Particle Track reconstruction using a recurrent neural network at the $\mu - 3e$ experiment

Bachelor thesis of
Sascha Liechti

06.04.2018

Supervised by
Prof. Nicola Serra
Dr. Patrick Owen

Abstract During the $\mu - 3e$ experiment we faced the challenge of reconstructing the paths of certain low momentum particles that curled back into the detector and cause additional hits. To face this, a recurrent neural network was used which found the right track for 87% of these particles.

Contents

1 Standard Model

1.1 Elementary particles and forces

The Standard Model(SM) describes all known elementary particles as well as three of the four known forces[1].
The elementary particles that make up matter can be split into two categories, namely quarks and leptons. There are 6 types of quarks and six types of leptons. The type of a particle is conventionally called flavour. The six quark flavours and the six lepton flavours are separated over 3 generations (each which two quarks and two leptons in it). Experimental evidence suggests that there exist exactly three generations of particles. Each particle of the first generation has higher energy versions of itself with the similar properties, besides their mass, (e.g. $e^- \rightarrow \mu^- \rightarrow \tau^-$)as in other generations. For each following generation, the particles have a higher mass than the generation before.

Table 1: Quarks in the Standard Model

Quarks				
	Particle		Q[e]	$\frac{mass}{GeV}$
1. Gen.	up	u	$-\frac{1}{3}$	0.003
	down	d	$\frac{2}{3}$	0.005
2. Gen.	strange	s	$-\frac{1}{3}$	0.1
	charm	c	$\frac{2}{3}$	1.3
3. Gen.	bottom	b	$-\frac{1}{3}$	4.5
	top	t	$\frac{2}{3}$	174

One category consists of quarks(q)(see Table 1). In this, we differentiate between up-type quarks, with charge $-\frac{1}{3}e$, and down-type, quarks with charge $\frac{2}{3}e$. Quarks interact with all fundamental forces.
Each quark carries a property called colour-charge. The possible color charges are red(r), green(gr), blue(bl) in which anti-quarks carry anti-colour. Quarks can only carry one colour, whilst every free particle has to be colorless[2]. In conclusion we cannot observe a single quark.
Free particles can achieve being colourless in two ways. Either by having all

[1] Strong, weak and electromagnetic forces
[2] Colour confinement

three colors present in the same amount (one quark of each color), which creates the characteristic group of baryons(qqq) and anti-baryons($\bar{q}\bar{q}\bar{q}$) or by having a color and its anticolor present, which creates the group of mesons($q\bar{q}$).

Table 2: Leptons in the standard model

Leptons				
	Particle		Q[e]	$\frac{mass}{GeV}$
1. Gen.	electron	e^-	-1	0.005
	neutrino	ν_e	0	$< 10^{-9}$
2. Gen.	muon	μ^-	-1	0.106
	neutrino	ν_μ	0	$< 10^{-9}$
3. Gen.	tau	τ^-	-1	1.78
	neutrino	ν_τ	0	$< 10^{-9}$

The other group consists of leptons(l)(see Table 2). They only interact through the weak and the electromagnetic force. Each generation consists of a lepton of charge -1 and a corresponding EM neutrally charged neutrino. The electron has the lowest energy of all charged leptons. This makes the electron stable while the higher generation particles decay to lower energy particles.

The leptons of one generation, namely the charged lepton and its corresponding neutrino are called a lepton family. A lepton of a family counts as 1 to its corresponding lepton family number whilst a anti-lepton counts as -1.

Table 3: Fundamental forces

Force	Strengh	Boson		Spin	Charge	$\frac{mass}{GeV}$
Strong	1	gluon	g	1	0	0
Electromagnetism	10^{-3}	photon	γ	1	0	0
Weak	10^{-8}	Z boson	Z	1	0	80.4
	10^{-8}	W boson	$W^\pm$	1	± 1	91.2

The particles of the SM interact through the 3 fundamental forces of the SM. In these interactions, particles called bosons are being exchanged which are the carriers of their respective force (see Table 3).
As mentioned above, only quarks can interact through the strong force, in

which they exchange gluons. Gluons are massless and EM neutrally charged. The strong force has the biggest coupling strengh of 1 (though it decreases with higher energies as a result of gluon-gluon self interaction loops, which interfere negatively in perturbation theory)[1]. A gluon carries colour charge and hence can change the colour of a quark but it conserves its flavour. The strong interaction has an underlying gauge symmetry of SU(3). Therefore, it can be derived that color charge is conserved through the strong interaction[3].
The electromagnetic(EM) force is propagated through the photon. It carries zero charge and no invariant mass. Exclusively charged particles can interact through the electromagnetic force. The coupling strength is $\alpha \approx \frac{1}{137}$, contrary to the strong force the coupling constant increases with higher energies[1]. This difference stems from the fact that photon-photon interaction loops are not allowed whereas gluon-gluon interaction loops are. In perturbation theory this results in only positive terms being added to the coupling strength. The underlying gauge symmetry is of SU(1). The electromagnetic force also conserves flavour.
The weak force has two types of bosons. The bosons of the weak force are the only bosons to have an inertial mass.
First we will discuss the EM neutrally charged Z boson. Even though the Z boson belongs to the weak force it, it also has an electromagnetic part additionally to the weak force part[4]. It follows directly, that the Z boson couples weaker to uncharged particles.
The other boson of the weak force is the W boson. In the classical SM, the only way particles can change flavour is through the weak force by emitting or absorbing W boson. It is important to notice that, besides of having an invariant mass, the W boson is the only boson with a non zero charge ($Q_{W^\pm} = \pm 1e$). In the gauge symmetry of the weak force the $W^\pm$ are actually the creation and annihilation operators of said symmetry[5].
An important characteristic of the weak force is that it exclusively couples to lefthanded(LH) particles and righthanded(RH) antiparticles (describing chirality states)[6].
The chirality operators for left- and righthandedness are:

LH: $\frac{1}{2}(1-\gamma^5)$, RH: $\frac{1}{2}(1+\gamma^5)$

As a consequence RH particles and LH anti-particles cant couple to the W

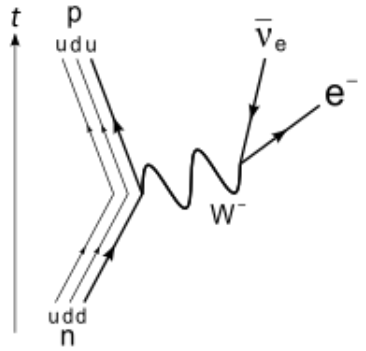
[3] E.g. through Gell-Mann matrices
[4] $Z \rightarrow EM_{part} + W^3$, [1]
[5] $W^\pm = W_1 \pm iW_2$
[6] In the ultrarelativistic limit helicity and chirality eigenstates are the same

boson at all. This also results in charged RH particles and LH anti-particles to couple to the Z boson only through the electromagnetic part of the itself, while uncharged RH particles and LH anti particles (e.g. RH ν, LH $\bar{\nu}$) don't couple with the EM force nor the weak force.

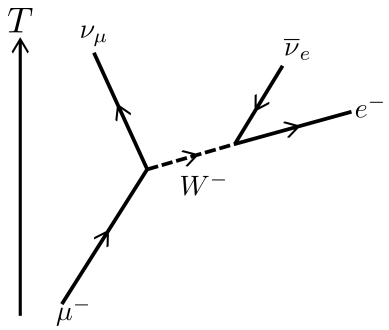
1.2 Interaction rules

Now we will establish the general rules for interactions in the SM.

Baryon number is conserved
As we already established before, the only interaction that can change flavour is the weak force through the W boson. We directly see that all other interactions baryon number has to be conserved. So any up-type quark can be changed to a down-type quark and backwards by emitting or absorbing a W boson. In the end however, there are still 3 quarks which form a baryon[7], even though it changed its type and charge. A well known example is the beta decay, where a down quark in a neutron decays into a an up quark to form now a proton(e.g. see Figure 1a). We easily see that the baryon number is conserved.



(a) Feynman diagram of the β-decay



(b) Feynman diagram of a μ-decay

Lepton family number is conserved
According to the SM lepton family number is conserved. As all interactions beside the W conserve particle flavour, it is easy to see that lepton family number is conserved.

[7]Pentaquarks($qqqq\bar{q}$) and other exotic states excluded

Whenever a lepton interaction with a W boson, it just changes a lepton to its corresponding lepton neutrino and or the other way around (e.g. see Figure 1b).

2 Physics beyond the SM

2.1 Neutrino Oscillation

Classically the SM considers neutrinos to be massless. While this assumption works well for a lot of cases, we know nowadays that at least two of the three neutrinos have to have mass[8]. Neutrinos are known to oscillate between all three states of flavour, as the eigenstates of flavour are not eigenstates of mass. As a consequence ν_e, ν_μ and ν_τ are not fundamental particle states but a mixture of the mass eigenstates ν_1, ν_2 and ν_3. They are connected through the PMNS matrix:

$$\begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} = \begin{pmatrix} U_{e_1} & U_{e_2} & U_{e_3} \\ U_{\mu_1} & U_{\mu_2} & U_{\mu_1} \\ U_{\tau_1} & U_{\tau_2} & U_{\tau_3} \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} \tag{1}$$

As a result neutrinos propagate as a superposition of all mass eigenstates. Additionally we can describe the PMNS matrix through three mixing angles θ_{12}, θ_{13} and θ_{23} and a complex phase δ [9]. The electron superposition looks then like this:

$$|\nu_e\rangle = U_{e_1} |\nu_1\rangle e^{-i\Phi_1} + U_{e_2} |\nu_2\rangle e^{-i\Phi_2} + U_{e_3} |\nu_3\rangle e^{-i\Phi_3} \text{ with } \Phi_i = E_i \times t$$

As a result lepton family number is not a conserved quantity anymore as neutrino flavour oscillates over time.
We can calculate the probability for a neutrino to transition from flavour α to β like:

[8]The mass difference between neutrinos is non zero: $m_i - m_j = \Delta m_{i,j} \neq 0$, $\forall j \neq i$
[9]Measurements: $\theta_{12} \approx 35°$, $\theta_{13} \approx 10°$, $\theta_{23} \approx 45°$ [2], [3]

$$
\begin{aligned}
P(\nu_\alpha \to \nu_\beta) = & 2Re\left(U_{\alpha_1} U^*_{\beta_1} U^*_{\alpha_2} U_{\beta_2} e^{-i(\Phi_1 - \Phi_2)}\right) \\
& + 2Re\left(U_{\alpha_1} U^*_{\beta_1} U^*_{\alpha_3} U_{\beta_3} e^{-i(\Phi_1 - \Phi_3)}\right) \\
& + 2Re\left(U_{\alpha_2} U^*_{\beta_2} U^*_{\alpha_3} U_{\beta_3} e^{-i(\Phi_2 - \Phi_3)}\right)
\end{aligned} \tag{2}
$$

An important thing to note is, that if any elements of the PMNS matrix are complex, this process is not invariant under time reversal ($t \to -t$)[10]
$P(\nu_\alpha \to \nu_\beta) \neq P(\nu_\beta \to \nu_\alpha)$.

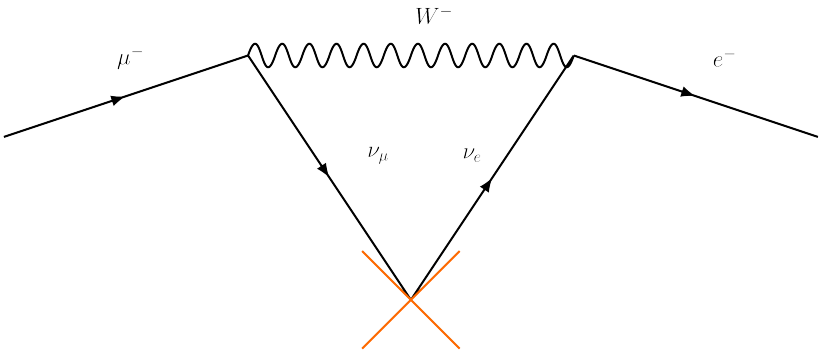


Figure 2: Process that violates lepton family number conservation through neutrino oscillation

Nowadays it's a well accepted fact that lepton family number gets violated through neutrino oscillation.
But why should flavour oscillation be exclusive to neutrinos?
Maybe there are ways for the EM charged leptons as well to directly transition to another lepton family[11]?

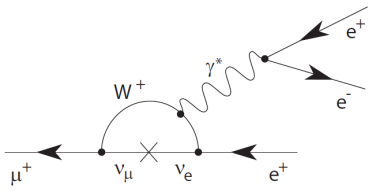
2.2 New physics

As a consequence of neutrino oscillation lepton flavour is a broken symmetry. The SM has to be adapted to include lepton flavour violation (LFV) and massive neutrinos. LFV is also expected for charged neutrinos.
Although it has yet to be determined how LFV violation exactly works to which scale it exists.

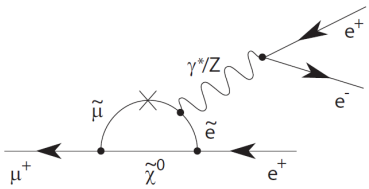
[10] The probability does not change if we add a complex phase to the PMNS matrix, just if one of the elements has a phase different from the others
[11] Maybe also possible for quarks?

This may raise the question on why charged LFV has never been observed yet. This is especially surprising as the mixing angles of the neutrinos have been measured to be big.
There are two reasons why charged LFV is strongly surpressed: The first is that charged leptons are much heavier than neutrinos and the other that the mass differences between neutrino flavour are tiny compared to the W boson mass.

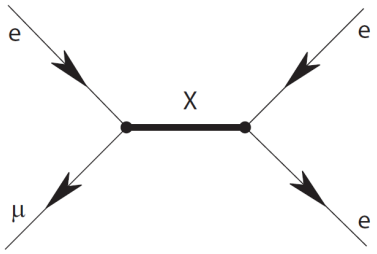
In the classical SM, charged LFV is already forbidden at tree level. Though it can be induced indirectly through higher order loop diagrams (using neutrino oscillation). By adding new particles beyond the SM, we generate new ways for LFV in the charged sector to happen. As LFV is naturally generated in many models beyond the SM, finding charged LFV is a strong hint for new physics.



(a) LFV through neutrino oscillation



(b) LFV by using supersymmetric particles



(c) LFV at tree level

One way charged LFV can occur is through super symmetric particles (see Figure 3b). By observing charged LFV supersymmetry would gain new importance.
Together with supersymmetric models, other extensions of the SM such as left-right symmetric models, grand unified models, models with an extended Higgs sector and models where electroweak symmetry is broken dynamically

are all good candidates to explain charged LFV and most importantly experimentally accessible in a large region of the parameter space.

3 $\mu \to eee$ decay

3.1 Kinematics

The two most prominent charged LFV decays are $\mu \to e\gamma$ and $\mu \to eee$. Here the latter is chosen as more diagrams beyond the SM contribute. Namely tree diagrams, Z penguin and box diagrams. This offers the possibility to test more models.

Possible ways for the decay $\mu \to eee$ to occur are shown in Figures 3a, 3b, 3c.

Still some simplifications are made as it is assumed that only the tree and the photon diagram are relevant. [4]

This gives us a Lagrangian of:

$$L_{LFV} = \left[\frac{m_\mu}{(\kappa+1)\Lambda^2}\overline{\mu_R}\sigma^{\mu\nu}e_L F_{\mu\nu}\right]_{\gamma-\text{penguin}} + \left[\frac{\kappa}{(\kappa+1)\Lambda^2}(\overline{\mu_L}\gamma^\mu e_L)(\overline{e_L}\gamma_\mu e_L)\right]_{\text{tree}} \tag{3}$$

If we neglect signal and background we can use momentum conservation as the decay happens rather quickly. As a result the total sum of all particle momenta should be equal to zero:

$$|\vec{p}_{tot}| = \left|\sum \vec{p}_i\right| = 0 \tag{4}$$

The particles resulting in the decay lie all in a plane. The resulting positrons and electrons are in the energy range of (0-53)MeV.

3.2 Background events

3.2.1

The event $\mu \to eee\nu\nu$ results in the same particles seen by the detector as the event we are searching for[12]. As a result it proves to be quite challenging to s$\mu \to eee\nu\nu$eparate the two.
By using momentum conservation, it becomes possible to differentiate the $\mu \to eee$ and the $\mu \to eee\nu\nu$ events. In the muon rest frame the total momentum is zero and the energy of the resulting particles is equal to muon rest energy.

[12]Neutrinos are invisible to our detector

By reconstructing the energy and momenta of the three e we can check if their momenta add up to zero and their energies equal the muon rest energy. If not we can assume that there are additional neutrinos. This differentiation between the two events is crucial for the experiment as the $\mu \to eee\nu\nu$ events pose the most serious background for $\mu \to eee$ decay measurements.
As a result, our detector needs a very good energy resolution to consistently make it possible to differentiate between the two events as neutrino energies and momenta are very small.

3.2.2 Michel decay

The biggest contributing background however stems from another decay called Michel decay, that is also allowed in the classical SM. As we use a beam of positive muons the corresponding Michel decay looks as follows: $\mu^+ \to e^+\nu\bar{\nu}$.
Contrary to the events before this one does not produce any em negatively charged particles. This makes these events easily distinguishable from our wanted events. As a result they only enter our data in form of a potential background through wrongly constructed tracks.

3.2.3 Radiative muon decay

This is the case where $\mu \to e^+\gamma\nu\nu$. If the photon produced in this event has high enough energies and creates a matter antimatter pair in the target region ($\gamma \to e^-e^+$), it can create a similar signature than the searched event. They contribute to the accidental background, as equal to the searched event no neutrinos are produced. To minimize these effects, the material in both the target and detector is minimized and a vertex constraint is applied.

3.2.4 BhaBha scattering

Another way how background can get produced is when positrons from muon decays or the beam itself scatter with electrons in the target material. Consequently they share a common vertex and together with an ordinary muon decay it can look similar as our searched $\mu \to eee$ event. This contributes to the accidental background.

3.2.5 Pion decays

Certain pion decays also lead to indistinguishable signature as our searched event, the most prominent being the $\pi \to eee\nu$ and $\pi \to \mu\gamma\nu$ decays. The later only produces a similar signature if produced photon converts through

pair production to an electron and a positron.
However as only a negligible portion will actually contribute to the background, as there is only a small branching fraction and the momenta and energy of the produced particles have to match up with the criteria mentioned in section 3.1.

3.2.6 Analysis of the background

The results of simulations indicate that the effect of purely accidental background contributions are small for a high enough energy resolvent detector. [4]
The most relevant background stems from the $\mu \to eee\nu\nu$ events. This problem can only be tackled by using a very precise total energy resolution of $\sigma_E = 1MeV$ at the aimed sensitivities.

4 Mu3e experiment

4.1 Requirements

The ultimate goal of this experiment is to observe a $\mu \to eee$ event. As we strive for a sensitivity of 10^{-16} , we should be able to observe this process if its branching ratio would be higher than our sensitivity. Otherwise we want to exclude a branching ratio $> 10^{-16}$ with a 90% certainty.
To get to this sensitivity, more than $5.5 \cdot 10^{16}$ muon decays have to be observed. To reach this goal within one year, a muon stopping rate of $2 \cdot 10^{9} Hz$ in combination with a high geometrical acceptance as well as a high efficiency of the experiment is required.

4.2 Phase I

Phase I of the experiment serves as an exploratory phase to gain more experience with the new technology and validate the experimental concept. At the same time it already strives to produce competitive measurements with a sensitivity of 10^{-15}. [13] This will be done, by making use of the already existing muon beams at PSI with around 1-1.5 $\cdot 10^{8} Hz$ of muons on target. The lowered sensitivity also allows for some cross-checks as the restrictions on the system are much more relaxed than in phase II.

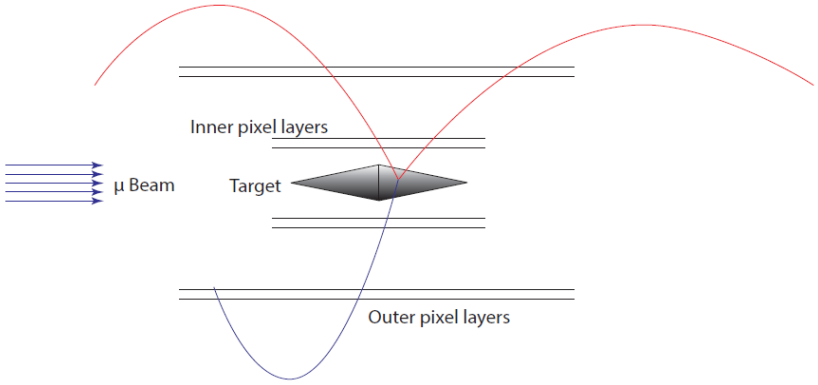
4.3 Phase II

Phase II strives to reach the maximum sensitivity of 10^{-16}. To achieve this in a reasonable time a new beamline will be used which delivers more than $2 \cdot 10^{9} Hz$ of muons.

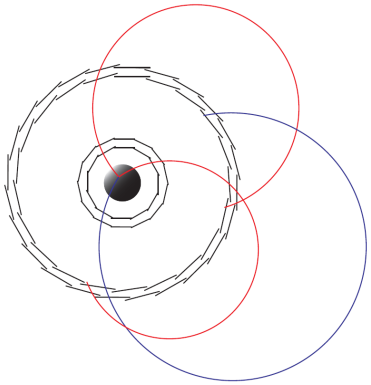
4.4 Experimental setup

The detector is of cylindrical shape around the beam. It has a total length of around $2m$ and is situated inside a $1T$ solenoid magnet with $1m$ of inner radius and a total length of $2.5m$. This form was chosen to cover as much phase space as possible. For an unknown decay such $\mu \to eee$, it crucial to have a high order of acceptance in all regions of phase space. There are only two kind of tracks that get lost. The first one are up- and downstream tracks and the second one are low transverse momenta tracks (no transversing of enough detector planes to be reconstructed).

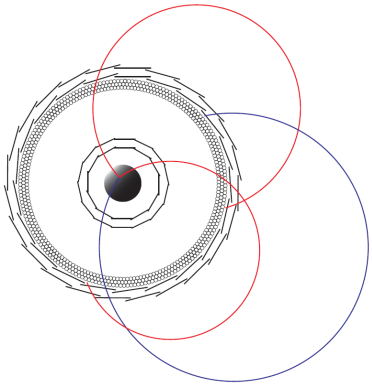
[13] Current experiments are in the 10^{-12} sensitivity range



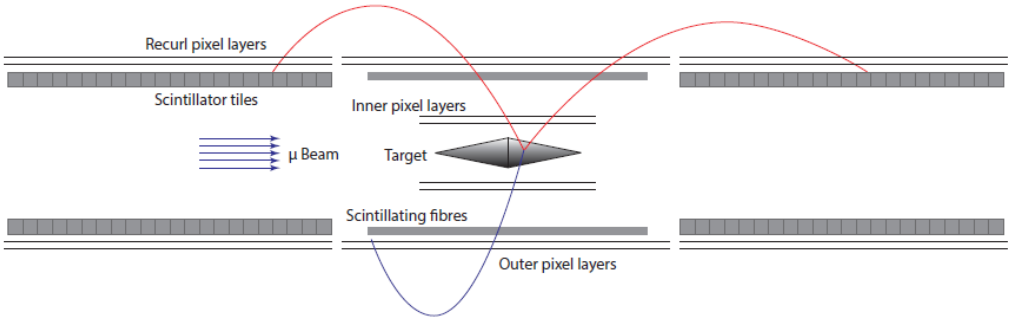
(a) Setup of the detector in the first part of phase I



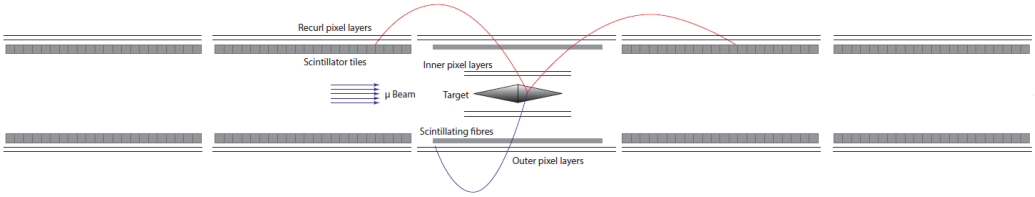
(b) Tracks in the detector in the first part of phase I



(c) Tracks in the detector in the second part of phase I and Phase II



(d) Setup of the detector in the second part of phase I



(e) Setup of the detector in phase II

As seen in figure 4e, the final version of the detector can be divided into 5 separate parts in the longitudinal direction. There is the central part with the target, two inner silicon pixel layers, a fibre tracker and two outer silicon layers. The forward and backward parts, called recurl stations, consist only of a tile timing detector surrounded by two silicon recurl layers. A big advantage pf this layout is that even a partially constructed detector (gradally over phase I to phase II parts get added) can give us competitive measurements.

The target itself is a big surfaced double cone with a surface length of $10cm$ and a width of $2cm$. The target was chosen specifically to be of this shape to facilitate separating tracks coming from different muons and hereby also helping to reduce accidental background.
The two inner detector layers, also called vertex layers, span a length $12cm$. The innermost layer consists of 12 tiles while the outer vertex layer consists of 18 tiles. The tiles are each of $1cm$ width, with the inner layer having an average radius of $1.9cm$, respectively $2.9cm$ [5], [6], [7]. They are supported by two half cylinder made up of $25\mu m$ thin Kapton foil mounted on plastic. The detector layers itself are $50\mu m$ thin and cooled by gaseous helium. The vertex detectors are read out at a rate of $20MHz$, giving us a time resolution of $20ns$.
After the vertex layers the particles pass through the fibre tracker (see Figure 4c, 4e). It is positioned around $6cm$ away from the center. Its main job is to provide accurate timing information for the outgoing electrons and positrons. It consist of three to five layers, each consisting of $36cm$ long and $250\mu m$ thick scintillating fibres with fast silicon photomultipliers at the end. They provide us a timing information of less than a $1ns$.
Next the outgoing particles encounter the outer silicon pixel detectors. They are mounted just after the fibre detector with average radii of $7.6cm$ and $8.9cm$. The inner layer has 24 and the outer has 28 tiles of $1cm$ length. The active area itself has a length of $36cm$. Similarly to the vertex detectors, they are mounted on $25\mu m$ thin Kapton foil with plastic ends.
The stations beam up- and downwards only consist of the outer pixel detector layers as well as a timing detector. While the silicon detector are the same as in the central station, the timing tracker was chosen to be much thicker than the fibre detector in the central station. It consists of scintillating tiles with dimensions of $7.5 \times 7.5 \times 5mm^3$. They provide an even better time resolution than the fibre tracker in the center. Incoming particles are supposed to be stopped here. The outer stations are mainly used to determine the momenta of the outgoing particles and have an active length of $36cm$ and a radius of around $6cm$.

4.5 The problem of low longitudinal momentum recurlers

As explained in section 4.4, the outgoing particles are supposed to recurl back into the outer stations of the detector to enable a precise measurement of the momentum. A problem arises if the particles have almost no momentum in the beam direction. Then they can recurl back into the central station and cause additional hits there. As the the central station is designed to let particles easily pass through, they can recurl inside the central station many more times without getting stopped. As we have a $20ns$ time window for the readout of the pixel detectors, we need a very reliable way to identify and reconstruct these tracks as recurling particles as otherwise they look exactly like newly produced particles coming from our target. As one can imagine this influences the precision of our measurements by a big margin. So finding a way to identify these low beam direction momentum particles consistently is of great importance as it is crucial for the experiment to reduce the background as much as possible.

There is already an existing software to reconstruct particle tracks. However it struggles to find the right tracks for a lot of the particles recurling back into the center station.
These recurlers will typically leave eight hits or more, four (one on each silicon pixel detector layer) when initially leaving the detector and another four when initially falling back in. It is possible for these recurlers to produce even more hits when leaving the detector again but for this thesis we will be only focusing on these 8 hit tracks.
The current reconstruction algorithm works by fitting helix paths with a χ^2 method onto the 8 hits.
However experience has shown that often the fit with the lowest χ^2 isn't necessarily the right track. If we increase the χ^2 limit value to some arbitrary limit, we get a selection of several possible tracks per particle. Without any additional tools however, it is impossible to figure out if the right track is in the selection[14] and if yes which one of them correct track is.

[14] Based on detector efficiency it is possible for a particle to leave less than 8 tracks and therefore not be reconstructed by the algorithm

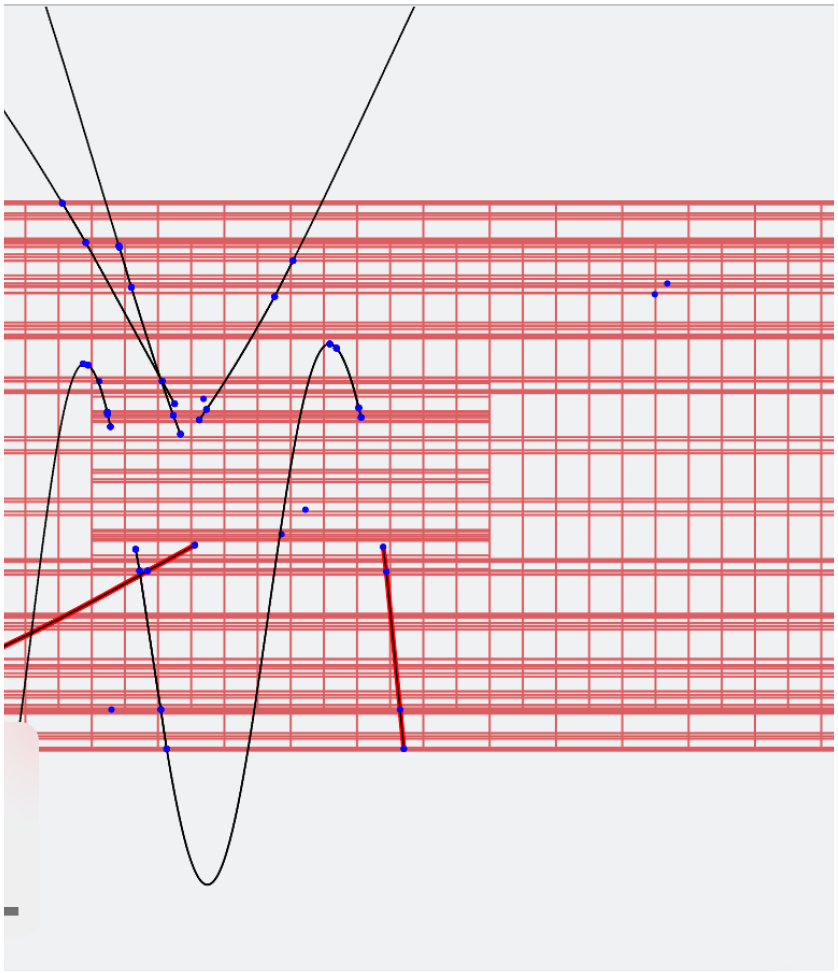


Figure 5: Particle recurling back into the center station

5 Machine learning

Machine learning has already proven itself to be very successful in resolving many problems in numerous other areas of science and also in the private sector. Based on these promising results, scientists are eager to study the potential of machine learning in physics.

References

[1] Mark Thomson. *Modern particle physics*. Cambridge University Press, 2013.

[2] S Abe, T Ebihara, S Enomoto, K Furuno, Y Gando, K Ichimura, H Ikeda, K Inoue, Y Kibe, Y Kishimoto, et al. Precision measurement of neutrino oscillation parameters with kamland. *Physical Review Letters*, 100(22):221803, 2008.

[3] P Adamson, C Andreopoulos, R Armstrong, DJ Auty, DS Ayres, C Backhouse, G Barr, M Bishai, A Blake, GJ Bock, et al. Measurement of the neutrino mass splitting and flavor mixing by minos. *Physical Review Letters*, 106(18):181801, 2011.

[4] A Blondel, A Bravar, M Pohl, S Bachmann, N Berger, M Kiehn, A Schöning, D Wiedner, B Windelband, P Eckert, et al. Research proposal for an experiment to search for the decay $\mu \to eee$. *arXiv preprint arXiv:1301.6113*, 2013.

[5] Heiko Augustin, Niklaus Berger, Sebastian Dittmeier, Carsten Grzesik, Jan Hammerich, Qinhua Huang, Lennart Huth, Moritz Kiehn, Alexandr Kozlinskiy, Frank Meier Aeschbacher, et al. The mupix system-on-chip for the mu3e experiment. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 845:194–198, 2017.

[6] Jonathan Philipp, Lennart Huth, Heiko Augustin, Raphael Philipp, Dirk Wiedner, Niklaus Berger, Mu3e Collaboration, et al. Das hv-maps basierte mupix teleskop. Technical report, Detector RD at DESY Test beam, 2015.

[7] Heiko Augustin, N Berger, S Bravar, Simon Corrodi, A Damyanova, F Förster, R Gredig, A Herkert, Q Huang, L Huth, et al. The mupix high voltage monolithic active pixel sensor for the mu3e experiment. *Journal of Instrumentation*, 10(03):C03044, 2015.