

**Department of Physics and Astronomy  
University of Heidelberg**

Bachelor Thesis in Physics  
submitted by

**Martin Braß**

born in Speyer (Germany)

**2014**

**Estimating Hadron Contamination in Pb-Pb  
Collisions at  $\sqrt{s_{NN}} = 2.76$  TeV for Low Transversal  
Momenta with TPC, TOF and TRD of the ALICE  
Experiment**

This Bachelor Thesis has been carried out by Martin Braß at the  
Physikalisches Institut in Heidelberg  
under the supervision of  
Prof. Dr. Johanna Stachel

# Abstract

Since it is important for many HFE analyses to know the amount of contamination in an electron sample, it is the aim of this thesis to give an estimation of the hadronic background in such a sample taken from Pb-Pb collisions at  $\sqrt{S_{NN}} = 2.76$  TeV, which were performed at the LHC and measured with the ALICE detector. The focus is on the most central collisions with a centrality of 0 – 20% and on particle tracks with low transverse momentum. Particle identification is done by using the sub detectors TPC, TOF and TRD of ALICE. For this purpose, measured TPC  $dE/dx$  distributions of protons, kaons pions and deuterons are used as templates and are fitted to the  $dE/dx$  distribution of the analyzed electron sample, to obtain the amount of contamination.

Finally the dependence between hadronic contamination and momentum - transverse momentum respectively - is obtained as numerical table and alternatively as a smooth function for practical purposes.

In the first chapter a short introduction to the Quark-Gluon plasma is given. The second one describes the different sub detectors used in this thesis. While the third chapter explains the “weighted least squares method”, the fourth one focuses on how the analyzed data are selected. In chapter five, a routine to estimate the hadronic background for different (transverse) momenta is developed and the usage of the templates as energy loss distributions is explained. The results can be found in chapter six.

# Zusammenfassung

Da es für viele HFE Analysen wichtig ist, die Kontamination eines Datensatzes aus Elektronen zu kennen, befasst sich diese Arbeit mit einer Abschätzung des hadronischen Untergrundes in einem solchen Datensatz, der in Pb-Pb Kollisionen bei  $\sqrt{S_{NN}} = 2.76$  TeV im LHC entstanden ist und mit dem ALICE Detektor gemessen wurde. Es werden Kollisionen mit 0 – 20% Zentralität und Teilchen mit niedrigem Transversalimpuls untersucht. Um die Teilchen zu identifizieren, werden die Subdetektoren TPC, TOF und TRD von ALICE verwendet. Dazu werden gemessene TPC  $dE/dx$  Verteilungen - so genannte “Templates” - von Protonen, Kaonen und Deuteronen genutzt, um diese an die  $dE/dx$  Verteilung des Elektronen Datensatzes zu fitten und somit die Kontamination von diesem abzuschätzen.

Das Ergebnis ist eine Tabelle mit der hadronischen Kontamination als Funktion des Impulses, beziehungsweise des Transversalimpulses. Als Alternative gibt es dies - aus praktischen Gründen - auch als stetige Funktion.

Eine kurze Einführung zum Quark-Gluonen Plasma findet sich im ersten Kapitel. Im zweiten werden die in dieser Arbeit eingesetzten Subdetektoren beschrieben. Während sich das dritte Kapitel mit der Methode der “gewichteten kleinsten Quadrate” befasst, wird die Auswahl der zu analysierenden Daten in Kapitel vier behandelt. Im fünften Kapitel wird eine Routine entwickelt, um den hadronischen Hintergrund für verschiedene (Transversal-) Impulse zu bestimmen und der Einsatz der gemessenen “Templates” wird erklärt. Die Ergebnisse werden in Kapitel sechs präsentiert.

# Contents

<b>1</b>	<b>Introduction</b>	<b>5</b>
1.1	Quark-Gluon Plasma . . . . .	5
1.2	Heavy Quarks as Probes of the QGP . . . . .	6
<b>2</b>	<b>Particle Identification with Different Detectors</b>	<b>8</b>
2.1	Gaseous Ionization Detectors and Energy Loss of Traversing Particles . . . . .	9
2.2	Setup of the ALICE Detector . . . . .	11
2.3	Time Projection Chamber . . . . .	12
2.4	Time of Flight Detector . . . . .	15
2.5	Transition Radiation Detector . . . . .	17
2.6	Particle Identification with TRD . . . . .	20
<b>3</b>	<b>Fitting with Weighted Least Squares Method</b>	<b>21</b>
<b>4</b>	<b>Data Samples and Applied Cuts</b>	<b>23</b>
4.1	Centrality . . . . .	23
4.2	TOF-Cuts and Data Sample . . . . .	25
<b>5</b>	<b>Estimating Contamination with different Fit Models</b>	<b>27</b>
5.1	Fitting multiple Gaussians . . . . .	27
5.2	Proton Distribution . . . . .	28
5.3	Fitting with Experimental Data as Template . . . . .	30
5.4	Contamination of Templates . . . . .	33
5.5	Cut on Active Volume . . . . .	35
5.6	A Template for Pions . . . . .	36
5.7	Contamination of the Pion Template . . . . .	39
5.8	Contamination at Energy Loss Crossing Lines . . . . .	40
<b>6</b>	<b>Results</b>	<b>43</b>
6.1	Contamination in Dependence of Momentum . . . . .	43
6.1.1	Statistical and Systematic Uncertainties . . . . .	44
6.1.2	Total Contamination . . . . .	46
6.2	Contamination in Dependence of Transverse Momentum . . . . .	47
6.3	Obtaining Contamination as a Smooth Function . . . . .	49
<b>7</b>	<b>Summary and Outlook</b>	<b>53</b>

# 1 Introduction

A Large Ion Collider Experiment (ALICE) at the Large Hadron Collider at CERN is exploring a new state of matter at high energy densities in Pb-Pb collisions. This state is called Quark-Gluon plasma and is expected to have occurred shortly after the Big Bang. Therefore exploring it may give a better understanding of the evolution of our early universe. Furthermore, the studies will improve our understanding of Quantum Chromodynamics. There are a lot of effects that influence the behavior of the plasma, like initial scattering, fluctuations in the density, hadronization and many others. Therefore different observables are needed to study the Quark-Gluon plasma. One of those is measuring heavy flavor particles, since properties of charm and beauty quarks shed light on the properties of the plasma. One important approach is measuring electrons from heavy flavor decays for instance. For this purpose it is important for Heavy-Flavor-Electron (HFE) analyses to know how strong the measured electron candidates are contaminated by other particles. For example the Cocktail-Method relies on the subtraction of the background and so it is important to know the fraction of electron candidates that are not electrons.

It is the aim of this thesis to estimate the relative amount of this contamination at low transverse momentum. In the following section a short introduction to the Quark-Gluon Plasma and why it is interesting to observe heavy flavor electrons is given. Chapter 2 contains an introduction to the detectors important for the present work and the concepts of particle identification with these detectors. Chapter 3 describes the applied fitting method.

The way of obtaining data and the main analysis with all results are presented in sections 4 - 6.

## 1.1 Quark-Gluon Plasma

According to experimental observations quarks are bound in hadrons and cannot exist freely in vacuum. This is described by the potential of the strong interaction between

two color charges, which for large distances between two interacting partners increases linearly with distance. Therefore an infinite amount of energy would be required to separate those particles. If the energy of the system is increased, quark-antiquark pairs form to create several color-neutral particles, so that charge carriers cannot be separated.

However, lattice calculations from Quantum Chromo Dynamics (QCD) predict a phase transition for matter at high energy densities around  $1 \text{ GeVfm}^{-3}$  and a high temperature[1]. This transition leads to a strongly interacting state of matter which is called Quark-Gluon Plasma (QGP). In this state confinement becomes canceled out and quarks are able to move freely inside the plasma. Nevertheless, the whole plasma has to be color-neutral[2]. Besides deconfinement it is predicted that in the phase of plasma the chiral symmetry of the Lagrangian is restored[3].

The hot and dense matter is expected to be created in Pb-Pb collisions where it evolves over a very short time. It expands and by that cools down so that the energy density decreases. Then the former deconfined quarks hadronize to form color-neutral hadrons.

## 1.2 Heavy Quarks as Probes of the QGP

Since the Quark-Gluon Plasma is not measurable directly, one has to observe particles produced in the heavy ion collision that traverse the plasma and interact with it. Heavy flavor quarks are interesting probes. A possible way to access them experimentally is observing leptons, produced in heavy flavor decays. Charm and beauty quarks are due to their high masses produced almost exclusively in the initial hard scatterings, so that they experience the whole evolution of the QGP while traversing it. However, the top quark is not an appropriate candidate, since it has a very short lifetime and decays before it has traveled through the plasma over a significantly long distance[2]. Beauty and charm are expected to interact measurably with the medium so that properties of the QGP can be deduced - like the dependence of color charge and mass in partonic energy loss inside the plasma. For example this can be achieved by comparing the in-medium energy loss of heavy-flavor hadrons to hadrons with light quarks[16]. There is inelastic energy loss due to gluon radiation caused by traversing partons similar to bremsstrahlung and radiative energy loss. Also elastic loss of energy caused by collisions contributes to the total energy loss[17].

Furthermore, measuring the production cross section of heavy-flavor (HF) quarks is a good test for perturbative QCD (pQCD)[3]. Therefore observing heavy-flavor hadrons or particles produced by them may give different insights into the probability and behavior

of the Quark-Gluon Plasma. Of special interest for this analysis are electrons from heavy flavor decays since their branching ratios are of order 10% [16]. On the other hand there is a slight disadvantage of observing electrons: since neutrinos are not detectable, reconstructing a full invariant mass from leptonic HF decays is impossible. Nevertheless it is worth observing HF electrons due to their large branching ratios and the corresponding high statistics. For this purpose this analysis deals with estimating the hadronic contamination in electron samples.

## 2 Particle Identification with Different Detectors

There are mainly two ways to identify particles. They can be distinguished by their masses or by their difference in interaction with different sub-detector materials (calorimetry), which is usually already incorporated in the design of most detectors[9]. For calorimetry one typically arranges sub-detectors as shown in fig. 2.1. It depends on the particle in which of the sub-detectors it interacts and produces a signal. For example an electron can be identified by leaving a track in the tracking system and interacting with the electromagnetic calorimeter which stops the electron in front of the hadronic calorimeter. Therefore it does not leave a signal in the hadronic calorimeter. This method is mainly used for photons and leptons[9].

Particle identification by mass is done by obtaining the momentum  $p = \beta\gamma mc$  and the velocity  $v = \beta c$  of a particle from observables that depend on  $v$  or  $p$ . The momentum of a charged particle can be obtained from the curvature of its trajectory in a magnetic field where the Lorentz force acts as centripetal force. The velocity can be determined indirectly by different methods[9]:

- Cherenkov radiation
- Time-of-flight
- Energy loss by ionization
- Transition radiation

Cherenkov radiation is emitted when a particle traverses a medium faster than the speed of light in this medium. The angle of the emitted radiation relative to the particle's track depends on its velocity[9]

$$\cos(\Theta_C) = \frac{1}{\beta n} \tag{2.1}$$

where  $n$  is the material's refractive index.

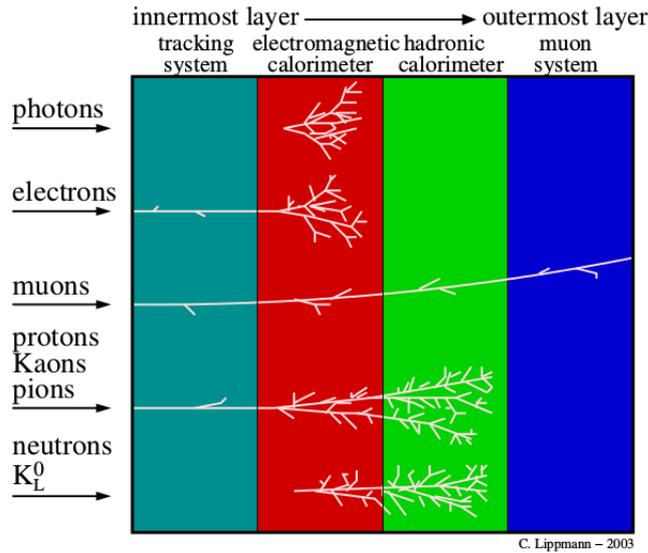


Figure 2.1: Concept of PID by difference in interaction[9]. Tracks of different particles leave signals in different sub detectors. This allows to identify the species of a particle corresponding to a track.

The other methods are explained in more detail in the section of the corresponding detectors.

## 2.1 Gaseous Ionization Detectors and Energy Loss of Traversing Particles

The concept of gaseous ionization detectors is to measure the energy loss per unit length  $\frac{dE}{dx}$  by ionization. Charged particles traversing the detector lose energy by ionizing the gas inside. This leads to  $N_I$  electron-ion-pairs over a distance  $x$  given by

$$N_I W = x \left\langle \frac{dE}{dx} \right\rangle \quad (2.2)$$

where  $W$  is the average energy for creating electron-ion-pairs[9]. Those electrons can further ionize the gas, which results in secondary electrons [4]. The ionization processes can be described by inelastic Coulomb collisions where the traversing particle loses a small amount of energy which excites or ionizes the involved atom.

Typically, the average energy loss can be described according to the Bethe-Bloch formula. However, the formula has to be modified to describe the signal of the detector,

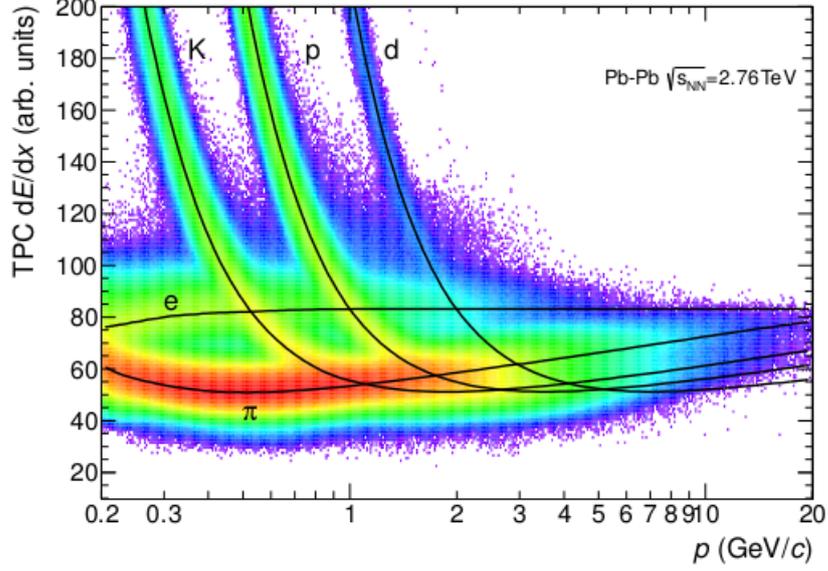


Figure 2.2: TPC signal of tracks of different particle species. The lines correspond to the average energy loss of the track of particles.[11]

since large energy losses, contained in the calculation of the average energy loss, do not contribute to the measured energy loss of a track. Instead they lead to a  $\delta$  electron which is knocked out of an atom and can be measured as a stand-alone track, so that they do not contribute to the measured mean energy loss of a track. Therefore [19] introduces an energy cut-off for calculating the average energy loss. The modified Bethe-Bloch formula can be described according to [8] as

$$\left\langle \frac{dE}{dx} \right\rangle \propto \frac{C_1}{\beta^2} \log(C_2 \beta \gamma - \beta^2 - C_3) \quad (2.3)$$

which depends on the velocity of a particle. Here  $C_i$  are detector dependent constants[8]. In fig. 2.2 one can see the characteristics of this formula. For low velocities the energy loss behaves like  $1/\beta^2$  and ends up in the region of minimum ionization. For higher  $\beta\gamma$  the relativistic rise dominates and follows  $\log(\beta\gamma)$  until it is suppressed by the density effect correction which is represented by  $C_3$  [9]. This region is the so called Fermi plateau.

For the primary ionization the Rutherford cross section describes the interaction for energies above the highest atomic binding energy[9]

$$\left( \frac{d\sigma}{dE} \right)_{\text{Ruth}} \propto \frac{1}{E^2}. \quad (2.4)$$

This shows that electrons from primary ionization with high energies are suppressed, but there is still a probability that they exist and can further ionize the gas so that secondary electrons are produced.

## 2.2 Setup of the ALICE Detector

In the ALICE detector the TPC and TRD are gaseous ionization detectors as described in the previous section. This section first gives an overview of the total setup and then investigates in further details the TPC, TRD and TOF.

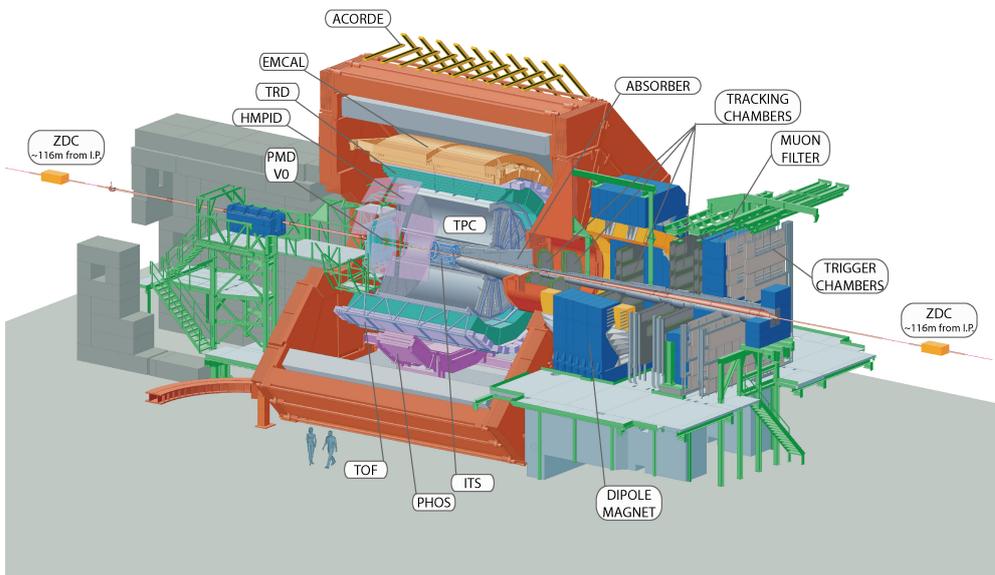


Figure 2.3: Setup of ALICE detector[18]

In fig. 2.3 the different sub detectors of ALICE are shown. The following passage focuses on detectors used in this analysis. The description is ordered from the inner to the outer sub detectors. Closest to the beam pipe the Inner Tracking System (ITS) is mounted. It is used to reconstruct the primary vertex, secondary vertices and to improve the resolution for momentum and angular reconstructions from TPC[8].

Next to the ITS is the Time Projection Chamber (TPC), which is the main tracking device of the experiment. Besides track reconstruction and momentum measurements it also provides PID by measurement of the specific energy loss of a particle traversing the gas in the TPC.

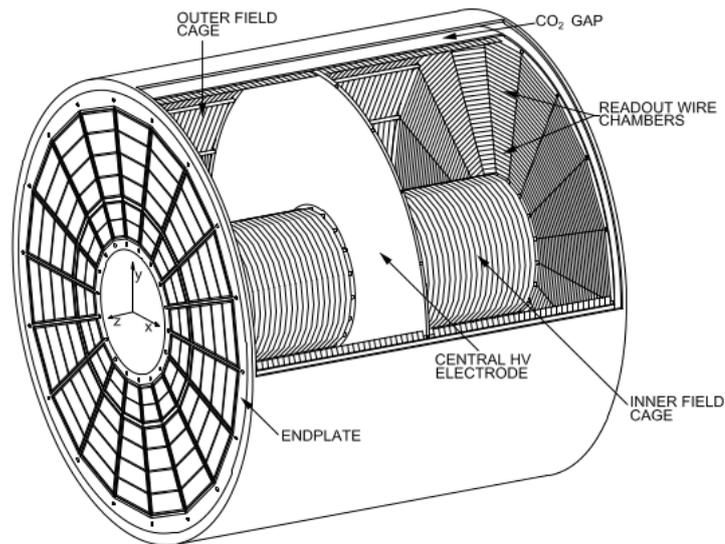


Figure 2.4: Setup of TPC[9]

The following detector is the Transition Radiation Detector (TRD) which is especially designed to distinguish electrons from pions over a momentum range from 1 GeV/c up to 100 GeV/c making use of the fact, that only electrons create transition radiation in this kinematic regime.

The TRD is followed by the Time of Flight (TOF) detector which is important for PID when the signals of pions, kaons and protons overlap in the TPC. Measuring the velocity via the time-of-flight and the track length yields an additional feature to discriminate between those particles.

There are also two electromagnetic calorimeters called EMCAL and PHOS. They can be used to identify photons and measure their 4-momentum. For electron identification they can be used as well.

Furthermore, there are other detectors which are important for event characterization and are described in [1] and [8] in more detail.

The magnetic field is created by a large solenoid and has a strength of 0.5 Tesla so that the observation of particles with low transverse momentum is possible[2].

## 2.3 Time Projection Chamber

The setup of the Time Projection Chamber can be seen in fig. 2.4. The field cage is a hollow cylinder whose axis is parallel to the beam axis. In 2010 - when the data for

this analysis were taken - it was filled with a Ne-CO<sub>2</sub>-N<sub>2</sub> gas mixture (90-10-5)[10]. The active volume has an inner radius of 85 cm, an outer radius of 250 cm and a length of 500 cm. The cage is separated into two halves by a thin membrane - the central electrode, which is at a potential of 100 kV. Together with a voltage dividing network on the surface of the outer and inner cylinder it creates an electric field of about 400 V/cm parallel to the beam axis[9]. Therefore electrons - created by ionization from traversing particles - drift towards the end-caps, where readout detectors are mounted. When the electrons reach the readout anode wires, an avalanche is produced and the signal gets amplified, which is read out by cathode pads via the induced potential change[10].

In front of the readout wires a gating grid is installed which prevents ions from entering the drift volume. It opens when the system is triggered so that electrons from ionization can pass. Since the produced ions drift with a much lower velocity than the electrons, they cannot enter the drift volume, so that they do not influence the signal in the read out region[9].

From the position of the readout pads a two dimensional projection of a particles trajectory on the TPC's end caps can be reconstructed from the deposited charge clusters from ionization by calculating the cluster's center of gravity.

The drift velocity can be calculated depending on the electric and magnetic fields as[4]

$$\mathbf{v} = \frac{e\tau}{m(1 + \omega^2\tau^2)} \left( \mathbf{E} + \frac{\omega\tau}{B} (\mathbf{E} \times \mathbf{B}) + \frac{\omega^2\tau^2}{B^2} (\mathbf{E} \cdot \mathbf{B})\mathbf{B} \right), \quad (2.5)$$

where  $\omega = eB/m$  is the Larmor frequency and  $\tau$  is the mean collision time. Since the magnetic field is parallel to the electric field the second term equals zero and the third is independent of  $B$ . In good approximation the electric field is homogeneous and therefore the drift velocity is constant. Therefore a measure of the drift time of electrons from ionization and the 2D projection of the track allow to reconstruct the three dimensional trajectory of a particle[2][4].

A helix-fit runs on the reconstructed particles track to get the parameters of the track - such as the curvature - from which one obtains the kinematics[8].

### **Shape of Energy Loss Distributions of Particles traversing the TPC-Detector**

The amount of energy that a track loses while traversing the Gas inside the TPC is a random variable with a distribution called straggling function[7] (or Landau distribution). This distribution can be obtained from the detector by measuring the energy losses of many tracks of the same particle species. However, since there are always dif-

ferent particles traversing the TPC, the measured energy loss spectrum will contain the sum of straggling functions corresponding to different particle species. Those straggling functions will be observed in small momentum or transverse momentum bins.

For the track of a single particle the measured energy loss is the sum over all clusters, made up by ionization. To understand the shape of the distribution for a single track, first consider the energy loss distribution  $\sigma(\Delta)$  of a single particle for a single collision with energy loss  $\Delta$ . The distributions for more collisions can be calculated iteratively by convolution[7]

$$\sigma^{*n}(\Delta) = \int_0^{\Delta} \sigma(E) \cdot \sigma^{*(n-1)}(\Delta - E) dE. \quad (2.6)$$

This formula describes the probability that a particle loses energy  $\Delta$  by producing a cluster by  $n$  collisions. Using the Rutherford cross section this leads to the Landau distribution  $L(\Delta)$ [7]. However, the measured distributions differ from this one, since the Rutherford cross section is not a good approximation for all detector gases[7].

The energy loss distributions - which can be described by straggling functions[7] - have a large tail that might be removed for practical purposes. Typically one applies a truncated mean cut, which consists of omitting a certain number of the clusters with highest energy loss of each track. A simple simulation of this is shown in figure 2.5. For illustration a signal out of many tracks - consisting of 100 clusters each - was created. For each cluster a random value from a Landau was sampled to simulate a random charge deposition. For a more accurate method see [7]. In the black histogram the total energy loss of many tracks was binned. In the blue one a truncated mean cut has been applied on the clusters. For every track the 40% highest clusters have been omitted and the rest has been summed up so that the large tail has been removed. Furthermore, one can see that there is a shift of the maximum.

With truncation the signal's shape is similar to a Gaussian[11]. Therefore PID is possible by fitting Gaussian distributions for every particle species to data. However, for being more accurate in [2] an alternative description of the signal is presented. There it is assumed that in good approximation the truncation can be described by a Landau distribution multiplied by a truncation factor depending on the energy loss  $\Delta$  of a track. The factor represents the probability that the track contains no clusters above a certain threshold, which is motivated by assuming introducing a thresholds and cutting all clusters above this threshold is equivalent to cutting of the 40% highest clusters of a track. With some statistical considerations one ends up at a falling exponential for the

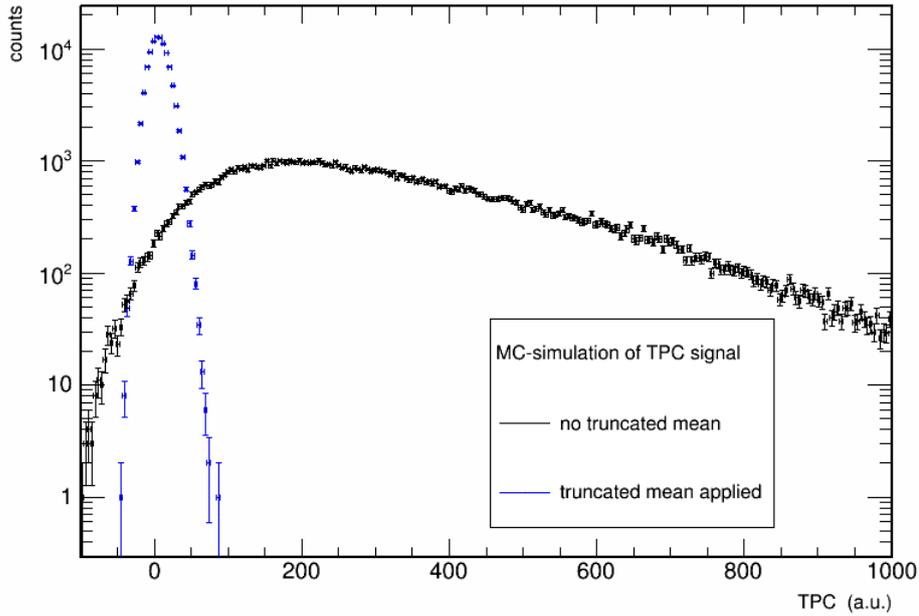


Figure 2.5: Monte Carlo simulation to show the effect of the truncated mean cut on the signal's shape: the large tail of the sampled Landau distribution (black) is removed and the shape becomes more like a Gaussian but with shifted maximum (blue).

truncation factor so that the signal is described as[2]

$$f(\Delta) = A \cdot L(\Delta; \mu, \sigma) \cdot \exp(-\alpha \cdot \Delta), \quad (2.7)$$

where  $\mu$  and  $\sigma$  describe the Landau's most probable value and the width of the peak. The normalization refers to  $A$  and  $\alpha$  is a free parameter that determines how strongly the Landau's tail is removed.

## 2.4 Time of Flight Detector

Particle identification by ionization becomes problematic when the lines from Bethe-Bloch curve for different particles cross each other. In this case additional information about the particles' velocities can be provided by the Time of Flight detector, which consists of an array of Multi-gap Resistive Plate Chambers for precise time measurements.

From the time-of-flight  $t$  and the length of the trajectory  $L$  one obtains an expression for the particle's mass depending on its momentum  $p$  by combining

$$\beta = \frac{L}{tc} \quad \beta = \frac{1}{\sqrt{\left(\frac{mc}{p}\right)^2 + 1}} \quad \Rightarrow \quad m = \frac{p}{c} \sqrt{\left(\frac{ct}{L}\right)^2 - 1}. \quad (2.8)$$

Together with the momentum measurement from the track reconstruction PID is possible.

To estimate the separation power between two particle species  $A$  and  $B$  with same momentum, one has to look at the time resolution[9]

$$|t_A - t_B| = \frac{L}{c} \left| \sqrt{1 + \left(\frac{m_A c}{p}\right)^2} - \sqrt{1 + \left(\frac{m_B c}{p}\right)^2} \right| \quad (2.9)$$

and expand it for  $p \gg mc$  so that

$$\sqrt{1 + \left(\frac{mc}{p}\right)^2} \approx 1 + \frac{m^2 c^2}{2p^2} \quad (2.10)$$

holds and therefore the separation power becomes

$$n = \frac{|t_A - t_B|}{\sigma_{\text{TOF}}} \approx \frac{Lc}{2p^2 \sigma_{\text{TOF}}} |m_A^2 - m_B^2|. \quad (2.11)$$

This shows that PID with TOF becomes problematic for particles with high momentum and where the masses differ little. This will be seen exemplary for electrons and pions in chapter 5. For the separation of pions and kaons according to[11] TOF works well up to 2.56 GeV/c and for protons up to 4 GeV/c, which will be enough for my purposes.

Although the TOF's signal has a tail, the detector's response function is modeled as Gaussian[8] and can be described by

$$g_i(t) \sim \frac{1}{\sigma} \exp\left(-\frac{(t - t_i^{\text{exp}})^2}{2\sigma^2}\right) \quad (2.12)$$

for the  $i$ -th mass hypothesis. The expected time  $t_i^{\text{exp}}$  is obtained during the reconstruction of the track by summing up the TOF increments  $\Delta t_k$  from all track segments  $k$  in the reconstruction process[8]. They are obtained by local estimates  $p_k$  for the momentum

and the track length increment  $\Delta l_k$  via

$$\Delta t_k = \sqrt{p_k^2 + m_i^2} \frac{\Delta l_k}{p_k}. \quad (2.13)$$

The width of the response function is given by  $\sigma^2 = \sigma_{\text{TOF}}^2 + \sigma_{\text{rec}}^2$  which includes the time resolution  $\sigma_{\text{TOF}}^2$  and the errors  $\sigma_{\text{rec}}^2$  on reconstructing  $L$  and  $p$ .

In this setup one can identify particles by  $n\sigma_i$ -TOF-cuts where a particle is said to be of the corresponding type to the  $i$ -th mass hypothesis when its TOF-signal is within a region of  $\pm n\sigma_i$  around the expected signal  $t_i^{\text{exp}}$ . In this analysis  $3\sigma$ -TOF-cuts will be applied since this reduces the background without losing too much statistics.

## 2.5 Transition Radiation Detector

The Transition Radiation Detector has a good separation power between pions and electrons over a momentum range from 1 to 100 GeV/c[14]. This will be helpful to estimate the pion contamination. As explained in a previous section, the separation power of TOF decreases with higher momenta. This will mainly affect pions and electrons in the data sample of this thesis and therefore TRD will be used for a better separation. On the other hand, in 2010, when the data this analysis uses, were taken, only seven out of 18 of the super-modules were installed. Therefore using the TRD will lead to a loss of statistics. This is one reason why the TRD will not be used for the current heavy flavor electron analyses, but for other samples used in this analysis, where statistics play a minor role, the TRD is the tool of choice.

A relativistic particle with Lorentz-factor  $\gamma \geq 1000$  has a chance to produce a photon while traversing the interface of two media with different dielectric constants. In the above mentioned momentum range only electrons have such a Lorentz-factor due to their low mass[14]. Therefore detecting transition radiation (TR) is characteristic of electrons.

In order to understand how this concept is realized, one may have a look at the set-up of a TRD layer as shown in fig. 2.6. It consists of a radiator, a drift region and the amplification region made up by anode wires and cathode pads. The drift chamber is filled with a gas mixture of Xe and CO<sub>2</sub> (85 – 15)[11]. Therefore the TRD is a gaseous ionization detector as described in section 2.1. Besides the ionization from traversing particles the TRD detects the TR photons which are absorbed by the gas due to the high atomic number of its components[11]. All clusters from ionization drift towards the

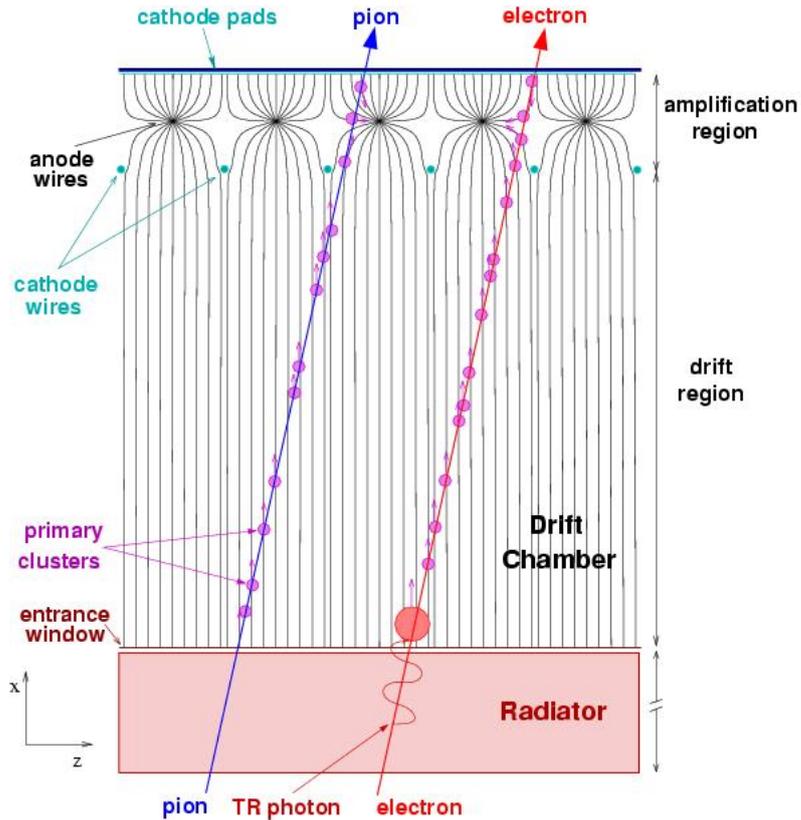


Figure 2.6: Schematic view of one layer of TRD[12]. A traversing particle produces charge clusters by ionization of the gas inside. Those drift towards the amplification region - due to an electric field - where the deposited charge is read out. Furthermore, electrons have a chance to produce a transition radiation (TR) photon. It is absorbed by the gas close to the radiator, so that a cluster with more charge than for an electron without a TR photon is produced.

anode wires where the signal is amplified and read out by the cathode pads. Their size is designed so that 2-3 pads detect one cluster whose position then is reconstructed by its center of gravity.

The deposited charge is collected and read out with a frequency of 10 MHz corresponding to a sampling interval of 100 ns. Since the drift time of the clusters is larger, the distance between the point where a cluster was produced and the read out region can be reconstructed, as can be seen in fig. 2.7. The average pulse height of the TRD signal in a layer for electrons and pions is shown in dependence of the drift time. For electrons there are two peaks. The first, which has to be close to the readout region due to the short drift time, is a result of the fact that near the anode wires charge moves from two sides towards them. So the signal from this region is much higher. The second peak,

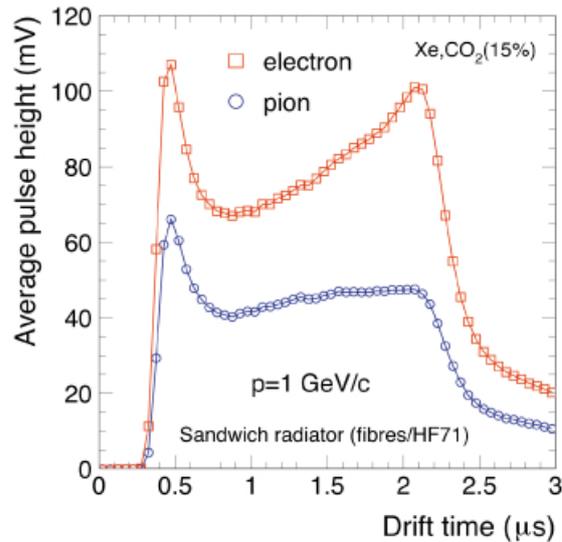


Figure 2.7: Average pulse height of the TRD signal in a layer for pions and electrons. It shows the dependence between the average signal produced by a cluster and the time the electrons from the cluster need to drift towards the read-out region. The larger the drift time the closer to the radiator a cluster has been produced.[13]

which has to correspond to clusters next to the radiator - that have a long drift time - arises due to the transition radiation produced by an electron. In comparison with fig. 2.6 one actually sees that TR photons are absorbed close to the radiator. Therefore the second curve in fig. 2.7 for the pions - which do not emit transition radiation - only shows the first peak and not the second one. The plateau is due to energy deposit from the particles while traversing the gas inside the drift chamber. With those information one is able to distinguish also other particle types by their characteristic energy loss as described in section 2.1.

To assure that the probability of producing a TR photon by an electron is high enough, one could increase the amount of interfaces in the radiator which is made of polypropylene fibers and foam[14], but this would lead to a much bigger radiator so that the TR photon might be absorbed by the radiator itself and would not be detected. To avoid this, six detector layers of the kind that can be seen in fig. 2.6 are stacked upon each other to increase the chance to detect TR photons.

## 2.6 Particle Identification with TRD

There are mainly four methods to identify particles with the TRD[11]:

- Truncated Mean Cut (TM)
- Neural network (NN)
- One-dimensional likelihood on total integrated charge LQ1D
- Two-dimensional likelihood on total integrated charge LQ2D

In this analysis only the last method will be applied. Therefore just both likelihood methods will be explained in the following.

For the LQ1D method one defines a likelihood[8] for a track being an electron or pion respectively by

$$L = \frac{P_e}{P_e + P_\pi} \quad \text{with} \quad P_e = \prod_{i=1}^N P(X_i|e) \quad \text{and} \quad P_\pi = \prod_{i=1}^N P(X_i|\pi) \quad (2.14)$$

where  $P(X_i|e)$  denotes the probability that the total deposited charge  $X_i$  in layer  $i$  belongs to an traversing electron.  $N$  is the total number of layers.

Since the energy loss of electrons and pions becomes more and more similar with higher momentum due to the relativistic rise of the Bethe-Bloch curve for pions, the total integrated charge may not only be a sufficient variable for the likelihood. Instead one can use additional information from fig. 2.7. Since the TRD is able to localize the cluster's distance to the radiator one can use the TR peak for electron identification more precisely by dividing the signal - shown in fig. 2.7 - into two halves for each layer. For every layer and every half, the probability that the signal in the corresponding half belongs to a traversing electron (a pion respectively) is determined and assembled to a likelihood similar to (2.14)[8]. This leads to the LQ2D method which uses the total integrated charge and the transition radiation to identify particles. A likelihood for protons, muons, and kaons is obtained in a similar way.

# 3 Fitting with Weighted Least Squares Method

The main tool of this analysis will be fitting analytic functions and measured histograms to data. Therefore this section explains the background of the applied fitting method which is deduced from the maximum likelihood method.

For a given set of data  $\{(y_i, x_i) | i = 1 \dots n\}$  one assumes a model  $y = f(x, \boldsymbol{\lambda})$  that describes the dependency between  $y$  and  $x$  depending on the parameters  $\boldsymbol{\lambda}$ . Let  $P(\{(y_i, x_i)\} | \boldsymbol{\lambda})$  be the probability to measure the data set  $\{(y_i, x_i)\}$  for given  $\boldsymbol{\lambda}$ . However, in the following  $P(\{(y_i, x_i)\} | \boldsymbol{\lambda})$  is interpreted as a likelihood of  $\boldsymbol{\lambda}$ . With a measured set we can fix  $(y_i, x_i) \forall i = 1 \dots n$  and maximize  $P$  with respect to  $\boldsymbol{\lambda}$ , which gives us the most likely parameters under assumption that the considered model describes the measured data correctly.

If the measurement of one data point is statistically independent of all others

$$P(\{(y_i, x_i)\} | \boldsymbol{\lambda}) = \prod_{i=1}^n P_i((y_i, x_i) | \boldsymbol{\lambda}) \quad (3.1)$$

holds, where  $P_i((y_i, x_i) | \boldsymbol{\lambda})$  is the probability to measure the pair  $(y_i, x_i)$ . Since the logarithm is a monotonic function, solving

$$\sum_{i=1}^n \frac{\partial \ln P_i((y_i, x_i) | \boldsymbol{\lambda})}{\partial \lambda_j} = 0 \quad \forall j \quad (3.2)$$

yields a local maximum of  $P(\{(y_i, x_i)\} | \boldsymbol{\lambda})$ , if the maximum exists.

Now consider that (3.1) holds and

$$P_i((y_i, x_i) | \boldsymbol{\lambda}) = \frac{1}{\sigma_i \sqrt{2\pi}} \exp\left(-\frac{(y_i - f(x_i, \boldsymbol{\lambda}))^2}{2\sigma_i^2}\right) \quad (3.3)$$

the  $y_i$  are normally distributed where  $f(x_i, \boldsymbol{\lambda})$  determines the expectation values and  $\sigma_i$  is the uncertainty of the measured  $y_i$ . For example  $x_i$  can label the bins of a histogram

and  $y_i$  is the number of counts in bin  $i$ . Then  $\sigma_i$  is the error on the counts in bin  $i$ . The assumption that the measured values are Gaussian distributed in every bin is a good approximation according to the central limit theorem if the number of counts in every bin is large. Otherwise a Poisson distribution has to be chosen.

Plugging (3.3) into (3.2) leads to

$$\frac{\partial}{\partial \lambda_j} \sum_{i=1}^n \frac{(y_i - f(x_i, \boldsymbol{\lambda}))^2}{\sigma_i^2} = 0 \quad \forall j, \quad (3.4)$$

which is also known as “weighted least squares” or “ $\chi^2$ -method”. It minimizes the weighted difference between measured data and values predicted by a model. If  $f$  is linear in  $\lambda_j \forall j$ , some simple algebraic conversions reduce (3.4) to solving an equivalent linear equation system. Otherwise,  $f$  can be linearized by a series expansion in  $\lambda_j$  around an initial value and (3.4) is solved by iterating the point around which it is expanded. The convergence of this process is only guaranteed by an appropriate choice of an initial point[21].

# 4 Data Samples and Applied Cuts

The analyzed data consist of Pb-Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV taken in 2010. This analysis concentrates on a centrality from 0 – 20%. Since it is focused on heavy-flavor electrons (HFE), the standard cuts from the HFE-group have been applied and are described briefly in the following.

To obtain a good signal quality, for every track a minimum number of 110 - out of 189 maximally detectable - produced clusters in the TPC is required, from which at least 80 have to be usable for PID. Furthermore, the ratio between the expected number of findable clusters - predicted by geometrical considerations - and actually found clusters has to be larger than 60%. In addition there have to be at least four clusters in the Inner Tracking System. Both cuts lead to an improvement of the track quality.

The ITS is also used to suppress electrons which are not from heavy flavor decays. Therefore for every track a signal in the first two layers of the ITS is required. This assures that electrons originate from processes close to the vertex and not from pair production by a photon at outer radii, so that the number of background electrons is reduced.

## 4.1 Centrality

Centrality describes the overlap between the two nuclei during a collision. There is a strong correlation between high multiplicity and large overlap of the nuclei. This quantity depends on the impact parameter  $b$  and is defined by

$$c(b) = \frac{\int_0^b \frac{d\sigma}{db'} db'}{\int_0^\infty \frac{d\sigma}{db'} db'} = \frac{1}{\sigma_{AA}} \int_0^b \frac{d\sigma}{db'} db' \quad (4.1)$$

where  $\sigma_{AA}$  is the total hadronic cross section in  $AA$ -collisions[11]. To replace the impact parameter  $b$  with quantities, which are easier to access experimentally, one can make an approximation under the assumption that the mean multiplicity at mid rapidity increases monotonically with the overlap region or, respectively, the zero degree energy

decreases[11]. Therefore one can replace the impact parameter with those quantities and can further switch between cross section and the number of observed particles - corrected for trigger effects and non hadronic background - which leads to the equation

$$c \approx \frac{1}{N_{ev}} \int_{N_{ch}}^{\infty} \frac{dn}{dN'_{ch}} dN'_{ch} \approx \frac{1}{N_{ev}} \int_0^{E_{ZDC}} \frac{dn}{dE'_{ZDC}} dE'_{ZDC} \quad (4.2)$$

where  $N_{ev}$  is the number of total events,  $N_{ch}$  is the largest multiplicity of charged particles and  $E_{ZDC}$  is the smallest zero degree energy[11].

For most central collisions there will be a small impact parameter which according to (4.1) leads to small centralities. Therefore centralities of 0 – 20% are called most central. From (4.2) one can deduce characteristics of samples with small centralities. For low  $c$  the integral has to be small, which is achieved by a high multiplicity  $N_{ch}$ . On the other hand low  $N_{ch}$  corresponds to centralities around 90 – 100%, which shows that more peripheral collisions behave similar to p-Pb collisions.

In this thesis data from 0 – 20% centrality are analyzed. Therefore one could expect, that there will be a larger background than in p-Pb collisions. However, the method which will be developed in the following can be applied for other centralities, too.

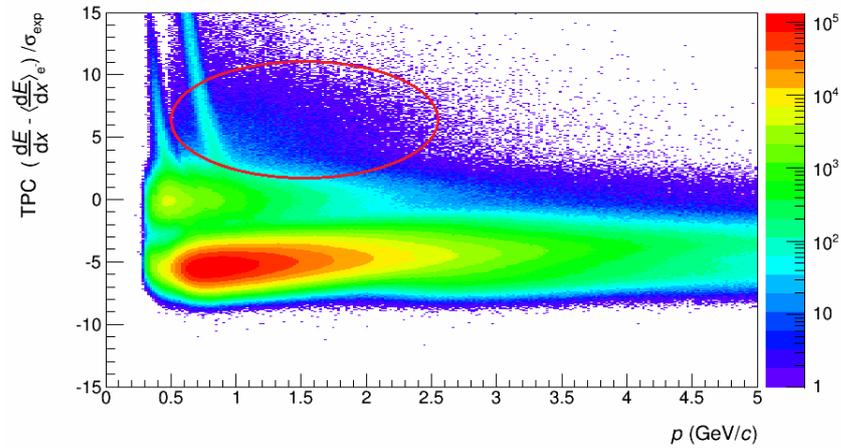


Figure 4.1: TPC signal after  $3\sigma$  TOF-cut around the electron expectation. Due to the high abundance of pions relative to electrons and the separation power of TOF - which decreases with increasing momentum - they dominate this sample. There is a large bulge - marked red - which will be discussed in chapter 5.

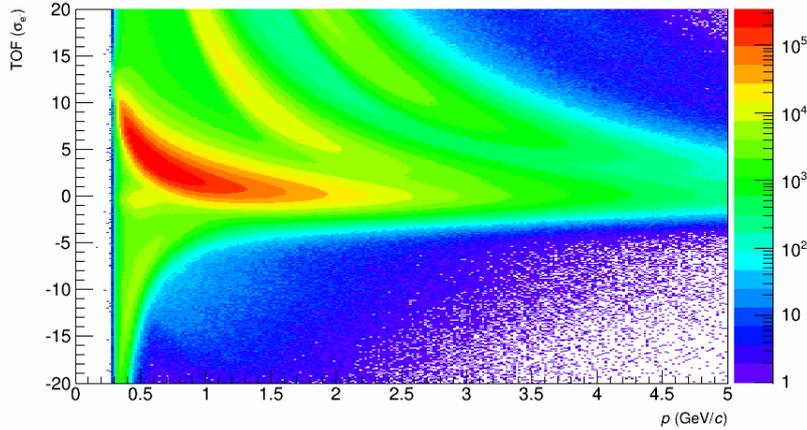


Figure 4.2: Time-of-Flight signal in  $\sigma_e$  plotted against the momentum. One can see the lines of different particle species: the large red area corresponds to pions. At a TOF signal of  $0 \sigma_e$  one can see the electron line starting to overlap with pions at  $p > 0.7 \text{ GeV}/c$ .

## 4.2 TOF-Cuts and Data Sample

The Time of Flight detector can be used for PID as described in section 2.4. Since a sample of electrons with reduced background is needed, a  $3\sigma_e$  cut around the electron expectation value is applied. Then the TPC signal in  $n\sigma_e$  and the total momentum - for later analysis the transverse momentum - are stored in a histogram as can be seen in fig. 4.1. Comparing this to the Bethe-Bloch curves from fig. 2.2 one can see the electron line around 0 above the pion line. Also the lines from kaons and protons can be seen clearly.

Taking into account that a TOF-cut has been applied one would expect by looking at the TOF signal in fig. 4.2 that only pions - represented by the huge red area - and electrons around  $0 \pm 3$  should be in this sample. However, there are TOF mismatches of other particles, so that a wrong time-of-flight is assigned to the particle's track. Therefore also protons, kaons and deuterons can be found in the TPC signal after the TOF-cut in fig. 4.1 has been applied.

The deuterons can hardly be seen, since they are covered by the large bulge above the electron line over a momentum range from 0.5 to 2.5  $\text{GeV}/c$  but they are expected to cross the electron line at a momentum of about 2  $\text{GeV}/c$ . Furthermore, protons cross at 1  $\text{GeV}/c$  and kaons around 0.5  $\text{GeV}/c$  (see fig. 2.2).

It is the aim of this analysis to estimate the contamination in these crossing regions for every particle species and the contamination from pions, which approach from below the electron line with increasing momentum. As can be seen in fig. 4.2, for low momentum the separation power between electrons and pions of the Time of Flight detector is quite good but for higher momenta the pion and electron distributions overlap more and more as expected from eq. (2.11).

Since in the HFE analyses electrons are selected by the TPC from the  $n\sigma_e$  interval  $[0, 3]$ , which keeps pion contamination low[15], the contaminations for the different particles in this specific interval will be estimated. However, the method which is described in this thesis could also be applied to any other interval. Furthermore, it is interesting to know what kind of particles make up the large bulge and how strongly they contaminate the electrons.

# 5 Estimating Contamination with different Fit Models

The contamination of the electron sample in the  $n\sigma_e$  interval from 0 to 3 is estimated for every transverse momentum bin for every particle species. The reason to choose the transverse momentum, is that other analyses - which rely on the estimation of contamination - observe transverse momentum. However, as will be seen, choosing the transverse momentum will make PID more complicated, so that the contamination has to be determined in dependence of the momentum and has to be converted afterwards.

For every transverse momentum bin in the histogram of fig. 4.1 the projection on the ordinate will be analyzed. Therefore different models will be fitted to the energy loss distributions so that they can be dealt with separately for each particle species.

## 5.1 Fitting multiple Gaussians

The simplest approach is fitting a Gaussian distribution for every particle species. For a transverse momentum above 0.6 GeV/c kaons are expected to be covered by pions.

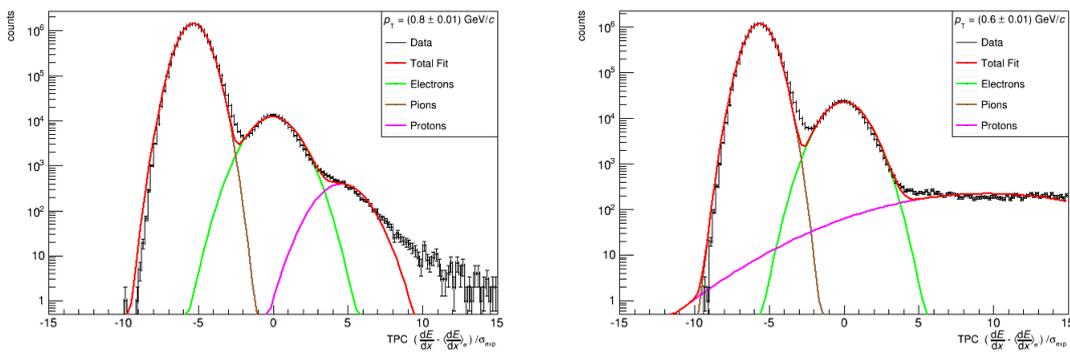


Figure 5.1: Fitting three Gaussian-Distributions to data in different  $p_T$ . Obviously this leads to poor results that have to be improved.

Therefore in fig.5.1 Gaussians have been fitted for electrons, pions and protons. In both plots one can see a dip between the pion and the electron peak, which leads to the assumption that a Gaussian does not describe pions well, because the Gaussian is symmetric around its mean value and a Landau is still asymmetric after truncation.

Furthermore, one might expect that - due to the choice of transforming the TPC signal in  $n\sigma$  electron - the Gaussian for electrons should have a mean value of 0 and a width of 1. This does not hold in this case, as can be seen in the left side of fig. 5.1. The fitted electron distribution is wider than the measured distribution and has a width of  $(1.248 \pm 0.003)$ , which is a result of the wrongly described right shoulder. The proton distribution does not describe the shoulder very well. Therefore the fit routine widens the electron Gaussian to compensate this.

In the right plot one can see that the description of protons by a Gaussian becomes worse for lower transverse momentum. To fit the shoulder the routine chooses a very large width for the Gaussian, which leads to a strong overestimation of the protons. At  $p_T = 0.6$  GeV/c no protons should have a TPC signal crossing the pion line as can be seen in fig. 4.1. Therefore one can conclude, that - in certain transverse momentum bins - fitting a Gaussian for protons can lead to false results. Hence, another distribution has to be found which describes protons more precisely.

## 5.2 Proton Distribution

To get a better description of the shape of the proton distribution one can use the good separation power for protons of the time-of-flight detector (see. sec. 2.4) and take a sample with a  $3\sigma_p$  TOF cut around the proton line. From this sample the distribution at  $p_T = (0.6 \pm 0.1)$  GeV/c is shown in fig. 5.2 together with the corresponding contaminated electron sample. As in the shown distributions different TOF-cuts have been applied, the distribution for protons has been scaled for better comparison, so that different numbers of total tracks in the two samples are taken into account. One can see that the large shoulder is reproduced well on the right, but close to the electron peak there is still background left, which must come from other sources. Furthermore, there are also misidentified pions and kaons in the proton sample, as can be seen on the left magenta peak, which is positioned around the expected energy loss for pions. The tail of this peak is made up of kaons which according to the Bethe-Bloch formula have to be next to the pions at this transverse momentum.

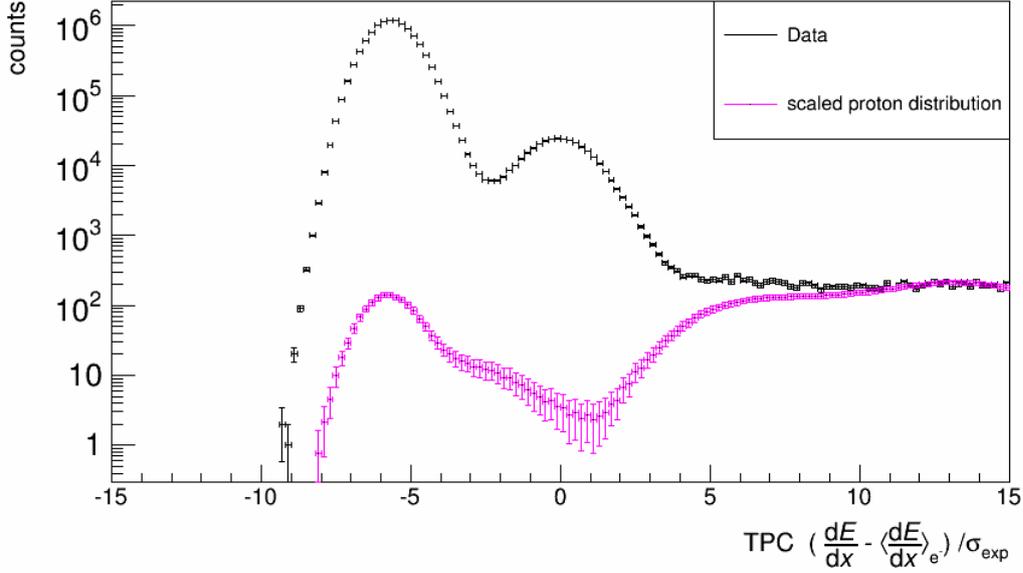


Figure 5.2: Proton distribution at  $p_T = (0.6 \pm 0.1)$  GeV/c selected by  $3\sigma_p$  TOF cut around the proton line, which is scaled for better comparison to the data for which contamination is analyzed.

However, as described in section 2.3 the response function of the TPC should be similar to a Gaussian. As can be seen clearly, this does not hold for this proton sample. This can be explained by the fact that all protons in this sample have the same transverse momentum but not the same absolute momentum. According to (2.3) the mean energy loss depends on  $\beta\gamma = p/m$  and therefore this sample contains protons with different average energy losses, which widens the distribution.

Taking the proton distribution for particles which are all in the same momentum bin leads to a much more Gaussian like shape as can be seen in fig. 5.3. It can be concluded that it is more reasonable to use the momentum for particle identification instead of  $p_T$  and this is why the analysis will be done in  $p$  and the results will be converted afterwards.

Now after changing to momentum one might suggest to use a Gaussian distribution for the protons, but another aspect is, that the proton contamination has to be estimated at momenta where the distributions of electrons and protons overlap. Independent of the shape of the proton distribution, it is nearly completely covered by the electron distribution around momenta near 1 GeV/c. Therefore the fit routine will get problems minimizing  $\chi^2$ , since the fit-parameters of electrons and protons will be strongly correlated. Every change that amplifies the protons could be replaced by amplifying

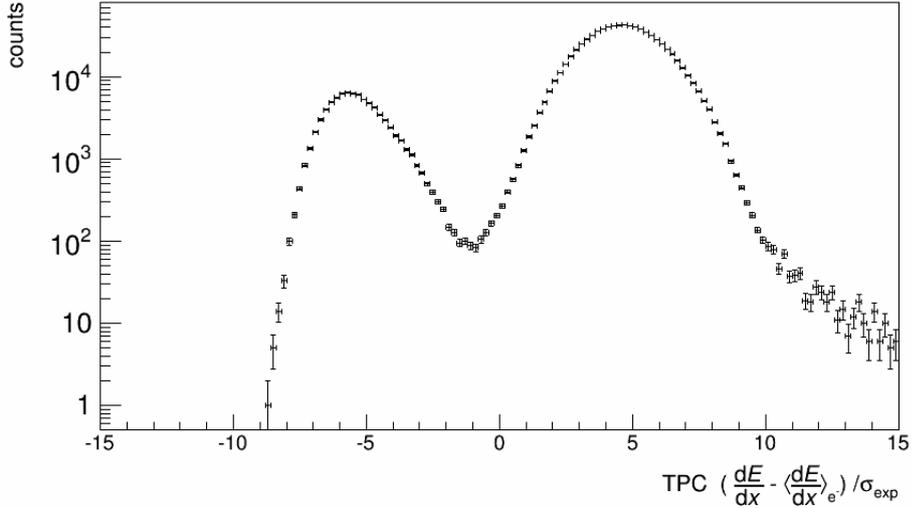


Figure 5.3: Proton distribution selected by  $3\sigma_p$  TOF cut around the proton line at  $p = 0.8$  GeV/c. Misidentified pions and kaons can be seen. The kaons are mostly covered by pions, but one can recognize them at the right shoulder of the left peak. The shape of the proton distribution - which refers to the right peak - is more Gaussian compared to fig. 5.2.

electrons, because they dominate the protons. On the other hand, the routine could interchange the two distributions and try to fit the proton distribution to the shape corresponding to electrons, so that the protons seem to dominate the electrons. In both cases the routine may converge, but will not yield reasonable results in terms of proton contamination.

To avoid these problems one might use experimental data describing the shape of the proton distribution.

### 5.3 Fitting with Experimental Data as Template

The idea behind this method is identifying the particles one wants to describe in another data set with different PID cuts and use the resulting distribution for the fit routine. Those templates can be obtained for kaons, protons and deuterons in the same way as in section 5.2 by applying  $3\sigma$  TOF-cuts respectively around the particles expectation values for every momentum bin. However, it must be assured that the same transformation of units is used as in the data which are to be analyzed. So the TPC  $n\sigma_{e^-}$  signal has to be obtained in dependency of the momentum, which can be seen for protons in fig. 5.4. In

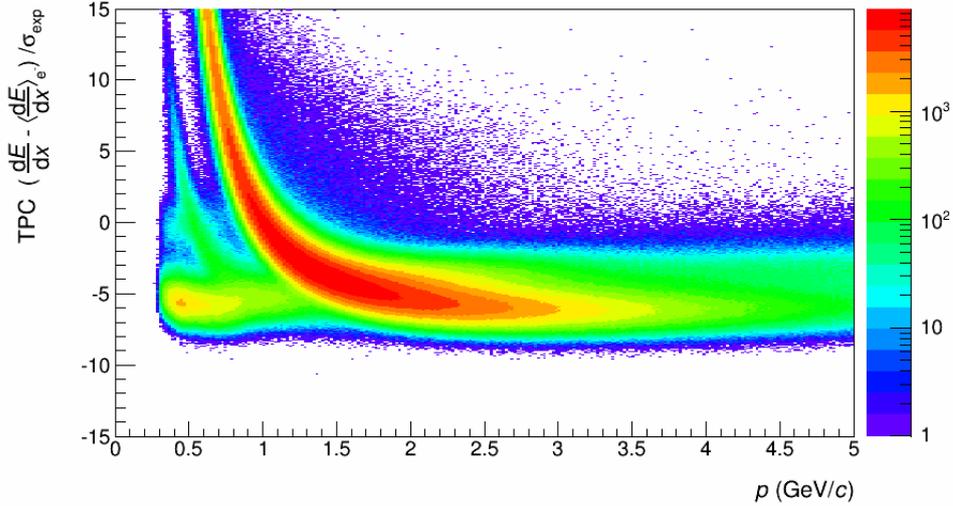


Figure 5.4: Selecting protons with a  $3\sigma_p$  TOF to get templates for their energy loss distributions in every momentum bin. An amount of misidentified particles remains.

this figure protons can be identified by the red area, but there are again misidentified kaons, electrons and pions as can be seen by the other lines. Protons will dominate the templates in the relevant momentum region, but due to the high multiplicity there will be a contribution from pions. In practice this will not affect the contamination in the  $dE/dx$  interval  $[0, 3]$  significantly, since there is - relative to the number of protons - less background in the template. The same effect is true for deuterons and kaons.

However, this method only works, if the shape of the particles' energy loss distributions of the TPC is not being changed by different TOF-cuts. The shapes of the templates have to be similar to the shapes of the misidentified particles in the analyzed data. One might think that, if a particle is misidentified by TOF, or identified correctly, is independent of the particle's energy loss in the TPC. Therefore misidentified particles are distributed the same way as in the template and therefore should have a similar shape. Then the only difference between the template and the distribution in data would be due to statistical fluctuations.

On the other hand, tracks with higher uncertainty in the reconstruction may be more likely to be misidentified by the TOF detector and therefore their TPC energy loss distribution may differ from the template. This has to be investigated further for a more precise estimate on contamination. However, in the following it is assumed that

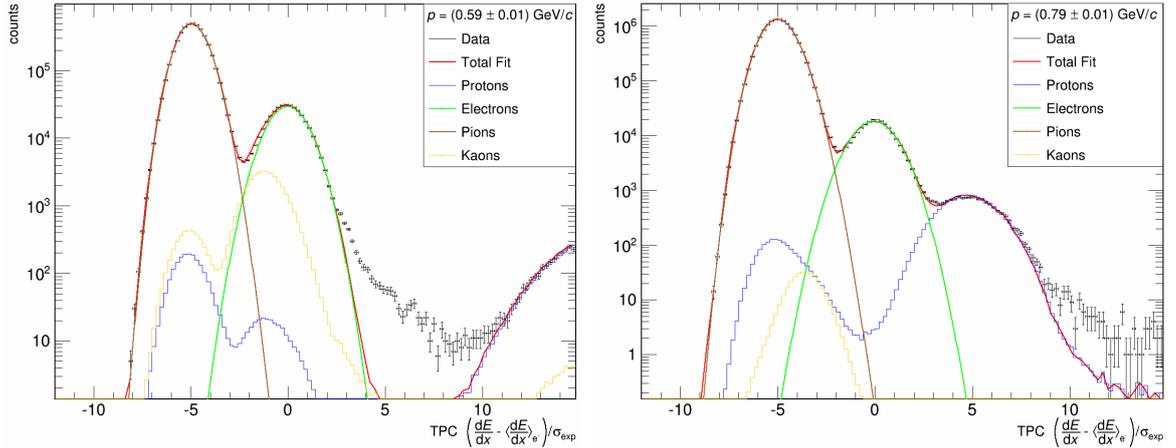


Figure 5.5: Fitting with templates for protons and kaons. Although the protons are described well, a shoulder on the right of the electron peak still remains.

differences between the templates' shapes and the energy loss distributions in the data sample are slight.

Since pions and electrons can't be separated well by TOF as discussed in section 2.4 their energy loss distributions have to be described in a different way. In the previous section it has also been shown that their distribution is asymmetric and since the contamination in the TPC  $\frac{dE}{dx}$  interval  $[0, 3]$  is estimated, an accurate description of the tail of the pion distribution is necessary. Therefore describing the pion distribution as a Gaussian is not sufficient. Instead, one can use the model of the truncated Landau from equation (2.7) which according to [2] works well for pions. Since the asymmetric part of the distribution of electrons is not relevant in the above mentioned interval, electrons are described by a Gaussian in the following.

For every  $p$ -bin it is now possible to get an energy loss distribution for each particle species. Those templates can be fitted to the data with the TOF-cut on electrons to get all particle distributions separately. Therefore the function

$$f(\Delta) = A_{\pi} \cdot L(\Delta; \mu_{\pi}, \sigma_{\pi}) \cdot \exp(-\alpha \cdot \Delta) + A_{e^{-}} \cdot G(\Delta; \mu_{e^{-}}, \sigma_{e^{-}}) + A_{p^{+}} \cdot P(\Delta) + A_{k} \cdot K(\Delta) + A_{d} \cdot D(\Delta) \quad (5.1)$$

will be fitted by varying the parameters  $A$ ,  $\mu$ , and  $\sigma$ , where  $\Delta$  refers to the total energy loss  $dE/dx$  after truncation,  $P(\Delta)$  describes the proton,  $K(\Delta)$  the kaon and  $D(\Delta)$  the deuteron template. In the following the  $A_x$  are called scaling factor. If a particle species is in a certain momentum region not expected to have energy loss in the  $dE/dx$

interval from -15 to 15, or if its distribution is totally covered by pions (see fig. 4.1), the corresponding scaling factor  $A_x$  is set to be zero. This leads to a higher stability of the fit routine. The result can be seen in Fig. 5.5.

While for  $p = 0.8$  GeV/c the right shoulder can be described by the proton template, for lower momentum this does not hold. This leads to two conclusions. On the one hand there must be other particles than protons which contribute to this shoulder and on the other hand if they can be seen at 0.6 GeV/c they may also exist at 0.8 GeV/c and higher momentum but are covered by the proton distribution. Therefore the amount of protons would be overestimated and the amount of other particles underestimated. Furthermore, the shape of the distribution from the unidentified particles is yet unknown and has to be determined for correct background estimation.

## 5.4 Contamination of Templates

The templates used in the previous section are not pure samples of one particle species as can be seen in fig. 5.5. Therefore, it is important to check if the templates are significantly contaminated. For this purpose the templates are obtained with the same cuts as described in sec. 5.3, but this time data are taken from Monte Carlo simulations.

By that it is possible to obtain two templates for each particle species. For the first template  $T_1(\Delta', p)$  a  $3\sigma$  TOF-cut around the expectation value of the species is applied. Here  $\Delta'$  refers to the value of the TPC  $(\frac{dE}{dx} - \langle \frac{dE}{dx} \rangle_e) / \sigma_{\text{exp}}$  and  $p$  to the momentum. For the second template  $T_2(\Delta', p)$  - additionally to the  $3\sigma$  TOF-cut - it is required that every track has to belong to the particular particle species according to Monte Carlo Truth.

Since the TPC  $dE/dx$  interval  $[0, 3]$  is investigated in this thesis, the contamination of the templates is estimated over this interval at fixed momentum. The relative difference

$$d(p) = 1 - \frac{\int_0^3 T_2(\Delta', p) d\Delta'}{\int_0^3 T_1(\Delta', p) d\Delta'} \quad (5.2)$$

corresponds to the relative contamination in  $[0, 3]$  of the template from Monte Carlo simulations. The uncertainty of  $d(p)$  can be obtained from error propagation. Since the templates are histograms, the integral in (5.2) is numerically represented as  $\sum_i T_x(\Delta_i, p)$ . Furthermore the uncertainty of  $T_x$  in bin  $i$  is given by  $\Delta(T_x(\Delta_i, p)) = \sqrt{T_x(\Delta_i, p)}$ .

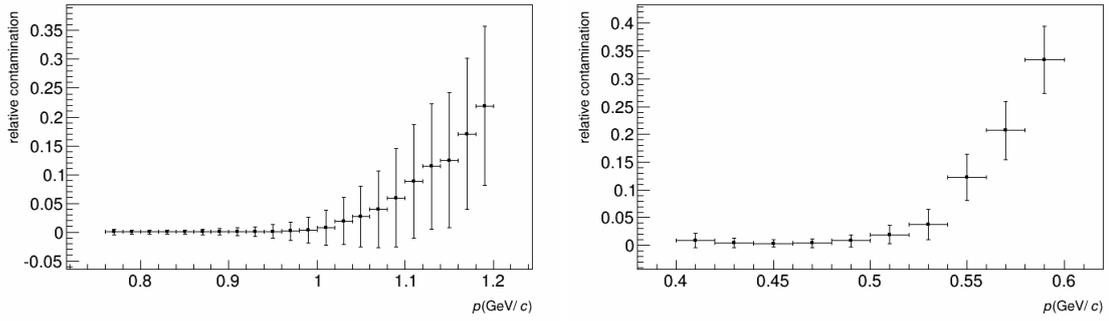


Figure 5.6: Relative contamination of templates from Monte Carlo simulations. The left plot shows the relative amount of particles which are in the TPC  $dE/dx$  interval  $[0, 3]$  of the proton template, but are not protons. The right plot shows the same for kaons.

Therefore, one obtains the uncertainty of the integral as

$$\begin{aligned} \Delta I_x &\equiv \Delta \left( \int_0^3 T_x(\Delta', p) d\Delta' \right) = \sqrt{\sum_i (\Delta(T_x(\Delta_i, p)))^2} \\ &= \sqrt{\sum_i T_x(\Delta_i, p)} = \sqrt{\int_0^3 T_x(\Delta', p) d\Delta'}. \end{aligned} \quad (5.3)$$

This leads to the uncertainty of  $d(p)$

$$\frac{\Delta d}{d} = \sqrt{\left( \frac{\Delta I_1}{I_1} \right)^2 + \left( \frac{\Delta I_2}{I_2} \right)^2} = \sqrt{\frac{1}{I_1} + \frac{1}{I_2}}, \quad \text{where} \quad I_x = \int_0^3 T_x(\Delta', p) d\Delta'. \quad (5.4)$$

The results are presented in fig. 5.6. One can see that for protons and kaons the contamination is low over a large momentum range, but rises suddenly for higher  $p$ . The same holds for the uncertainties. Looking at (5.4), one finds that the error increases with decreasing  $I_x$ . This means the uncertainty increases with decreasing number of tracks in the interval  $[0, 3]$ . From this observation one could expect that the contamination of the template is not significant until the number of tracks in the particular interval is small. Another consideration leads to the same conclusion. Assuming that the total amount of contamination of the kaon template is nearly constant over the displayed momentum range, the relative amount in the observed interval will increase, if there are fewer kaons in this interval. By looking at the TPC signal in fig. 4.1 one finds that indeed the amount of kaons in the TPC  $dE/dx$  interval  $[0, 3]$  decreases for momenta higher than 0.55 GeV/c. The same holds for protons.

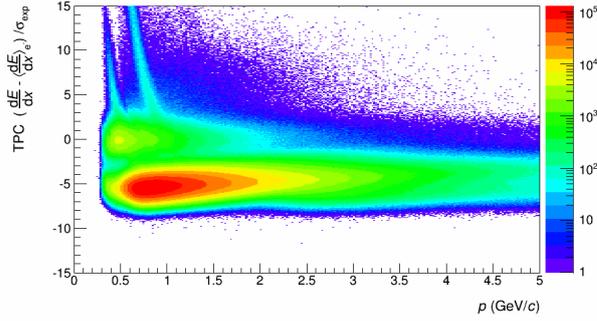


Figure 5.7: TPC signal after cut on active volume. The large background above the electron line is not removed.

Therefore, one can expect that the contamination of the template is not significant until there is only a small amount of the given particle species in the observed  $dE/dx$  interval. This will be confirmed when the amount of protons and kaons in the electron sample is determined. Then one will see that the momenta regions, where the contamination of the templates would be significant, are those regions where the amount of kaons - protons respectively - is small in the particular interval of the electron sample.

Since the TOF signal is implemented quite well in Monte Carlo simulation, one can assume from the previous discussion, that the templates - obtained from experimental data - are not contaminated significantly over the momenta regions, where they are used to describe the contamination of the electron sample.

## 5.5 Cut on Active Volume

Since the data in fig. 5.5 are taken at low momentum, the unidentified particles cannot be deuterons. Furthermore, the distributions of protons and kaons can be seen in the plot and the shoulder in the left histogram cannot be described by them. According to the Bethe-Bloch formula shown in fig. 2.2 no other particle species is expected to have an energy loss of this amount at  $p = 0.6$  GeV/c. Therefore one may suppose that the shoulder is a result of detector effects.

In most central Pb-Pb collisions - corresponding to very high multiplicity - there may be particles traversing the TPC close to each other and therefore some of the resulting clusters may overlap and cannot be resolved and assigned to the corresponding tracks correctly. Those tracks would have higher energy losses than expected for the

corresponding particle species also after truncation. Assuming that there is a significant amount of such tracks, this could explain the observed shoulder.

It would be reasonable if those overlapping clusters occur at the edges of the multi wire proportional chambers of the TPC where the spatial resolution is not as good as in the rest of the detector and where the amplification of the signal differs. In this case data might be improved by using cuts on the active volume in the TPC. However, as can be seen in fig. 5.7 this does not reduce the large background that corresponds to the observed shoulders. Therefore overlapping clusters may not only be near the edges of the multi wire proportional chambers.

## 5.6 A Template for Pions

As the active volume cut did not yield the desired results, another solution has to be found, which describes the shoulder instead of removing it. Due to the high multiplicity of pions it is reasonable to assume that statistically the majority of overlapping clusters belong to tracks corresponding to pions. However, the time-of-flight detector has difficulties separating electrons and pions, but by looking at fig. 4.2 one may suggest that cutting on the TOF interval  $[-3, -1]$  will lead to a sample where pions are more strongly separated from electrons as after a  $3\sigma$  TOF-cut.

As one can see in the right plot of fig. 5.8 the cut actually removed the bulge but not kaons and protons. The left plot shows a pure pion sample within a  $3\sigma_e$  TOF-cut around the electron expectation value from Monte Carlo Truth, in which the bulge is represented well. Therefore one can conclude that the observed shoulders in fig. 5.5 should be described by a distribution for pions. However, both models that have been used - Gaussian and Landau times falling exponential - do not describe the shoulder, because they do not include the effect of the overlapping clusters.

As described in section 2.6 the TRD provides methods to separate electrons and pions. Besides the standard cuts and the  $3\sigma_e$  TOF-cut the TRD will be used to reject electrons in the data set from which a template for pions will be obtained. Therefore the LQ2D method is used and all tracks with an electron probability larger than 10% are rejected. The resulting distribution has less statistic significance than the sample that is analyzed as discussed in section 2.5. However, it is enough for using it as template. It is displayed as the brown histogram in fig. 5.9.

Similar to the templates selected by the TOF detector, the usage of the TRD may influence the shape of the template for pions. By rejecting electrons as described above

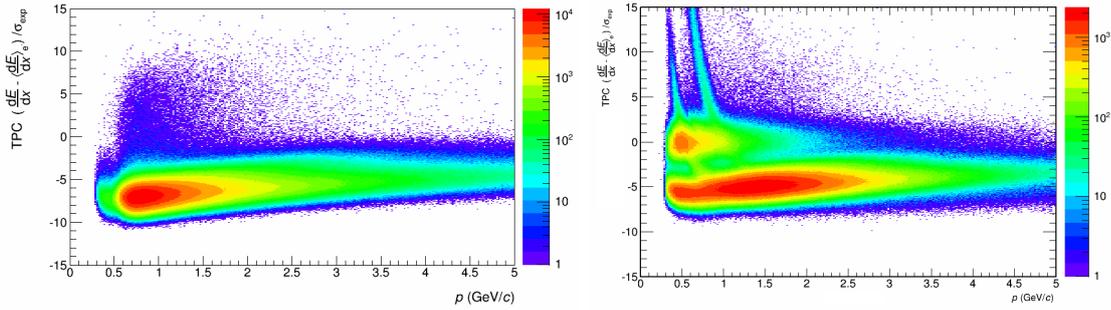


Figure 5.8: Left: Pure pion sample within a  $3\sigma_e$  TOF cut around the electron expectation value from Monte Carlo Truth. This sample also shows a bulge similar to the one in the electron sample in fig. 4.1. Right: Separating electrons from pions with a TOF-cut on  $[-3, -1]\sigma_e$  removes the large bulge, but also reduces statistics strongly.

there is still a probability that some electrons remain within the template. On the other hand some pions may be rejected.

If the tracks of pions - with overlapping clusters in the TPC - also have overlapping clusters in the TRD, the energy loss of a track in the TPC and in the TRD may be correlated. Hence, there can be a correlation between the LQ2D likelihood and the TPC signal of the track, so that the probability that a pion is rejected or an electron is not rejected is correlated with the energy loss in the TPC of the corresponding track. This could influence the shape of the energy loss distribution in the template and one has to further investigate this for less uncertainty in the estimate on contamination.

In equation (5.1) the modified Landau distribution is replaced by the pion template with the scaling factor as free parameter. Fitting this function to data results to fig. 5.9. One can see that due to the pions' tail there is much more pion contamination in the interval  $[0, 3]$  than expected from a truncated Landau model. Looking at the pion peak one finds that the template differs a little from the data represented by the black line. This difference is due to statistical fluctuations in the template and the data set and will be discussed later.

However, looking at the right plot, one finds that the pion tail does not fully describe the shoulder on the right of the electron peak. This may have different reasons. The pions contained in the tail of the distribution are expected to have tracks with overlapping clusters in the TPC (see sec. 5.5). This could influence the matching of the tracks when leaving the TPC and entering the TRD. Therefore those tracks may be suppressed if the TRD is used for the pion template.

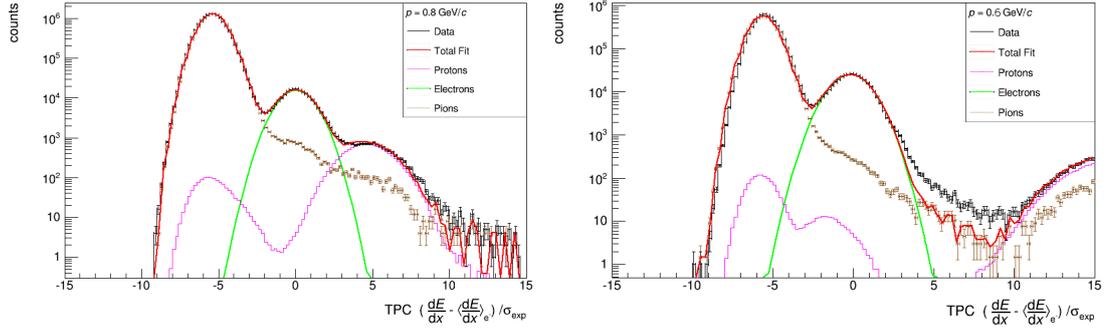


Figure 5.9: Fitting a Gaussian distribution for electrons and templates for all other particle species. The left plot shows the fitted distributions at  $p = 0.8$  GeV/c, the right plot at  $p = 0.6$  GeV/c. In the right plot one can see, that the tail of the pion distribution does not fully describe the shoulder next to the electron peak.

Another explanation for the underrepresented pion tail may be that the tracks with overlapping clusters in the TPC also have overlapping clusters in the TRD, because the two tracks - which correspond to the overlapping clusters - may traverse the TRD close to each other. This would lead to a TRD signal with higher energy loss than expected for pions. By that the two-dimensional likelihood (LQ2D), that one of those tracks corresponds to an electron, would be larger than the likelihood for the track corresponding to a pion. This would be crucial at momenta below 1 GeV/c, because in this kinematic regime the separation power of the TRD is not as good as for higher momenta. As a result of this, pion tracks with overlapping clusters may be identified as electrons by the TRD and therefore are rejected from the pion template.

In fact, the effect of having a template with underestimated tail can be seen for momenta below 0.9 GeV/c. This leads to an underestimation of the pion contamination of the electron sample.

To avoid this, the TPC  $dE/dx$  range - on which the templates are fitted - can be changed from the total range  $([-15,15])$  to a smaller range from -1 to 15. This would assure that the pion tail is fitted to the shoulder next to the electrons and therefore the amount of pions is not underestimated anymore. Instead it will be overestimated, because an electron contamination in the pion template still remains, as discussed in this section. In section 6.1.1 it will be reconsidered how to obtain uncertainties due to the discussed difficulties, but until then the fit-range will still be from -15 to 15, taking into account that pions will be underestimated.

## 5.7 Obtaining an Estimate on the Contamination of the Pion Template Caused by Electrons

As discussed in the previous section, the contamination of the pion template due to electrons - which have not been rejected by the TRD - influence the estimate of the amount of pions in the electron sample. Since the TRD response is not implemented well in Monte Carlo simulations, the method to approximate the contamination of the template as described in section 5.4 does not work in the case of the pion template. Instead one might estimate the order of magnitude of the contamination made up by electrons as follows.

It would be a reasonable assumption that, if there were no electrons in the template, the tail of the pion distribution (see fig. 5.9 brown markers) is monotonically decreasing. Thus one may remove step by step some electrons from the contaminated template until monotony is broken. Then the amount of removed electrons would be slightly higher than the actual contamination of the template. Therefore this amount could serve as an upper limit for the electron contamination of the pion template.

This can be done by taking into account that electrons should have a Gaussian distribution centered at zero with width equals one, because of the choice of transforming the TPC signal as described in the previous sections. Therefore, one can increase the scaling factor  $A$  of this Gaussian step by step and subtract it from the pion template until the tail of the pion distribution is no longer monotonically. Then the relative amount of subtracted electrons  $\Delta T_\pi$  - which refers to the relative contamination of the pion template  $T_\pi(\Delta)$  - can be obtained as

$$\Delta T_\pi = \frac{\int_0^3 A \cdot \exp(-\Delta'^2) d\Delta'}{\int_0^3 T_\pi(\Delta') d\Delta'}, \quad (5.5)$$

where it is integrated over the observed TPC  $dE/dx$  interval  $[0, 3]$ . This is shown for the pion template at  $p = 0.6$  GeV/c in fig. 5.7. The blue histogram

$$h(\Delta) = T_\pi(\Delta) - A \cdot \exp(-\Delta^2) \quad (5.6)$$

corresponds to the pion template after subtracting a Gaussian. The scaling factor  $A$  of the Gaussian is chosen to be the smallest factor, so that

$$h(\Delta) - h(\Delta') + 3\sigma < 0, \quad \text{where} \quad \sigma^2 = h(\Delta) + h(\Delta') \quad (5.7)$$

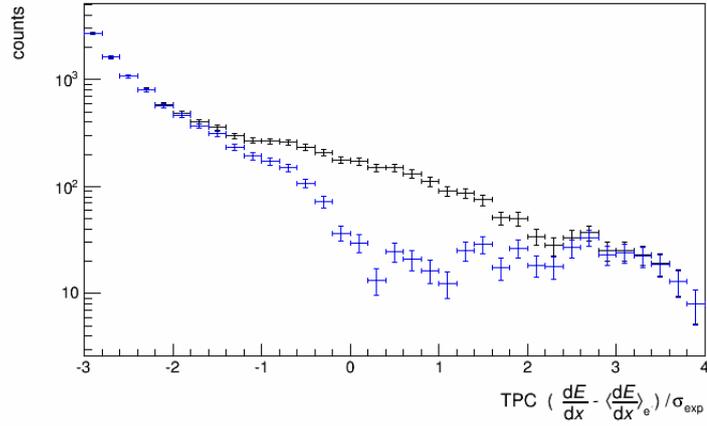


Figure 5.10: The black histogram shows the pion template with electron contamination. The blue histogram shows the pion template subtracted by a Gaussian which has been scaled step by step until the pion tail is no longer monotonically.

for at least one  $\Delta \in [0, 2]$  and at least one  $\Delta' \in [\Delta, \Delta + 0.4]$ . This means that monotony is broken for at least one point. Furthermore, including the propagated uncertainty of the histogram  $\sigma$  with a factor of three assures, that monotony is not broken due to statistical fluctuations. Restricting  $\Delta'$  to be smaller than  $\Delta + 0.4$  makes sure that bins with very small statistics are not taken into account.

This method is applied to all momentum bins up to 2 GeV/c, so that one obtains an approximation of the relative contamination of the pion template  $\Delta T_\pi$  for each momentum bin. The average of these

$$\langle \Delta T_\pi \rangle = (10 \pm 2)\% \quad (5.8)$$

gives an estimate of the contamination of the pion template caused by electrons and will be propagated to the uncertainty of the hadronic background of the analyzed electron sample. The uncertainty of  $\pm 2\%$  refers to the standard deviation of the average.

## 5.8 Contamination at Energy Loss Crossing Lines

Since pions dominate the sample, the fit with a template always converges, because the scaling factor as free parameter can be obtained from the ratio between the pion peaks

from data and template. For other particles this is not possible in regions where the lines of the energy loss of different particles cross (see fig. 2.2 and 4.1). As described in section 5.3, the fit routine will lead to false results or does not converge. Therefore it would be a great advantage if one could fix the scaling factor of distributions in crossing regions.

This can be motivated by the assumption - made in section 5.3 - that the  $3\sigma$  TOF-cuts around the expectation of a particle species should not influence the shape of the energy loss distributions significantly. The scaling factor is only needed to correct for different numbers of particle tracks in the template and the electron sample. Hence, the scaling factor  $A_x$  represents the fraction between the number of tracks of the corresponding particle species in the data sample  $N_{x,D}$  and in the template  $N_{x,T}$  (see [20]). One can assume, that while the number of tracks - both in the template and the electron sample - varies with momentum, the fraction does not, because the variation cancels out. By that, it is reasonable to fix the scaling factor.

In a first step the scaling factor of the proton distribution is determined in momentum bins between 0.6 and 0.82 GeV/c where the distribution of protons is not covered by distributions from other particles. The scaling factor is plotted against the momentum in fig. 5.11 Since the scaling factor is nearly constant, it can be fixed with the average

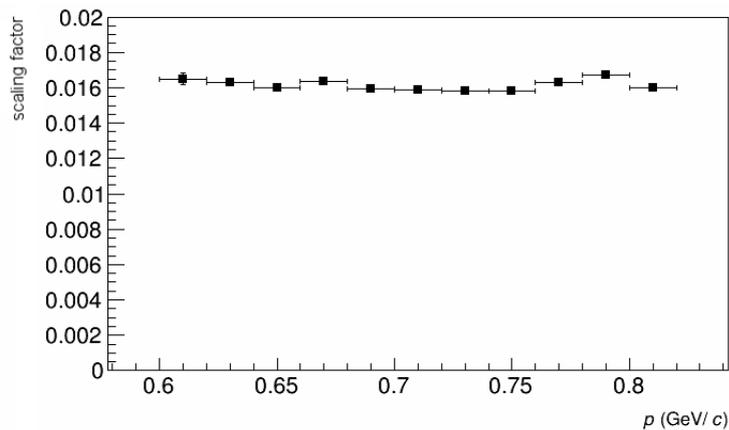


Figure 5.11: Scaling factor of proton distribution at momenta where no lines from energy loss of other particles cross. The scaling factor is in good approximation constant. Errors obtained from the fit routine are too small to be seen.

scaling factor - taken from the factors in fig. 5.11 - for fitting in higher momentum bins, where the proton distribution is covered by electrons. The standard deviation of the shown sample can be propagated to the statistical uncertainty of the contamination.

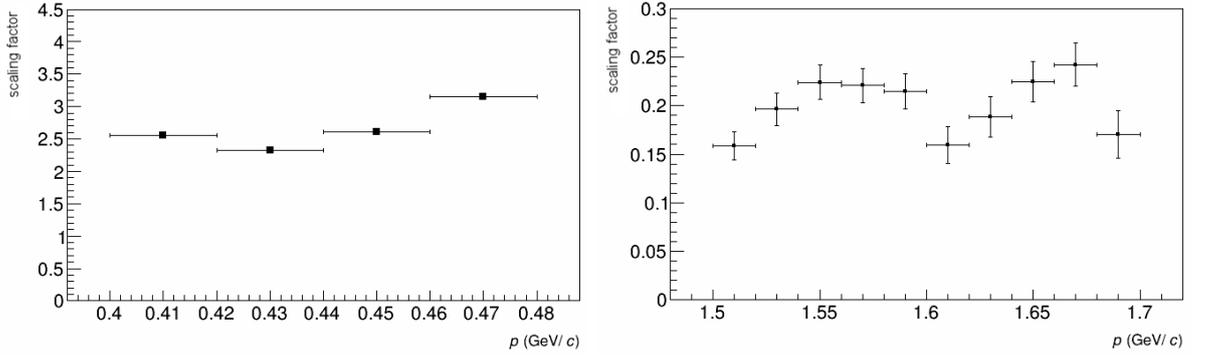


Figure 5.12: Scaling factors of kaon (left) and deuteron (right) templates from fitting in different momentum bins, where the corresponding energy loss distributions are not covered by the electron peak.

A similar approach for kaons and deuterons is desirable. However, as can be seen in fig. 5.12 this is more complicated than for protons. There are only four momentum bins that can be used for obtaining the scaling factor. For  $p < 0.4$  GeV/c there is not enough statistics in the main sample to fit the kaon template and for  $p > 0.48$  GeV/c the kaons are covered by the electron distribution. This is also the reason for the rise of the scaling factor at 0.47 GeV/c. In this momentum bin the scaling factor is overestimated, because only a small part of the template is not covered by electrons and contributes to the total fit.

The average scaling factor could only be taken from those four bins. This results in a larger standard deviation than for the protons.

The momentum range where deuteron templates can be fitted is larger, but there are not many deuterons in the main sample. Therefore statistics are low and fluctuations influence the value for the scaling factor obtained from the fit routine. This can be seen in the large uncertainties for the scaling factor which are returned from the routine. However, within these uncertainties the scaling factor can be assumed to be constant and fixed for fits at higher momenta.

# 6 Results

With the final setup from section 5 the fit routine can be applied to all momentum bins from 0.4 up to 3 GeV/c. The fit function consists of a Gaussian for electrons, a template for pions with varying scaling factor as free parameter and templates for kaons, protons and deuterons with fixed scaling factors. With overall four free parameters this function leads to fast and stable fits.

One could use the same routine for higher momenta, but then the Gaussian for electrons has to be fixed, since the pion distributions will overlap with the electrons more and more with increasing momentum (see fig. 4.1). For momenta higher than 2 GeV/c it is already necessary to introduce limits for the width as fit parameter of the Gaussian. Those limits are taken to be 0.9 for the lower limit and 1.1 for the upper limit, since - due to the choice of transforming the TPC signal in  $n\sigma$  electrons - the width of the electron distribution is expected to be 1. Otherwise the routine would yield not reasonable results.

## 6.1 Contamination in Dependence of Momentum

The fit routine can be applied to every momentum bin. By that one obtains the energy loss distributions for all of the observed particle species separately and it is possible to estimate the corresponding contamination. Therefore, every distribution (represented in a histogram) is integrated over the TPC  $dE/dx$  interval  $[0, 3]$ , which yields the amount of particles of the corresponding type times the width of the TPC  $dE/dx$  bins. The result is divided by the integral with same borders over the data sample, so that the factor corresponding to the width of the bins cancels out and one obtains the relative amount of contamination. For more clarity the fractions of contamination are presented in two plots in fig. 6.1.

As expected the peaks of the fractions are at positions where the energy loss curves from electrons and the corresponding other particle species cross. The kaon contamina-

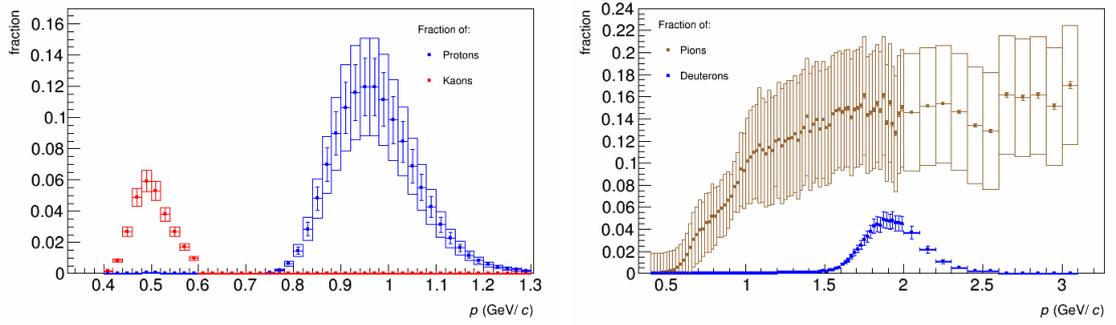


Figure 6.1: Fraction of the contaminating particles in the observed interval  $[0,3] n\sigma_e$  in dependence of momentum. Left: Fraction of protons and kaons. Right: Fraction of pions and deuterons. The momentum binning has been widened for  $p > 2$  GeV/c to obtain higher statistics. The boxes indicate the systematic uncertainty.

tion shows a maximum around 0.5 GeV/c with about 6% and the protons just below 1 GeV/c with approximately 12 – 13%.

Due to the decreasing separation power between electrons and pions with increasing momentum, the TOF detector cannot distinguish them and therefore the pion contamination increases rapidly between 0.5 and 1 GeV/c.

### 6.1.1 Statistical and Systematic Uncertainties

The statistical uncertainty for the fraction of pions is calculated by error propagation of the uncertainty on the estimation of the templates scaling factor from the fit routine. Therefore the error is very small and hardly visible in the plots.

For the other particle species the scaling factor has been fixed, so that there is no uncertainty from the fit routine. Since the fixed value was the average of the results from different fits, the standard deviation was used for error propagation instead.

The systematical uncertainties include two aspects. The first takes into account, that the templates are contaminated themselves as it was shown in section 5.3 and in fig. 5.5.

For proton, kaon and deuteron templates this is not critical. Even the amount of pions in those templates is small enough and does not contribute relevantly to the estimated contamination of protons, kaons and deuterons. However, the pion template contains a relevant amount of protons and kaons, since the TRD rejected electrons only. This leads to an overestimation of pions and an underestimation of the other particles.

Therefore, another pion template is created and this time the TRD uses the LQ2D method to reject also kaons and protons with a likelihood larger than 10%. The problem with this method, is that none of those particles produces transition radiation, so that the TRD can distinguish them only by their energy loss. This becomes problematic in the momentum regions where particles have similar energy losses. Therefore, using the TRD to cut out protons and kaons in those regions will also cut out pions, so this pion template will underestimate the amount of pions. However, the fit routine compensates the cut out pions with a higher scaling factor of the proton or kaon template respectively, which leads to an overestimation of the other particles.

With this information, the systematic uncertainty due to the kaon and proton contamination in the pion template can be calculated. Fitting with the new pion template yields another estimation of contamination similar to the one shown in fig. 6.1, but with less pion and more proton and kaon contamination. The routine can only be run up to 1.2 GeV/c, since then the TRD cannot distinguish protons and pions. However, it is possible to get the differences  $(\Delta f_x)_{\text{dif}}$  in the fractions between the first fit and this one for each kind of particle (except deuterons).

This difference approximates the uncertainty due to the kaon and proton contamination in the pion template. For this reason, the discussed over- and underestimation of proton and kaon contamination is included in the systematic uncertainties (of protons and kaons) which are set to be  $(\Delta f_x)_{\text{dif}}$ .

This will be also taken into account for calculating the systematic uncertainty of the pion contamination, but in this case the electron contamination of the pion template has to be included, too. For this purpose the averaged relative electron contamination  $\langle \Delta T_\pi \rangle$  - estimated in section 5.7 - has to be propagated as follows

$$((\Delta f_\pi)_{\text{cont}})^2 = ((\Delta f_\pi)_{\text{dif}})^2 + (\langle \Delta T_\pi \rangle \cdot f_\pi(p))^2 \quad (6.1)$$

As the second template works only up to 1.2 GeV/c, for higher momentum  $(\Delta f_\pi)_{\text{dif}}$  has been chosen to be the maximum of  $|(\Delta f_\pi)_{\text{dif}}|$  in the range  $p \leq 1.2$  GeV/c.

The second aspect - contained in the systematic uncertainties - is the fact that fit routine does not fully describe the pion tail as can be seen in fig. 5.9. This leads to an underestimation of the pions. To approximate the uncertainty due to this underestimation, the contamination is estimated once again for all momentum bins below 0.9 GeV/c with the pion template, which is obtained by rejecting protons, kaons and electrons.

This time the fit-range is chosen to be the TPC  $dE/dx$  interval  $[-1, 15]$  as described in section 5.6. Then the contaminations of all particle species are summed up to  $f(p)_{\text{over}}$ .

Since the results from the beginning of this section - shown in fig. 6.1 - underestimate the total contamination, the difference to the overestimated contamination  $f(p)_{\text{over}}$  gives a good approximation to the systematical uncertainty. However,  $f(p)_{\text{over}}$  is determined for  $p < 0.9$  GeV/c, because the pion template - used for this purpose - only works for low momenta as described previously. Hence, the systematic uncertainty of the pion contamination  $f_\pi$  is calculated as

$$(\Delta f_\pi)_{\text{sys}} = \sqrt{((\Delta f_\pi)_{\text{cont}})^2 + (\langle |f(p)_{\text{over}} - f(p)_{\text{under}}| \rangle)^2}, \quad (6.2)$$

where

$$\langle |f(p)_{\text{over}} - f(p)_{\text{under}}| \rangle = \sum_p \frac{|f(p)_{\text{over}} - f(p)_{\text{under}}|}{N} \approx 1.8\% \quad (6.3)$$

is the average difference between the overestimated and the underestimated fraction (note: this is an absolute difference). It is summed over all  $N = 25$  momentum bins from 0.4 GeV/c to 0.9 GeV/c.

The systematic uncertainties of the proton, kaon and deuteron contamination are neither influenced by the electron contamination of the pion template nor by the effect of the underestimation of the pion tail. Therefore, the systematic uncertainties of the contamination of this particle species are set to be the differences  $(\Delta f_x)_{\text{dif}}$  between the respective contaminations which are estimated by using different pion templates as discussed previously.

## 6.1.2 Total Contamination

For most Heavy Flavor analyses it is more important to know the total hadronic contamination than the contamination of single particle species. Therefore, the problem that the pion template itself is contaminated by protons and kaons becomes less crucial, because they also contribute to the total hadronic contamination. Thus the contamination of the pion template caused by electrons and the underestimated pion tail are the only relevant systematic uncertainties of the total hadronic contamination, which is calculated as

$$(\Delta f_{\text{tot}})_{\text{sys}} = \sqrt{((\Delta f_\pi)_{\text{dif}})^2 + (\langle |f(p)_{\text{over}} - f(p)_{\text{under}}| \rangle)^2}, \quad (6.4)$$

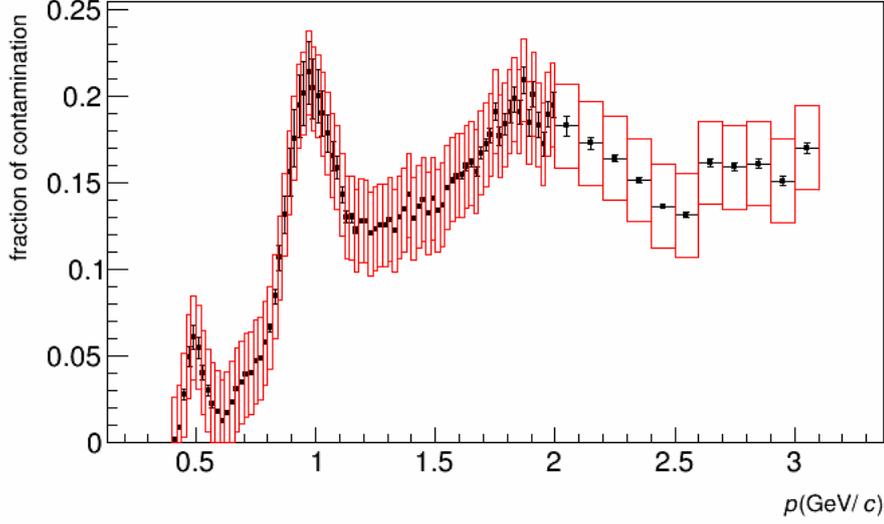


Figure 6.2: Total hadronic contamination in dependence of the momentum. The red boxes indicate the systematic uncertainty.

where  $(\Delta f_\pi)_{\text{dif}}$  is the same as in equation (6.2) and  $\langle |f(p)_{\text{over}} - f(p)_{\text{under}}| \rangle$  is given by (6.3). For momenta above 1 GeV/c  $(\Delta f_{\text{tot}})_{\text{sys}}$  has a value about 2.4% (note: this is an absolute uncertainty).

The total hadronic contamination with the systematic uncertainties calculated as in (6.4) can be seen in fig. 6.2.

## 6.2 Contamination in Dependence of Transverse Momentum

Since for some analyses the contamination in dependence of the transverse momentum is important, in this chapter the estimated contamination is converted from  $p$  to  $p_T$ . This is done by integrating the fraction  $f_x(p)$  over the observed pseudorapidity range  $-0.8 < \eta < 0.8$  weighted with the phase space density  $\rho(\eta)$  of  $\eta$

$$f(p) \rightarrow \tilde{f}(p_T) = \int_{-0.8}^{0.8} d\eta f(p(\eta, p_T)) \rho(\eta). \quad (6.5)$$

This neglects, that the separation power of the detectors - used for obtaining the templates and data - depends on the pseudorapidity. This dependence would have influence on the contamination, so that  $f \equiv f(p(\eta, p_T), \eta)$  becomes directly dependent of  $\eta$ . However, in the following it is assumed, that this effect is negligible and therefore one considers only the dependence of  $p$ ,  $p_T$  and  $\eta$  which is given by

$$p(\eta, p_T) = p_T \cosh(\eta). \quad (6.6)$$

The phase space density  $\rho(\eta)$  can be obtained from the phase space density of the rapidity  $\rho(y)$ . Therefore it is assumed that the density of rapidity is nearly constant in the observed region. For more precise results one could actually use a measured pseudorapidity distribution. One could also check the difference between the assumed constant rapidity distribution and a measured one to obtain an estimate of the uncertainty caused by assuming  $\rho(y) = \text{const.}$

To obtain  $\rho(\eta)$  from  $\rho(y)$  the Jacobian is needed

$$\rho(\eta) = \rho(y) \frac{dy}{d\eta}, \quad (6.7)$$

where

$$\frac{dy}{d\eta} = \left( \sqrt{m^2 + p_T^2 \cosh^2(\eta)} + p_T \sinh(\eta) \right)^{-1} \left( \frac{p_T^2 \cosh(\eta) \sinh(\eta)}{\sqrt{m^2 + p_T^2 \cosh^2(\eta)}} + p_T \cosh(\eta) \right). \quad (6.8)$$

From this one can normalize  $\rho(\eta)$  and compute the integral in (6.5) numerically. The results can be seen in fig. 6.3.

In comparison to the momentum dependence of contamination in fig. 6.1 the distributions for kaons, protons and deuterons are widened and smaller than before. This can be confirmed by comparing the proton template in fig. 5.2 with the one in fig. 5.3. In  $p_T$  bins the templates are much wider than in  $p$  bins as discussed in section 5.2. Therefore the proton distribution will overlap for more  $p_T$  bins with the electron distribution than for  $p$  bins. However, the total contamination of protons over all bins has to be the same in  $p$  and  $p_T$ , so a wider distribution yields a smaller maximum. The same is true for kaons and deuterons. Also for pions one can see that contamination has spread over a larger interval.

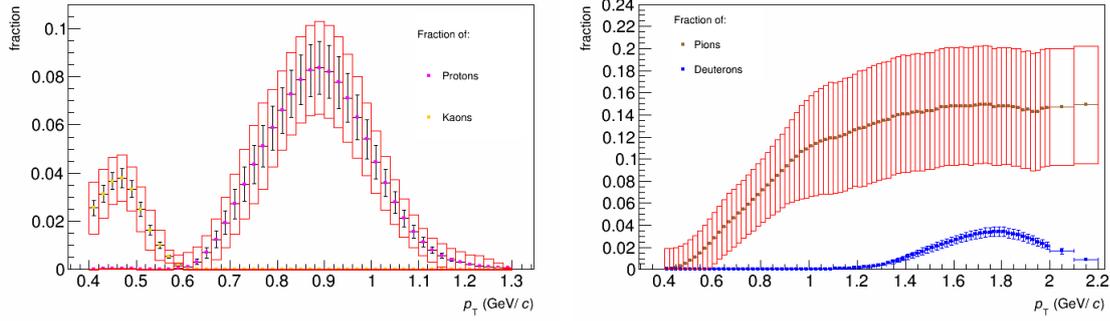


Figure 6.3: Fraction of the contaminating particles in the observed interval  $[0,3] n\sigma_e$  in dependence of transverse momentum. The red boxes indicate the systematic uncertainty.

Statistical uncertainties from the contamination in  $p$  are propagated as

$$(\Delta \tilde{f}_x)(p_T) = \left( \int_{-0.8}^{0.8} d\eta ((\Delta f_x)(\eta, p_T) \cdot \rho(\eta))^2 \right)^{1/2}. \quad (6.9)$$

Since the systematic uncertainties of  $f_x(p_i)$  and  $f_x(p_k)$  are correlated for momentum bins  $p_i$  and  $p_k$  which are close to each other, the systematic uncertainty has to be propagated as

$$(\Delta \tilde{f}_x)_{\text{sys}}(p_T) = \int_{-0.8}^{0.8} d\eta (\Delta f_x)_{\text{sys}}(\eta, p_T) \cdot \rho(\eta). \quad (6.10)$$

### 6.3 Obtaining Contamination as a Smooth Function

For practical purposes one might prefer to have the amount of contamination as a smooth function of momentum or transverse momentum respectively. A function would be easier to implement in other analyses for instance. Since it is very hard to calculate an analytic expression for the contamination, the function is approximated by fitting a generalized normal distribution to the obtained data in fig. 6.1. Although the generalized normal distribution might not be physically correct, it leads to a good description of data and can be used in analyses if one needs to implement the estimated contamination in a simple but efficient way. It is a good approximation, because the difference between the fitted function and data is not significant compared to the systematic uncertainties.

The generalized normal distribution can be defined as

$$f(x) = \frac{\exp\left(\frac{-y^2}{2\sigma^2}\right)}{\sigma\sqrt{2\pi}(\alpha - \kappa(x - \mu))}, \quad \text{where } y = \begin{cases} \frac{x-\mu}{\alpha} & \text{if } \kappa = 0 \\ -\frac{1}{\kappa} \log\left(1 - \frac{\kappa(x-\mu)}{\alpha}\right) & \text{else} \end{cases} \quad (6.11)$$

The Greek letters refer to free parameters which will be obtained by the fit routine. Fitting this function to the estimated results for protons, kaons and deuterons leads to good approximations of the corresponding contamination as can be seen in fig. 6.4. The fit provides the necessary parameters of the generalized normal distribution in (6.11). The parameters are presented in table 6.1, so that it is possible to reproduce the contamination for protons, kaons and deuterons as smooth function.

To describe the contamination due to pions, a modified Fermi distribution has been chosen

$$f(x) = N \cdot \left(1 - \frac{1}{\exp\left(\frac{x-\mu}{T}\right) + 1}\right). \quad (6.12)$$

The parameters  $N$ ,  $\mu$ , and  $T$  are obtained from fitting this distribution to the fraction of pion contamination, which can be seen in fig. 6.4 and table 6.2.

In this setup one is able to describe the total hadronic contamination as the sum of three generalized normal distributions (6.11) and the modified Fermi distribution (6.12) using the estimated parameters. For this purpose one can assume a systematic uncertainty of 2.4% (note: this is an absolute value) on the estimated fraction of contamination for every momentum (see sec. 6.1.2).

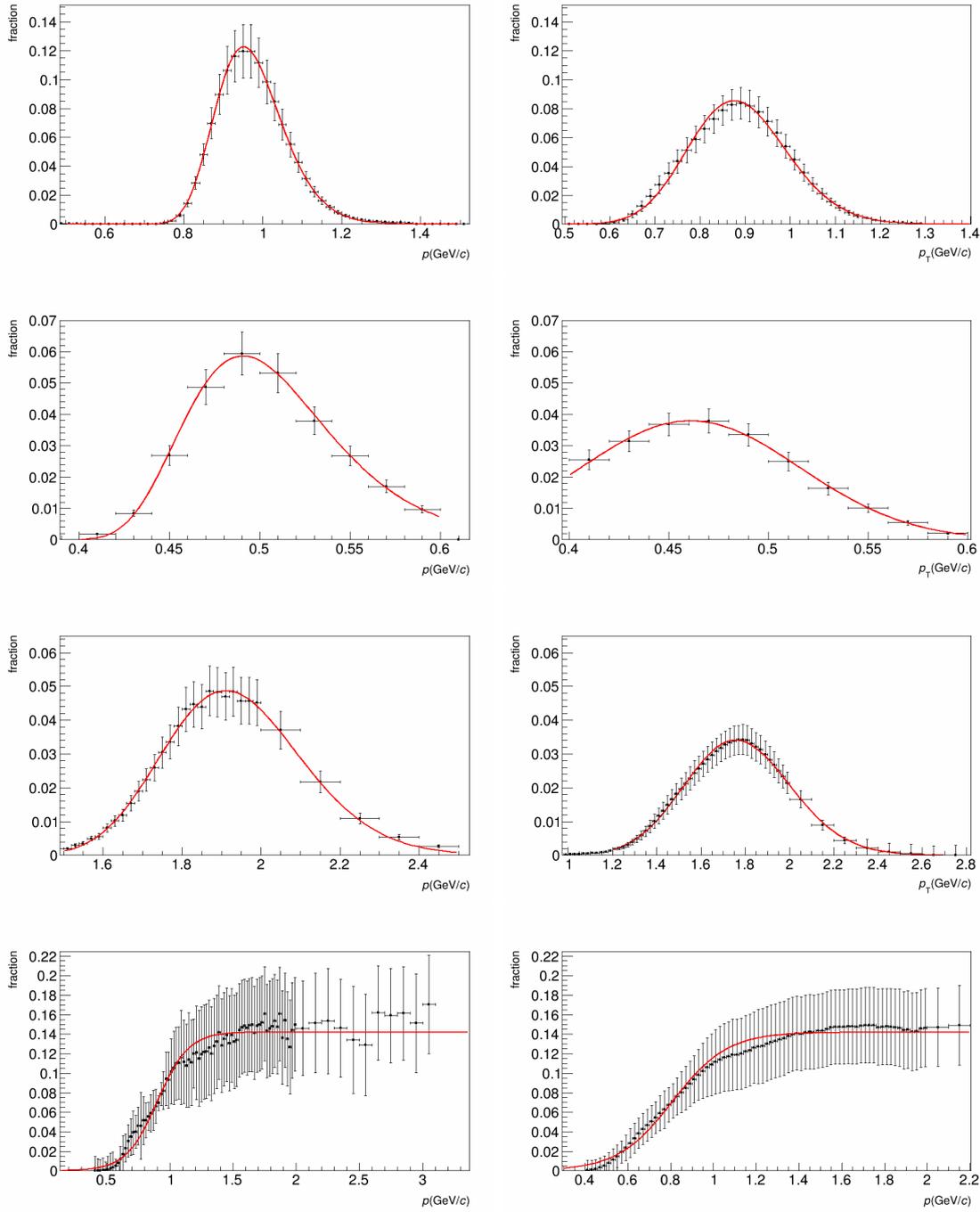


Figure 6.4: Fitting generalized normal distributions to the estimated contamination of protons (1st row), kaons (2nd row), deuterons (3rd row) and fitting a modified Fermi distribution to the contamination of pions (4th row). The left plots show the fitted contamination in dependence of momentum; the right ones in dependence of transversal momentum.

$p :$				
species	$\mu$	$\alpha$	$\kappa$	$\sigma$
protons	$9.66 \cdot 10^{-1}$	8.3	-17	$1.03 \cdot 10^{-2}$
kaons	$5.02 \cdot 10^{-1}$	18	-120	$2.4 \cdot 10^{-3}$
deuterons	1.93	21	-13	$8.5 \cdot 10^{-3}$

$p_T :$				
species	$\mu$	$\alpha$	$\kappa$	$\sigma$
protons	$8.85 \cdot 10^{-3}$	11.7	-8.8	$9.3 \cdot 10^{-3}$
kaons	$4.6 \cdot 10^{-1}$	26	1.52	$2.1 \cdot 10^{-3}$
deuterons	1.77	29	-6.1	$8.1 \cdot 10^{-3}$

Table 6.1: Parameters of generalized normal distributions of equation (6.11), which has been fitted to data represented in fig. 6.4. The distributions approximate the fraction of protons, kaons and deuterons contaminating the electron sample. The first table shows the parameters for the contamination in dependence of momentum; the second table shows the parameters for the contamination in dependence of transverse momentum.

pions	$N$	$\mu$	$T$
$p$	$1.42 \cdot 10^{-1}$	0.9	$1.2 \cdot 10^{-1}$
$p_T$	$1.42 \cdot 10^{-1}$	0.8	$1.3 \cdot 10^{-1}$

Table 6.2: Parameters of modified Fermi distribution of equation (6.12) to describe pion contamination in dependence of momentum and transverse momentum.

# 7 Summary and Outlook

In this thesis a routine to estimate the hadronic contamination in an electron sample has been developed. Particle identification is done by fitting measured energy loss distributions of different particle species to data of the electron sample. Those distributions - called templates - are obtained by selecting the particular particle species by cuts on the TOF signal. Particularly the distribution of pions is selected with the TRD.

The routine scales the templates to fit them to data. Since the scaling is nearly constant over a large momentum range for protons, kaons and deuterons, the scaling factor can be fixed to its average value. This makes it possible to estimate the contamination in momentum intervals, where the energy loss distributions of electrons and other particles overlap completely. Furthermore, fixing the scaling factor leads to a more stable and faster fit routine.

Data from Monte Carlo simulations are used to check how strongly the templates are contaminated themselves. In fact the contamination of the proton and the kaon template is negligible. Since deuterons are not implemented in those Monte Carlo simulations, the contamination of their template cannot be checked this way. The same holds for the pion template, because the TRD response is not implemented sufficiently well in Monte Carlo simulations. However, the uncertainty on the hadronic contamination due to an impure pion template is approximated by using templates with different cuts to obtain an overestimated and an underestimated fraction of hadronic contamination.

The main result of this thesis is the estimated hadronic contamination in dependence of momentum and transverse momentum (fig. 6.1 and fig. 6.3). While the contamination depending on momentum is obtained directly from the routine, the contamination depending on transverse momentum is obtained by conversion from  $p$  to  $p_T$  and integrating over the observed pseudorapidity range.

Furthermore, the momentum dependence of the contamination (transverse momentum respectively) can be approximated by smooth functions, which have been fitted to the estimated contamination. Therefore, Heavy Flavor electron analyses can make use of

the knowledge about the contamination in a simple way. They just have to implement the functions as described in chapter 6.3.

However, there is still something to improve. The contamination of the pion template could be checked via Monte Carlo simulations. Therefore the TRD response has to be implemented and the simulations have to reproduce the actual TRD signal well.

Furthermore, using the TRD leads to a rejection of pions in the corresponding template as discussed in chapter 5.6. It seems to be the tracks that make up the pion tail, which are more strongly rejected than other pions. This leads to an underestimation of contamination. Therefore it could be checked if there is a correlation between the TRD signal and the TPC signal of the rejected tracks. This could imply that the rejected tracks - which have a higher energy loss in the TPC than expected for pions - would also have a higher energy loss in the TRD because of overlapping clusters in both detectors.

On the other hand, one could also check whether the pion tracks - that contain overlapping clusters - are rejected because a TRD signal for the tracks is required. This could imply that the matching of those tracks when leaving the TPC and entering the TRD is problematic, due to a bad resolution which could be a consequence of the overlapping clusters.

It is also possible that the pion template could be improved by having more tracks making up the pion distribution. This can be achieved by analyzing data, which have been taken after 2010, when more supermodules of the TRD were installed.

As a conclusion, one can say that a major improvement of the estimated hadronic contamination has to include a correction of the pion template. However, removing the pion tail in the electron sample, could also improve the results. The pion template may then be replaced by the truncated Landau distribution.

## Acknowledgments

I would like to thank Prof. Dr. Johanna Stachel for making it possible to work out my Bachelor thesis in the ALICE group. Many thanks to Martin Völkl for his great support and guidance during the whole time I was working in this group and to Michael Winn for his great explanations and having a look at my work. Thanks to Dr. Yvonne Pachmayer, Jochen Klein and Sebastian Klewin for their valuable advice. Furthermore I would like to thank Lennart Volz and Manuel Wittner for their inspiring ideas. Also many thanks to Nicolas Schmidt and Heinrike Braß for having a look at my thesis.

# Bibliography

- [1] ALICE Collaboration, “ALICE Physics Performance Report Volume I Part 1”, J. Phys. G: Nucl. Part. Phys. 30 (2004) 1517–1763
- [2] M. Völkl, Masterthesis “Study of the Transverse Momentum Spectra of Semielectronic Heavy Flavor Decays in pp-Collisions at  $\sqrt{s} = 7$  TeV”, 2012
- [3] ALICE Collaboration, “Beauty production in pp collisions at  $\sqrt{s} = 2.76$  TeV measured via semi-electronic decays”, Phys.Lett B738, pp. 97-108, arXiv:1405.4144 [nucl-ex] May 16th 2014
- [4] Particle Data Group, “Review of Particle Physics”, IOP Publishing, 2010
- [5] A. Kalweit, “Particle Identification in the ALICE Experiment”, J.Phys G38, p. 124073, arXiv:1107.1514 [hep-ex], Jul 7th 2011
- [6] N. Brambilla et. al., “QCD and strongly coupled gauge theories: challenges and perspectives”, Eur.Phys.J. 10 C74, p. 2981, arXiv:1404.3723 [hep-ph], May 18th 2014
- [7] H. Bichsel “A method to improve tracking and particle identification in TPCs and silicon detectors”, Nucl.Instrum.Meth. A562, pp. 154–197, March 2006
- [8] ALICE Collaboration, ALICE Physics Performance Report Volume II Part 1
- [9] C. Lippmann, “Particle Identification”, Nucl.Instrum.Meth. A666, pp. 148-172, arXiv:1101.3276v4 [hep-ex] Jun 12th 2011
- [10] J. Alme et al., “The ALICE TPC, a large 3-dimensional tracking device with fast readout for ultra-high multiplicity events.”, Nucl.Instrum.Meth. A622, pp. 316-367, Oct 2010
- [11] ALICE Collaboration, “Performance of the ALICE Experiment at the CERN LHC”, Int.J.Mod.Phys. A29, p. 1430044, arXiv:1402.4476v2 [nucl-ex] Mar 14th 2014

- [12] [http://aliceinfo.cern.ch/Public/Objects/Chapter2/DetectorComponents/TRD\\_electronId.gif](http://aliceinfo.cern.ch/Public/Objects/Chapter2/DetectorComponents/TRD_electronId.gif) access on Sep 24th 2014 12:40am
- [13] [http://www-alice.gsi.de/trd/tests/trd\\_prin.jpg](http://www-alice.gsi.de/trd/tests/trd_prin.jpg) access on Sep 24th 2014 12:40am
- [14] J. Klein “Commissioning of and Preparations for Physics with the Transition Radiation Detector in A Large Ion Colider Experiment at CERN”, Diploma thesis 2008
- [15] Abelev et. al., “Measurement of electrons from semileptonic heavy-flavour hadron decays in pp collisions at  $\sqrt{s} = 7$  TeV”, Physical Review D 86, 112007 (2012)
- [16] Abelev et. al., ALICE Collaboration, “Measurement of electrons from semileptonic heavy-flavor hadron decays in pp collisions at  $\sqrt{s} = 2.76$  TeV”, arXiv:1405.4117 [nucl-ex], May 16th 2014
- [17] Abelev et al., “Suppression of high transverse momentum D mesons in central Pb–Pb collisions at  $\sqrt{s_{NN}} = 2.76$  TeV”, JHEP 1209, p. 112, arXiv:1203.2160 [nucl-ex], Oct 12th 2012
- [18] J.Schukraft, “Heavy-ion physics with the ALICE experiment at the CERN Large Hadron Collider” Royal Society of London Philosophical Transactions Series A 370, Feb 2012, pp. 917-932, arXiv 1109.4291 [hep-ex]
- [19] W. Blum, L. Rolandi, “Particle detection with drift chambers”, Springer-Verlag, second printing 1994
- [20] R. Barlow “Fitting using finite Monte Carlo samples”, Computer Physics Communications 77 (1993) 219—228
- [21] [itp.tugraz.at/LV/boeri/NUM\\_METH/SKRIPT/kapitel4.pdf](http://itp.tugraz.at/LV/boeri/NUM_METH/SKRIPT/kapitel4.pdf) access on Nov 5th 2014 9:43pm

# Erklärung

Ich versichere, dass ich diese Arbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Heidelberg, den 6.11.2014,