Department of Physics and Astronomy
University of Heidelberg

Bachelor Thesis in Physics
submitted by

Sebastian Schenk

born in Gifhorn (Germany)

2013

A Vertex Fit for Low Momentum Particles in a Solenoidal Magnetic Field with Multiple Scattering

This Bachelor Thesis has been carried out by Sebastian Schenk at the
Physikalisches Institut in Heidelberg
under the supervision of
Dr. Niklaus Berger

Zusammenfassung

Da Neutrinooszillationen beobachtet wurden, ist es offenkundig, dass Lepton Flavour Erhaltung im Standardmodell im Bereich neutraler Leptonen verletzt ist. Viele neue Modelle enthalten auch Lepton Flavour Verletzung im Bereich geladener Leptonen, eine experimentelle Beobachtung gelang bis jetzt jedoch nicht.

Das Mu3e Experiment wird nach dem Lepton Flavour verletzenden Myon Zerfall $\mu \rightarrow eee$ suchen, welcher im Standardmodell stark unterdrückt ist. Dazu nutzt es ein neuartiges Design bestehend aus vier dünnen zylindrischen Lagen von Pixeldetektoren in einem homogenen Magnetfeld. Aufgrund von geringen Impulsen der Elektronen ist die Vielfachstreuung in der ersten Detektorlage bedeutsam.

Im Kontext des Mu3e Experiments habe ich einen neuen linearisierten dreidimensionalen Vertex Fit untersucht, der die Vielfachstreuung als einzige Fehlerquelle berücksichtigt. Auf Basis einer Simulation wurde das Rekonstruktionsverhalten im Hinblick auf Effizienz, Vertex Auflösung, Impuls- und Algorithmus-Konfigurations-Abhängigkeiten, Unterdrückung von Untergrund sowie invariante Massenauflösung ausgewertet.

Der neue Ansatz für die Vertex Rekonstruktion funktioniert und erreicht eine Vertex Auflösung von $200\,\mu$m. Desweiteren kann möglicher Untergrund durch die Rekonstruktion von einem gemeinsamen Vertex hervorragend unterdrückt werden. Die Ergebnisse zeigen, dass die Vertexrekonstruktion elementar ist, um die geplante Sensitivität eines Verzweigungsverhältnisses von 1×10^{-16} zu erreichen.

Abstract

Since neutrino oscillations have been observed, it is evident that lepton flavour conservation as predicted by the original Standard Model is violated in the neutral lepton sector. Many new physics models include lepton flavour violation in the charged sector as well, but an experimental discovery is still missing.

The Mu3e experiment will search for the lepton flavour violating muon decay $\mu \rightarrow eee$, which is heavily suppressed in the SM. It uses a novel design involving four thin cylindrical layers of pixel detectors in a solenoidal magnetic field. Due to low momentum electrons, multiple scattering in the first detector layer is significant.

In the context of the Mu3e experiment, I investigated a new linearised three-dimensional vertex fit, treating multiple scattering as the only source of uncertainty. Its performance is evaluated in terms of reconstruction efficiency, vertex position resolution, particle momentum and algorithm configuration dependencies, background suppression and invariant mass resolution on the basis of a simulation.

The new approach of vertex reconstruction works and achieves a vertex position resolution of $200\,\mu$m. Moreover, accidental background can be greatly suppressed by the reconstruction of a common vertex. The results show that the linearised vertex fit is essential to achieve the proposed branching ratio sensitivity of 1×10^{-16}.

Contents

List of Figures

Chapter 1

Introduction

The recent discovery of a Higgs boson by ATLAS and CMS at the Large Hadron Collider in 2012 can be characterized as the "final ingredient" of the most successful theory of particle physics: the Standard Model [1].

Despite its great success in predicting and explaining many experimental results in particle physics there are still some phenomena lacking a consistent and satisfying explanation by the Standard Model, e.g. neutrino oscillations or the missing unification with gravity. This motivates many new theories in particle physics, in turn predicting several other new phenomena often referred to as *physics beyond Standard Model*. However, none of these phenomena has been observed so far. Thus, other experimental signs for new physics theories are still missing.

The current experimental frontier in the search for these signs can be separated into two different strategies. One approach is to increase energies in particle interactions and directly search for the appearance of new particles. This is the approach of the Large Hadron Collider.

Another approach is to search for hints at low energies but with high intensities, i.e. small deviations from known quantities as given by the Standard Model. In order to achieve high precision, high statistics and high intensities are required.

This is the approach of the Mu3e experiment. Mu3e will search for the rare lepton flavour violating muon decay $\mu \rightarrow eee$, which is heavily suppressed in the Standard Model. The goal of Mu3e is to search for $\mu^+ \rightarrow e^+e^-e^+$ with a branching ratio sensitivity of 1×10^{-16} [2]. The current experimental limit for the branching ratio is given by the SINDRUM experiment with 1.0×10^{-12} at 90% confidence level [3].

A high-intensity muon beam is stopped and the decay electrons are detected by pixel sensors with high precision. The upper limit for each particle momentum is given by half the muon mass, leading to low momenta and consequently highly bent tracks due to Lorentz forces in a solenoidal magnetic field.

One obvious characteristic of the $\mu \rightarrow eee$ decay is the fact that the decay products originate in a common vertex. For primary vertex reconstruction, the helical particle trajectories have to be extrapolated through the innermost detector layer into the vertex region. However, low momenta lead to strong deflection of the particles due to multiple Coulomb scattering when traversing the detector material. Therefore, the reconstruction of a common vertex is not trivial.

In this thesis a linearised three-dimensional vertex fit for low momentum particles in a solenoidal magnetic field is described. This is a completely new approach in vertex reconstruction because of the assumption that the particle momenta and corresponding track parameters are precisely known due to measurements with modern pixel sensors. Moreover, multiple Coulomb scattering in the first detector layer will be treated as the only source of uncertainty. The performance of this vertex reconstruction algorithm will be analysed in the context of the Mu3e experiment by performing a computer simulation.

The first part of this thesis in chapter 2 will provide a short theoretical introduction to the Standard Model of particle physics in general and the lepton flavour violating muon decay $\mu \rightarrow eee$ in particular. In chapter 3 the kinematics of $\mu \rightarrow eee$ will be derived and the resulting challenges for the Mu3e experiment will be described, concluding with the proposed detector design. Chapter 4 contains a detailed derivation and description of the track model and the vertex reconstruction algorithm itself. Finally, the implementation of the computer simulation and corresponding results on the performance of the algorithm are presented in chapters 5 and 6.

Chapter 2

Theoretical Background

The Standard Model of particle physics is sometimes characterized as *the theory of almost everything* [4], because of its great ability to explain most experimental results in particle physics and for having predicted many new phenomena that were later experimentally observed.

Lepton flavour – a quantum number associated with elementary particles called leptons – is predicted to be a conserved quantity in particle interactions described by the Standard Model [5]. Nevertheless, recent experiments showed that this conservation is violated by neutrino oscillations. A violation in charged lepton interactions on the other hand has not been observed so far. Therefore, the search for the lepton flavour violating muon decay $\mu \to eee$ is a well motivated approach for tackling this obvious asymmetry. Furthermore, many new physics models predict lepton flavour violation in the charged sector [6].

Keeping this in mind, I will introduce the Standard Model of particle physics in general and rare lepton flavour violating muon decay channels in particular in this chapter to provide a theoretical basis for the rest of my thesis.

2.1 The Standard Model of Particle Physics

The Standard Model of particle physics (henceforth referred to as SM) is a quantum field theory, describing all known fundamental elementary particles and the corresponding interactions between them (see e.g. [5]). It contains three generations of matter particles, i.e. fermions, and the gauge bosons mediating the different interactions.

In detail there are twelve elementary matter particles - six leptons and quarks respectively - and four gauge bosons as seen in Figure 2.1.

The recently discovered Higgs particle is separated, being neither matter particle, nor mediating gauge interactions.

The SM describes just three of the four known fundamental interactions,

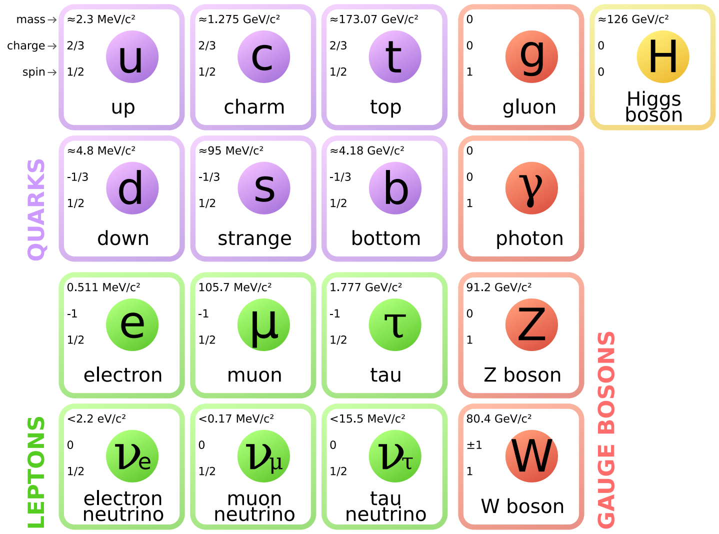


Figure 2.1: Standard Model of elementary particles and their respective quantum numbers. The fermions are separated into quarks and leptons in three generations of matter [7].

namely the strong, the weak and the electromagnetic one. Since it is the weakest one on typical energy scales in particle interactions, the gravitational force can be neglected[1].

In general, every elementary particle in the SM is characterized by its mass and several quantum numbers, e.g. electric charge (in units of elementary charge e), spin, weak hypercharge, color charge and flavour quantum numbers [5]. Some of these characteristic numbers are shown in Figure 2.1.

Each generation of leptons[2] i can be associated with a specific lepton flavour quantum number L_i, i.e. one can define the electronic, muonic and tauonic lepton flavour L_e, L_μ and L_τ respectively. In detail, matter is always associated with the lepton flavour $L_i = 1$, while antimatter is associated with $L_i = -1$, e.g. an electron has lepton flavour $L_e = 1$ and its antimatter partner the positron has lepton flavour $L_e = -1$.

The sum of all lepton flavours is defined to be the lepton number

$$L := \sum_i L_i \,.$$

A lepton this thesis will deal with is the muon μ. It is a second generation particle with a mass of about $m_\mu = 105.67\,\mathrm{MeV}/c^2$, a charge of $q_\mu = -e$, a spin[3] of 1/2 [8] and a lepton flavour of $L_\mu = 1$.

---

[1] A consistent and satisfactory theory for unifying gravity with all other fundamental interactions is still missing.

[2] The generations are $(e,\, \nu_e)$, $(\mu,\, \nu_\mu)$ and $(\tau,\, \nu_\tau)$ as seen in Figure 2.1.

[3] The spin is always given in units of the reduced Planck constant $\hbar$.

Each lepton flavour quantum number L_i is predicted to be conserved[4] in every particle interaction by the laws of the original Standard Model, i.e.

$$L'_i = L_i \quad \forall \ i \in \{e, \mu, \tau\} \ , \tag{2.1}$$

where the primed quantity denotes the final state after the interaction.

An example for lepton flavour conservation is the so called Michel decay $\mu^+ \to e^+ \nu_e \bar{\nu_\mu}$, which is the common decay of a muon via weak interaction. According to Equation 2.1 lepton flavour is conserved in this decay, because on the one hand the initial state yields $L_e = 0$ and $L_\mu = -1$ and on the other hand the final state yields $L'_e = 0$ and $L'_\mu = -1$.

2.2 Charged Lepton Flavour Violation

Since neutrino oscillations have been observed, it is evident that lepton flavour conservation as predicted by the original SM is violated in the neutral lepton sector. This indicates that lepton flavour might be violated in the charged lepton sector as well.

Table 2.1 shows possible lepton flavour violating decay channels of the muon with an experimental upper bound for the branching ratio.

Decay channel	B_{exp}	CL	Experiment
$\mu^+ \to e^+ \gamma$	$< 5.7 \times 10^{-13}$	90 %	MEG [9]
$\mu^+ \to e^+ e^+ e^-$	$< 1.0 \times 10^{-12}$	90 %	SINDRUM [3]
$\mu^- Au \to e^- Au$	$< 7 \times 10^{-13}$	90 %	SINDRUM II [10]

Table 2.1: Possible lepton flavour violating decay channels of the muon with an experimental upper bound for the branching ratio B given at a specific confidence level CL.

The one channel I will concentrate on is the decay of an antimuon into one electron and two positrons

$$\mu^+ \to e^+ e^- e^+ \ .$$

Lepton flavour is violated in this decay, because in the initial state the corresponding lepton flavours are $L_e = 0$ and $L_\mu = -1$ while in the final state $L'_e = -1$ and $L'_\mu = 0$. This is a clear violation of Equation 2.1.

The specific lepton flavour violating decay channel $\mu^+ \to e^+ e^- e^+$ can be induced by neutrino mixing as seen in Figure 2.2, but does not occur in the description by the SM.

By reference to the processes in the loop it can be explained why this decay channel is greatly suppressed in the SM.

---

[4] This implies conservation of the lepton number L as well.

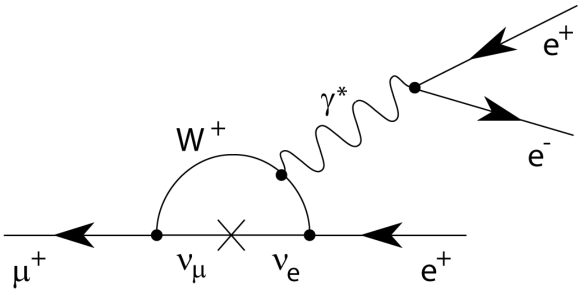


Figure 2.2: Feynman diagram for lepton flavour violating decay $\mu \rightarrow eee$ on loop-level in the SM via neutrino mixing [2].

The channel's branching ratio B is scaling with the fourth order of the mass ratio, i.e. $B \propto \left(\frac{\Delta m_\nu}{m_W}\right)^4$ [11, 12]. The mass differences Δm_ν due to neutrino mixing are extremely small compared to the weak scale, considering a W boson mass of $m_W \approx 80.39\,\mathrm{GeV}$ [8]. Thus, the branching ratio will be $B < 10^{-50}$ [12], eluding any practical measurement.

Many new physics models include LFV reactions in the charged sector and any observation of $\mu \rightarrow eee$ would be an obvious sign for physics beyond the Standard Model (henceforth referred to as BSM).

There are basically two different approaches to introduce charged LFV processes in BSM models.

On the one hand LFV can be introduced already on the tree-level by adding new heavy particles, coupling to both electrons and muons. For example these particles could be Higgs or doubly charged Higgs particles, R-parity violating scalar neutrinos or new heavy vector bosons [2]. For these models the high observed suppression of LFV reactions is induced by the high masses of the new intermediate particles. A generic Feynman diagram for this process can be seen in Figure 2.3.

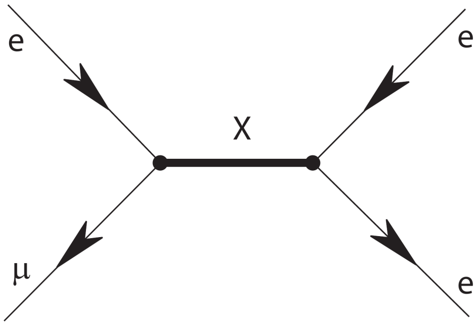


Figure 2.3: Generic Feynman diagram for lepton flavour violating decay $\mu \rightarrow eee$ on tree-level with new heavy particle X, coupling to both electrons and muons [2].

On the other hand LFV can be introduced on the loop-level. One example for such a family of BSM models would be the so-called supersymmetric (SUSY) extension of the SM, introducing new supersymmetric particles in the loop as seen in Figure 2.4. In this model every particle has a supersymmetric partner particle, its superpartner, with opposite spin statistics.

The LFV is induced by slepton[5] mixing instead of neutrino mixing as in the SM contribution seen in Figure 2.2.

Since the supersymmetric particles can have similar mass, reducing the suppression due to the mass difference, the branching ratio for this decay channel could be much higher [6]. Consequently, the process would be experimentally accessible.

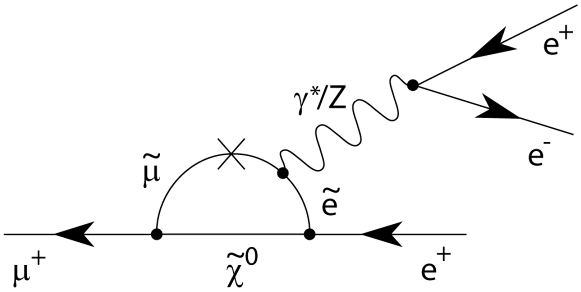


Figure 2.4: Feynman diagram for lepton flavour violating decay $\mu \to eee$ on loop-level, introducing supersymmetric particles [2].

In summary the search for the lepton flavour violating muon decay $\mu \to eee$ is very sensitive to many new physics models and an actual observation of this decay would be a clear sign for physics beyond the Standard Model.

---

[5] Supersymmetric partner particles of leptons.

Chapter 3

The Mu3e Experiment

The Mu3e experiment is designed to search for the decay $\mu \rightarrow eee$ with a branching ratio sensitivity of about $B = \mathcal{O}(10^{-16})$ [2] – four orders of magnitude more sensitive than the latest experimental result published by the SINDRUM collaboration [3].

In this chapter the kinematics of $\mu \rightarrow eee$, possible backgrounds and the resulting challenges will be discussed, leading to an overview of the Mu3e detector baseline design.

3.1 Muon Decay Kinematics

A very essential part for reaching the proposed branching ratio sensitivity is the suppression of possible background signals. Consequently, it is necessary to precisely know the kinematics of the muon decay channel $\mu \rightarrow eee$ to distinguish a desired signal from the background.

3.1.1 Signal Topology

The signal decay $\mu \rightarrow eee$ occurs instantaneously and the origin of the three electron tracks is located in one single vertex.

Since it is a three body decay, the system's invariant mass corresponds to the muon mass, i.e. the Minkowski norm of the four-momenta $\mathbf{P}_i$ of each electron i and the four-momentum conservation yields

$$m_\mu^2 c^4 = \left(\sum_{i=1}^{3} \mathbf{P}_i \right)^2 \quad . \tag{3.1}$$

Because the muon is stopped in the target before decaying, the muon's rest frame and the laboratory frame are approximately identical[1]. Consequently, the three momentum conservation of this three body decay results

---

[1] They are exactly identical, if the muon's velocity vanishes completely during the stopping process.

in

$$\sum_{i=1}^{3} \mathbf{p_i} = 0 \, . \tag{3.2}$$

These equations are leading to two important results. On the one hand Equation 3.2 tells us that the initial momenta $\mathbf{p_0}$ of all tracks are located in the same geometrical plane while on the other hand Equation 3.1 implies that the upper bound of the kinetic energy of one single decay electron is equivalent to half the muon mass.

Because the Mu3e experiment will use a solenoidal magnetic field $\mathbf{B}$ for momentum measurements, the decay electrons will move on helical trajectories in the detector due to Lorentz forces. Figure 3.1 shows a possible topology of a signal decay projected on the plane perpendicular to $\mathbf{B}$.

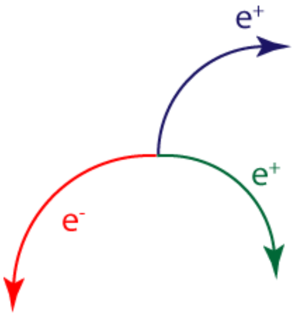


Figure 3.1: Topology of signal decay $\mu \to eee$ in a magnetic field $\mathbf{B}$ projected on the plane perpendicular to $\mathbf{B}$ [13].

3.1.2 Background

Since $\mu \to eee$ is a very rare decay, backgrounds from various sources have to be taken into account. This becomes even more clear when the known lepton flavour conserving decay channels of the muon with their specific branching ratios shown in Table 3.1 are considered.

Decay channel	B	Designation
$\mu^+ \to e^+ \nu_e \bar{\nu_\mu}$	$\approx 100\,\%$	Michel decay
$\mu^+ \to e^+ \nu_e \bar{\nu_\mu} \gamma$	$(1.4 \pm 0.4)\,\%$	Radiative decay
$\mu^+ \to e^+ \nu_e \bar{\nu_\mu} e^+ e^-$	$(3.4 \pm 0.4) \times 10^{-5}$	Internal conversion

Table 3.1: Possible lepton flavour conserving decay channels of the muon with the specific branching ratio B. This table is adapted from [8].

Since there is no irreducible background in this scenario, i.e. background that can not be distinguished from signal by increasing the precision of measurements, any background is due to common muon decays as seen in Table 3.1. These background signals can be divided into two separate parts. On the one hand there is some accidental background, while on the other hand internal conversion contributes to background, too.

Accidental Background

There are several possible coincidences in phase-space, leading to so-called accidental background. That means three electron tracks with appropriate charges from different independent physical processes can originate from the same vertex region and thus fake a signal.

One possible coincidence would be the combination of two common Michel decays $\mu^+ \to e^+\nu_e\bar{\nu_\mu}$ (see chapter 2), producing two positrons, and one (fake) electron from a recurling track, i.e. a track that makes several turns in the magnetic field of the detector. The topology of this process can be seen in Figure 3.2.

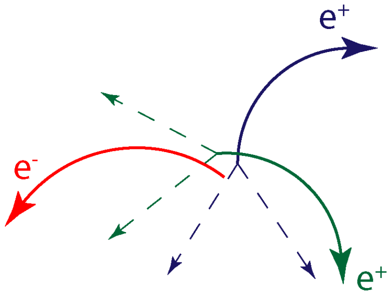


Figure 3.2: Topology of accidental background signal from Michel decay and recurling track [13].

Another possible contribution to accidental background is the combination of an internal conversion decay $\mu^+ \to e^+\nu_e\bar{\nu_\mu}e^+e^-$ and a common Michel decay as seen in Figure 3.3.

Since accidental background does not originate from the same vertex and does not satisfy energy and momentum conservation (see Equation 3.1 and Equation 3.2), a very high momentum, timing and vertex resolution is needed for distinguishing it from the signal.

Internal Conversion Background

The internal conversion decay is the radiative decay channel of the muon $\mu^+ \to e^+\nu_e\bar{\nu_\mu}e^+e^-$ and occurs with a branching ratio of about 3.4×10^{-5} [8] as listed in Table 3.1.

In this radiative decay the emitted photon from $\mu^+ \to e^+\nu_e\bar{\nu_\mu}\gamma$ almost instantly converts into an electron-positron pair. The Feynman diagram of

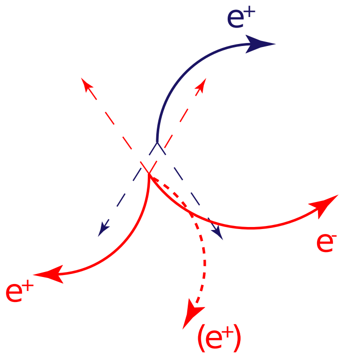


Figure 3.3: Topology of accidental background signal from Michel decay and internal conversion [13].

this process can be seen in Figure 3.4.

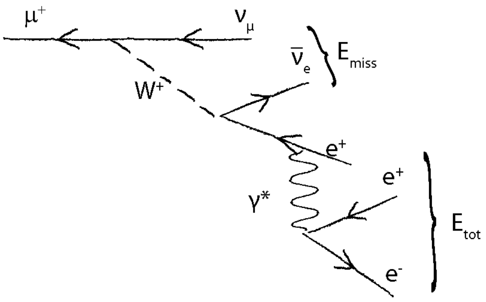


Figure 3.4: Feynman diagram of internal conversion decay $\mu^+ \to e^+ \nu_e \bar{\nu_\mu} e^+ e^-$ with missing energy labelled.

Measurements of a final state nearly identical to the signal can occur, because there are three electron tracks with appropriate charges and undetected neutrino tracks, originating from the same vertex. Therefore, it is necessary to take energy and momentum conservation of the process $\mu \to eee$ into consideration, i.e. Equation 3.1 and Equation 3.2 have to be used to reconstruct the momentum carried away by the neutrinos. This momentum or energy respectively is denoted by E_{miss} in Figure 3.4.

Consequently, a high momentum resolution is needed to reduce this kind of background signal.

3.2 Multiple Scattering

Charged particles travelling trough a medium are repeatedly interacting with the material's nuclei via Coulomb interaction and thus will be deflected.

This effect is consequently named multiple Coulomb scattering or multiple scattering for short.

To describe the impacts of this effect, a geometrical sketch is very useful and can be seen in Figure 3.5. In detail one can clearly see how the particle is changing its direction multiple times, resulting in an overall deflection angle θ_{plane} and an offset y_{plane} with respect to the particle's original trajectory. In this description momentum is always assumed to be conserved and possible energy losses are neglected.

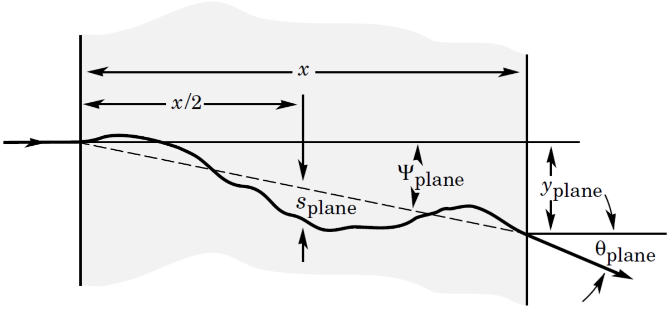


Figure 3.5: Geometrical sketch for description of multiple Coulomb scattering [8].

The scattering angle distribution θ is well described by a simple Gaussian distribution with a mean of $\bar{\theta} = 0$ and a standard deviation given by the so-called Highland formula [8]:

$$\theta_0 = \frac{13.6\,\text{MeV}}{\beta c p} z \sqrt{\frac{x}{X_0}} \left[1 + 0.038 \ln \frac{x}{X_0} \right] , \tag{3.3}$$

where z, βc and p denote the particle's charge number, velocity and momentum respectively. The travelled thickness x in the material is given in units of the radiation length X_0, which is a material constant. This multiple scattering model is only valid for small scattering angles θ projected onto a plane perpendicular to the initial track.

Since θ_0 is inversely proportional to the particle's momentum p, multiple Coulomb scattering is dominant for low particle momenta. An example for multiple scattering of electrons with different momenta in material with $x/X_0 = 0.001$ is shown in Table 3.2.

p [MeV/c]	10	25	40	50
θ_0 [°]	1.820	0.727	0.454	0.363

Table 3.2: Example for multiple scattering of electrons with different momenta p with $x/X_0 = 0.001$.

3.3 Detector Concept

There are several challenges that have to be taken into account concerning the detector design. The aforementioned background reduction requires high timing, vertex and momentum resolution. The suppression of multiple scattering requires low detector material. These issues are resolved by the detector baseline design, described in the following section.

In order to reach the proposed branching ratio sensitivity, in excess of 10^{16} muon decays have to be observed. Consequently, a high intensity muon beam is required, providing up to about 2×10^9 muons per second [2]. About 83 % of these muons will be stopped in a hollow aluminium double cone target with a wall thickness of $30\,\mu$m to $80\,\mu$m [2]. Once a muon is stopped, it will decay due to its finite lifetime and the corresponding decay products, i.e. electrons and positrons, will be detected[2].

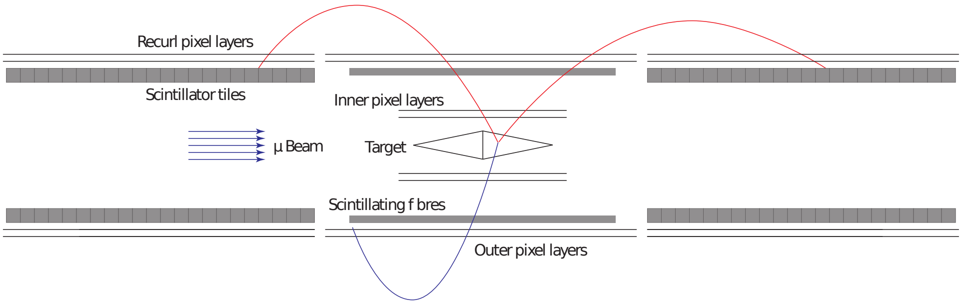
The detector itself is an ultra thin silicon pixel tracker based on two double layers of HV-MAPS[3] [14] positioned around the target. Since the layers in the double pairs have a distance of 1 cm, the effect of deflection by multiple scattering between two layers exceeds their pixel size. The whole setup is located in an homogeneous solenoidal magnetic field aligned in direction of the detector's rotational axis. The two outer pixel layers are supplemented with scintillating tiles and a scintillating fibre tracker for precise timing measurements. Furthermore, these layers are extended along the beam axis for precise momentum measurements with the help of recurling tracks [2]. The two inner pixel layers will be used for vertex reconstruction.

An overview of the whole detector concept in both a longitudinal and a transverse view is illustrated in Figure 3.6.

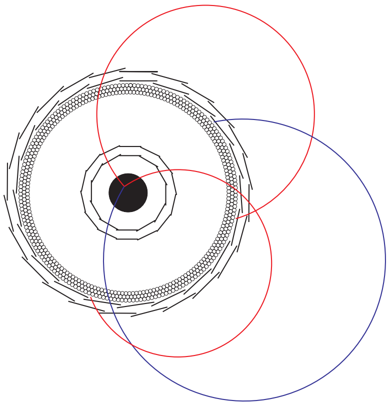
---

[2] Neutrinos will not be detected directly. In fact, they will be reconstructed from momentum reconstruction of the electrons.

[3] High-Voltage Monolithic Active Pixel Sensor.



(a) Longitudinal view



(b) Transverse view

Figure 3.6: Sketch of Mu3e detector baseline design with possible electron trajectories in a longitudinal and transverse view.

Chapter 4

Vertex Fitting

One obvious characteristic of the muon decay $\mu \to eee$ is the fact that it is originated in a common vertex. Consequently, accidental background can be greatly suppressed by requiring the three electron tracks to be compatible with a single vertex.

Due to multiple Coulomb scattering in the first detector layer, the reconstruction of a common vertex is not trivial, because the three scattered tracks will not intersect in a single space point, see Figure 4.1. Moreover, the bending of the tracks in the magnetic field causes the problem to be highly non-linear.

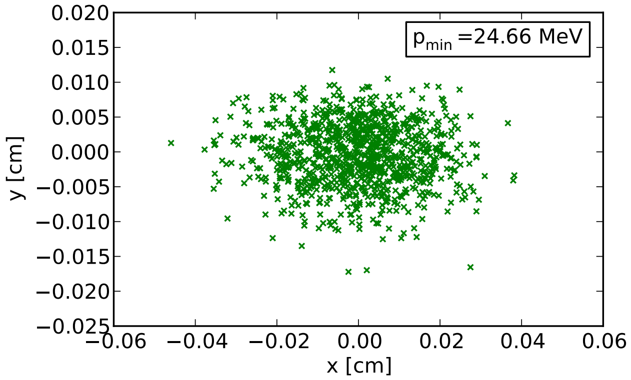


Figure 4.1: Effects of multiple scattering in the first detector layer. The track with lowest momentum is chosen from a signal, scattered in the first detector layer and is then reconstructed into the vertex region, by propagating it to the point of closest approach to the original vertex. The original vertex is at the coordinate origin. This procedure is done 1000 times.

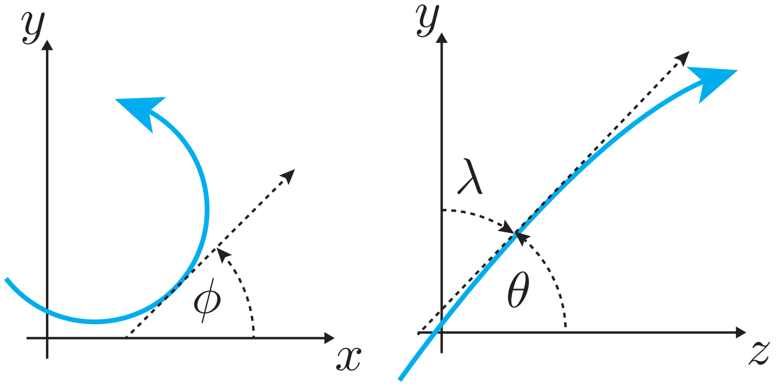
In this chapter a linearised vertex fitting method will be described (a more compact description is given in [15]). This method assumes that multiple scattering of particles in the first detector layer will be the only source of uncertainties while position and momentum determination is assumed to be perfectly accurate. It takes the physical parameters of all three tracks as an input. The parameters are the particle's charge q, the hit position $\mathbf{x}_h$ of each particle in the first detector layer and its momentum $\mathbf{p}_h$ at this specific position. The resulting output is the best estimate for a possible common three-dimensional vertex position $\mathbf{x}_v$ of all three particle tracks.

This chapter will include some remarks on the chosen coordinate system, a derivation of the particle track model used, an introduction to finding an initial guess for the possible vertex position and finally the linearised reconstruction algorithm itself.

4.1 Coordinate System

Since the Mu3e detector baseline design has a cylindrical shape as described in chapter 3, a cylindrical coordinate system (r, ϕ, z) is the most appropriate choice for describing spatial information. Moreover, the z-axis is chosen to coincide with the direction of the homogeneous magnetic field $\mathbf{B}$. Consequently, the x-y-plane will be referred to as the transverse plane while the z-axis will be referred to as the longitudinal direction.

In this coordinate system, the different track angles can be defined by reference to Figure 4.2. For the rest of this thesis Φ, ϕ or φ will always denote an azimuthal angle in the transverse plane while Θ, θ or ϑ will be the corresponding polar angle in longitudinal direction. Furthermore, θ is defined to be the corresponding[1] angle to the dip angle λ as seen in Figure 4.2, while λ again is defined to be the angle between the track's projected tangent vector in the x-y-plane and the track tangent $\mathbf{T}$.



(a) Transverse plane (b) Longitudinal direction

Figure 4.2: Track angles ϕ and θ in both transverse plane and longitudinal direction. The particle track is illustrated as the blue curling arrow.

---

[1] That is $\lambda + \theta = \frac{\pi}{2}$.

The coordinate origin is chosen to coincide with the middle of the double cone target. This configuration is illustrated in Figure 4.3.

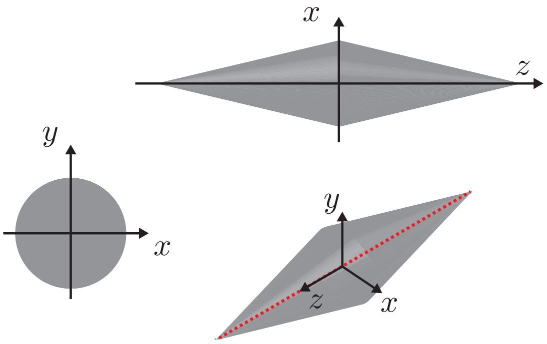


Figure 4.3: Global coordinate system with double cone target. The homogeneous magnetic field **B** is aligned in direction of the z-axis.

4.2 Track Model

Because the vertex of a possible decay signal has to be fitted, the trajectory of each particle measured in the detector layers has to be reconstructed very precisely. Hence, an appropriate track model is needed.

In this section, the equations of motion for electrons in a solenoidal magnetic field are derived and parametrized in an appropriate way for describing each trajectory. This derivation can be found in [11] in a similar way.

In general particles with charge q moving in an electromagnetic field $\mathbf{E}$ and $\mathbf{B}$ with velocity $\mathbf{v}$ are influenced by the Lorentz force:

$$\mathbf{F} = q\left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right) \; . \tag{4.1}$$

Assuming there is no electric field present, i.e. $\mathbf{E} = 0$, the Lorentz force as given in Equation 4.1 simplifies to

$$\mathbf{F} = q\left(\mathbf{v} \times \mathbf{B}\right) \; . \tag{4.2}$$

Due to the properties of the vector product $\mathbf{v} \times \mathbf{B}$ the force is always perpendicular to the particle trajectory. Consequently, the equation of motion $\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = \mathbf{F}$ predicts momentum conservation for its absolute value $p = \gamma m v$ [11]. The corresponding equation of motion is then given by

$$m\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}t^2} = q\left(\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} \times \mathbf{B}\right) \; . \tag{4.3}$$

However, the time variable t is not a good choice for describing the track model and it is more useful to have Equation 4.3 parametrized in units of the geometrical path length s along the track curve. This proceeds as follows and is also described in [11]. It holds that

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s}\frac{\mathrm{d}s}{\mathrm{d}t} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s}\beta c \tag{4.4}$$

$$\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}t^2} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}\beta c = \frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}s^2}\beta^2 c^2 \ , \tag{4.5}$$

where $\frac{\mathrm{d}s}{\mathrm{d}t} = v = \beta c$ is used.

Inserting these relations into Equation 4.3 finally yields

$$\frac{\mathrm{d}^2\mathbf{x}}{\mathrm{d}s^2} = \frac{q}{pc}\left(\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}s} \times \mathbf{B}\right) \ . \tag{4.6}$$

This is the final equation of motion in purely geometrical quantities for the idealized track model. The solution to Equation 4.6 is a helical trajectory.

Solving Equation 4.6 for $\mathbf{x}(s)$ will provide the particle trajectory in terms of the geometrical path length s. A general solution for the particle position vector $\mathbf{x}$ in an arbitrary magnetic field $\mathbf{B}$ is given in [16] and will be used for the rest of this thesis:

$$\mathbf{x}(s) = \mathbf{x}_0 + \frac{\gamma}{\kappa}(\theta - \sin\theta)\mathbf{H} + \frac{\sin\theta}{\kappa}\mathbf{T}_0 + \frac{\alpha}{\kappa}(1 - \cos\theta)\mathbf{N}_0 \ , \tag{4.7}$$

where the index 0 indicates the corresponding quantity at the initial geometrical path length $s = 0$ and $\mathbf{x}_0$ denotes the specific reference point. Moreover, the following shorthand notation is used:

$$\begin{aligned}
\mathbf{H} &= \frac{\mathbf{B}}{\|\mathbf{B}\|} \\
\mathbf{T} &= \frac{\mathbf{p}}{\|\mathbf{p}\|} \\
\mathbf{N} &= \frac{\mathbf{H} \times \mathbf{T}}{\alpha} \\
\alpha &= \|\mathbf{H} \times \mathbf{T}\| \\
\gamma &= \langle \mathbf{H}, \mathbf{T} \rangle \\
\kappa &= -\|\mathbf{B}\| \frac{q}{\|\mathbf{p}\|} \\
\theta &= \kappa s \ .
\end{aligned}$$

Here $\mathbf{p}$ denotes the particle momentum, $q = \pm 1$ is the particle's electric charge[2] and $\mathbf{B}$ is the magnetic field vector. The quantity κ is called the signed three dimensional track curvature.

The tangent vector $\mathbf{T}$ of the track is then given in terms of the geometrical path length s by the expression [16]:

$$\mathbf{T}(s) = \frac{\partial \mathbf{x}}{\partial s} = \gamma(1 - \cos\theta)\mathbf{H} + \cos\theta\mathbf{T}_0 + \alpha\sin\theta\mathbf{N}_0 \ . \tag{4.8}$$

---

[2] Again given in units of the elementary charge e.

Finally, the particle helical track itself and its propagation in the solenoidal magnetic field is fully specified in geometrical quantities by the expressions of $\mathbf{x}(s)$ and $\mathbf{T}(s)$ given in Equation 4.7 and Equation 4.8.

For a homogeneous magnetic field aligned along the z-axis we have $\mathbf{B} = B\hat{\mathbf{e}}_z$ and consequently $\mathbf{H}$ simplifies to $\mathbf{H} = \hat{\mathbf{e}}_z$.

4.3 Initial Vertex Finding

Once each positron or electron from $\mu \to eee$ reaches the first detector layer, it is scattered due to multiple Coulomb scattering while traversing the material (see Equation 3.3). Consequently, the reconstruction of the scattered tracks back to the vertex region shows that no spacial point in this region exactly matches the condition for a common vertex (see Figure 4.4) as already mentioned in the beginning of this chapter.

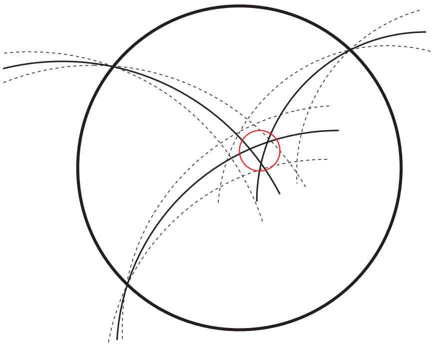


Figure 4.4: Illustration of multiple scattering of an event with three particles in the first detector layer. No exact common vertex is found at first.

Since we want to use an iterative linearised method for fitting the vertex position $\mathbf{x}$, an initial guess for a possible vertex position $\mathbf{x}_{v,0}$ is necessary.

There are several different methods for constructing this initial guess. I will use the method of circle-circle intersection in the transverse plane as described in the following. In general this method can be divided into several steps:

1. Choose two particle tracks with highest momenta p_1 and p_2.
2. Search for the point of closest approach on the two tracks.
    (a) Search for both track circle intersection points in the transverse plane.
    (b) Use the one from (a) with smallest longitudinal, i.e. z-distance between both tracks.

3. Search for the point on the remaining third track that is closest to the just determined circle-circle intersection point.

4. Initial vertex position $\mathbf{x}_{v,0}$ is set to the mean of all determined points.

Using this method, some variations can be considered. For example, in step 2. (b) one could not only use the smallest longitudinal distance between both tracks but also check whether the geometrical path length s between the chosen circle-circle intersection is minimal, because there are some pathological cases where 2. (b) without another crosscheck fails at finding an appropriate initial vertex position. These details will be discussed in the description of my simulation setup in chapter 5.

The mathematical method for calculating the circle-circle intersection points in the transverse plane can be found in section A.1 in the appendix.

4.4 Linearised Fitting Algorithm

For the actual vertex reconstruction, the particle trajectories are extrapolated from the innermost detector layer into the vertex region and a common intersection point is to be fitted.

The curvature of low momentum particle trajectories in the magnetic field due to the Lorentz force can not be neglected and the problem of vertex reconstruction becomes non-linear. No analytical solution can be found. In order to solve this problem a linearised iterative fitting method is used and will be described in the following.

Since spatial uncertainties are neglected in this scenario, multiple scattering will dominate the uncertainties in extrapolating the particle tracks from the first detector layer inwards to the vertex region. The tracks are now forced to intersect with a chosen vertex position. This can be accomplished by introducing multiple scattering angles Φ in a transverse and Θ in a longitudinal plane for each particle. These angles are used to bend the track in the "right" direction for intersecting with the chosen vertex position by changing its corresponding track angles [15]. Furthermore, Φ and Θ together with the expected scattering angle are used to define a χ^2 function which is minimal at the best estimate for a possible vertex position.

The specific scattering angles will be introduced first before the corresponding χ^2 function is defined in terms of Φ_i and Θ_i and finally minimized.

4.4.1 Transverse Scattering

By taking a look at the scattering of a single track in the transverse plane illustrated in Figure 4.5, I will first define some characteristic quantities repeatedly used in this section.

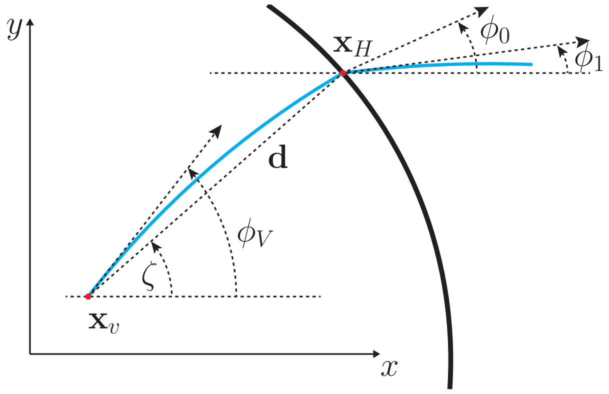


Figure 4.5: Illustration of scattering in the transverse plane for definition of all relevant quantities, describing the vertex fitting method.

With reference to Figure 4.5 the multiple scattering angle Φ in the transverse plane, bending the track in the “right” direction, can be defined as

$$\Phi := \phi_1 - \phi_0 \ , \tag{4.9}$$

where and ϕ_0 and ϕ_1 denote the reconstructed track angle before and after the scattering process, respectively [15]. ϕ_1 is directly obtained by measuring the momentum and hit position of the track in the first detector layer. However, ϕ_0 is not known explicitly and has to be calculated by using some further considerations, done in the following.

Again referencing to Figure 4.5, let $\mathbf{d}$ be the distance vector between the vertex position $\mathbf{x}_v$ and the hit position in the first detector layer $\mathbf{x}_H$, i.e.

$$\mathbf{d} := \mathbf{x}_H - \mathbf{x}_v \ . \tag{4.10}$$

Furthermore, let Φ_{0V} be the difference between the azimuthal angles at the hit and the vertex position, i.e.

$$\Phi_{0V} := \phi_0 - \phi_V \ . \tag{4.11}$$

This angle is the so called bending angle.

If the particle’s momentum p or the corresponding 3D track radius R is known, the bending angle Φ_{0V} can be obtained from any two points even if no other track angle is known[3]. Choosing $\mathbf{x}_H$ and $\mathbf{x}_V$ to be these two points, Φ_{0V} is implicitly given by a transcendental equation (see [15] and [17]):

$$\sin^2 \frac{\Phi_{0V}}{2} = \frac{d^2}{4R^2} + \frac{z^2}{R^2 \Phi_{0V}^2} \sin^2 \frac{\Phi_{0V}}{2} \ , \tag{4.12}$$

where $d = \|\mathbf{d}\|_{12}$ denotes the magnitude of the distance vector in the transverse plane and $z = (\mathbf{d})_3$ is the third component of the distance vector.

---

[3] This applies to our case, because ϕ_0 is unknown.

Consequently, Φ_{0V} will be highly dependent on the vertex position, i.e. $\Phi_{0V} = \Phi_{0V}(\mathbf{x}_v)$. Equation 4.12 can be solved for Φ_{0V} numerically, e.g. using the Newton-Raphson method with appropriate starting values. Possible starting values are given in section A.2 in the appendix.

From this calculation, the missing azimuthal angle ϕ_0 before scattering can be obtained as

$$\phi_0 = \frac{\Phi_{0V}}{2} + \zeta \,, \tag{4.13}$$

where $\zeta := \angle\mathbf{d}$ denotes the azimuthal angle of the distance vector $\mathbf{d}$ [15].

Using this calculation for ϕ_0 given in Equation 4.13, Equation 4.9 can be rewritten as

$$\Phi = \phi_1 - \zeta - \frac{\Phi_{0V}}{2} \,. \tag{4.14}$$

4.4.2 Longitudinal Scattering

Analogously to the definition of Φ in Equation 4.9 the polar scattering angle Θ_i can be defined as

$$\Theta := \theta_1 - \theta_0 \,. \tag{4.15}$$

Because the polar angle θ is invariant under motion in a solenoidal magnetic field and assuming that no energy is lost during scattering, it can be expressed by using the bending angle Φ_{0V} calculated above [15]:

$$\cos\theta_0 = \frac{z}{R\Phi_{0V}} \,. \tag{4.16}$$

Consequently, Equation 4.15 can be rewritten again as

$$\Theta = \theta_1 - \arccos\frac{z}{R\Phi_{0V}} \,. \tag{4.17}$$

Given Equation 4.14 and Equation 4.17 the scattering angles, bending the track in the "right" direction, can be calculated in geometrical terms without knowledge of any specific track angle.

4.4.3 Definition and Minimization of χ^2

Since there are three particle tracks in the final state of $\mu \to eee$, each particle will be given the index i.

For the specific scattering angles Φ and Θ described above a χ^2 function can be defined [15]:

$$\chi^2(\mathbf{x}_v) := \sum_{i=1}^{3}\left[\frac{\Phi_i^2(\mathbf{x}_v)}{\sigma_{\Phi,i}^2} + \frac{\Theta_i^2(\mathbf{x}_v)}{\sigma_{\Theta,i}^2}\right] \,, \tag{4.18}$$

which has to be minimized in terms of the fitted vertex position $\mathbf{x}_v$. Here $\sigma_{\Phi,i}^2$ and $\sigma_{\Theta,i}^2$ are basically the variances of Φ_i and Θ_i, which are given by the multiple scattering model in Equation 3.3.

In order to solve the problems of non-linear equations, the track model is linearised around the initial vertex position $\mathbf{x}_{v,0}$, i.e. $\mathbf{x}_v = \mathbf{x}_{v,0} + \Delta\mathbf{x}_v$. Consequently, a two-dimensional Taylor series expansion of the multiple scattering angles Φ_i and Θ_i around $\mathbf{x}_{v,0}$ up to first order yields

$$\begin{aligned} \Phi_i(\mathbf{x}_v) &= \Phi_i(\mathbf{x}_{v,0}) + \langle \Delta\mathbf{x}_v, \nabla\Phi_i(\mathbf{x}_{v,0})\rangle \qquad (4.19) \\ \Theta_i(\mathbf{x}_v) &= \Theta_i(\mathbf{x}_{v,0}) + \langle \Delta\mathbf{x}_v, \nabla\Theta_i(\mathbf{x}_{v,0})\rangle \ , \qquad (4.20) \end{aligned}$$

where $\langle\cdot,\cdot\rangle$ denotes the standard scalar product in $\mathbb{R}^3$ [15].

The minimization of χ^2 given in Equation 4.18 then yields

$$\begin{aligned} \nabla\chi^2(\mathbf{x}_v) &= \sum_{i=1}^{3} \left[\frac{\nabla\Phi_i^2(\mathbf{x}_v)}{\sigma_{\Phi,i}^2} + \frac{\nabla\Theta_i^2(\mathbf{x}_v)}{\sigma_{\Theta,i}^2} \right] \\ &= 2\sum_{i=1}^{3} \left[\frac{\Phi_{i,0}\nabla\Phi_i}{\sigma_{\Phi,i}^2} + \frac{\Theta_{i,0}\nabla\Theta_i}{\sigma_{\Theta,i}^2} \right] \qquad (4.21) \\ &+ 2\sum_{i=1}^{3} \left[\frac{\langle\Delta\mathbf{x}_v, \nabla\Phi_i\rangle\nabla\Phi_i}{\sigma_{\Phi,i}^2} + \frac{\langle\Delta\mathbf{x}_v, \nabla\Theta_i\rangle\nabla\Theta_i}{\sigma_{\Theta,i}^2} \right] \ , \end{aligned}$$

where $\Phi_{i,0} := \Phi_i(\mathbf{x}_{v,0})$ and $\Theta_{i,0} := \Theta_i(\mathbf{x}_{v,0})$ is used as short notation.

To minimize Equation 4.21 with respect to the small vertex position correction $\Delta\mathbf{x}_v$, the system of three linear equations has to be solved:

$$\nabla\chi^2(\mathbf{x}_v) = 0 \ , \qquad (4.22)$$

which can be expressed more explicitly in the following way:

$$\mathbf{F}\Delta\mathbf{x}_v + \mathbf{C} = 0 \qquad (4.23)$$

with index notation

$$\begin{aligned} F_{kj} &= \sum_{i=1}^{3} \left[\frac{\partial_j\Phi_i\partial_k\Phi_i}{\sigma_{\Phi,i}^2} + \frac{\partial_j\Theta_i\partial_k\Theta_i}{\sigma_{\Theta,i}^2} \right] \\ C_k &= \sum_{i=1}^{3} \left[\frac{\Phi_{i,0}\partial_k\Phi_i}{\sigma_{\Phi,i}^2} + \frac{\Theta_{i,0}\partial_k\Theta_i}{\sigma_{\Theta,i}^2} \right] \ . \end{aligned}$$

Here the short notation $\partial_j = \frac{\partial}{\partial x_j}$ is used. The specific derivatives $\partial_j\Phi_i$ and $\partial_j\Theta_i$ can be calculated using Equation 4.14 and Equation 4.17. This calculation and the final expressions for $\partial_j\Phi_i$ and $\partial_j\Theta_i$ are given in section A.3 in the appendix.

The solution of Equation 4.23 is then given by

$$\Delta\mathbf{x}_v = -\mathbf{F}^{-1}\mathbf{C} \ . \qquad (4.24)$$

The method described above can be used iteratively [15]. Therefore, the fitted vertex position $\mathbf{x}_v$ can be improved by using the sequence

$$\mathbf{x}_{v,n+1} = \mathbf{x}_{v,n} + \Delta\mathbf{x}_{v,n} \qquad \text{with} \quad n \in \mathbb{N} \ .$$

Chapter 5

Simulation

In order to test and evaluate the performance and functionality of the linearised vertex fitting method described in chapter 4, I implemented the corresponding algorithm in a computer simulation using PYTHON [18] and SCIPY [19]. Moreover, much of the code that is internally used in the simulation is taken from or based on the Mu3e tracking library TRACKY developed by Moritz Kiehn [11].

For the purpose of testing the vertex fitting method, a complete detector simulation is not needed and consequently several simplifications can be made.

In the following chapter the general configuration and the basic modules used for the final simulation will be described in detail.

5.1 Event Generation

The first part of the simulation is to create the signal decay $\mu \to eee$ in phase space, i.e. the physical parameters of the resulting electrons or positrons respectively.

As described in subsection 3.1.1 the decay topology has to satisfy specific physical constraints such as four-momentum conservation as seen in Equation 3.1, which have to be taken into account when randomly generating an event in phase space.

To satisfy these constraints an implementation of the phase space generator RAMBO[1] is used [20]. This implementation is linked in the TRACKY library.

Using this method, one has the opportunity to choose the spatial position $\mathbf{x}_0$ of the decay signal, while the electric charges q_i and momenta $\mathbf{p}_{0,i}$ of all three tracks are randomly generated in phase space, satisfying the properties of the signal topology. This provides a complete set of seven physical

---

[1] Short for Random Momenta Booster.

parameters $(q, \mathbf{p}_0, \mathbf{x}_0)$ for each decay particle, determining its further motion in the magnetic field by Equation 4.3.

Moreover, this implementation can be used to generate actual signal events as well as background events.

5.1.1 Signal

Since the initial position $\mathbf{x}_0$ can be chosen freely there are several possible configurations, which will be briefly named and described in the following.

`Origin` Every event from a sample is created in the coordinate origin, i.e. $\mathbf{x}_0 = 0$. The coordinate origin is chosen to be the middle of the hollow double cone target as described in chapter 4.

`Uniform` The incoming muon beam profile is taken into account and is assumed to be uniformly distributed in the plane perpendicular to the beam direction.

Consequently, a uniform distribution is convoluted with the target's double hollow cone geometry to get the spacial distribution of the decays on the target. Every $\mathbf{x}_0$ of an event is drawn from this distribution.

`Gauss` The incoming muon beam profile is assumed to follow a rotationally symmetric two-dimensional Gaussian distribution with $\mu = 0\,\mathrm{cm}$ and $\sigma = 1\,\mathrm{cm}$ [21] in a plane perpendicular to the beam direction. Furthermore, the beam is assumed to be collimated, i.e. this distribution is cut off at a beam radius $r = 1\,\mathrm{cm}$.

Again, this distribution is convoluted with the target geometry to get the spacial distribution of muon decays on the target. Every $\mathbf{x}_0$ of an event is drawn from this distribution.

This is the most realistic configuration.

The exact mathematical description and calculation of the spatial distributions for the possible configurations is given in section A.4 in the appendix.

5.1.2 Background

The simulation of accidental background is implemented in a similar way.

Because accidental background (see chapter 3) is going to be simulated, four-momentum conservation is usually not satisfied, but this will be detected by the momentum measurement. For the purpose of vertex fitting only events with appropriate momenta are used, i.e. only the initial position $\mathbf{x}_0$ of each particle track has to be perturbed, while the generated momentum $\mathbf{p}$ remains unperturbed and satisfies the required signal selection. This is done in the same way as described in the section before with the extension that the initial position $\mathbf{x}_{0,i}$ for each track is generated individually and independently from one another.

For this purpose I implemented two possible configurations: `Uniform` and `Gauss` which were already described in the previous subsection.

5.2 Detector Geometry

For the purpose of my simulation the general detector geometry as described in chapter 3 is simplified. I will concentrate on the first detector layer only with radius set to $R = 1.9\,\mathrm{cm}$ in a first approximation to the real detector configuration [2]. Moreover, the detector layer is assumed to have a perfect cylindrical shape with rotational symmetry along the z-axis and finite extension.

5.3 Track Propagation

The three generated particles will propagate independently in the solenoidal magnetic field on helical trajectories as given by the expressions for the position $\mathbf{x}(s)$ and tangent $\mathbf{T}(s)$ in Equation 4.7 and Equation 4.8 respectively. The propagation will be done in purely geometrical terms for the track model. The magnetic field strength is set to $\|\mathbf{B}\| = 1\,\mathrm{T}$ [2].

Since the detector is assumed to have a perfect cylindrical shape, an analytical expression for the hit position in the first detector layer can be derived by using the intersection of a helical track with a cylindrical surface.

This method is already implemented in TRACKY [11].

5.4 Multiple Scattering

Once the hit position of the propagated track in the first detector layer is calculated, multiple scattering is simulated by changing the track parameters.

The detector layer is assumed to be a thin scatterer, so that momentum and energy conservation can be assumed [11]. Consequently, only the track angles ϕ and λ are influenced by multiple scattering. The deflection angles caused by multiple scattering are drawn from Gaussian distributions with $\mu = 0$ and σ from Equation 3.3 and are added to the corresponding track angles ϕ and λ. However, in Equation 3.3 the particle track is assumed to be perpendicular to the scattering plane. Thus, the traversed thickness x of the medium has to be adjusted by the inclination angle α, using the simplified assumption that the particle travels perpendicularly through an effective thickness x_{eff} of the medium as seen in Figure 5.1.

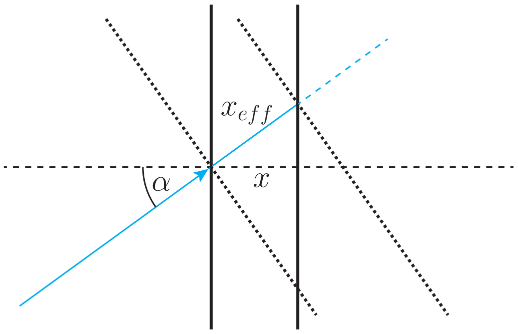


Figure 5.1: Simplification of an effective thickness x_{eff} for multiple scattering of a track with inclination angle α.

The effective thickness is obtained as

$$x_{eff} = \frac{x}{\cos\alpha} \ . \tag{5.1}$$

Moreover, the ratio of the thickness x and the material's radiation length X_0 is set to $x/X_0 = 0.001$. The particle momentum p can be calculated using either the current track state vector at the hit position or the initial one, because the absolute momentum is always assumed to be conserved.

5.5 Pixel Structure

The simulation of a finite pixel structure of the detector layer is described in [11] and is already implemented in Tracky. It is used in a similar way in this simulation.

The finite size of the pixel in the detector layers is simulated using Gaussian perturbations along the detector layer plane. This plane is defined by the unit vectors $\hat{\mathbf{e}}_\phi$ and $\hat{\mathbf{e}}_z$ from a common cylindrical coordinate system (r, ϕ, z) originated at the true hit position of a track. The hit position inside a pixel is assumed to be uniformly distributed over its surface. The resulting variance in each pixel direction is $\sigma^2 = p^2/12$ with one-dimensional pixel length p in the specific direction.

Finally, the pixel resolution is simulated by drawing samples from a normal distribution with variance σ^2 as defined above and adding them to the true hit position along the specific unit vector.

5.6 Initial Vertex Finding

The way I implemented the initial vertex finding algorithms is based on the corresponding description given in chapter 4, i.e. the intersection of two circles in the transverse plane is searched for. However, there are different methods to choose the “better” one of both intersection points in the transversal plane;

`z-Dist` Once both intersection points of two circles in the transverse plane are found, the one where the tracks have the smallest distance along the z-axis is chosen.

`s-Dist` Instead of choosing the intersection point with the smallest distance in z-direction, this method chooses the intersection point where the geometrical path length s between the intersection point and the hit position for the track with highest momentum is minimal.

`Mid-Target` In contrast to the two methods mentioned before this one does not use intersections of two circles in the transverse plane. Instead, the initial vertex position $\mathbf{x}_{v,0}$ is chosen to be exactly in the middle of the hollow cone target, i.e. $\mathbf{x}_{v,0} = (0,0,0)^T$.

This method might be sufficient because of the target's compact geometry.

There are still some special cases to be taken into consideration for `z-Dist` and `s-Dist` respectively. Since intersection points of two circles are calculated, it might occur that there is no solution at all because of some special topology (see section A.1 in the appendix). If this applies to our case, the specific point where both tracks or circles approach each other the closest is chosen.

Furthermore, another constraint or simplification is made: If the initial vertex position $\mathbf{x}_{v,0}$ is found to be somewhere outside of the first detector layer, `Mid-Target` is applied.

5.7 Termination of Iteration Method

In general there are two different approaches to decide when to terminate the iterative vertex fitting procedure as described in chapter 4. These two approaches will be named and briefly described in the following:

`Maxiter` A fixed number of maximum iterations i_{max} is set before running the procedure. If the number of iterations i exceeds i_{max}, the procedure is terminated and the fitted vertex position $\mathbf{x}_{v,i_{max}}$ of the last iteration i_{max} is returned.

`Tol` An accuracy limit ϵ is set before running the procedure. The corresponding accuracy is calculated using the change of χ^2 given in Equation 4.18 in each iteration step i, i.e. the procedure is terminated if

$$\left|\chi^2(\mathbf{x}_{v,i}) - \chi^2(\mathbf{x}_{v,i-1})\right| \leq \epsilon \; .$$

If this condition is satisfied $\mathbf{x}_{v,i}$ is returned.

Furthermore, the methods `Maxiter` and `Tol` can be combined by setting a tolerance limit but terminating the procedure early if this limit is not reached within a fixed number of maximum iterations.

5.8 Configuration List

To discuss and evaluate results obtained by the computer simulation I want to define a list C of fixed configurations.

This list shall contain the number of generated events n, the event and background generation distribution G, the finite pixel size p, the method for initial vertex finding V_0 and the method for terminating the fit procedure T, i.e.

$$C := (n, G, p, V_0, T) \ ,$$

where

$$\begin{aligned} n &\in \mathbb{N} \\ G &\in \{\texttt{Origin}, \texttt{Uniform}, \texttt{Gauss}\} \\ p &\in \mathbb{R}_0^+ \\ V_0 &\in \{\texttt{z-Dist}, \texttt{s-Dist}, \texttt{Mid-Target}\} \\ T &\in \{\texttt{Maxiter}, \texttt{Tol}\} \ . \end{aligned}$$

This configuration list C will later be used to determine which configuration was used for a simulation of the vertex reconstruction procedure.

Chapter 6

Results

In this chapter the results of my computer simulation as described in chapter 5 will be presented. These results will provide an overview of the performance and functionality of the linearised vertex fitting method.

I will begin with some specific results for one fixed configuration set and then compare this configuration to possible other sets. Furthermore, I will give a description of the possibilities to suppress background signals by using the results of the vertex fit procedure.

6.1 Performance of Single Configuration

The performance and functionality of the linearised vertex fitting procedure itself can be evaluated in several different illustrations and methods, described in the following.

Residuals The residuals Δ of a reconstructed parameter ξ_{fit} are calculated, i.e. $\Delta := \xi_{fit} - \xi_{true}$, and the corresponding distribution is plotted.

Assuming normally distributed measurement errors and a linear error propagation model, Δ will follow a normal distribution as well. The mean of this distribution should vanish, if the fitting procedure is working properly. A non-zero mean indicates a bias in the reconstruction algorithm.

The standard deviation of the distribution can be used as an estimate of the resolution, i.e. a small standard deviation is a sign for the reconstructed parameter ξ_{fit} describing the true parameter ξ_{true} well.

Pull Distributions So called pull distributions are used for evaluating the error description of the reconstruction algorithm. It makes use of the fact that the linear transformation of a normally distributed random variable $X \sim \mathcal{N}(\mu, \sigma^2)$ to another random variable $Z := \frac{X-\mu}{\sigma}$ will be standard normally distributed, i.e. $Z \sim \mathcal{N}(0, 1)$.

The pull of a reconstructed parameter ξ_{fit} is then defined to be $(\xi_{fit} - \xi_{true})/\sigma_\xi$ where σ_ξ is the standard deviation of the reconstructed parameter. In the case of a correct error description the pull distribution is a standard normal distribution $\mathcal{N}(0,1)$. The standard deviation σ of the pull distribution is again an estimate for the error description. A deviation $\sigma < 1$ is a sign for an overestimation of the parameter error, while $\sigma > 1$ on the other hand indicates underestimation with respect to the assumed error [11].

χ^2 Distribution Since the minimized $\chi^2(\mathbf{x}_v)$ function defined in Equation 4.18 is given in terms of independent standard normally distributed random variables, it will follow a χ^2 distribution:

Let Z_i with $i \in \{1, ..., k\}$ be independent, standard normally distributed random variables. Then the random variable $Q := \sum_{i=1}^{k} Z_i^2$ is defined to be χ_k^2 distributed with $k \in \mathbb{N}$ degrees of freedom, shortly noted by $Q \sim \chi_k^2$ [22, 23].

The only free parameter of the χ_k^2 distribution is the number of degrees of freedom k. In the context of a fitting procedure, k can be obtained by

$$k = n_m + n_c - n_p \ , \tag{6.1}$$

where n_m denotes the number of independent measurements, n_c is the number of additional constraints and n_p is the number of independent adjustable parameters [24].

This definition of $Q \sim \chi_k^2$ applies to our case, because the definition of the random variable $\chi^2(\mathbf{x}_v)$ given in Equation 4.18 satisfies[1] the definition mentioned above and consequently setting $Q \equiv \chi^2(\mathbf{x}_v)$ yields $\chi^2(\mathbf{x}_v) \sim \chi_k^2$.

The only missing parameter to fully describe the distribution is k. However, k can be obtained by using Equation 6.1: This scenario yields three particle tracks with two quantities (the specific angles Φ and Θ) each to be fitted, thus $n_m = 6$. Furthermore, to accomplish the fit we have three adjustable parameters, the coordinates of the vertex position $\mathbf{x}_v$, thus $n_p = 3$. Consequently, the number of degrees of freedom k of the expected χ_k^2 distribution is $k = 3$.

Finally, the $\chi^2(\mathbf{x}_v)$ random variable defined in Equation 4.18 for the fitted vertex position theoretically will follow a χ_3^2 distribution:

$$\chi^2(\mathbf{x}_v) \sim \chi_3^2 \ . \tag{6.2}$$

To describe the performance and functionality of the fit in detail I will use a fixed configuration to generate a sample with signal decays only:

$$C = \left(10^5, \texttt{Gauss}, 0, \texttt{s-Dist}, \texttt{Tol} = 10^{-8}\right) \ .$$

---

[1] This holds, because a normally distributed random variable X with a mean of μ and variance of σ^2, i.e. $X \sim \mathcal{N}(\mu, \sigma^2)$, can be linearly transformed into a standard normally distributed random variable Z by $Z = \frac{X-\mu}{\sigma}$, i.e. $Z \sim \mathcal{N}(0,1)$. In this case: $Z_{1,i} = \frac{\Phi_i}{\sigma_{\Phi,i}} \sim \mathcal{N}(0,1)$ and $Z_{2,i} = \frac{\Theta_i}{\sigma_{\Theta,i}} \sim \mathcal{N}(0,1)$, respectively.

For a detailed description of this quantity see section 5.8.

6.1.1 Reconstruction Efficiency

The first method to judge the performance of the vertex reconstruction algorithm is the reconstruction efficiency, i.e. the amount of data for which the vertex reconstruction did not work properly. This can be separated in two possible scenarios:

Fit not converged When a certain number of iterations[2] is exceeded, the fitting procedure is terminated and the fit is declared to be not converged, i.e. no values for the fitted vertex position $\mathbf{x}_v$ and the corresponding $\chi^2(\mathbf{x}_v)$ are returned.

Fit outside detector Sometimes the fit might diverge in a way that the fitted vertex position will be outside the first detector layer. In this case the fitting procedure looses its functionality.

Whenever one of these cases applies, the data is discarded and not taken into account in the following analysis. The amount data for which the reconstruction did not work in this configuration is illustrated in Figure 6.1.

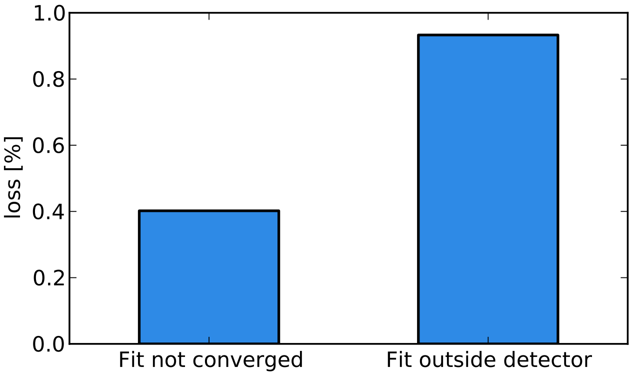


Figure 6.1: Reconstruction efficiency of a signal sample with 10^5 events.

In general the reconstruction procedure seems to work and only about 1.4 % of vertices of the generated events are not reconstructed properly. In many of these events, for which the vertex fitting procedure fails, the particle with the lowest momentum does not reach the second detector layer, so that its track would not be reconstructed in the experiment anyway. In this context and considering realistic conditions for track reconstruction, the reconstruction efficiency would be even better than in this simulation scenario.

---

[2] Here this number is arbitrarily set to $i_{max} = 50$.

The discarded data is not taken into consideration in the following analysis except for the efficiency and rejection studies.

6.1.2 Vertex Position Residuals

The residuals of the reconstructed vertex position $\mathbf{x}_v$ for each spatial dimension are shown in Figure 6.2.

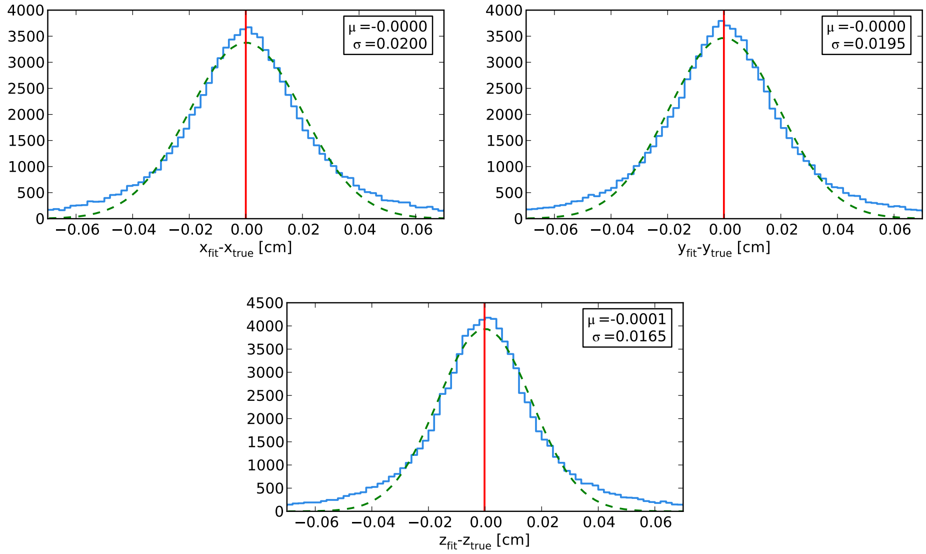
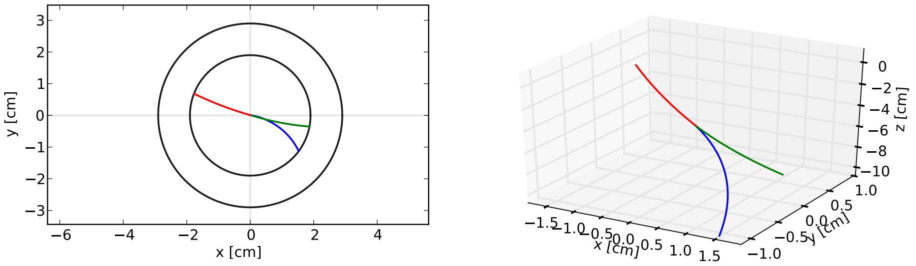


Figure 6.2: Residuals of the reconstructed vertex position $\mathbf{x}_v$ in each spatial dimension x, y and z. The dotted green lines are fitted normal distributions. The red vertical lines are the fitted means of the distributions.

To each distribution of the residuals $\Delta = \mathbf{x}_{v,fit} - \mathbf{x}_{v,true}$ a normal distribution is fitted, yielding parameters for the mean μ and variance σ^2 of the distribution which can also be seen in Figure 6.2.

Each dimension of Δ shows small deviations from a fitted normal distribution and especially the distributions' tails are clearly non-Gaussian. Nevertheless, the core seems to be well modelled by a Gaussian distribution. Each mean μ nearly vanishes and confirms that there is no bias in the reconstruction algorithm. The fitted width of the distributions are almost identical, but it is the smallest in the z-direction. Moreover, the widths yield an estimate for a vertex resolution $\delta\mathbf{x}_v$, i.e. with what spatial resolution a possible vertex can be reconstructed. These resolutions are $\delta x_v = 200\,\mu\text{m}$, $\delta y_v = 200\,\mu\text{m}$ and $\delta z_v = 170\,\mu\text{m}$.

The non-Gaussian tails of the distributions shown in Figure 6.2 can be understood when taking the single event topology into account. Since all events are generated in a flat phase space, topologies ill suited for vertex fitting might occur. An example for the topology of an event producing the corresponding data in one of the tails is shown in Figure 6.3.



(a) Transverse view with two inner detector layers.

(b) Three-dimensional view without detector layers.

Figure 6.3: Topology of an event producing data in the tails of the vertex position residuals Δ. The particle tracks are almost *back-to-back*.

In this scenario the particle tracks are almost *back-to-back*, leading to a loss of one constraint in reconstructing $\mathbf{x}_v$. The idea of fitting multiple scattering angles Φ and Θ to bend the tracks in the direction intersecting with $\mathbf{x}_v$ does not work in this case, because bending the tracks by using Φ and Θ does not affect a possible estimate for $\mathbf{x}_v$ along the track direction as seen in Figure 6.4. Thus, another degree of freedom is gained and the reconstruction does not work very precisely.

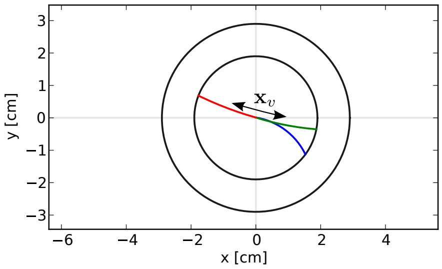


Figure 6.4: Topology of an event producing data in the tails of the vertex position residuals Δ in a transverse view. The estimate for $\mathbf{x}_v$ gains another degree of freedom along the track direction.

This observation can be studied by transforming the vertex position residuals Δ into new coordinates (a, b, c). The orthonormal basis of this coordinate set is defined as follows:

Let $\mathbf{p}_{max}$ be the initial momentum vector of the track with highest momentum in an event and define $\hat{\mathbf{e}}_a := \mathbf{p}_{max} / \left\| \mathbf{p}_{max} \right\|$. This is the axis for which the vertex reconstruction seems to fail. Furthermore, let $\hat{\mathbf{e}}_b :=$

$(\mathbf{p}_i \times \mathbf{p}_j) \,/\, \|\mathbf{p}_i \times \mathbf{p}_j\|$ with arbitrary $i \neq j$ be the normal vector of the decay plane and $\hat{\mathbf{e}}_c := \hat{\mathbf{e}}_a \times \hat{\mathbf{e}}_b$ lie in the decay plane.

The vertex position residuals Δ can now be projected on the new set of basis vectors. The corresponding distributions are shown in Figure 6.5. The vertex resolution in the a-direction is the worst, i.e. the corresponding distribution width is three times bigger than the width of the distributions in the orthogonal directions and has the biggest non-Gaussian tails.

The vertex reconstruction indeed does not work very precisely along a track direction with high momentum.

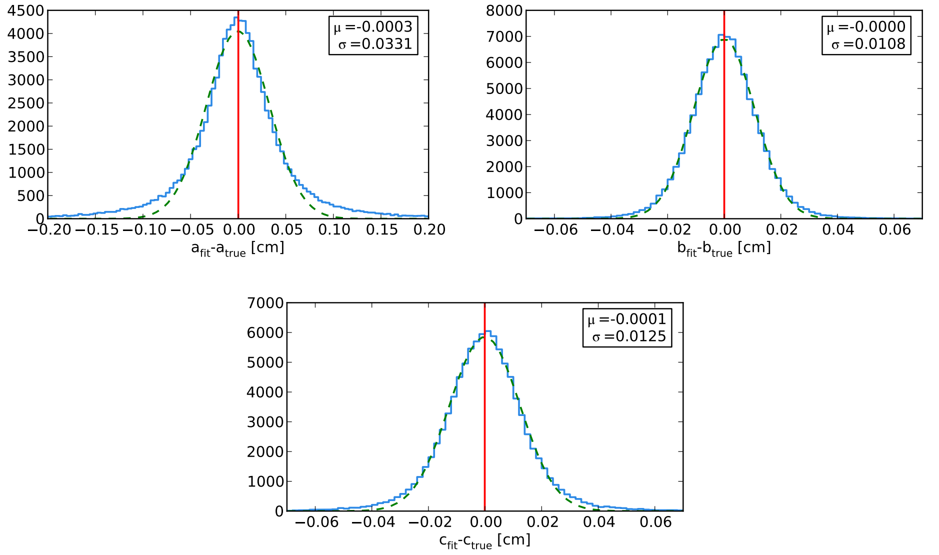


Figure 6.5: Residuals of the reconstructed vertex position $\mathbf{x}_v$ in new coordinates a, b and c. The dotted green lines are fitted normal distributions. The red vertical lines are the respective fitted means of the distributions.

6.1.3 Pull Distributions

Since the multiple scattering angles Φ and Θ are fitted by minimizing the $\chi^2(\mathbf{x}_v)$ function for reconstructing the vertex position (see Equation 4.18), the error description can be analysed by plotting the pull distributions for these angles.

Because the multiple scattering model described in chapter 3 is used, the mean of Φ and Θ should vanish. The standard deviations σ_Φ and σ_Θ of the reconstructed angles are estimated by using the Highland formula in Equation 3.3 for the track momenta and directions measured in the first detector layer.

The distribution of the reconstructed multiple scattering angles Φ and

Θ can be seen in Figure 6.6. The mean of each distribution vanishes as expected and its core is well described by a normal distribution.

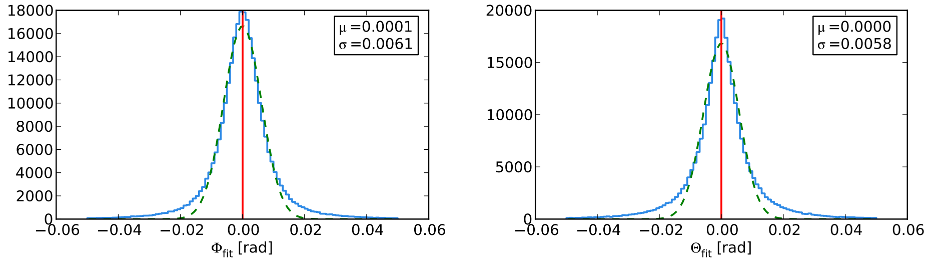
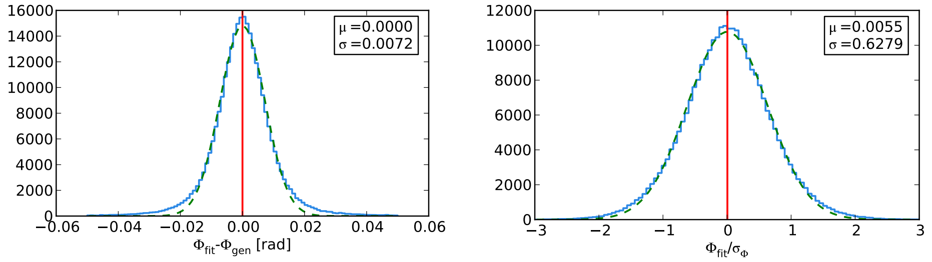
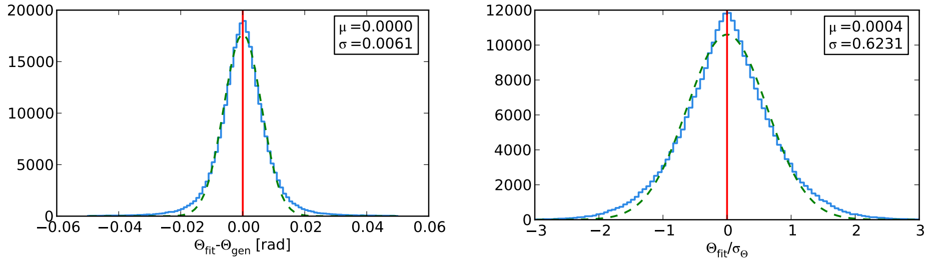


Figure 6.6: Distribution of reconstructed multiple scattering angles Φ and Θ. The dotted green lines are fitted normal distributions. The red vertical lines are the respective fitted means of the distributions.



(a) Residuals of Φ

(b) Pulls of Φ



(c) Residuals of Θ

(d) Pulls of Θ

Figure 6.7: Residuals and pull distributions for reconstructed multiple scattering angles Φ and Θ. Φ_{gen} and Θ_{gen} denote the true multiple scattering angles generated by the simulation. The dotted green lines are fitted normal distributions. The red vertical lines are the respective fitted means of the distributions.

Similar properties are expected for the residuals and pull distributions of Φ and Θ, which are shown in Figure 6.7. For the residuals the respective angles Φ_{gen} and Θ_{gen} generated by the simulation are taken as the true values for the multiple scattering angles.

All distribution cores are well described by a normal distribution with vanishing mean, which is expected. However, the fitted standard deviation σ_{fit} of the pull distributions is smaller than the expected width of $\sigma = 1$, indicating an overestimation of the reconstructed parameter error or reconstructed angles, which are too small. This conjecture can be checked by taking the residuals of the absolute values of Φ and Θ into consideration, i.e. the distributions of $|\Phi_{fit}| - |\Phi_{gen}|$ and $|\Theta_{fit}| - |\Theta_{gen}|$. These can be seen in Figure 6.8. The figure shows that the mean of both distributions does not vanish but is shifted to negative values. Thus, the reconstructed multiple scattering angles are systematically too small in comparison to the corresponding generated values.

This effect can be explained with the vertex reconstruction algorithm itself. Two multiple scattering angles per track have to be fitted, i.e. six angles in total. However, only three angles can be constrained by three tracks [25]. The remaining angles are not constrained at all. Therefore, the χ^2 minimization shifts the reconstructed angles to smaller values and the pull width is smaller than expected. This issue could be fixed by adding additional constraints, i.e. requiring the tracks to be in one plane. However, the influences of an additional constraint can not be estimated, yet.

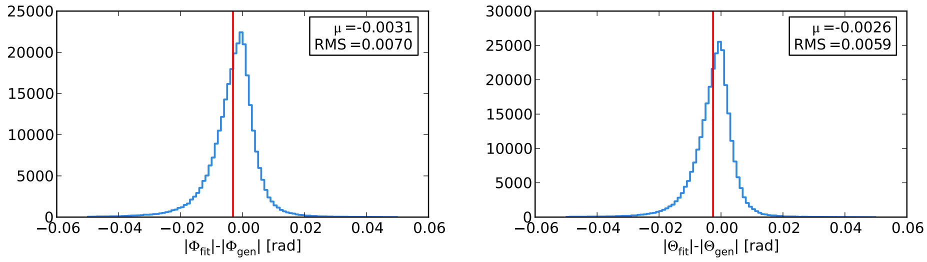


Figure 6.8: Residuals of the absolute values of the reconstructed multiple scattering angles Φ and Θ. The calculation of mean illustrated by the red vertical line and RMS is based on the data between the 0th and 99th percentile.

The reconstruction of the multiple scattering angles Φ and Θ in the first detector layer influences the estimate for the track angles ϕ and λ as well. This can be illustrated by taking their residuals with respect to the true track angles ϕ_{true} and λ_{true} into account. The reference point for the true track angles is chosen to be the true vertex position, i.e. the initial track angles of a signal are considered. One the one hand, ϕ and λ can be calculated at the track position extrapolated to the point of closest approach to an arbitrary estimate for a vertex position. This arbitrary estimate is chosen to be the coordinate origin. On the other hand, ϕ and λ at the reconstructed vertex position $\mathbf{x}_v$ for a track forced to intersect with $\mathbf{x}_v$ can be evaluated. The corresponding distributions are shown in Figure 6.9. As expected, the distributions for the tracks after reconstruction of $\mathbf{x}_v$ are much more narrow, especially for the azimuthal track angle ϕ. The RMS shown in Figure 6.9 is always calculated for data between the 5th and 95th percentile. Therefore, the reconstruction of the vertex position significantly improves the estimate of the track angles ϕ and λ.

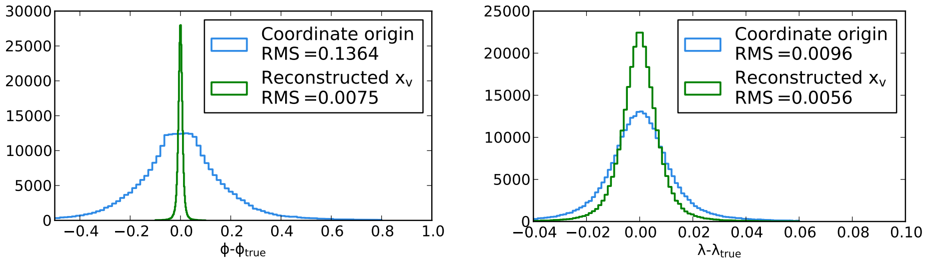


Figure 6.9: Residuals of the reconstructed track angles ϕ and λ with respect to the track angles angles ϕ_{true} and λ_{true} at the true vertex position. The distributions are shown for tracks propagated to the point of closest approach to the coordinate origin and for tracks bent and extrapolated to the reconstructed vertex position $\mathbf{x}_v$. The RMS is always calculated for data between the 5th and 95th percentile.

6.1.4 χ^2 Distribution

The distribution of $\chi^2(\mathbf{x}_v)$ given in Equation 4.18 for the reconstructed vertex positions $\mathbf{x}_v$ is shown in Figure 6.10.

From considerations in the beginning of this chapter it is known that the distribution of $\chi^2(\mathbf{x}_v)$ given in Equation 4.18 will be a χ^2 distribution with $k = 3$ degrees of freedom:

$$\chi^2(\mathbf{x}_v) \sim \chi^2_3 \, .$$

Consequently, the distribution's mean will be expected at $\mu_{theo} = 3$. The calculated mean of the simulated data yields $\mu \approx 2.89$ and thus fits well to the expected value. This mean is based on the acquired data between its 0th and 99th percentile, i.e. it is based on 99 % of the data sample.

Figure 6.10 shows that $\chi^2(\mathbf{x}_v)$ follows indeed a χ^2_3 distribution. Thus, the reconstruction algorithm is working and $\chi^2(\mathbf{x}_v)$ is well-defined in terms of the minimization.

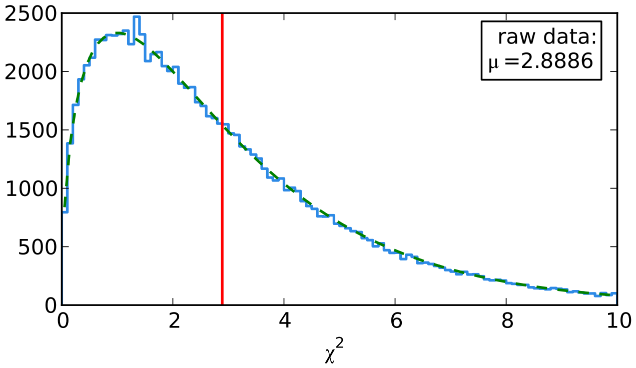


Figure 6.10: Distribution of $\chi^2(\mathbf{x}_v)$ for reconstructed vertex positions $\mathbf{x}_v$. The red vertical line shows the calculated mean μ of the distribution. The green dashed line is a fit of a scaled probability density function of a χ^2_3 distribution.

6.1.5 Momentum Dependence

The effects of multiple scattering for particles with different momenta p are already shown in chapter 3. According to Equation 3.3 the dependence of the variance of the multiple scattering distribution on the particle momentum is:

$$\theta_0^2 \propto p^{-2} \, .$$

This dependence might affect the vertex reconstruction, too.

After an event went through the fitting procedure the lowest particle momentum p_{min} of this event is calculated and the corresponding results are classified on the basis of p_{min} into specific momentum bins. Due to the kinematics the largest possible value for the smallest momentum corresponds to $m_\mu/3$, i.e. $p_{min} \approx 35\,\text{MeV/c}$.

First the distribution of the vertex position residuals Δ in every spatial dimension as a function of the respective momentum bin p_{min} is shown. For each momentum bin, the mean of the distribution vanishes as expected. The width of the distribution can again be understood as a vertex resolution $\delta\mathbf{x}_v$ (see subsection 6.1.2).

The relation between the momentum bin p_{min} and the vertex resolution $\delta\mathbf{x}_v$ is shown in Figure 6.11. Obviously, low momenta cause a large width in the distribution. This matches the expectation, because low momentum particle tracks are highly bent and strongly deflected due to multiple scattering.

The distribution of $\chi^2(\mathbf{x}_v)$ for reconstructed vertex positions $\mathbf{x}_v$ seems not to be affected as shown in Figure 6.12. All distributions still follow the expected χ^2_3 distribution.

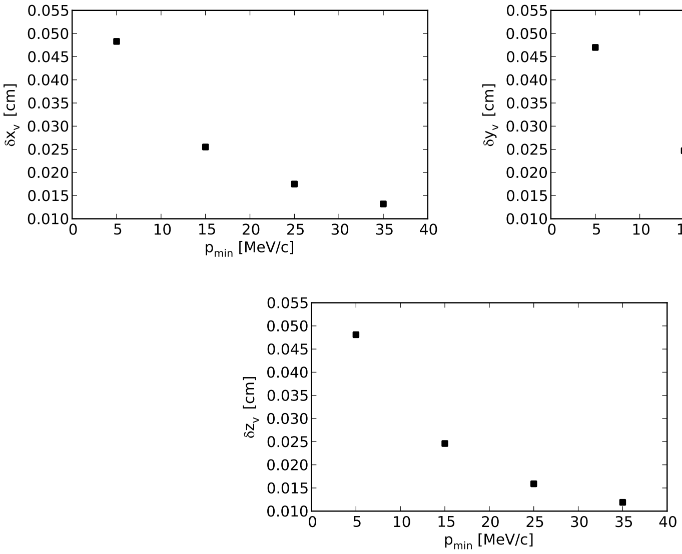


Figure 6.11: Dependence of the vertex resolution $\delta\mathbf{x}_v$ in every spatial dimension on the momentum bin p_{min}.

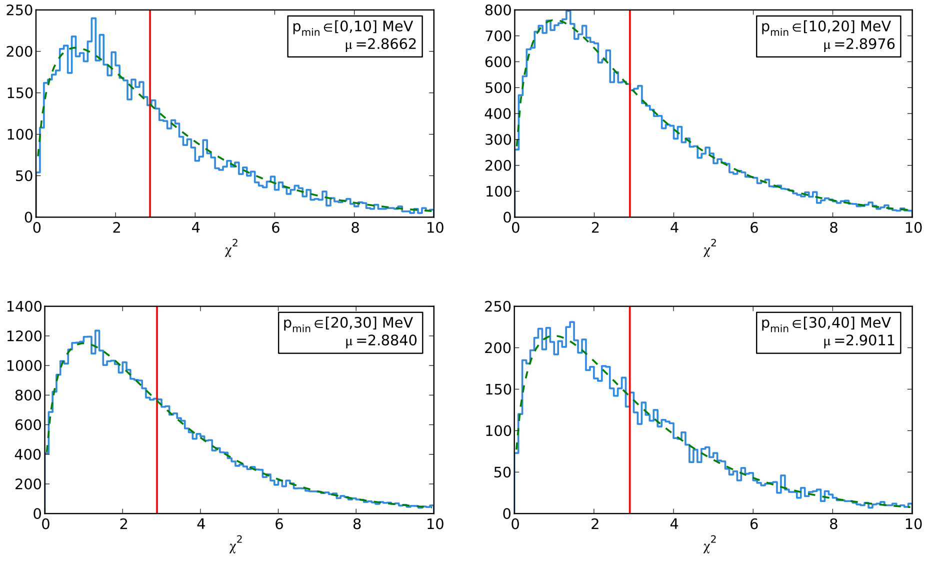


Figure 6.12: Distribution of $\chi^2(\mathbf{x}_v)$ classified into lowest momentum bins. No big differences occur.

6.2 Configuration Dependencies

Since the performance of the vertex reconstruction algorithm was discussed in terms of a fixed configuration set in the previous section, it might behave different under configuration changes.

These changes in the configuration set can be classified into two groups. On the one hand there are internal changes, i.e. choices that can be arbitrarily made on the software side, e.g. which initial vertex finding procedure to use. On the other hand there are external changes, i.e. constraints that are given by the setup of the experiment, e.g. the pixel structure or the thickness of the detector layers.

Several changes in the configuration set and the resulting behaviour will be discussed in the following section. The configuration from section 6.1 will be used, but this time with 10^4 generated events in total and only one element of the configuration set will be changed at a time. The other elements remain unchanged, if it is not mentioned explicitly.

Effects of changes in the configuration set will only be discussed in detail if they are significant. The plots in the following section are shown without error bars, because they are mainly used to show correlations qualitatively.

6.2.1 Internal

Initial Vertex Finding

Since the multiple scattering angles Φ and Θ are linearised around an initial vertex position $\mathbf{x}_{v,0}$, it is necessary to choose appropriate values for $\mathbf{x}_{v,0}$.

In section 6.1 the method `s-Dist` was used for establishing an initial guess for the possible vertex position.

`Mid-Target` Effects can only be observed in the reconstruction efficiency shown in Figure 6.13.

The reconstruction efficiency is significantly worse. The vertices of almost 10 % of the generated signals can not be reconstructed properly, because the fit does not converge at all or starts to diverge towards the outside of the first detector layer. This means that almost 10 % of the acquired data is lost in the context of vertex reconstruction.

The remaining data for which the fit did converge yield the same results as in the previous section. There are no significant changes in the vertex resolution, the pull distributions or the χ^2 distribution of the reconstructed vertices.

The effects on the reconstruction efficiency can be explained due to the geometry of the target. The distance between the initial vertex position and the true vertex position will always be in a range between the target's radius R and half of its length L. If the muon beam is assumed to be Gaussian as

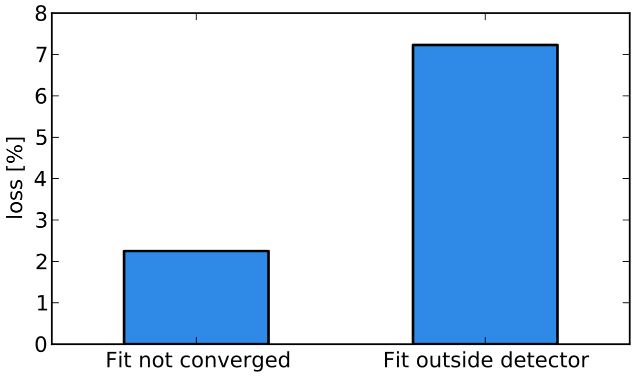


Figure 6.13: Reconstruction efficiency of a signal sample with 10^4 events with initial vertex finding method `Mid-Target`.

it is done in this configuration, the true vertex position will be more likely at the tips of target so the distance will be strongly shifted to values close to $L/2$. This is a bad choice for an initial vertex position in some parts of the generated phase-space.

`z-Dist` This method yield the same results like `s-Dist`. No significant effects on the reconstruction efficiency, vertex position residuals, pull distributions or χ^2 distribution can be observed.

Termination of Fitting Procedure

Because an essential part of the Mu3e experiment is the fast online data analysis, it might be interesting to know what results are returned by the reconstruction algorithm, if it only has a limited number of iterations.

`Maxiter` After i_{max} iterations $\chi^2(\mathbf{x}_v)$ is calculated and compared to the one from the iteration before as a crosscheck, i.e. $\chi^2_{i_{max}} - \chi^2_{i_{max}-1}$ is calculated. Its dependence on i_{max} is shown in Figure 6.14. Since signal samples with 10^4 events each are analysed, the shown value of $\chi^2_{i_{max}} - \chi^2_{i_{max}-1}$ in Figure 6.14 is the calculated mean of the values in the sample, based on data between the 0th and 90th percentile. For really small i_{max} the difference is of the order of magnitude of $\mathcal{O}(10^2)$, but then almost two orders of magnitude are achieved per maximum iteration step.

This can also be seen in an example χ^2 distribution for $i_{max} = 5$ shown in Figure 6.15. The mean is again calculated on a basis of 99 % data and yields a high deviation from the expected mean at $\mu_{theo} = 3$. The $\chi^2(\mathbf{x}_v)$ values causing the high mean are consequently not minimal.

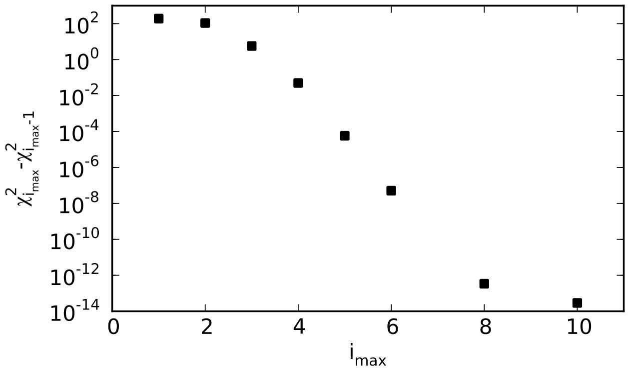


Figure 6.14: Mean of the difference between the values of $\chi^2(\mathbf{x}_v)$ in the last two iterations of a signal sample with 10^4 events as a function of the maximum iterations i_{max}.

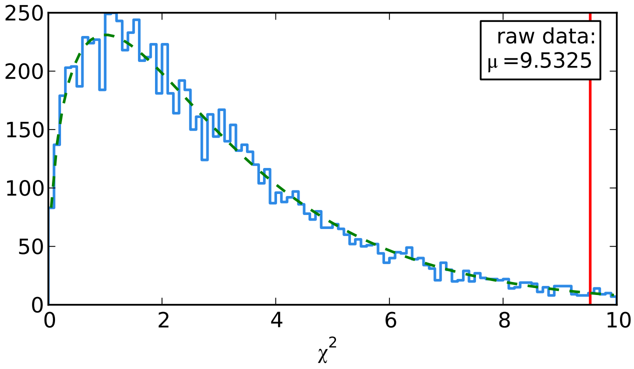


Figure 6.15: Distribution of $\chi^2(\mathbf{x}_v)$ for reconstructed vertex positions $\mathbf{x}_v$ with $i_{max} = 5$. The red vertical line shows the calculated mean μ of the distribution. The green dashed line is a fit of a scaled probability density function of a χ^2_3 distribution.

Another effect of using `Maxiter` can be seen in the vertex resolution $\delta\mathbf{x}_v$ (see subsection 6.1.2). Since the observable effect on $\delta\mathbf{x}_v$ is similar in every spatial dimension it is sufficient to only look at the three-dimensional resolution given by its norm:

$$\delta r := \|\delta\mathbf{x}_v\| = \sqrt{\delta x_v^2 + \delta y_v^2 + \delta z_v^2}\ . \tag{6.3}$$

δr as a function of i_{max} is illustrated in Figure 6.16. For small i_{max} the three-dimensional width of the vertex position residuals is three times bigger than for high values of i_{max}. An almost constant plateau for the vertex resolution seems to start at the fourth iteration.

For the purpose of online data analysis and with reference to Figure 6.14 and Figure 6.16, maximum iterations $i \geq 4$ should be chosen.

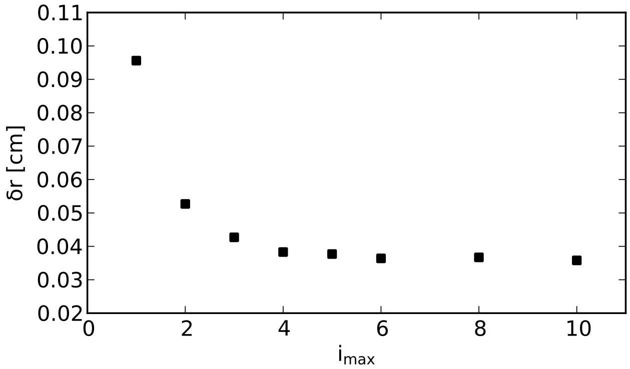


Figure 6.16: Dependence of the three-dimensional vertex resolution δr on the maximum iterations i_{max}.

6.2.2 External

Muon Beam Profile

In section 6.1 the muon beam was assumed to follow a two-dimensional normal distribution in order to describe the vertex reconstruction in detail in an almost realistic way.

This assumption will be substituted by a uniformly distributed muon beam and a signal generation in the coordinate origin in the following.

`Uniform` Distributing the true vertex positions uniformly on the target yield the same results as a Gaussian muon beam profile does.

No significant effects on the reconstruction efficiency, vertex position residuals, pull distributions or χ^2 distribution can be observed.

`Origin` Effects can only be observed on the reconstruction efficiency.

When every signal is generated in the coordinate origin, the reconstruction efficiency is slightly better than the one of the Gaussian beam profile shown in Figure 6.1, because less vertices are reconstructed outside the first detector layer. The reconstruction efficiency of a sample created in the coordinate origin is shown in Figure 6.17.

Less vertices are reconstructed outside the first detector layer, because the true vertex positions have the maximum distance to the detector material compared to a vertex somewhere on the target's surface. This also holds for the initial vertex in a small region around the true vertex position.

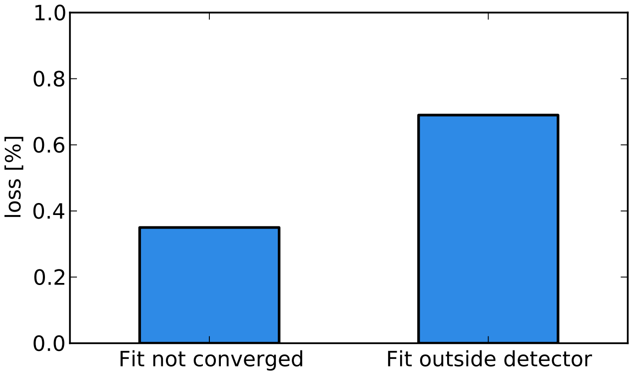


Figure 6.17: Reconstruction efficiency of a signal sample with 10^4 events created in the coordinate origin.

The vertex resolution, pull distributions and χ^2 distribution of reconstructed vertices remain unchanged.

Pixel Size

Since multiple scattering is assumed to be the dominant source of spatial uncertainties, the finite pixel resolution of the first detector layer has been neglected so far. Nevertheless, the fitting procedure might behave different when this assumption is loosened and a finite pixel structure is introduced according to the description in chapter 5.

Therefore, the distribution of the vertex position residuals Δ in every spatial dimension as a function of the pixel size p can be investigated. The standard deviation σ of the distribution can be understood as a vertex resolution $\delta\mathbf{x}_v$ (see subsection 6.1.2). The following configuration was used for each dataset:

$$C_p = \left(10^4, \texttt{Origin}, p, \texttt{s-Dist}, \texttt{Tol} = 10^{-8}\right) .$$

Again the observable effect on the vertex position resolution $\delta\mathbf{x}_v$ is similar in every spatial dimension it is sufficient to only look at the three-dimensional resolution δr defined in Equation 6.3.

The observed dependence of the resolution δr on the pixel size p is shown in Figure 6.18.

This figure shows clearly that in the range of a realistic pixel size of $p = 80\,\mu$m used in the Mu3e experiment [2] δr is nearly constant. The reconstruction algorithm seems not to be affected by the pixel size in this range. Moreover, an effect from the pixel structure is only visible for pixel sizes above $p > 200\,\mu$m.

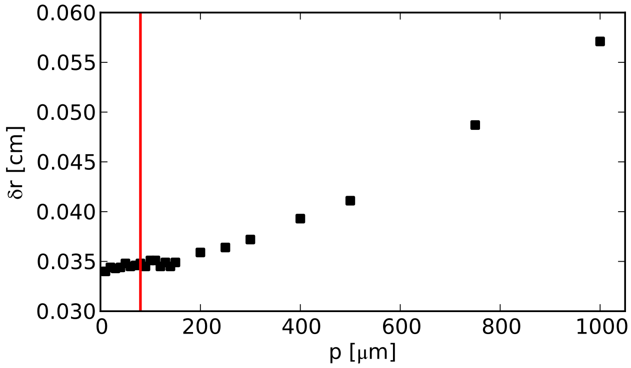


Figure 6.18: Dependence of the three-dimensional vertex resolution δr on the pixel size p in the first detector layer. The vertical red line shows the realistic pixel size of $p = 80\,\mu\text{m}$ to be used in the experiment [2].

Material Thickness

The dominant effects of multiple scattering are not only affected by the momentum of the particle traversing the material. Equation 3.3 claims for the standard deviation of the multiple scattering distribution:

$$\theta_0 \propto \sqrt{\frac{x}{X_0}}\left[1 + 0.038 \ln \frac{x}{X_0}\right] .$$

It depends on the ratio of the traversed material thickness x and its radiation length X_0. Since multiple scattering angles Φ and Θ are fitted, this dependence affects the vertex reconstruction algorithm, too.

The following configuration is used for each dataset.

$$C = \left(10^4, \texttt{Gauss}, 0, \texttt{s-Dist}, \texttt{Tol} = 10^{-8}\right) .$$

The effects of the material thickness can already be seen in the reconstruction efficiency, i.e. how much data is lost due to divergence or non-convergence of the fit. This dependency is illustrated in Figure 6.19 and shows that there is a constant plateau of about 1 % data loss in a thickness range between 10^{-7} and 10^{-3}. Thicknesses exceeding this range strongly influence the data loss, so that a thickness $x/X_0 \leq 10^{-3}$ is preferable.

The vertex position residuals yield that three-dimensional vertex resolution $\delta r = \|\delta \mathbf{x}_v\|$ (see subsection 6.1.2) is affected by the material as well. The dependence of δr on the material thickness in units of the radiation length x/X_0 is shown in Figure 6.20.

When going to lower x/X_0 the resolution δr gets constantly better up to very low δr as expected, because the generated tracks are less deflected by multiple scattering and thus can be reconstructed more precisely. However,

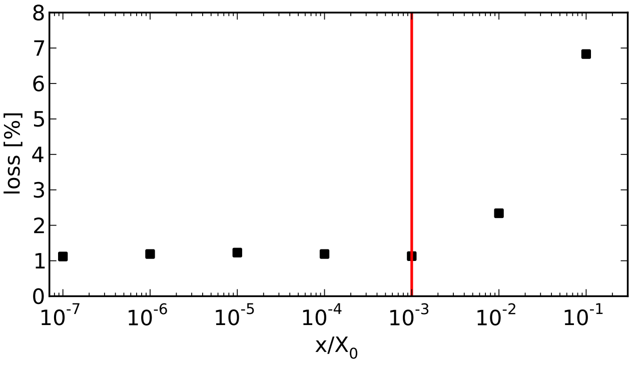


Figure 6.19: Reconstruction efficiency as a function of the material thickness x in units of its radiation length X_0. The vertical red line shows a realistic value for the purpose of the Mu3e experiment.

the more significant changes can be seen when going to higher x/X_0, strongly influencing the vertex resolution to worse values. In conclusion, the influence below a material thickness of $x/X_0 = 10^{-3}$ is not as strong as in the opposite direction, so that $x/X_0 = 10^{-3}$ is not an optimal but appropriate value for the material in terms of vertex reconstruction.

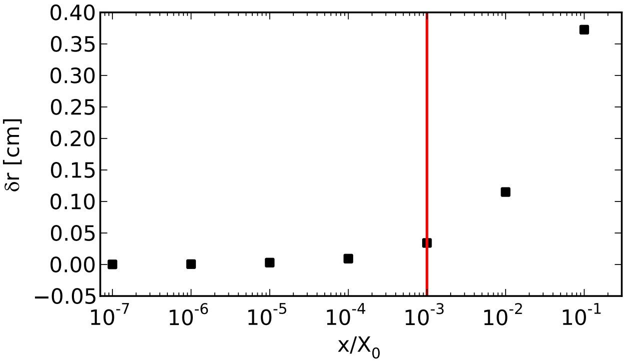


Figure 6.20: Dependence of the three-dimensional vertex resolution δr on the material thickness x in units of its radiation length X_0 of the first detector layer. The vertical red line shows a realistic value for the purpose of the experiment.

6.3 Background Suppression

For studying the suppression of possible backgrounds a signal sample S and a background sample B are generated according to chapter 5 and given to the vertex reconstruction routine both having the following configuration:

$$C = \left(10^4, \texttt{Gauss}, 0, \texttt{s-Dist}, \texttt{Tol} = 10^{-8}\right) .$$

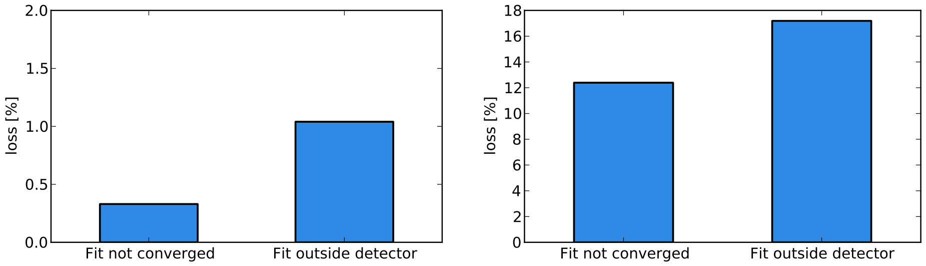
The suppression can be judged by taking the reconstruction efficiency and the χ^2 distribution of B into account as well as by setting up so called ROC[3] curves.

6.3.1 Reconstruction Efficiency

For both the signal sample S and the background sample B, the reconstruction efficiency of the algorithm can be studied in a similar fashion to the signal sample in subsection 6.1.1.

An illustration of the reconstruction efficiency is shown in Figure 6.21. The amount of events where the fit did not converge or converged outside the detector are clearly different for both samples. Almost 30 % of the events from the background sample cause the algorithm to diverge or to not converge. In contrast to the background sample, about 1.5 % of the signal sample can not be taken into account in the analysis.

The reconstruction algorithm does not seem to work very well on accidental background. This is desired in terms of background suppression.



(a) Signal sample (b) Background sample

Figure 6.21: Reconstruction efficiencies of a signal and background sample with 10^4 events each.

6.3.2 χ^2 Distribution

Another approach in distinguishing background from a possible signal is to take the χ^2 distribution into account. As seen in subsection 6.1.4, the

---

[3] Short for receiver operating characteristic, indicating their origin in the study of data transmission.

distribution of $\chi^2(\mathbf{x}_v)$ for a pure signal sample follows a χ^2 distribution with three degrees of freedom, i.e. $\chi^2(\mathbf{x}_v) \sim \chi^2_3$.

Figure 6.22 shows the corresponding distribution of $\chi^2(\mathbf{x}_v)$ for a pure background sample.

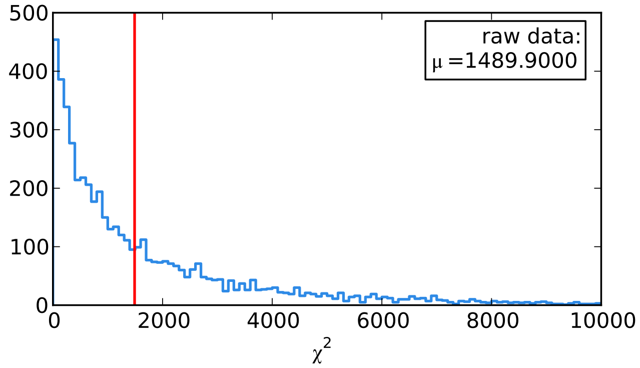


Figure 6.22: χ^2 distribution of an accidental background sample. The red vertical line shows the calculated mean of the distribution.

It is evident that $\chi^2(\mathbf{x}_v)$ does not follow a χ^2 distribution at all. The scale of values $\chi^2(\mathbf{x}_v)$ is much larger than the ones from a pure signal sample as shown in Figure 6.10.

Consequently, the value of $\chi^2(\mathbf{x}_v)$ is a helpful criterion for distinguishing a possible signal from accidental background via vertex reconstruction.

6.3.3 ROC Curves

ROC curves are graphical plots originally used in signal detection theory. They show the performance of a binary classifier system[4] as its discrimination threshold is varied [26].

For the purpose of vertex reconstruction, this threshold could for example be a certain χ^2-cut which corresponds to the value of $\chi^2(\mathbf{x}_v)$, calculated for a reconstructed vertex position $\mathbf{x}_v$ using Equation 4.18. If the value $\chi^2(\mathbf{x}_v)$ for an event does not exceed the value of the χ^2-cut, the event is classified as *selected*, denoted with the index *sel*, i.e. the event would be taken into consideration for possibly being a signal. This acts as a binary classifier system. $S_{sel} \subseteq S$ and $B_{sel} \subseteq B$ are the sets of events from a sample classified as *selected*.

With these sets different quantities can be defined as function of a certain cut c.

---

[4] A system that classifies a given set of objects into two groups on the basis of whether the object has a certain property or not.

Efficiency The efficiency can be defined on the basis of the signal sample S via

$$\varepsilon(c) := \frac{|S_{sel}(c)|}{|S|} \ , \tag{6.4}$$

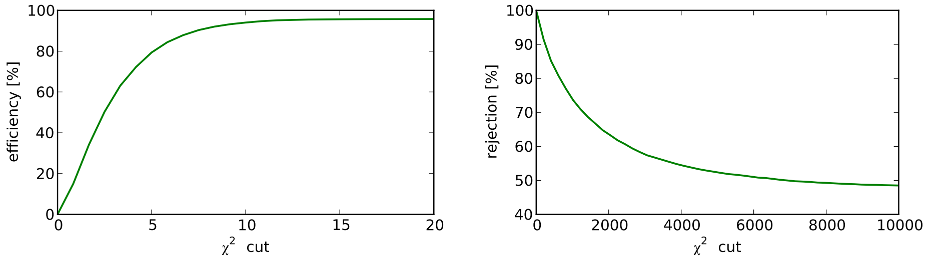
where $|\cdot|$ denotes the cardinality of the set. It is a measure for how many signals from a sample will be classified as an actual signal according to a specific cut criterion, i.e. $\varepsilon = 1$ ideally.

Rejection The rejection r can be defined as the opposite of the efficiency ε on the basis of a background sample B:

$$r(c) := 1 - \frac{|B_{sel}(c)|}{|B|} \ . \tag{6.5}$$

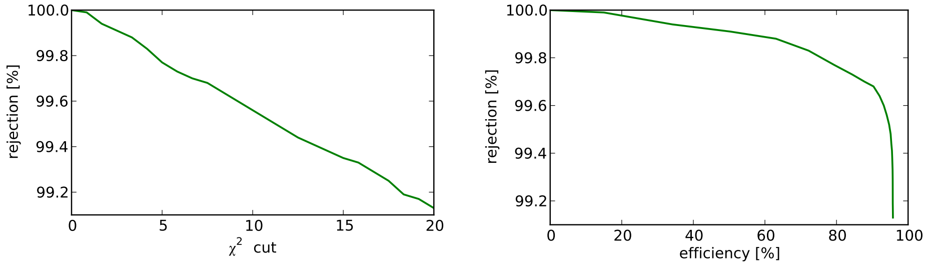
It is a measure for how many background will be suppressed and classified as actual background according to a certain cut criterion, i.e. $r = 1$ ideally.

The ROC curves for the configuration C defined above with the $\chi^2(\mathbf{x}_v)$ value as a threshold are shown in Figure 6.23.



(a) Efficiency

(b) Rejection



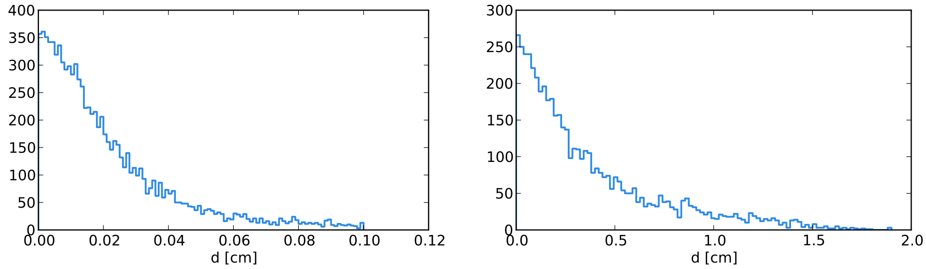
(c) Rejection excerpt

(d) Rejection vs efficiency

Figure 6.23: ROC curves for configuration C with $\chi^2(\mathbf{x}_v)$ value as threshold.

The efficiency has a fast rising edge with looser χ^2 cut and a flat plateau at about 96 % efficiency. The missing 4 % can be explained due to non-converging fitting procedures, fitted vertex positions outside the first detector layer (see subsection 6.1.1) or really high values of $\chi^2(\mathbf{x}_v)$. The rejection shows an inverse behaviour on a much larger χ^2 scale. Combining efficiency and rejection shows that a high efficiency and a high rejection can be achieved at the same time by finding an appropriate sweet-spot on the corresponding curve. Thus, the vertex reconstruction algorithm seems to be a helpful constraint in distinguishing signal from accidental background.

Since the muon decays are distributed on the target's surface, another choice for a threshold would be the distance d from the reconstructed vertex position to the target. The distribution of d for the signal and the background sample is shown in Figure 6.24.



(a) Signal sample

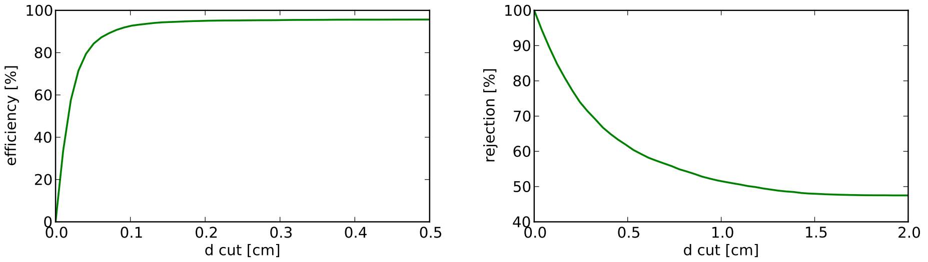
(b) Background sample

Figure 6.24: Distribution of the distance d of the reconstructed vertex position to the target's surface for a signal and a background sample. Note the different scale of the abscissa.

Because the scale of d for the signal sample is one order of magnitude smaller than for the background sample, it is also an appropriate choice for a threshold in ROC curves. These can be seen in Figure 6.25.

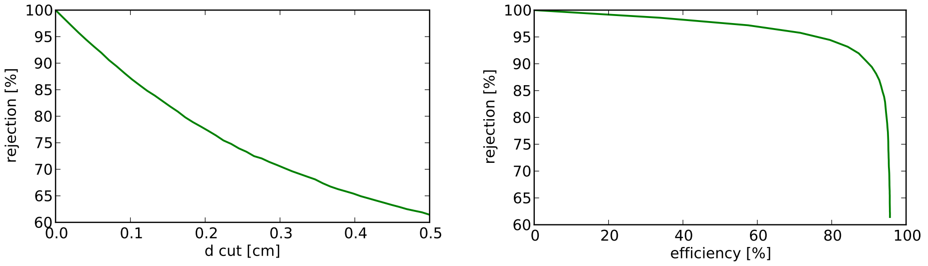
In general the ROC curves with $\chi^2(\mathbf{x}_v)$ and d as threshold show the same behaviour. However, choosing d as a threshold yields a much faster rejection drop in the range of a still rising efficiency, which might be difficult for finding an appropriate sweet-spot for high rejection at high efficiency.

Since the methods with $\chi^2(\mathbf{x}_v)$ and d as a threshold are very similar, the ROC curves could be optimized by combining both, e.g. vary the $\chi^2(\mathbf{x}_v)$ and d threshold at the same time.



(a) Efficiency

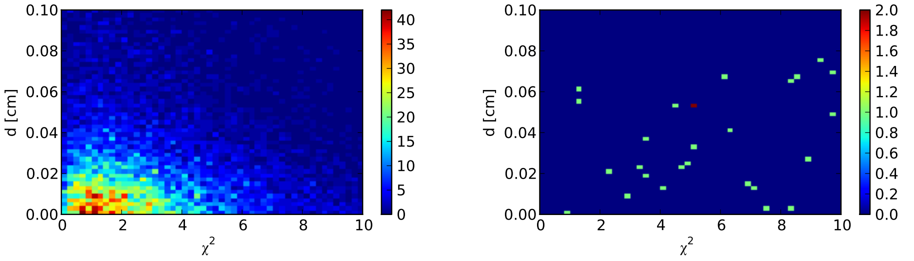
(b) Rejection



(c) Rejection excerpt

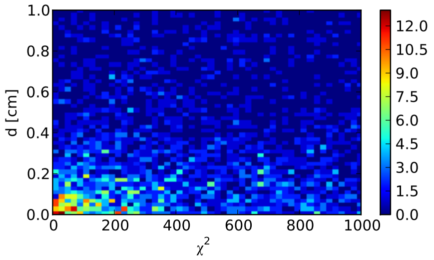
(d) Rejection vs efficiency

Figure 6.25: ROC curves for configuration C with distance d from the reconstructed vertex position to the target's surface as threshold.



(a) Signal sample

(b) Background sample



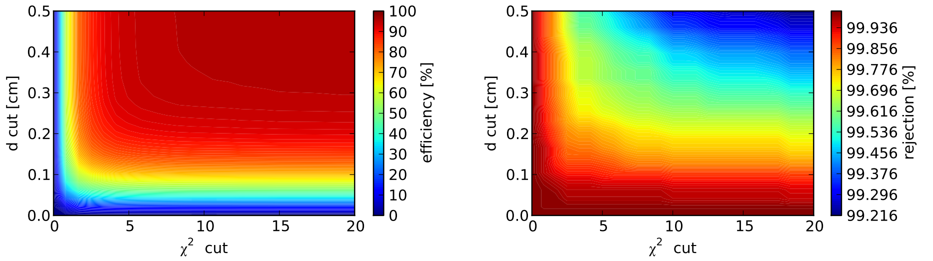
(c) Background sample on a large scale

Figure 6.26: Correlation between $\chi^2(\mathbf{x}_v)$ for reconstructed $\mathbf{x}_v$ and its distance d to the target surface for a signal and a background sample.

The correlation between $\chi^2(\mathbf{x}_v)$ and d for a signal and a background sample is illustrated in Figure 6.26 as a two-dimensional histogram. This figure shows that for most of the signal events $\chi^2(\mathbf{x}_v)$ and d are small at the same time, especially at the median of the χ^2_3 distribution. In contrast to the signal, the background sample does not show any clear correlation between $\chi^2(\mathbf{x}_v)$ and d on an appropriate scale.

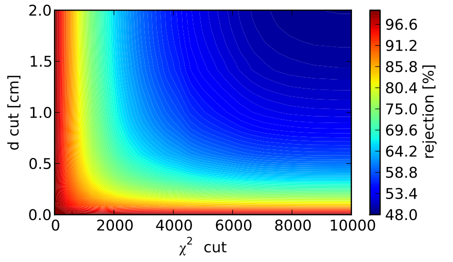
The correlation between $\chi^2(\mathbf{x}_v)$ and d for signal motivates the study of the efficiency and rejection in two dimensions, which might be helpful for finding the sweet-spot in rejection and efficiency. The two-dimensional ROC curves for efficiency and rejection are shown in Figure 6.27.

With appropriate choices for $\chi^2(\mathbf{x}_v)$ and d in the combined two-dimensional ROC curves the efficiency and rejection can indeed be improved. By comparing the correlation between $\chi^2(\mathbf{x}_v)$ and d from a signal sample to a background sample, one could for example find an appropriate sector of an ellipse with the major axis in χ^2 and d direction for choosing cut criteria to distinguish signal from background in the corresponding two-dimensional ROC-curves.



(a) Efficiency

(b) Rejection excerpt



(c) Rejection

Figure 6.27: Efficiency and rejection curve in two-dimensions, i.e. the $\chi^2(\mathbf{x}_v)$ and d threshold are varied at the same time.

6.4 Invariant Mass Resolution

The kinematics of the $\mu \rightarrow eee$ decay described in chapter 3 show that the system's invariant mass at the moment of the decay will be the muon mass (see Equation 3.1). However, when the particle tracks propagate helically towards the first detector layer only their absolute momentum will be conserved, while the angles between the different momentum vectors $\mathbf{p}_i$ will change[5].

Once the angles between the momentum vectors change, the net momentum appears not to be conserved or in other words the system's center of mass will not stay at rest with respect to the detector frame. This energy can only be taken from the muon mass. Thus, the invariant mass of a system in which the particles are propagated anywhere but to the original vertex position will always be smaller than the muon mass. Furthermore, the invariant mass of the particles at the reconstructed vertex position will be identical to the muon mass, if the reconstructed and original vertex coincide precisely.

The distributions of the invariant masses for propagation to the point of closest approach of an arbitrary spatial position or for bending the tracks directly to the reconstructed vertex position are compared. In this scenario, the arbitrary position is chosen to be the coordinate origin and the configuration from section 6.1 is used.

Figure 6.28 shows the residuals for both distributions with respect to the muon mass m_μ.

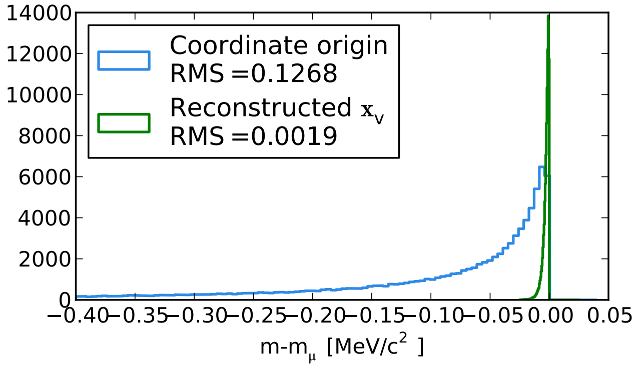


Figure 6.28: Distribution of invariant mass residuals with respect to the muon mass m_μ for signals propagated to the coordinate origin and the reconstructed vertex position. The RMS is always calculated on a data basis in range between the 5th and the 95th percentile.

The distribution with the reconstructed vertex position is much more

---

[5] This holds for the process of multiple scattering, too.

narrow than the other one for an $\mathbf{x}_v$ in the origin. This is expected, because the reconstructed vertex position will hopefully coincide with the true vertex position and thus the center of mass will be almost at rest. The muon mass will be obtained by calculating the system's invariant mass.

This is a useful aspect for background suppression, too. For every possible background, the upper bound for the invariant mass will also be m_μ. If an event is measured for which the invariant mass is smaller than the muon mass, it can be background with neutrinos carrying away momentum or it can be an actual signal. An additional estimation for distinguishing both scenarios would be reconstruction of $\mathbf{x}_v$, forcing the particle tracks to intersect with $\mathbf{x}_v$ and calculate the invariant mass of the particle tracks at $\mathbf{x}_v$. For an actual signal the calculation will much more likely yield the muon mass, while for background the neutrino momentum is still missing.

The mass resolution is highly influenced by badly reconstructed signal events, e.g. a mismeasurement of a track momentum or direction affects momentum balance and leads to a measurable missing momentum [2]. Rejecting these events will consequently improve the invariant mass resolution. For this purpose the acoplanar momentum vector $\mathbf{p}_{acoplanar}$ is defined as done by the SINDRUM collaboration [3]. $\mathbf{p}_{acoplanar}$ is the vectorial sum of all particle momenta projected into the decay plane, which is defined by the three tracks.

Due to multiple scattering the momentum vectors at the reconstructed vertex position will not necessarily lie in a single plane. Thus, a decay plane is defined first by its normal vector

$$\hat{\mathbf{n}} = \frac{\mathbf{p_1} \times \mathbf{p_2} + \mathbf{p_2} \times \mathbf{p_3} + \mathbf{p_3} \times \mathbf{p_1}}{\|\mathbf{p_1} \times \mathbf{p_2} + \mathbf{p_2} \times \mathbf{p_3} + \mathbf{p_3} \times \mathbf{p_1}\|} ,$$

where $\mathbf{p}_1$, $\mathbf{p}_2$ and $\mathbf{p}_3$ denote the momentum vector of the respective particle track.

In a similar way to reconstruct the invariant mass, the effects of vertex reconstruction on $\mathbf{p}_{acoplanar}$ can be studied. The distribution of its magnitude is shown with respect to the coordinate origin and the reconstructed vertex position $\mathbf{x}_v$ in Figure 6.29. For the signal decay $\mathbf{p}_{acoplanar}$ should vanish at the decay position due to its kinematics as seen in Equation 3.2.

Again the distribution with respect to the reconstructed vertex position is much more narrow, i.e. reconstructing vertices can compensate mismeasurements of momenta.

Since the distributions of the invariant mass difference and the acoplanar momentum show a similar behaviour, their correlation is shown in Figure 6.30 with respect to reconstructed vertices. The acoplanar momentum reconstruction seems to be dominant in reconstructing the invariant mass in the decay plane. If $p_{acoplanar}$ is small, the invariant mass difference is close to zero in most cases. Finally, the invariant mass resolution could be improved by applying an appropriate cut for $p_{acoplanar}$.

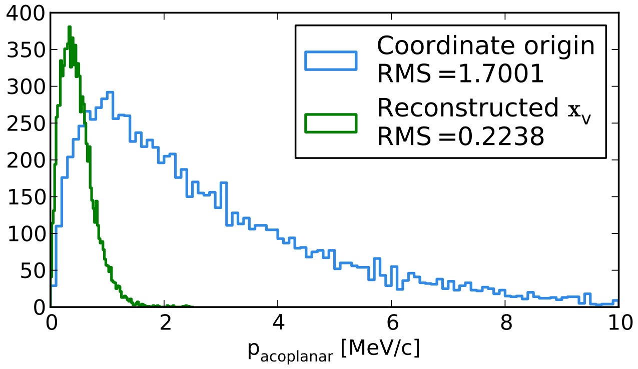
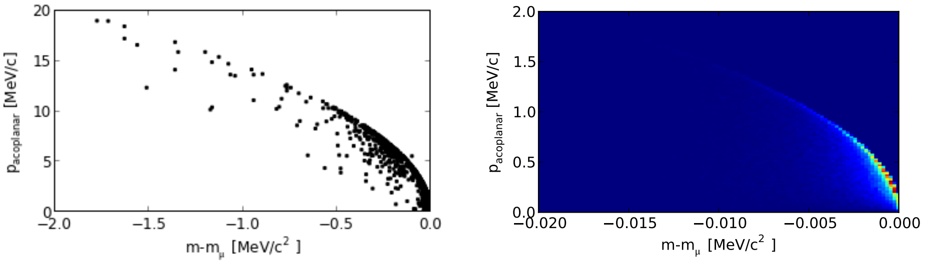


Figure 6.29: Distribution of the absolute acoplanar momentum $p_{acoplanar}$ for signals propagated to the coordinate origin and the reconstructed vertex position. The RMS is always calculated on a data basis between the 5th and the 95th percentile.



(a) Overview (b) Two-dimensional histogram excerpt

Figure 6.30: Correlation between invariant mass difference and acoplanar momentum.

Chapter 7

Conclusion

Since neutrino oscillations have been observed, it is evident that lepton flavour conservation as predicted by the original formulation of the Standard Model is violated in the neutral lepton sector. Many new physics models include lepton flavour violation in the charged sector as well, but an experimental discovery is still missing.

In order to search for physics beyond the Standard Model the Mu3e experiment is specifically designed to search for the lepton flavour violating muon decay $\mu \rightarrow eee$ with a branching ratio sensitivity of about $B = \mathcal{O}(10^{-16})$ [2] – four orders of magnitude more sensitive than the latest experimental result published by the SINDRUM collaboration [3].

The proposed branching ratio sensitivity requires high precision for timing, momentum and position measurements. This precision is achieved by a novel experimental design involving four thin cylindrical layers of pixel detectors based on HV-MAPS [14] positioned around a double hollow cone target. For precise momentum measurements a homogeneous solenoidal magnetic field is aligned along the rotational axis of the detector. Due to Lorentz forces the decay electrons move on helical trajectories in the magnetic field. The track curvature is defined by the momentum of the respective particle.

The decay electrons of $\mu \rightarrow eee$ are limited in their kinetic energy by an upper bound of approximately 53 MeV, leading to low momentum particle tracks. Consequently, they are strongly deflected due to multiple Coulomb scattering when traversing the detector layers. The effects of multiple scattering in the detector material are the dominant source of uncertainties in track reconstruction.

A useful characteristic of $\mu \rightarrow eee$ to distinguish it from accidental background is the fact that the decay products originate in a common vertex, which has to be reconstructed. Due to multiple scattering and bent particle tracks this reconstruction is not trivial.

In this thesis a linearised three-dimensional vertex fit for low momentum particles in a solenoidal magnetic field was presented and its performance in the context of the Mu3e experiment was investigated. This is a completely new approach in vertex reconstruction, because multiple scattering is taken into consideration as the only source of measurement uncertainties while the remaining track parameters such as momentum or hit position in the detector layers are assumed to be known exactly.

The performance was evaluated using a computer simulation written in PYTHON. A full detector simulation was not needed and several simplifications were made. Only the first detector layer was taken into consideration and assumed to have a perfect cylindrical shape. For evaluating the general performance of the vertex reconstruction the layer's finite pixel structure was neglected, i.e. spatial uncertainties were avoided. A Gaussian model was used to simulate multiple Coulomb scattering as the only source of uncertainties in track reconstruction.

The performance of the linearised vertex reconstruction algorithm was studied in terms of reconstruction efficiency, vertex position resolution, particle momentum and algorithm configuration dependencies, background suppression and invariant mass resolution.

The performance analysis of a signal sample showed that the linearised vertex reconstruction algorithm works quite well. The proper reconstruction of vertices only fails at the 1 % level. For the remaining fitted vertex positions $\mathbf{x}_v$ the minimized $\chi^2(\mathbf{x}_v)$ function is well defined and follows the expected χ^2_3 distribution with three degrees of freedom. The vertex position residuals of the signal sample showed that there is no spatial bias in the algorithm and yielded a total vertex position resolution of about $200\,\mu$m in each spatial dimension. This matches the expectation from early sensitivity studies [2]. The pull distributions of the reconstructed multiple scattering angles Φ and Θ showed that $\chi^2(\mathbf{x}_v)$ function shifts them to smaller values than expected or generated respectively because of the lack of constraints. However, this is an intrinsic problem of the fitting procedure by construction and could be resolved by adding additional constraints, e.g. requiring the particle tracks to be in the same plane. However, the effects of additional constraints again have to be studied in terms of background suppression. As expected, the vertex reconstruction works best for high momentum particles, because they are less bent in the magnetic field and less deflected by multiple scattering in the first detector layer.

The investigation of algorithm configuration dependencies showed that the vertex reconstruction is more affected by external changes than by internal ones.

On the internal side the method by which the reconstruction procedure is terminated, affected the vertex reconstruction the most. For proper reconstruction efficiency and appropriate vertex position resolution at least four iterations of the vertex reconstruction procedure should be executed, if fast online data analysis is considered. The method of initial vertex finding did not have any significant influence on the vertex reconstruction at all except for minor changes in the reconstruction efficiency. For the purpose of the Mu3e experiment, the methods described in this thesis seem to be sufficient. Nevertheless, improved initial guesses of an initial vertex position might be considered, e.g. the target's surface could be used to minimize the distance from the initial vertices to the surface or the point of closest approach of two helical tracks could be calculated.

On the external side only the pixel size and the thickness of the detector layers had a strong influence on the performance of the vertex reconstruction whereas the muon beam profile did not induce any effects. It turned out that a pixel size above $200\,\mu$m significantly reduces the vertex position resolution. A realistic pixel size of $80\,\mu$m as to be used by the Mu3e experiment [2] is an appropriate choice in terms of vertex resolution, because in this range the reconstruction is independent of the pixel size. Moreover, the simplification of a vanishing pixel structure and neglecting spatial resolution for the purpose of vertex reconstruction was justified. Multiple scattering in the first detector layer was indeed the dominant source of uncertainties compared to spatial ones. The material thickness yielded similar results - a relative thickness of $x/X_0 = 0.001$ is again appropriate for having high vertex resolution.

Another essential achievement of the vertex reconstruction is the suppression of accidental background. In contrast to the signal samples considered before, the vertex fitting procedure does not work properly on accidental background samples, which is a desired aspect in terms of background suppression. The suppression was investigated using ROC curves with varying $\chi^2(\mathbf{x}_v)$ and distance to the target surface d thresholds for reconstructed vertices $\mathbf{x}_v$. The ROC curves showed that especially the value of $\chi^2(\mathbf{x}_v)$ is a helpful threshold for discriminating signal against background. High rejection at the 99 % level and high efficiency at the 95 % level can be achieved at the same time using $\chi^2(\mathbf{x}_v)$ as a discrimination threshold. This is a much better sweet-spot than using the distance d as a discrimination threshold. Furthermore, this sweet-spot might be even improved due to correlations between $\chi^2(\mathbf{x}_v)$ and d for signal, e.g. by choosing an appropriate sector of an ellipse with the major axis in $\chi^2(\mathbf{x}_v)$- and d-direction as a cut criterion. The intrinsic check of the assumption of a common vertex is a powerful tool for discriminating signal against accidental background.

Finally, the invariant mass resolution was investigated. Due to the kinematics of $\mu \to eee$ the system's invariant mass at the moment of the decay will be always the mass of the muon. However, when the decay electrons propagate in the magnetic field the angles between the different momentum vectors will change. The center of mass will not stay at rest with respect to the detector frame and the calculation of the system's invariant mass will yield values smaller than the muon mass. The exact muon mass will only be obtained, if the momentum vectors at the the true decay position are taken into the calculation. Thus, reconstructing the vertex will hopefully yield the muon mass. The corresponding distribution showed that the invariant mass resolution is significantly improved by the vertex reconstruction which is helpful for background suppression, because the upper bound for any possible background is also the muon mass. However, if the vertex is reconstructed properly, calculating the invariant mass will much more likely yield the muon mass for actual signal than for background. Similar results were seen for the acoplanar momentum, i.e. the sum of all particle momenta projected into the decay plane. The acoplanar momentum reconstruction is highly correlated to the invariant mass reconstruction. An appropriate cut in this quantity might consequently lead to an improvement of the invariant mass resolution.

In this thesis, it was shown that the linearised vertex reconstruction method for particles in a solenoidal magnetic field suffering from multiple scattering does work and that the assumption of multiple scattering as the only source of uncertainty is justified in the context of the Mu3e experiment. A vertex resolution of $200\,\mu$m can be achieved. The constraint of a common vertex for three particle tracks is a very helpful tool for discriminating signal against background and for reconstructing the invariant mass. For the real experimental situation the track reconstruction including a finite momentum resolution from more than one detector layer has to be taken into account and its influences on the vertex reconstruction have to be investigated. Furthermore, since the Mu3e experiment will run at high intensities, the vertex reconstruction has to be further improved in terms of stability and processing speed, e.g. by running it on a graphics processing unit.

Appendix A

Mathematical Derivations

A.1 Transverse Intersection Point

The following derivation for two intersection points of two circles is taken from [15]. All quantities in this derivation are given and evaluated in the transverse plane only.

In the transverse plane charged particle tracks in a solenoidal magnetic field are described by circles. Let R_1 and R_2 be the radii and $\mathbf{c}_1$ and $\mathbf{c}_2$ be the two center points of the circles in the transverse plane.

Both circles have two intersection points, if

$$|\mathbf{c}_1 + \mathbf{c}_2| < R_1 + R_2 \quad \wedge \quad |\mathbf{c}_1 + \mathbf{c}_2| > |R_1 - R_2| \quad .$$

Let $\mathbf{u}$ be the normalised vector, connecting both center points:

$$\mathbf{u} := \frac{\mathbf{c}_2 - \mathbf{c}_1}{\|\mathbf{c}_2 - \mathbf{c}_1\|} \ .$$

Furthermore, let $\mathbf{v}$ be the orthonormal vector to $\mathbf{u}$ in the transverse plane, i.e. $\|\mathbf{u} \times \mathbf{v}\| = 1$.

The positions of the two intersection points represented by the coordinates (u, v) are then given by

$$u = \frac{(\mathbf{c}_2 - \mathbf{c}_1)^2 + R_1^2 - R_2^2}{2\,\|\mathbf{c}_2 - \mathbf{c}_1\|} \tag{A.1}$$

$$v = \pm\sqrt{R_1^2 - u^2} \ . \tag{A.2}$$

If the circles are separated from one another, i.e. no intersection points exist, the center point of the line connecting the circle center points is taken. If one circle contains the other, the smallest distance between both is taken. In the context of helical particle tracks this scenario is not very likely.

A.2 Solution of Transcendental Φ_{0V} Equation

The possible starting values for numerical methods are taken from [17].

A solution to Equation 4.12 only exists, if $d^2/(4R^2) < 1$.

There two possible scenarios, requiring different starting values of Φ_{0V} for solving Equation 4.12 numerically.
If $z^2 \leq d^2$, $\Phi_{0V} = \pm 2 \arcsin \frac{d}{2R}$ is to be used.
If $z^2 > d^2$, $\Phi_{0V} = \sqrt{\frac{z^2+d^2}{R^2}}$ is the more appropriate choice.

A.3 Spatial Derivatives of Φ and Θ

The derivatives of the multiple scattering angles Φ and Θ with respect to the vertex position $\mathbf{x}_v$ are calculated in the following section. This calculation can be found in a similar way in [15]. The following notation will be used:

$$\mathbf{x} = \mathbf{x}_H - \mathbf{x}_v \ ,$$

where $\mathbf{x}_H$ denotes the hit position of a particle track in the first detector layer.

At first the derivatives of the bending angle Φ_{0V} with respect to d and z have to be calculated from Equation 4.12, yielding:

$$\frac{\partial \Phi_{0V}(\mathbf{x}_v)}{\partial d} = \frac{\Phi_{0V}^2 d}{(\Phi_{0V}^2 R^2 - z^2)\sin\Phi_{0V} + 4z^2 \frac{\sin^2(\Phi_{0V}/2)}{\Phi_{0V}}} \tag{A.3}$$

$$\frac{\partial \Phi_{0V}(\mathbf{x}_v)}{\partial z} = \frac{4z \sin^2(\Phi_{0V}/2)}{(\Phi_{0V}^2 R^2 - z^2)\sin\Phi_{0V} + 4z^2 \frac{\sin^2(\Phi_{0V}/2)}{\Phi_{0V}}} \ . \tag{A.4}$$

The derivatives $\frac{\partial \Phi_{0V}(\mathbf{x}_v)}{\partial x}$ and $\frac{\partial \Phi_{0V}(\mathbf{x}_v)}{\partial y}$ are obtained by substituting:

$$\frac{\partial d}{\partial x} = \frac{x}{d} \tag{A.5}$$

$$\frac{\partial d}{\partial y} = \frac{y}{d} \ . \tag{A.6}$$

Finally, the derivatives with respect to $\mathbf{x}_v$ are obtained, using

$$K = (\Phi_{0V}^2 R^2 - z^2)\sin\Phi_{0V} + 4z^2 \frac{\sin^2(\Phi_{0V}/2)}{\Phi_{0V}} \ .$$

First Derivatives of Φ

$$\frac{\partial \Phi}{\partial x_v} = \frac{x}{2K}\Phi_{0V}^2 - \frac{y}{d^2} \tag{A.7}$$

$$\frac{\partial \Phi}{\partial y_v} = \frac{y}{2K}\Phi_{0V}^2 + \frac{x}{d^2} \tag{A.8}$$

$$\frac{\partial \Phi}{\partial z_v} = \frac{2z}{K}\sin^2\frac{\Phi_{0V}}{2} \tag{A.9}$$

First Derivatives of Θ

$$\frac{\partial\Theta}{\partial x_v} = \frac{xz\Phi_{0V}}{K\sqrt{\Phi_{0V}^2 R^2 - z^2}} \tag{A.10}$$

$$\frac{\partial\Theta}{\partial y_v} = \frac{yz\Phi_{0V}}{K\sqrt{\Phi_{0V}^2 R^2 - z^2}} \tag{A.11}$$

$$\frac{\partial\Theta}{\partial z_v} = \frac{1}{\sqrt{\Phi_{0V}^2 R^2 - z^2}}\left[\frac{4z^2\sin^2\frac{\Phi_{0V}}{2}}{K\,\Phi_{0V}} - 1\right] \tag{A.12}$$

A.4 Spatial Distribution of μ Decays

The geometry of a hollow double cone of maximal radius R and length L can be described in terms of the radius r as a function of the position z on the rotational symmetry axis as seen in Figure A.1:

$$r(z) = \begin{cases} -\frac{2R}{L}\,|z| + R & z \in \left[-\frac{L}{2}, \frac{L}{2}\right] \\ 0 & \text{otherwise} \end{cases} . \tag{A.13}$$

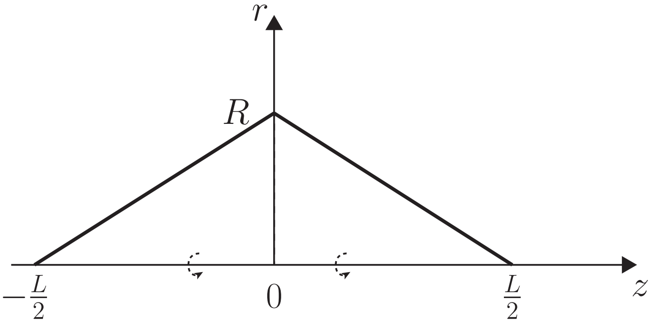


Figure A.1: Radius r of the double cone target as a function of longitudinal position z.

This expression can be rewritten as a piecewise linear function:

$$r(z) = \begin{cases} \frac{2R}{L}z + R & z \in [-\frac{L}{2}, 0[\\ -\frac{2R}{L}z + R & z \in [0, \frac{L}{2}] \\ 0 & \text{otherwise} \end{cases} . \tag{A.14}$$

The inverse function is then given by

$$z(r) = \begin{cases} \frac{L}{2R}r - \frac{L}{2} & z \in [-\frac{L}{2}, 0[\\ -\frac{L}{2R}r + \frac{L}{2} & z \in [0, \frac{L}{2}] \\ 0 & \text{otherwise} \end{cases} \quad . \tag{A.15}$$

These will be useful functions for generating the decay position on the target.

Now we want to consider two different beam distribution profiles (a uniform and a Gaussian one) and generate the muon decays on the target's geometry correspondingly.

A.4.1 Uniform Distribution

Considering a uniformly distributed muon beam profile, we want to draw samples from the resulting distribution $F(x)$ on the target without an explicit function but using a random number generator, drawing uniformly distributed samples $U \in [0, 1]$.

Consequently, we will make use of the so called *inversion method* as described in [27] in the following way:

"Let F be a continuous distribution function on $\mathbb{R}$ with inverse F^{-1} defined by

$$F^{-1}(u) = \inf \{x : F(x) = u, 0 < u < 1\} \ .$$

If U is a uniform $[0, 1]$ random variable, then $F^{-1}(U)$ has distribution function F. Also, if X has distribution function F, then $F(X)$ is uniformly distributed on $[0, 1]$" [27].

That is if we draw uniformly distributed samples $U \in [0, 1]$, $X = F^{-1}(U)$ will follow the desired distribution.

This applies to our case, because we have some a radial distribution $r(z)$ as given in Equation A.14. For simplicity I will calculate a normalised distribution $p(z) = \gamma r(z)$ with normalisation factor γ, which is given by the condition $\int_{-\infty}^{\infty} p(z)\mathrm{d}z = 1$. This yields $\gamma = \frac{2}{RL}$.

In the next step the distribution function $U = F(Z) = \int_{-\infty}^{Z} p(t)\mathrm{d}t$ is to be calculated and inverted, i.e. $Z = F^{-1}(U)$.

A straight forward calculation finally yields:

$$Z(U) = L \begin{cases} \sqrt{\frac{U}{2}} - \frac{1}{2} & 0 \leq U \leq \frac{1}{2} \\ \frac{1}{2} - \sqrt{\frac{1-U^2}{2}} & \frac{1}{2} < U \leq 1 \\ 0 & \text{otherwise} \end{cases} \quad , \tag{A.16}$$

with U uniformly distributed in $[0, 1]$. Note that these are only the solutions that are physically and geometrically relevant for the problem.

Finally, to generate a complete random three-dimensional position information for the muon decay on the target we can use a common cylindrical coordinate system (r, ϕ, z). The z-coordinate can be drawn from Equation A.16. The corresponding radius $r(z)$ can be calculated using Equation A.14 and finally a random azimuthal angle ϕ can be uniformly drawn from $[0, 2\pi]$ because of the rotational symmetry with respect to the z-axis.

A.4.2 Gaussian Distribution

Considering a two-dimensional Gaussian distributed muon beam profile, the approach for generating a sample from the resulting distribution on the target is different from the one described above.

Again, because of the rotational symmetry a common cylindrical coordinate system (r, ϕ, z) is appropriate. In fact the two-dimensional Gaussian distribution can be described by a one-dimensional Gaussian distribution in the radial coordinate r, which probability density function is described by

$$f(r) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(\frac{r-\mu}{\sigma}\right)^2\right) \tag{A.17}$$

with mean μ and standard deviation σ.

Because the muon beam itself will be collimated to a diameter of 2σ, the real distribution will be

$$p(r) = \chi_{[\mu-\sigma,\mu+\sigma]}(r) f(r) \;, \tag{A.18}$$

where χ denotes the indicator function on $[\mu - \sigma, \mu + \sigma]$.

Consequently, a sample for the r-coordinate can be drawn from Equation A.18 by choosing its absolute value $|r|$. Furthermore, a random azimuthal angle ϕ can be uniformly drawn from $[0, 2\pi]$. Because the corresponding z-coordinate given in terms of $z(r)$ in Equation A.15 is ambiguous for a given r, one of both segments has to be randomly chosen. This is accomplished with a $50 - 50\%$ chance.

Finally a complete random three-dimensional position information (r, ϕ, z) is given.

Bibliography

[1] J. Mureika, "Q&A: Prof. Jonas Mureika on the Higgs Boson." The Buzz: University News, Online Ressource. Visited on 09.09.2013.

[2] A. Schöning *et al.*, "Research Proposal for an Experiment to Search for the Decay $\mu \rightarrow eee$," Dec 2012.

[3] U. Bellgardt *et al.*, "Search for the decay $\mu^+ \rightarrow e^+e^+e^-$," *Nuclear Physics B*, vol. 299, no. 1, pp. 1 – 6, 1988.

[4] R. Oerter, *The Theory of Almost Everything: The Standard Model, the Unsung Triumph of Modern Physics*. Penguin Group US, 2006.

[5] B. Povh, K. Rith, C. Scholz, and F. Zetsche, *Teilchen und Kerne*. Springer-Lehrbuch; SpringerLink: Bücher, Berlin, Heidelberg: Springer Berlin Heidelberg, 8 ed., 2009.

[6] Y. Kuno and Y. Okada, "Muon decay and physics beyond the standard model," *Rev. Mod. Phys.*, vol. 73, pp. 151–202, 2001.

[7] Wikipedia Contributors, "Standard Model of Elementary Particles." Online Ressource, Jun 2013. Visited on 27.06.2013.

[8] J. Beringer *et al.*, "Review of Particle Physics," *Phys. Rev. D*, vol. 86, p. 010001, Jul 2012.

[9] J. Adam *et al.*, "New constraint on the existence of the $\mu^+ \rightarrow e^+\gamma$ decay," 2013.

[10] W. Bertl, SINDRUM II Collaboration, *et al.*, "A search for $\mu - e$ conversion in muonic gold," *Eur. Phys. J. C*, vol. 47, p. 337, 2006.

[11] M. Kiehn, "Track Fitting with Broken Lines for the MU3E Experiment," Diploma Thesis, University of Heidelberg, 2012.

[12] W. J. Marciano, T. Mori, and J. M. Roney, "Charged Lepton Flavor Violation Experiments*," *Annual Review of Nuclear and Particle Science*, vol. 58, no. 1, pp. 315–341, 2008.

[13] N. Berger, "The Mu3e Experiment," May 2013. Presented at the 1st Conference on Charged Lepton Flavor Violation, Lecce, Italy.

[14] I. Perić, "A novel monolithic pixelated particle detector implemented in high-voltage CMOS technology," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 582, no. 3, pp. 876–885, 2007.

[15] A. Schöning and S. Schenk, "Linearised Vertex 3D Fit in a Solenoidal Magnetic Field with Multiple Scattering." Internal note, Jul 2013.

[16] A. Strandlie and W. Wittek, "Derivation of jacobians for the propagation of covariance matrices of track parameters in homogeneous magnetic fields," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 566, no. 2, pp. 687 – 698, 2006.

[17] A. Schöning, "A Three-Dimensional Helix Fit with Multiple Scattering using Hit Triplets." Internal note, Sep 2012.

[18] Python Software Foundation, "Python Programming Language." Visited on 31.07.2013.

[19] E. Jones, T. Oliphant, P. Peterson, *et al.*, "SciPy: Open source scientific tools for Python," 2001–.

[20] R. Kleiss, W. J. Stirling, and S. D. Ellis, "A new Monte Carlo treatment of multiparticle phase space at high energies," *Computer Physics Communications*, vol. 40, no. 2, pp. 359–373, 1986.

[21] P. R. Kettle, "Towards High Intensity Stopped Muon Beams," Oct 2011. Presented at an internal Mu3e informal meeting, PSI, Switzerland.

[22] H. O. Lancaster and E. Seneta, *Chi-Square Distribution.* John Wiley & Sons, Ltd, 2005.

[23] E. W. Weisstein, "Chi-Squared Distribution." Mathworld - A Wolfram Web Ressource. Visited on 03.09.2013.

[24] R. Frühwirth, M. Regler, R. Bock, H. Grote, and D. Notz, *Data analysis techniques for high-energy physics*, vol. 11. Cambridge University Press, 2000.

[25] N. Berger and A. Schöning, 2013. Private Communication.

[26] Wikipedia Contributors, "Receiver operating characteristic." Online Ressource, Aug 2013. Visited on 05.09.2013.

[27] L. Devroye, *Non-uniform Random Variate Generation*, ch. II, pp. 27–29. Springer-Verlag, 1986.

Acknowledgements

An dieser Stelle schulde ich vielen Menschen meinen Dank, die mich während des Entstehens dieser Arbeit unterstützt haben.

Zuerst möchte ich mich bei Dr. Niklaus Berger bedanken, der mich überhaupt erst auf dieses Forschungsthema aufmerksam gemacht und mich während meiner gesamten Arbeits- und Schreibphase hervorragend betreut hat.

Außerdem möchte ich mich bei Prof. Stephanie Hansmann-Menzemer bedanken, die sich sofort bereit erklärt hat, diese Arbeit als Zweitgutachterin zu beurteilen.

Mein Dank gilt weiterhin Prof. André Schöning für die vielen anregenden Denkanstöße während meiner Arbeit. Ein besonderer Dank gilt außerdem Moritz Kiehn, der jederzeit während der gesamten Arbeits- und Schreibphase zur Verfügung stand, um Probleme zu diskutieren.

Bei allen Mitgliedern der Mu3e Arbeitsgruppe möchte ich mich für die sehr angenehme Arbeitsatmosphäre bedanken.

Ein großer Dank gilt selbstverständlich auch meiner Familie für ihre große Unterstützung während meines gesamten Bachelorstudiums.

Erklärung:

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

Heidelberg, den 12. September 2013 ....

Sebastian Schenk