Track Propagation Methods for the Correlation of Charged Tracks with Clusters in the Calorimeter of the \textit{PANDA} Experiment

T. Nasawad, T. Simantathammakul, C. Herold, T. Stockmanns, J. Ritman, and C. Kobdaj, on behalf of the \textit{PANDA} collaboration

\begin{tabular}{l}
1 School of Physics, Suranaree University of Technology, 111 University Avenue, Nakhon Ratchasima, 30000 Thailand \\
2 Forschungszentrum Jülich GmbH, Wilhelm-Johnen-Straße, 52428, Jülich, Germany \\
E-mail: ThanachotC3@gmail.com, tawanchut1311@gmail.com
\end{tabular}

\textbf{Abstract}: To classify clusters of hits in the electromagnetic calorimeter (EMC) of \textit{PANDA} (antiProton ANnihilation at DArmstadt), one has to match these EMC clusters with tracks of charged particles reconstructed from hits in the tracking system. Therefore the tracks are propagated to the surface of the EMC and associated with EMC clusters which are nearby and below a cut parameter. In this work, we propose a helix propagator to extrapolate the track from the Straw Tube Tracker (STT) to the inner surface of the EMC instead of the GEANE propagator which is already embedded within the PandaRoot computational framework. The results for both propagation methods show a similar quality, with a 30\% gain in CPU time when using the helix propagator. We use Monte-Carlo truth information to compare the particle ID of the EMC clusters with the ID of the extrapolated points, thus deciding upon the correctness of the matches. By varying the cut parameter as a function of transverse momentum and particle type, our simulations show that the purity can be increases by 3-5\% compared to the default value which is a constant cut in the \textit{PANDA} simulation framework PandaRoot.

\textbf{KEYWORDS}: Detector modeling and simulations I and II, Particle identification method
1 Introduction

The PANDA experiment[1] is one of the core projects at the upcoming Facility for Antiproton and Ion Research (FAIR) [2, 3]. It has been designed as a fixed target detector to measure signals from the collision between antiprotons and protons or heavier nuclei, where exotic particles are created in high multiplicities through gluon-rich processes. The central part of the physics program of PANDA are the spectroscopy of charmonium states and the investigation of open charm production, meson spectroscopy, and the search for exotic matter like glueballs and hybrids including excitations in the charmonium range which have been predicted by quantum chromodynamics (QCD) calculations [4].

PANDA uses a cooled antiproton beam in the momentum range from 1.5 to 15.0 GeV/c provided by the High Energy Storage Ring (HESR) with an excellent momentum spread of \( \Delta p/p < 4 \cdot 10^{-4} - 5 \cdot 10^{-5} \) depending on the operation mode.

PANDA consists of several specific detectors which can be classified into two groups: the Target Spectrometer (TS) and the Forward Spectrometer (FS). In the TS area, a superconducting solenoid magnet generating a field of \( B = 2 \) T surrounds the interaction point together with a Micro Vertex Detector (MVD), a Straw Tube Tracker (STT), a Cherenkov detector (DIRC), a Time-of-Flight detector (TOF) an Electromagnetic Calorimeter (EMC) and a large angle tracking Gas Electron Multiplier (GEM). In the FS area, a dipole magnet generating a field of up to \( B = 2 \) Tm is used for small angle tracks together with six planes of straw tube trackers (FTS), Cherenkov detectors, a forward EMC and muon counters [6].

In this work we concentrate on the TS part to propagate trajectories of electrically charged particles from the STT to the inner surface of the EMC. Since the specific parts and subsystems are currently under construction, we have to rely on a special simulation software package called
PandaRoot [7] to simulate events, reconstruct them and develop efficient algorithms for charged particle identification. PandaRoot is the offline simulation and event reconstruction software for the PANDA detector based on the FairRoot framework [8].

Reconstruction in the TS is based on the information from all tracking devices in the TS. The track reconstruction can be separated into two steps. First, the track finder assigns hits to tracks. Various different algorithms are provided by the framework. For this work an idealized pattern recognition based on Monte Carlo information was used to exclude any side effects from the track finding algorithm. The track parameters from the ideal pattern recognition are smeared and used as an input for the second stage, the track fitting. For this stage a Kalman filter is used. The Kalman filter considers the interaction with the detector material, i.e. multiple scattering and energy loss and uses the detailed magnetical fields inside the detector. A detailed description of this algorithm which has been developed by the BaBar Collaboration can be found in [9].

When a particle hits a crystal in the EMC, electromagnetic showers are produced and often spread to several nearby modules. This contiguous area of modules is called an EMC cluster. The point \( \bar{y}_i \) representing the center of the \( i \)th EMC cluster is then identified by the EMC reconstruction software. From the detector design [12, 13], a space of 15 cm between the STT and the EMC makes an extrapolation of the reconstructed track from the STT to the inner surface of the EMC necessary. Currently, this is done by the GEANE algorithm [14] taking into account magnetic field, detector geometry and material interactions. The propagator projects the track \( t \) onto a single point \( \bar{x}_t \) on the EMC which is then compared to the positions of all \( n \) EMC clusters \( \{\bar{y}_i, i \in 1 \ldots n\} \). The squared distance from \( \bar{x}_t \) to the closest point \( \bar{y}_i \) is the EMC quality \( q_t \) of this track,

\[
q_t = \min \{(\bar{x}_t - \bar{y}_i)^2\}. \tag{1.1}
\]

Matching of a track with the such determined closest EMC cluster is then decided by comparing \( q_t \) to an EMC cut \( c \) such that the matching of this track with the cluster is accepted if \( q_t \leq c \). At present, PandaRoot uses a constant value of \( c \) for all particles and momentum ranges.

In this paper, we aim at two improvements of PandaRoot. First, we test a simpler and faster method of propagation to possibly replace GEANE, which takes into account solely the effect of the magnetic field. Second, we determine EMC cuts for different particle types and as a function of transverse momentum to improve the matching between particle tracks and EMC clusters. As the latter one influences the curvature of each track, we expect that it also has impact on the EMC quality. For instance, low-momentum tracks with high curvature will be subject to larger deviations and therefore require a larger EMC cut whereas high-momentum tracks which run almost in a straight line can be matched with a lower EMC cut, therefore gaining a purer signal.

The paper is organized as follows: In section 2, we describe and compare three different propagation methods between the STT and the EMC including linear propagation, helix propagation for particles in a homogeneous magnetic field, and GEANE. Section 3 is dedicated to the determination of an EMC cut which depends on particle type and transverse momentum \( p_T \). We calculate cuts on the basis of single-particle box generator events in specific \( p_T \) bins such that a certain percentage of matches which can be selected as 85\%, 90\% or 95\% is achieved. Results are subsequently tested for purity and completeness with a dual-particle model (DPM) background generator. We close with a summary and conclusion in section 4.
Figure 1: The figures show (a) the location of Straw Tube Tracker (STT) and Electromagnetic Calorimeter (EMC) near the interaction point, (b) the front view to exhibit the 15 cm gap between STT and EMC.

2 Method of Propagation

The EMC is one of the main components of PANDA for measuring the energy via the creation of electromagnetic showers in the material. After an incident particle hits the grid of PbWO4 crystals, it will be deflected and produce photons. These photons will then interact with the electromagnetically charged particles in the material leading to pair production of electrons, which will again result in new photons and so on until all of the energy of the incident particle has been deposited in the material.

From Fig. 1, we see that after a $p\bar{p}$ collision, the produced particles which have passed the tracking system have to cross a space of 15 cm before they are reaching the EMC. Along this distance, they are subject to the Lorentz force due to the homogeneous magnetic field of about $B = 2$ T in $z$-direction. This force causes charged particles to move on a helix path with a constant velocity in $z$-direction and a circular motion in the $x$-$y$ plane. To connect the reconstructed curved tracks of charged particles that are obtained from measurements in the tracking systems to EMC clusters is a nontrivial task which requires a way to propagate the track from the outer surface of the STT, where its position and momentum vector are given from a track finding and fitting algorithm, to the inner surface of the EMC. The such projected track creates a point upon intersecting the EMC surface which can be compared to the measured EMC clusters.

In this section, we present and compare three different methods of propagation, namely the presently used GEANE propagator, a linear propagator, and a helix propagator. We hereby do not only focus on the quality of these but also determine the average CPU time that each algorithm
GEANE is written in Fortran. It allows the user to extrapolate a trajectory of a charged particle in terms of mean values and errors in forward and backward direction in dense matter. It is integrated into the GEANT3 [15] system in both the simulation and the reconstruction part. Later, GEANE has been integrated into the FairRoot framework [16], adding a new feature which allows for low-density materials as, for example, in gaseous detectors. This algorithm, however, uses a high CPU time to process as it takes several complex effects into account, namely, energy loss, Coulomb multiple scattering, and the magnetic field.

The most simple method of propagation is a linear propagator, assuming a linear track, \( \vec{x}_{\text{lin}}(t) \), from the end point \( \vec{x}_e \) of the reconstructed track on the STT surface to the EMC with the constant velocity \( \vec{v} \), the particle then moving on a straight line parametrized by time \( t \),

\[
\vec{x}_{\text{lin}}(t) = \vec{x}_e + t \cdot \vec{v} .
\]

(2.1)

Hereby, the effect of the magnetic field is completely neglected and we may therefore expect that it will yield higher precision with increasing particle momentum, though probably be less useful for particles with low momentum.

A more sophisticated, yet still simple propagation method is the helix propagator which describes charged particle tracks by using the momentum and position vectors from the tracking system and also takes into account the magnetic field between STT and EMC. The particles in the magnetic field will move on a helix due to the Lorentz force with a circular motion in the transversal plane with a synchroton radius \( R \) of

\[
R[m] = \frac{p_T[GeV/c]}{0.3B[T]} ,
\]

(2.2)

where \( p_T \) denotes the transverse momentum and \( B \) magnetic field strength. We describe the helix by the equation

\[
\vec{x}_{\text{hel}}(t) = (R \cos(\omega t), R \sin(\omega t), v_z t) ,
\]

(2.3)

with a constant angular frequency \( \omega \).

### 2.1 EMC quality

We determine the accuracy of these propagators by generating events in the so-called box generator [17], where one particle of a particular type is created with a value of \( p_T \) within a predefined range. The emission angle is chosen such that the particle hits the barrel part of the EMC detector. Ideally, we will then have a single track and a single EMC cluster and the distance from the latter one to the propagated point is then the EMC quality \( q_t \) as given in Eq. 1.1. The smaller the value of \( q_t \), the higher the accuracy of the propagation method.

In some cases which we consider as outliers the EMC quality \( q_t \) can be extremely large due to wrong track fitting by the Kalman filter. Since these extreme values do not affect the median as strongly as the mean value, we use the median to describe our data instead of the mean. Fig. 2 shows the median of the EMC quality as function of transverse momentum for 10,000 events where a single pion is produced in each \( p_T \) bin. For all three propagators, we find that the EMC quality strongly increases with decreasing \( p_T \) as is to be expected.
From Lorentz’ law, the smallest transverse momentum for particles to reach the EMC surface is given by $(p_T)_{\text{min}} = R_E q B$ where $R_E = 31.6$ cm is half the radius of the EMC barrel. Thus, the lower bound of the transverse momentum is 0.19 GeV/c. So for the barrel EMC, the cases with $p_T$ lower than 0.19 GeV/c will give a high EMC quality $q_t$. Note that the 0.19 GeV/c is only valid for primary particles coming from the interaction point.

We also see that the linear propagator gives the worst EMC quality because it oversimplifies the motion of a particle as a linear track. For high momenta, it yields better results as the particles will move on practically straight lines in the short distance between the STT and the EMC. Helix propagator and GEANE propagator give similar results. However, square roots of the EMC qualities obtained by GEANE are better than those of the helix propagator for about 19% on average. The difference is especially significant at low $p_T$. For all relevant momenta, the difference between these does not seem to account for more than the size of a single crystal (with a surface area of $2 \times 2$ cm$^2$).

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure2.png}
\caption{Median of the EMC quality for different propagators as function of transverse momentum.}
\end{figure}

We created events where a single pion was produced.

\subsection{CPU time}

In a next step, we determine the CPU time that each of the propagation methods requires on average, see Fig. 3. As can be seen, both the linear propagator and the helix propagator require significantly less CPU time than the GEANE propagator for a single event, amounting to actually two orders of magnitude. For the particle identification (PID) macro which is heavily used in PandaRoot, this saves about 30% of computation time.
Figure 3: Average CPU time for different propagation methods, two orders of magnitude in CPU time can be gained by replacing GEANE with the helix propagator.

2.3 Verification matches

PandaRoot stores in each data-object one or more unique identifiers to the Monte-Carlo truth object they were generated from. By comparing the unique identifiers one can decide whether a track and an EMC cluster were created by the same particle or not which allows us to check the matching between tracks and clusters that we are trying to achieve.

From Eq. 1.1, the EMC quality is the distance between the propagated point on the EMC surface and the closest EMC cluster. We want to verify that these two were indeed created by the same particle. Fig. 4 shows results of the EMC quality over a $p_T$ range from 0 – 4 GeV. Red points represent events where the matched pair of track and EMC cluster originate from the same MCTrack and are therefore dubbed correct matches. The blue points are incorrect matches, here the EMC cluster and the corresponding propagated track originate from different MCTracks. We want to determine the EMC quality by considering only primary particles crossing the space between STT and EMC. The green points in Fig. 4 show the EMC quality of secondary particles. They show an especially high abundancy in the range of small $p_T$ below 0.5 GeV.

Fig. 5 shows the percentage of events from helix and GEANE propagators which are classified into correct and incorrect matches and also events with more than one track. Both helix and GEANE provide similar results, we can therefore conclude that the helix propagator is a suitable substitute for the GEANE propagator, yielding practically the same accuracy but giving us a significant gain in computational speed.

3 EMC cut as function of particle type and transverse momentum

3.1 Determining an EMC cut

A maximum for the EMC quality to allow matching between the propagated track and the EMC cluster is given by the EMC cut. In the current version of PandaRoot, this is simply a constant
Figure 4: EMC quality as function of transverse momentum for single-pion events where we distinguish between correct matches, incorrect matches and matches from secondary particles.

Figure 5: The percentage of the classified EMC quality for single-pion events as function of $p_T$ for helix and GEANE propagators.

number equal to 2,500 cm$^2$ which corresponds to a maximum distance of 50 cm between a track point and the EMC cluster. Results from the previous section, especially Fig. 2, however, suggest that this should rather be a function of $p_T$ as especially for $p_T < 1$ GeV, a strong dependence of the EMC quality on the transverse momentum can be observed. We also expect a different behavior of the EMC quality depending on the particle types, as mass, and subjectivity to specific interaction ways clearly influence their behavior in the material. To calculate the EMC cut for different particle types and as function of transverse momentum, we consider correct matches by plotting relative frequency histogram and cumulative relative frequency of the EMC quality.
Figure 6: (a) The relative frequency histogram and (b) the cumulative relative frequency of EMC quality for pion in $p_T$ range 0.4 - 0.5 GeV.

We propose that the EMC cut is the upper bound of the EMC quality such that 85%, 90%, 95% or 99% of all matches are accepted. The percentage will be later free to choose, so the user can decide whether he requires a higher purity, i.e. a low number of incorrect matches or a higher completeness, i.e. a high number of correct matches. The relative frequency histogram (Fig. 6 (a)) shows the percentage with which a given EMC quality bin occurs in a set of single-particle events. The cumulative relative frequency (Fig. 6 (b)) at a given $p_T$ bin is the sum of the relative frequency from 0 up to this respective $p_T$ value, it is therefore necessarily a monotonically increasing function, approaching 1. The data that we used to calculate the EMC cut is the EMC quality of the correct matches in specific intervals of transverse momenta. From these histograms, we are going to determine for each $p_T$ bin the EMC quality where the cumulative frequency reaches 85%, 90%, 95% or 99%, respectively. These values will then be used as the corresponding EMC cuts for these percentages. To determine these as precisely as possible, we fit the discrete histogram by a polynomial of sufficiently high order,

$$P(x) = a_0 + a_1 q_t + a_2 q_t^2 + \ldots + a_{n-1} q_t^{n-1},$$  \hspace{1cm} (3.1)

where $a_i$ are the fit parameters.

We consider four types of charged particle, namely kaons ($K$), muons ($\mu$), pions ($\pi$), and protons ($p$) with transverse momenta in the range of 0 – 4 GeV. We generate sets of events for each particle type and transverse momentum bin with a width of $\Delta p_T = 0.1$ GeV. For each of these, we perform the analysis detailed in the previous section to determine the EMC cut from the cumulative frequency histogram.

Fig. 7 shows the EMC cut for an acceptance of 85% as function of $p_T$. The little bump around 1.5 GeV suggest that the EMC cut function is to be fit by a combination of power law and Gaussian function,

$$c(p_T) = E_0 + (a p_T)^{-k} + \frac{E_A}{\sqrt{2\pi}\sigma}e^{-\frac{(p_T-\mu)^2}{2\sigma^2}}.$$  \hspace{1cm} (3.2)

Here, the first term, $E_0$ is the offset of the EMC cut. The second term is the power law function where $a$ is a scaling factor and $k$ a positive constant. The last term is the Gaussian function with
amplitude $E_A$. We performed this procedure for all particle types and all three percentages of acceptance, results are shown in Fig. 8. Note that we continued the functions by constants for both $p_T > 4$ GeV and $p_T < 0.19$ GeV. Parameters of the fit are shown in Table 1.

Table 1: The table of fit parameters for all particle types and acceptances.
3.2 Testing the EMC cut

After determining the EMC cut as function of $p_T$ for charged particles, we test our result using MC truth matched data, first for single-particle events. We define two quantities, purity and completeness, to check the validity of our results. The purity is defined as the ratio of the number of correct matches and the total number of all matches.

$$\text{Purity} = \frac{\text{Number of correct matches}}{\text{Total number of matches}}.$$  \hspace{1cm} (3.3)

Here, a match is defined as a pair of track and EMC cluster which has been decided to belong to the same particle by the EMC cut. The completeness is the percentage of correct matches from all possible matches. The latter one is the maximum number of tracks and EMC clusters that should ideally be matched.

$$\text{Completeness} = \frac{\text{Number of correct matches}}{\text{Possible correct matches}}.$$  \hspace{1cm} (3.4)

The completeness corresponds very well with the acceptance percentage as can be seen on the left hand side of Fig. 9. The right hand side of Fig. 9 shows the relation between completeness and
Figure 9: (a) The Completeness vs the acceptance percentage and (b) The purity vs the Completeness.

purity for different particle types and different EMC cuts including the previously used constant cut of $c = (50 \text{ cm})^2$. From this figure, we can draw several conclusions: First, for muons the purity is found to be 100% for all values of the completeness, including the constant EMC cut which gives a general completeness of almost 100%. This is due to the fact that muons as heavy leptons interact only electromagnetically (and through weak decay) and usually require specific detectors to be absorbed and detected as they pass most detector materials unaffected. Second, for particles that interact via the strong nuclear force, we see that a higher completeness amounts to a lower purity and vice versa. This is due to the obvious fact that by including more matches you also include more wrong ones. Among the strongly interacting particles, the highest purity is achieved for protons which are the heaviest of these particles. Finally, we can quantify the difference in performance of our $p_T$ and particle species dependent cut to the previously used constant one. Comparing it to a percentage cut of 99% which yields an up to 1% similar completeness, we are able to increase the purity by 2 – 3%. If we are willing to accept lower completeness, we are able to increase the purity by 5 – 6%.

When we increase the percentage of acceptance and therefore the completeness, the purity will decrease because more incorrect matches will be included. This figure clearly shows that a constant EMC cut provides lower purity than our EMC cut depending on transverse momentum and particle types except for muons.

Only for muons, the purity remains at 100% for all values of the completeness. We see that particles which interact also via the strong interaction have a slightly lower purity which is also dependent on the completeness. Among these, the highest purity is obtained for the protons which are the heaviest of these particles.

As the EMC is used for particle identification, we have no information whatsoever about which type of particle causes a specific hit in the calorimeter material. We therefore suggest to use the
The results of EMC cut using power-law function

<table>
<thead>
<tr>
<th>Matches</th>
<th>Purity</th>
<th>Completeness</th>
</tr>
</thead>
<tbody>
<tr>
<td>85.14</td>
<td>85.83</td>
<td>88.03</td>
</tr>
<tr>
<td>38.37</td>
<td>38.30</td>
<td>38.28</td>
</tr>
<tr>
<td>94.06</td>
<td>94.65</td>
<td>97.03</td>
</tr>
</tbody>
</table>

Figure 10: Matches, purity and completeness achieved with an EMC cut for kaons in the DPM background generator.

 generally largest EMC cut, namely that of kaons, for all tracks and clusters in general. We tested this assumption with the DPM background generator, where particles of several types and high multiplicities are created in each single event. Matches represents the percentage of matched to all propagated tracks.

\[
\text{Matches} \equiv \frac{\text{Number of matches}}{\text{Number of total events}}. \tag{3.5}
\]

Fig. 10 shows that the number of events with \( q_t < c \) increases when we increase the acceptance percentage. As one would expect, we found that this yields generally a lower purity because more secondary particles are collected. Furthermore, the purity decreases when the proportion of correct matches is higher. On the other hand, the completeness is relatively high because we used the biggest possible cut but it still increases with higher proportion of correct matches.

4 Conclusion

We tested different types of propagators to connect tracks obtained from hits in the tracking system to clusters of EMC hits. We found that a simple helix propagator, taking into account only the effect of the bending of the particle trajectory in the homogeneous magnetic field yields results of similar quality to the GEANE propagator which is currently used in PandaRoot. The helix propagator, however, requires significantly less computational time, we gain about 30% of time in the particle identification macro. So the helix propagator can be used as an alternative to extrapolate the track from the tracking system to clusters of EMC hits when one is not seriously concerned about the energy loss parameters during the extrapolation.

From our initial observation that the EMC quality depends on transverse momentum, we motivated the determination of a \( p_T \) dependent EMC cut to decide whether to reject or accept a match between the propagated point and the closest EMC cluster. We calculated the EMC cut for different particle types and percentages of acceptance as a function of transverse momentum and
then performed fits with a combination of power law and Gaussian function. In subsequent test, we found that the percentage of completeness is equal to the acceptance percentage, that means our function is suitable to be used as an EMC cut and yields better results than the constant value that has so far been used in PandaRoot.

Compared to the previously used constant EMC cut, we were able to show that our $p_T$ dependent cut is able to provide a higher purity even in the case of a similar value of completeness. By lowering the completeness to values of 85%, we achieve approximately 5 – 6% higher purities. It is therefore desirable to have an option in PandaRoot which lets the user select the acceptance of percentage and therefore to control the purity of the track matching.

**Acknowledgments**

C. H., T. N. and T. S. would like to thank Prof. James Ritman for giving us the opportunity to conduct research at Forchungzentrum Jülich and for covering our expenses during the researching stay. T. N. and T. S. are supported by Development and Promotion of Science and Technology Talents Project (DPST) scholarships. This work is funded by Suranaree University of Technology (SUT) and by the Office of the Higher Education Commission under NRU project of Thailand. C.H. and C.K. acknowledge support from SUT-CHE-NRU (Grant No. FtR.15/2559) project.

**References**


[8] https://fairroot.gsi.de


