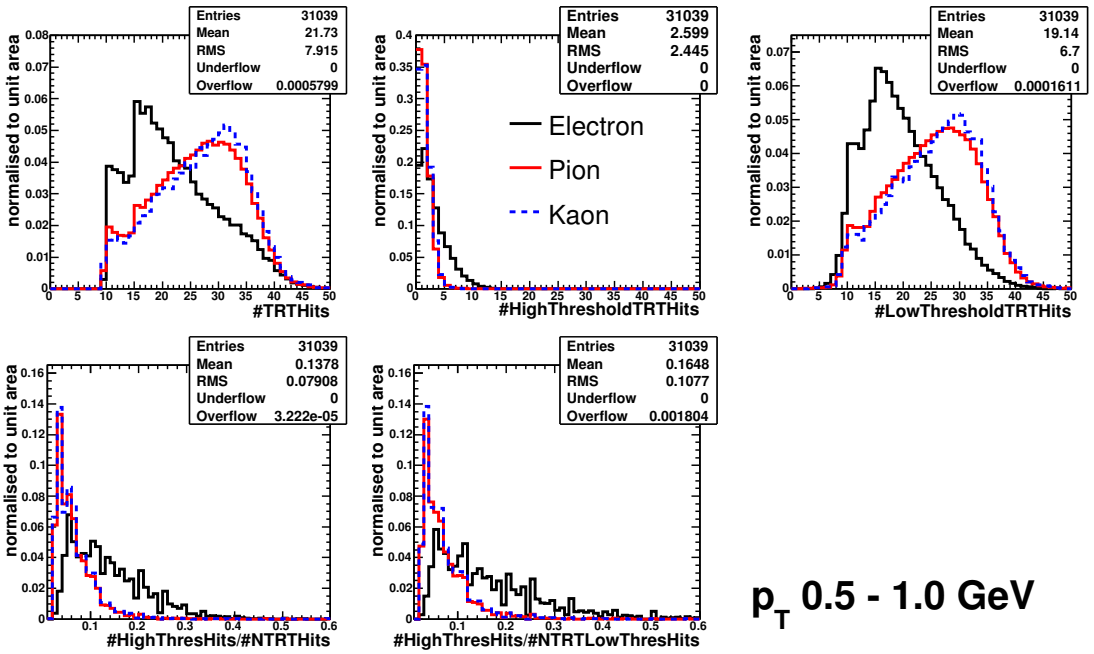

Figure A.12: *Several histograms of TRT variables ($0.5 < p_T < 1.0$ GeV). All histograms show TRT variables for electrons pions and kaons. From upper left to lower right: number of total TRT hits, number of TRT hits above high threshold, number of low threshold TRT hits, ratio of high threshold hits over total and ratio of high threshold hits over low threshold hits for each track. The statistic box at the top of each plot shows the values for electrons.*

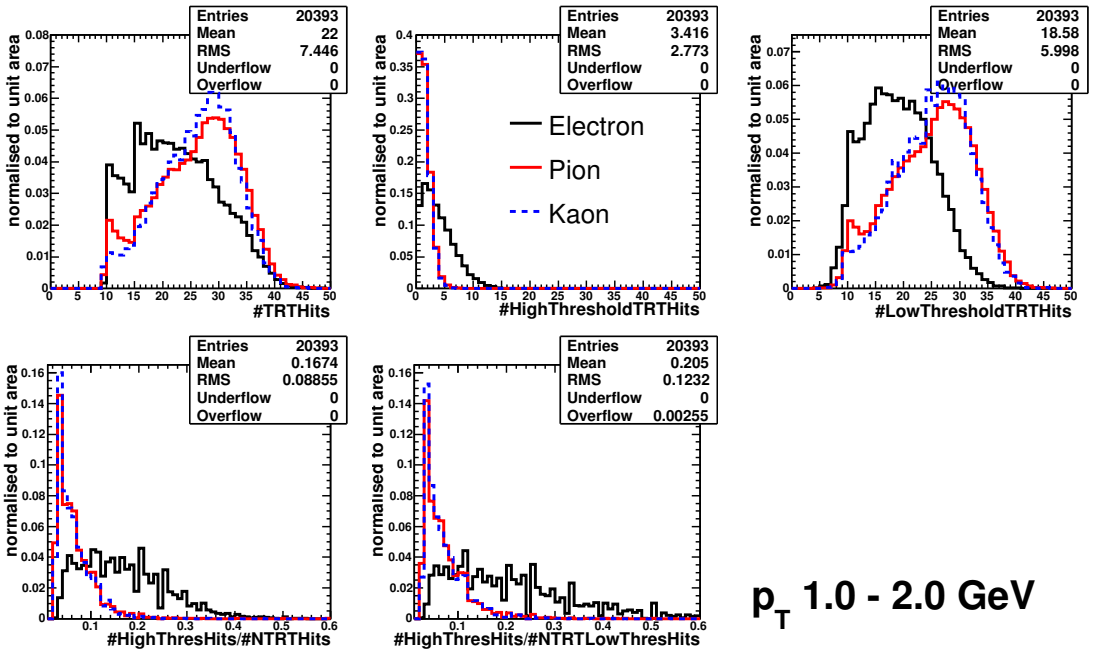

Figure A.13: *Several histograms of TRT variables (1.0< p_T < 2.0 GeV).*

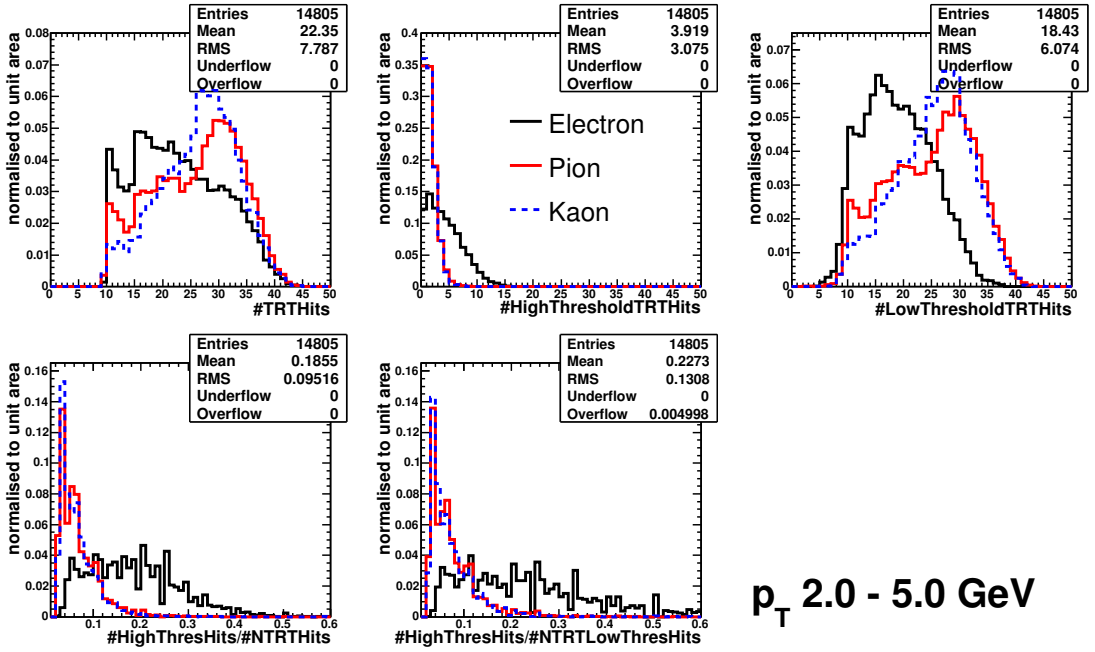

Figure A.14: *Several histograms of TRT variables (2.0< p_T < 5.0 GeV).*

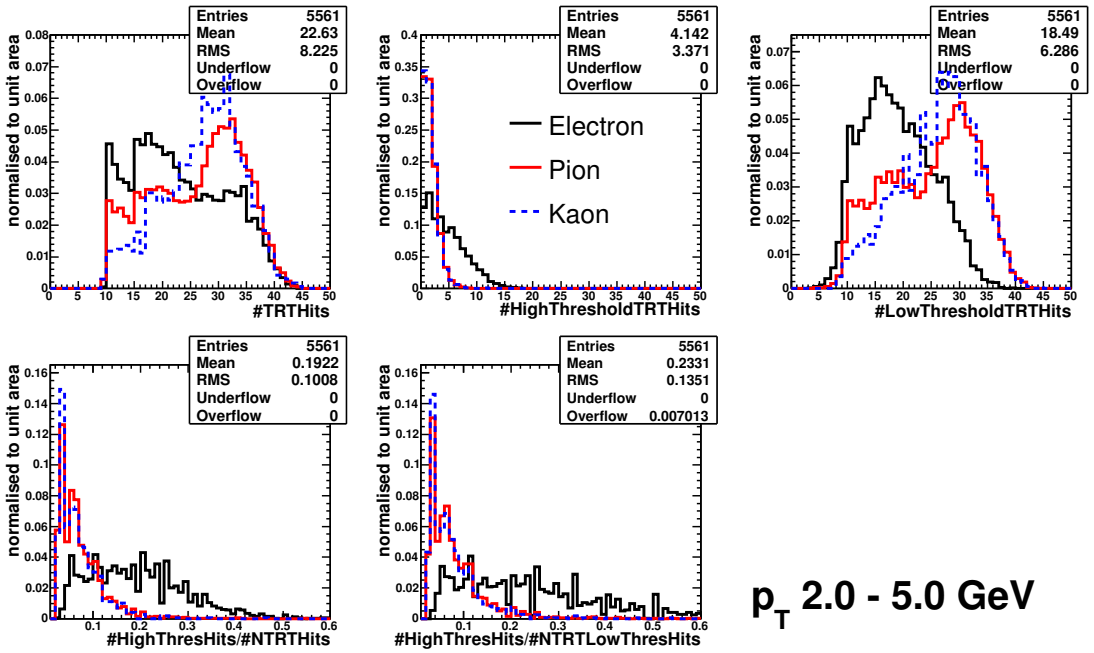

Figure A.15: *Several histograms of TRT variables (5.0< p_T < 10.0 GeV).*

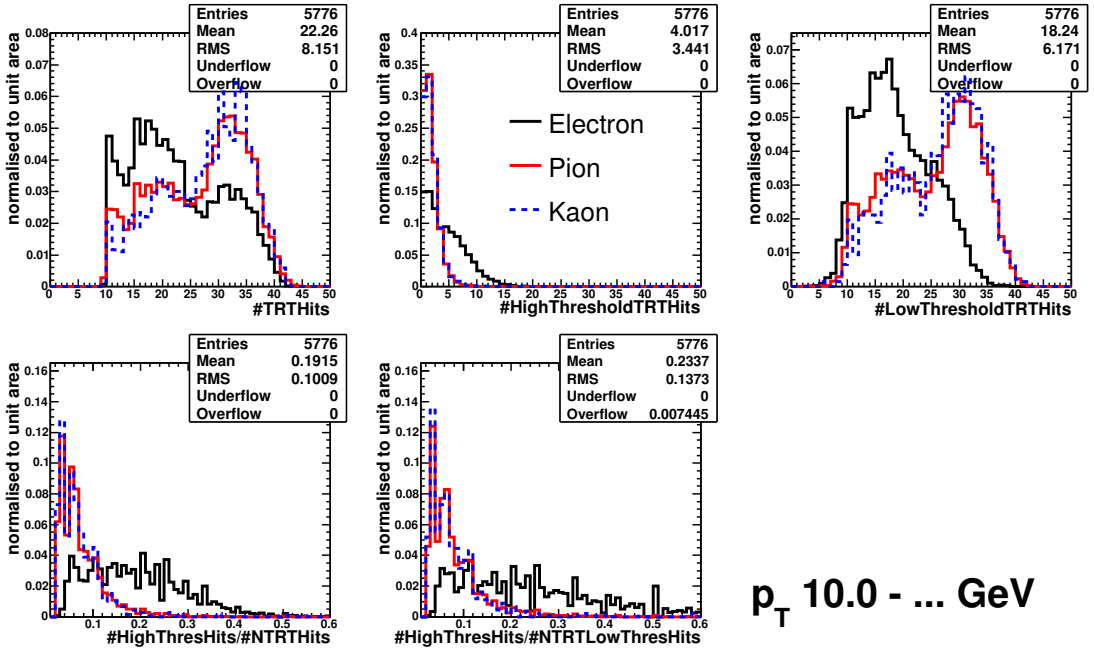

Figure A.16: *Several histograms of TRT variables (p_T > 10.0 GeV).*

A.4.4 Multivariate Analysis for Discrimination

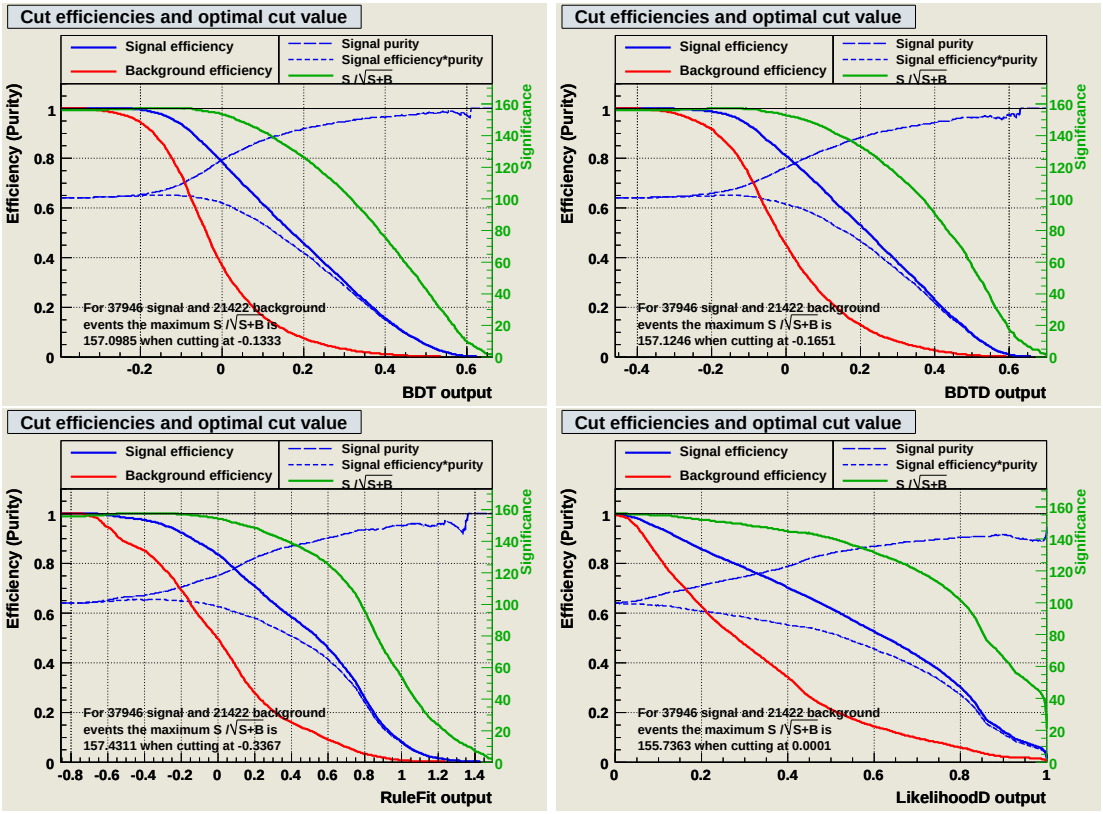

Figure A.17: *These plots contain a lot of informations. Most important information, if the classifier response is directly used as discrimination variable, is the peak of the significance, which shows the best cut value to distinguish signal from background. Additionally the signal purity, the signal and background efficiency and the signal purity times the signal efficiency are plotted. The results of the same classifier as in fig. 7.9 are pictured. As the BDTD classifier has been chosen for further studies only this plots is briefly discussed (upper right plot): The significance plot proposes a BDT response cut at -0.165. Due to the constraint that for vetoing conversion tracks in τ candidates a very high purity (> 90 %) is needed a cut on the BDT response larger than 0.22 has to be required. Nevertheless to get a value much better comparable with the Inner Detector PID the BDT response has to be translated into an electron probability.*

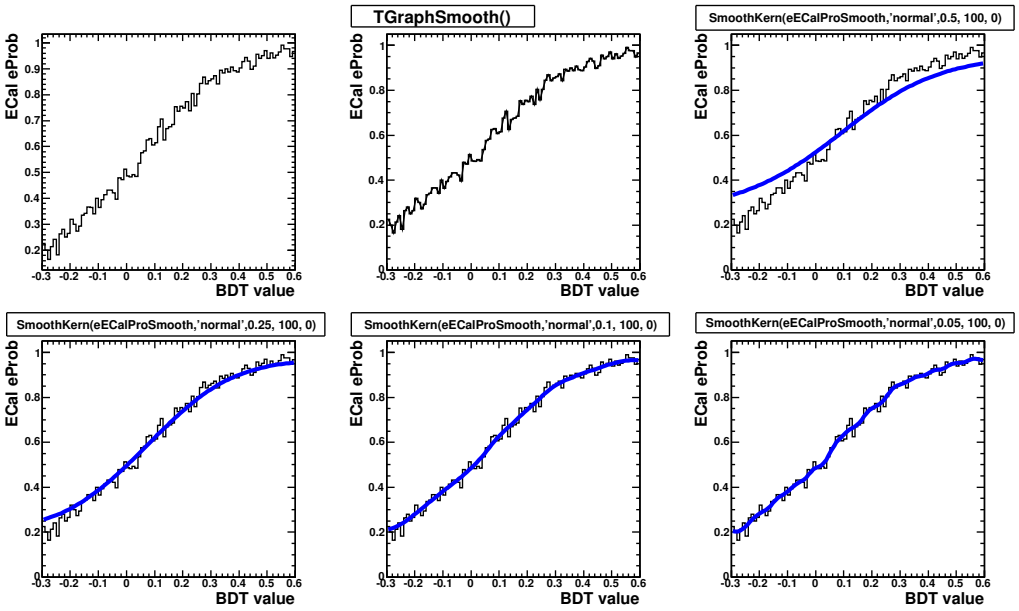

Figure A.18: *Smoothing the ECal eProb histogram for classification. To avoid fluctuations in the ECal eProb distribution depending on the bin size of the "look up" histogram, as described in Chap. 7.4.4, it has to be smoothed. The upper left plot shows the binned "look up" distribution. Upper middle plot is the result of using a standard ROOT TGraphSmooth method. To get a reasonable distribution a so-called ROOT SmoothKern method has to be used. The results of with several input variables are shown in the other four histograms.*

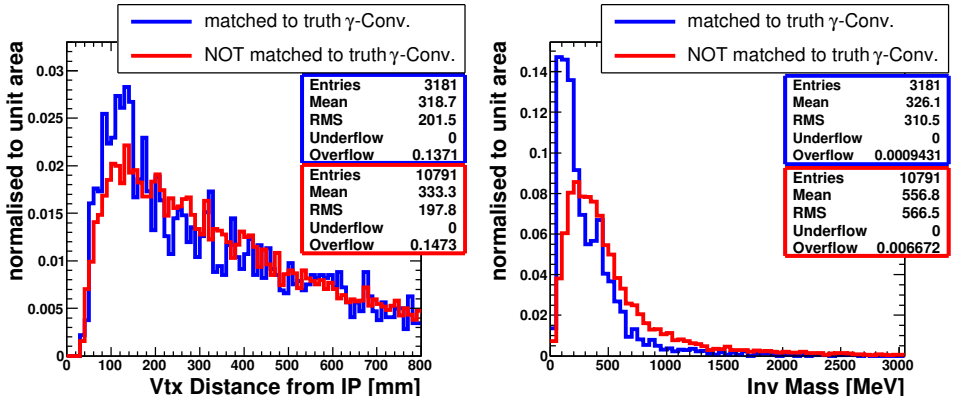

Figure A.19: *To decide which VxCandidate should be kept if finding several VxCandidates build with the same tracks additional discrimination parameters have to be used. Therefore many parameters have been tested. In the end the discrimination parameter is the unfitted invariant mass of the track pair. Are more than two VxCandidates found built with the same tracks the VxCandidate with the lowest invariant mass is kept. The left plot shows the distance of the secondary vertex from interaction point.*

Table A.6: *Cut flow table of VxCandidate cleaning during optimisation.*

Cuts	Eff.	Pur.							
no cuts									
no cuts	34.4	47.4							
preselection cuts									
within a τ cone $\Delta R < 0.3$	36.1	58.7							
chi2Prob > 0.00001	29.0	67.9							
cuts	0.01		0.05		0.10		0.15		0.2	
eProb	29.0	67.9	29.0	67.9	28.7	71.2	28.1	74.4	27.7	76.4
ECal eProb	29.0	67.9	29.0	67.9	29.0	67.9	29.0	67.9	28.9	68.4
eProb + ECal eProb	29.0	67.9	29.0	67.9	28.7	71.2	28.1	74.4	27.6	76.8
(eProb · ECal eProb)$_{Trk}$	29.0	67.9	28.8	71.8	28.1	76.8	27.4	80.7	26.1	84.1
(eProb · ECal eProb)$_{Trk_1}$ · (eProb · ECal eProb)$_{Trk_2}$	29.0	70.0	28.2	78.4	26.3	83.9	23.9	86.6	21.1	89.0
(eProb · ECal eProb) + sup. comb.	21.2	67.3	21.1	71.3	20.6	76.7	20.0	81.1	19.1	84.9
cuts	0.25		0.30		0.35		0.40		0.45	
eProb	27.3	77.9	27.0	78.8	26.7	79.9	26.4	80.8	21.4	85.4
ECal eProb	28.4	70.3	27.8	71.8	26.8	73.2	25.5	75.1	23.1	78.2
eProb + ECal eProb	26.8	79.5	26.0	81.4	24.8	83.4	23.3	85.3	16.9	89.5
(eProb · ECal eProb)$_{Trk}$	24.6	86.7	22.8	89.3	21.0	91.2	17.7	91.8	15.5	92.8
(eProb · ECal eProb)$_{Trk_1}$ · (eProb · ECal eProb)$_{Trk_2}$	19.4	90.8	17.4	92.5	15.7	93.2	14.1	94.1	12.7	94.8
(eProb · ECal eProb) + sup. comb.	17.9	87.7	16.5	90.1	14.9	92.0	12.5	92.7	10.8	93.7
cuts	0.5		0.55		0.60		0.65		0.70	
eProb	21.0	86.4	20.4	88.1	19.9	89.2	19.4	89.9	18.8	90.4
ECal eProb	21.2	80.8	19.9	83.1	18.5	84.6	16.5	86.7	14.2	88.5
eProb + ECal eProb	15.3	90.9	14.1	92.7	12.7	93.7	11.2	94.8	9.5	95.6
(eProb · ECal eProb)$_{Trk}$	14.0	93.8	12.9	94.6	11.5	95.1	10.0	95.7	8.6	96.7
(eProb · ECal eProb)$_{Trk_1}$ · (eProb · ECal eProb)$_{Trk_2}$	11.5	95.1	10.1	95.8	8.8	96.8	7.6	97.4	6.1	97.7
(eProb · ECal eProb) + sup. comb.	9.6	95.0	8.8	95.8	7.8	96.5	6.9	97.1	6.0	97.5
cuts	0.75		0.80		0.85		0.90		0.95	
eProb	18.0	91.2	17.2	92.1	16.3	93.0	15.2	94.0	13.3	95.1
ECal eProb	12.3	90.5	10.0	92.6	7.7	93.8	4.1	95.4	0.7	100
eProb + ECal eProb	7.8	96.3	6.1	97.4	4.4	97.7	2.1	98.3	0.4	100
(eProb · ECal eProb)$_{Trk}$	7.0	97.1	5.4	97.8	3.7	98.3	1.7	98.0	0.3	100
(eProb · ECal eProb)$_{Trk_1}$ · (eProb · ECal eProb)$_{Trk_2}$	4.7	98.2	3.4	98.0	1.7	97.9	0.5	98.7	0	-
(eProb · ECal eProb) + sup. comb.	4.9	97.5	3.8	97.8	2.7	98.3	1.3	98.2	0.2	100

Table A.7: *Truth ID of VxCandidates in the Tau Cone. Compared with the electron purity of 60 % of the unpurified VxCandidates the purity decreases. This is due to the large number of charged pions within the τ decay environment. Thus the conversion reconstruction tool picks up more pions as conversion tracks and the purity decreases. To achieve a sufficient photon conversion electron purity additional identification criteria are needed.*

Truth ID	Entries	Fraction	Combined
e^-	1566 ± 39.57	24.72 ± 0.62	
e^+	1593 ± 39.91	25.15 ± 0.63	49.87 ± 1.25
π^+	1340 ± 36.61	21.16 ± 0.58	
π^-	1436 ± 37.89	22.67 ± 0.60	43.83 ± 1.18
K^+	182 ± 13.49	2.87 ± 0.21	
K^-	217 ± 14.73	3.43 ± 0.23	6.30 ± 0.45
total:	6334		

Table A.8: *Truth ID of VxCandidates in TauCone after applying a cut on eProb > 0.9 and the mentioned preselection cuts. The electron purity of the VxCandiates within the τ cone is slightly worse than the electron purity of all identified conversions (94%), but due to the much higher pion background within the τ lepton decay environment the results are very promising.*

Truth ID	Entries	Fraction	Combined
e^-	858 ± 29.29	45.04 ± 1.54	
e^+	847 ± 29.10	44.46 ± 1.53	89.50 ± 3.07
π^+	78 ± 8.83	4.09 ± 0.46	
π^-	96 ± 9.80	5.04 ± 0.51	9.13 ± 0.98
K^+	9 ± 3.00	0.47 ± 0.16	
K^-	9 ± 3.00	0.47 ± 0.16	0.94 ± 0.31
total:	1905		

Table A.9: *Table of the track multiplicity of τ candidates before and after conversion veto without any applied track selection criteria. Upper part lists the total number of entries of 10k $Z^0 \rightarrow \tau\tau$ events. The lower part shows the total fraction of a specific decay mode with respect to the number of reconstructed tracks of all τ candidates.*

Track Multiplicity:	0	1	2	3	4	5	more	total
1 Prong	69	4648	2464	1838	1120	754	1247	12140
1 Prong after veto	152	5669	2734	1619	837	455	674	12140
3 Prong	7	522	407	1373	734	404	636	4083
3 Prong after veto	24	605	423	1436	729	372	494	4083
QCD backg.	508	1265	2030	2379	2345	1918	5921	16366
QCD backg. after veto	602	1419	2216	2556	2402	1958	5213	16366
1 Prong	0.21	14.26	7.56	5.64	3.44	2.31	3.83	37.25
1 Prong after veto	0.47	17.40	8.39	4.97	2.57	1.40	2.07	37.25
3 Prong	0.02	1.60	1.25	4.21	2.25	1.24	1.95	12.53
3 Prong after veto	0.07	1.86	1.30	4.41	2.24	1.14	1.52	12.53
QCD backg.	1.56	3.88	6.23	7.30	7.20	5.89	18.17	50.22
QCD backg. after veto	1.85	4.35	6.80	7.84	7.37	6.01	16.00	50.22

Table A.10: *Table of the track multiplicity of τ candidates before and after conversion veto with applied track selection criteria. Upper part lists the total number of entries of 10k $Z^0 \rightarrow \tau\tau$ events. The lower part shows the total fraction of a specific decay mode with respect to the number of reconstructed tracks of all τ candidates.*

Track Multiplicity:	0	1	2	3	4	5	> 5	total
1 Prong	593	9202	1471	702	124	29	19	12140
1 Prong after veto	774	9505	1252	493	85	14	17	12140
3 Prong	39	992	771	1995	209	58	19	4083
3 Prong after veto	72	982	779	2002	186	43	19	4083
QCD backg.	1683	3318	3800	2892	1902	1163	1608	16366
QCD backg. after veto	1734	3358	3805	2893	1897	1159	1520	16366
1 Prong	1.82	28.24	4.51	2.15	0.38	0.09	0.06	37.25
1 Prong after veto	2.38	29.17	3.84	1.51	0.26	0.04	0.05	37.25
3 Prong	0.12	3.04	2.37	6.12	0.64	0.18	0.06	12.53
3 Prong after veto	0.22	3.01	2.39	6.14	0.57	0.13	0.06	12.53
QCD backg.	5.16	10.18	11.66	8.87	5.84	3.57	4.93	50.22
QCD backg. after veto	5.32	10.30	11.68	8.88	5.82	3.56	4.66	50.22

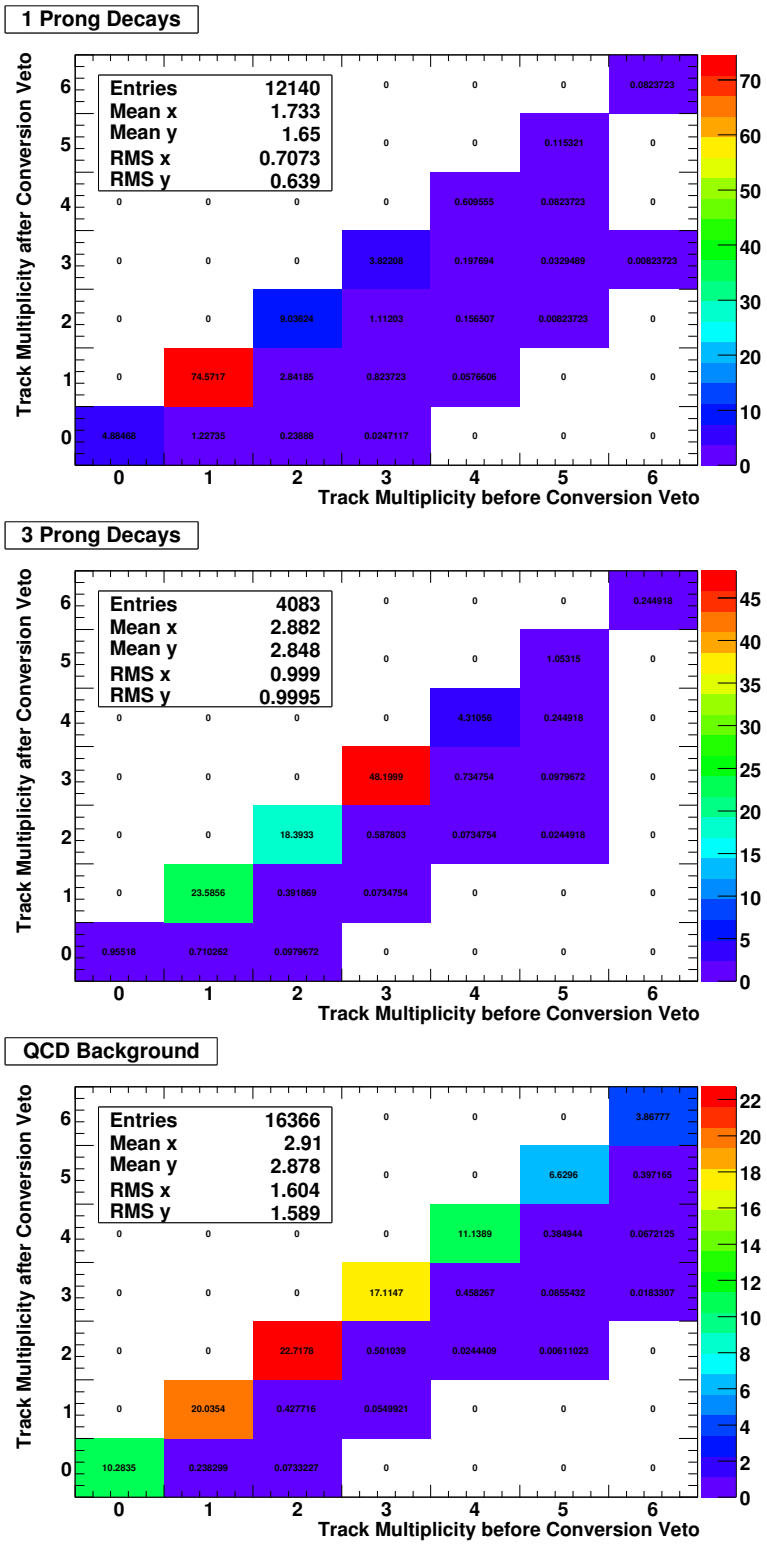

Figure A.20: *Fraction of tracks per τ candidate before and after explicit conversion veto.*

Danksagung

An dieser Stelle möchte ich mich gerne bei alle bedanken, die direkt und indirekt zum Ergebnis dieser Diplomarbeit beigetragen haben. An erster Stelle bedanke ich mich bei Dr. Philip Bechtle, der mir die Möglichkeit zu einer spannenden und herausfordernden Diplomarbeit bot und mir immer mit Rat und Tat zur Seite stand. Dr. David Côté möchte ich danken für dessen Hilfe bezüglich Softwareentwicklung, die ich gerne in Anspruch genommen habe, und dafür, dass er immer die Zeit fand, wann und wo auch immer, Probleme und Ideen meiner Arbeit mit mir zu erörtern.

Desweitern möchte ich mich bei Prof. Dr. Peter Schleper bedanken, der sich bereit erklärt hat, diese Arbeit als Zeitgutachter zu bewerten.

Ebenfalls möchte ich mich bei Dr. Karsten Köneke und Mark Terwort bedanken, die immer und immer wieder die speziellen Features von Photon Konversionen mit mir diskutieren durften.

Nochmals einen ganz herzlichen Dank an die gesamte ATLAS DESY Gruppe für die freundliche und angenehme Arbeitsatmosphäre und die konstruktiven Diskussionen aller Art.

Nicht zuletzt möchte ich meinen Eltern danken, dass sie es mir ermöglichten, Physik in Hamburg zu studieren. Danke.

In ganz besondere Weise danke ich Dörthe Ludwig, die ebenfalls zahlreiche, grundlegende Diskussionen zu jeder Tages- und Nachtzeit nicht gescheut hat und mich in jeder Situation unterstützt und zu mir gehalten hat. Vielen Dank!