

Track Finding for the PANDA Detector Based on Hough Transformations

Anna Alicke^{1,*}, Tobias Stockmanns^{1,**}, and James Ritman^{2,3,1,***}

¹Forschungszentrum Jülich, Institut für Kernphysik

²GSI Helmholtzzentrum für Schwerionenforschung

³Ruhr-Universität Bochum, Experimentalphysik4, Lehrstuhl I

Abstract. The PANDA experiment at FAIR (Facility for Antiproton and Ion Research) in Darmstadt is currently under construction. In order to reduce the amount of data collected during operation, it is essential to find as many true tracks as possible and to be able to distinguish them from false tracks. Part of the preparation for the experiment is the development of a fast online track finder. This work presents an online track finding algorithm based on Hough transformations, which is comparable in quality and performance to the currently best offline track finder in PANDA. In contrast to most track finders the algorithm can handle the challenge of extended hits delivered by PANDA's central Straw Tube Tracker and thus benefit from its precise spatial resolution. Furthermore, optimization methods are presented that improved the ghost ratio as well as the speed of the algorithm by 70 %. Due to further development potential in terms of track finding for secondary particles and speed optimization on GPUs, this algorithm promises to exceed the quality and speed of other track finders developed for PANDA.

1 Tracking with PANDA

The PANDA experiment at the Facility for Antiproton and Ion Research (FAIR) in Darmstadt is a proton-antiproton fixed target experiment currently under construction. The collisions at energies up to 15 GeV/c enable research in various physics fields. The main focus of research is meson spectroscopy, baryon production, nucleon structure and hypernuclei as well as exotic states like glueballs and quark molecules [1]. The experiment is divided into a target spectrometer surrounding the interaction point with a coverage of nearly 4π and a forward spectrometer to detect forward boosted tracks. The target spectrometer consist of a Micro-Vertex-Detector (MVD), a Straw Tube Tracker (STT) and Gas Electron Multiplier (GEM) planes. The MVD and GEM detectors provide precise point-like position information. In contrast, the STT provides the position of the anode wire and the drift time of the electrons to the anode. This results in a cylinder (isochrone) around the center of the straw tube. To benefit from the STT's good hit-resolution better than $150\mu\text{m}$ it is necessary to use the additional isochrone information, since tracks of the charged particles must pass tangentially

*e-mail: a.alicke@fz-juelich.de

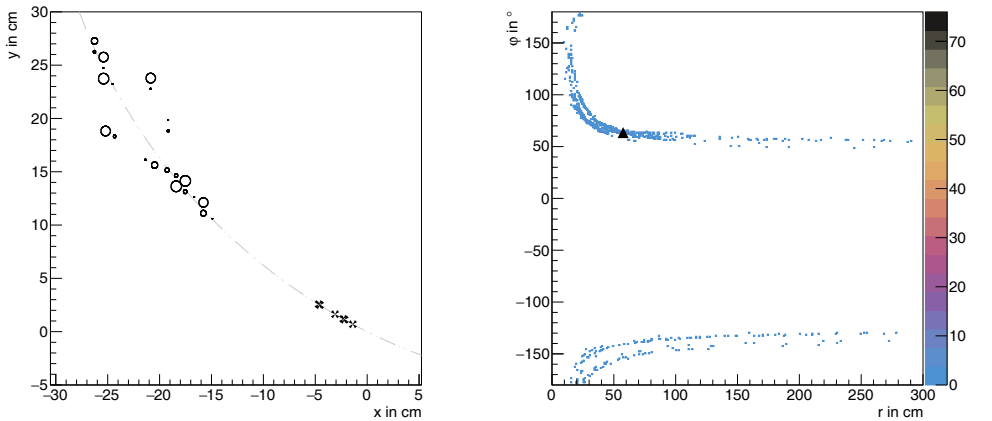
**e-mail: t.stockmanns@fz-juelich.de

***e-mail: j.ritman@fz-juelich.de

to the isochrones. Running at an average event rate of up to 20 MHz, not only a precise but also a fast online track reconstruction is essential for a meaningful physics analysis in PANDA [2].

2 Track Finding based on Hough transformations

A tracking algorithm is presented which uses the Apollonius problem as basis for a Hough transformation. By design the Apollonius problem enables the isochrone information to be used because it is the mathematical description for finding a circle that is tangent to three other circles [3]. Each combination of three hits is chosen to calculate the resulting Apollonius circles. All circles of all hit combinations are then filled into a three-dimensional histogram: the Hough space. Hits belonging to the same track are described by the same track parameters and therefore fill the same bin in the histogram. The maximum of the Hough space is then chosen as the track parameters [4]. As a simplification the track finder is currently designed for tracks originating from the interaction point (IP), so-called primary tracks, which reduces the necessary dimension of the Hough space to two. This means that two hits and the interaction point are chosen for the calculation. As parameters for the Hough space the mid point of the circle in polar coordinates is chosen.



(a) Visualization of one particle track. The black crosses represent the MVD hits and the black circles represent the hits of the STT. The resulting track is shown as a gray dashed line

(b) Visualization of the Hough transformation. The density of entries in the Hough space is displayed with a color scheme from blue to black. The black triangle indicates the maximum in the Hough space.

Figure 1: Visualisation of the Hough transformation for one particle track.

Figure 1 shows a particle track with the resulting Hough space. The black circle and cross symbols in Fig. 1a visualize the STT and MVD hit points, respectively. The circle found by the Hough transformation is shown as a gray dashed line. The entries of the Hough space are displayed in Fig. 1b in the color scheme from blue to black. The black triangle represents the maximum entry in the Hough space. The distribution of the points in the Hough space shows an hyperbolic behaviour in cartesian coordinates which leads to the bands of points

shown in Figure 1b. The reason for that is that the centers of all circles which are tangent to one circle and a point result in two hyperbolas in the middle between the point and the circle. The four intersection points between the two hyperbolas from the first isochrone and the two hyperbolas from the second isochrone are then filled into the Hough space. Due to the fact that we use the interaction point as one of the points for the Apollonius calculation and the other hits are relatively close together compared to the radius of the Apollonius circle, we get a band in a hyperbolic structure. Further information can be found in [5].

The key requirements for a tracking algorithm are to have the highest possible efficiency to find a track and to produce as few fake tracks as possible. To quantify these requirements, the efficiency and ghost ratio are introduced. The ghost ratio is the ratio of all wrong tracks found to all reconstructable tracks. The results shown in this paper are based on a simulated data sample with 100,000 events at a beam momentum of 7 GeV/c generating 250,014 reconstructable primary tracks. The data sample was simulated with a background generator for anti-proton – proton interactions based on the FRITIOF model (FTF) [9]. Tests of the algorithm showed that applying the Hough transformation to smaller sets of hits (tracklets) found by preselection algorithms and merging the tracklets afterwards significantly improved the speed as well as the quality of the algorithm. It has the advantage that the number of hit combinations can be drastically reduced. Therefore, the main steps of the total algorithm are:

1. Preselection of tracklets: Create smaller groups of hits belonging together.
2. Hough transformation for each tracklet to generate track parameters for the tracklets.
3. Merging of tracklets to particle tracks.

In a first step hits are grouped into smaller tracklets. The basis for a preselection of the STT hits is a Cellular Automaton [7], which uses the neighborhood relations of the dense packing of the STT. Neighboring hits, which unambiguously result from one track are collected in tracklets, as illustrated in Figure 2.

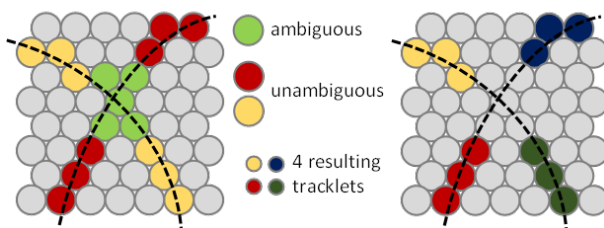


Figure 2: Principle of the Cellular Automaton. The red and yellow hits can be unambiguously identified as tracklets. The green hits are ambiguous and therefore not included. This results in the fact that a particle track can be divided into many smaller tracklets.

The MVD and GEM hits do not provide a dense packing like the STT and therefore the Cellular Automaton cannot be used. A second preselection method was introduced which uses the assumption that hits belonging to the same track have a continuous azimuth angle. By filling all azimuth angles in a histogram, different tracks can be preselected by dividing the data in subsectors depending on gaps in the azimuthal angle distribution (see Fig. 3). The preselected tracklets are then used for a Hough transformation for each tracklet.

Due to the different preselection methods particle tracks are divided into smaller tracklets. To decide if two tracklets belong to the same particle track different merging parameters were investigated. The first one is the relative overlapping area of two circles (dA). The second

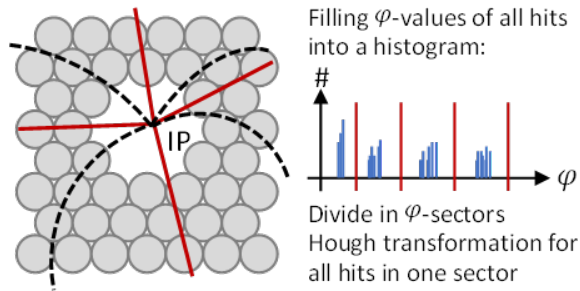


Figure 3: Principle of the Segmentation Algorithm. The azimuthal angles of the hits are filled into a histogram. Based on gaps in the histogram the data set is divided into subsectors.

and third one describe the difference in position (d) and direction ($d\varphi$) when extrapolated to a plane in the middle between both tracklets. Receiver operating characteristics (ROC) curves were created to find optimal cut values, with a high true positive (TP) rate (tracklets belonging to the same track are merged) and a small false positive (FP) rate (tracklets not belonging to the same track are merged). For the merging procedure the false positive rate must be minimized. Based on the ROC analysis a set of parameters ($dA > 0.48$, $d < 8$ cm, $d\varphi < 20^\circ$) were chosen with the highest accuracy ($\frac{TP+TN}{TP+TN+FP+FN} = 93.6\%$), a false positive rate of 1.9% and a true positive rate of 76.1%.

At this stage of development the algorithm had a finding efficiency of 88.0%, which is better than the other track finders developed for PANDA. The number of ghosts, clones and the runtime, however, were not competitive. The next section shows optimization methods, that improved the ghost ratio by 70%, the clone ratio by 47% and the runtime by 75% at a loss of finding rate of only 6.1%.

3 Optimization methods

To optimize the efficiency of the tracking algorithm ghost reduction is performed after merging. Ghost tracks occur when the algorithm combines hits to one track that don't belong to the same particle track. Tracks that are particularly difficult to reconstruct are those with low transverse momentum and therefore a small radius of curvature in the magnetic field. Due to interactions with the detector material, they lose energy and curl inside the detector. The result is a larger region of fired tubes in the x-y-plane of the detector. To identify and exclude these curling tracks, the neighborhood of the hits is an interesting parameter. The neighborhood was investigated for STT hits and GEM hits. Due to the dense packing of the STT, the neighborhood is clearly defined. For the GEM, however, a radius in the x-y-plane was chosen. All hits with a distance in x-y smaller than the radius are defined as neighbored GEM hits. Another ghost reduction parameter is the average distance between hits in the x-y-plane. This parameter is motivated by the fact that a real track, in contrast to a ghost track, must be continuous and therefore has a smaller distance between the hits. A final parameter of interest is the number of hits per detector type, e.g. it is unlikely that a primary particle creates only GEM hits because the particle also has to pass the MVD. For all investigated parameters ROC curves were created, which are shown in Figure 4. The chosen cut values are represented as small circles on the ROC curves in Figure 4 and are summarized in Table 1.

The cut values for the ghost reduction were chosen with an emphasis on the true positive rate causing also an increase of the false positive rate. The reason for this is, that a high

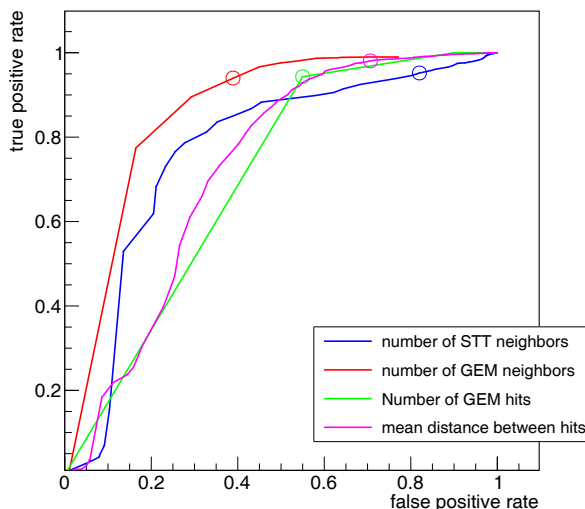


Figure 4: ROC curves for the parameters for ghost reduction. The chosen cut values are displayed as small circles on each ROC curve.

track finding rate is more important than a low ghost ratio. Combining all parameters, it was possible to reduce the number of ghosts by 22.7 % points with an efficiency loss of 5.4 % points.

In a next step the selection method for the maxima in the Hough space was optimized. This is the method to decide which maximum is chosen if there are many bins with the same number of entries. A method was chosen that determines the sum of distances of each hit to the track resulting from a certain maximum. The maximum with the smallest average distance is chosen.

The efficiency optimization mainly includes a reduction of the ghost ratio, which always causes a reduction of the finding rate. The optimization therefore consisted of reducing the ghost ratio as much as possible with as little loss of finding rate as possible. Thereby, the cut parameters for the ghost reduction are tunable depending on the need for higher efficiency or lower ghost ratio. After the optimization procedure the track finding algorithm reaches a ghost ratio of 9.9 %, with an only slightly reduced finding rate of 83.1 %. These results already exceed the quality parameters of PANDA’s best track finder: the BarrelTrackFinder [8].

In a next step the speed of the algorithm was optimized. The aim was to find a leaner and more efficient data structure, where only the relevant data are transferred between methods. In particular those data with a high access rate must be easily accessible. For the Hough space creation a high access rate of the given hits and in particular the isochrone information of the STT is necessary. The hits are stored in a data structure consisting of an array of double values. However, the usual data structure for hit points does not contain the isochrone information of the STT. A type cast is needed to get access to this additional information. This type cast is time consuming and was solved by including a map that directly links each hit to the corresponding isochrone.

Table 1: Chosen cut values for the ghost reduction parameters with achieved true positive and false positive rate.

parameter	cut value	true positive rate [%]	false positive rate [%]
mean number of STT neighbors	3.65	95.2	82.0
mean number of GEM neighbors	3.5	94.0	38.9
number of GEM hits	6	94.3	55.0
mean distance between hits	4.9	98.1	70.7

Additionally, the Hough space class itself was improved. Originally it was implemented as a histogram class, which has the disadvantage that a two-dimensional grid of bins is created. Depending on the number of hits in a tracklet this leads to a poorly filled histogram. For the Hough space it is not necessary to store the information of not or poorly filled bins. Only the maximum bins are important. It is sufficient to only map the data to an integer representing a region in the Hough space. This has the advantage that only the relevant data are stored and the access to the data is much faster. In total it was possible to speed up the code by a factor of 2.8.

In a last optimization step the merging method was improved. The calculation of the merging parameters was replaced by a simpler method that directly uses the distance of the maxima in the Hough space.

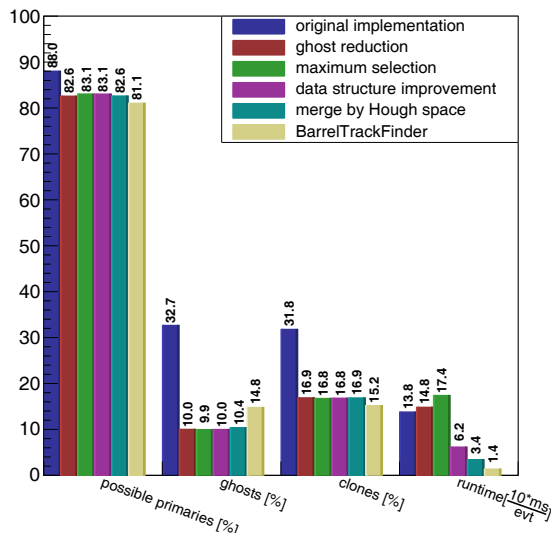


Figure 5: Contribution of each optimization step. In total the ghost ratio could be reduced by a factor of three, the clone ratio by a factor of two. The improvements lead to a small efficiency loss of 5.4 % points. The runtime of the algorithm could be improved by a factor of four. A comparison to the BarrelTrackFinder shows that the quality parameters and runtime are similar.

4 Summary and Outlook

The effect of the individual optimization steps on the quality parameters and the runtime are shown in Fig. 5. The ghost ratio could be reduced by a factor of three, the clone ratio by a factor of 1.9. The improvements lead to a small efficiency loss of 5.4 % points. The runtime of the algorithm could be improved by a factor of four. For all presented results an Intel Core i7-10750H CPU operating at 2.6 GHz clock frequency and 16 GB RAM was used. The maximum memory used by the algorithm stayed for all optimization steps at about 13 GB. A comparison to the BarrelTrackFinder, which is currently the best track finder in PANDA, showed that this implementation is comparable to the BarrelTrackFinder in terms of quality, speed and memory. The advantage of this algorithm is the possibility to extend the algorithm to find secondary particles, which is not possible for the BarrelTrackFinder. Additionally, the speed of the algorithm still has potential for development since it is easily parallelizable. For this purpose, a parallelization on GPUs is currently under development. For an online version of the algorithm an estimate of the required computing resources is necessary to achieve the needed runtime for processing the generated amount of data but also keep a high efficiency. Such an estimation will be given after the optimization of the algorithm on GPUs, since the GPU calculation is expected to have an essential impact to the speed of the track finder.

References

- [1] G. Barucca *et al.* for the \bar{P} ANDA Collaboration, *PANDA Phase One* [e-Print:2101.11877 [hep-ex]] (Jan 28, 2021).
- [2] W. Erni *et al.* for the \bar{P} ANDA Collaboration, *Technical design report for the \bar{P} ANDA (AntiProton Annihilations at Darmstadt) Straw Tube Tracker*, Eur. Phys. J. A **49**, 25 (2013). DOI:10.1140/epja/i2013-13025-8.
- [3] H. S. M. Coxeter *The Problem of Apollonius*, The American Mathematical Monthly, **75**:1, 5-15, (1968) DOI: 10.1080/00029890.1968.11970941.
- [4] J. Illingworth and J. Kittler, *The Adaptive Hough Transform*, IEEE Transactions on Pattern Analysis and Machine Intelligence **PAMI-9**: 690-698 (1987).
- [5] Herten, Andreas (Jul 2015), *GPU-based Online Track Reconstruction for \bar{P} ANDA and Application to the Analysis of $D \rightarrow K\pi\pi$* pp. 109.
- [6] A. Capella, U. Sukhatme, C.-I. Tan, and J. Tran Thanh Van, *Dual parton model*, Phys. Rep. **236**, 225-329 (1994).
- [7] S. Wolfram, *Theory and applications of cellular automata: including selected papers, 1983-1986*, Singapore: World scientific (1986).
- [8] R. Karabowicz, *Global Track Finder for PANDA experiment*, in GSI Scientific Report 2010 - GSI, Darmstadt, Germany, p. 117 (2011).
- [9] B. Andersson, G. Gustafson and Pi. Hong, *The FRITIOF model for very high energy hadronic collisions*, in Zeitschrift für Physik, C (Particles and Fields) **57**:3, Germany, (1993) doi:10.1007/BF01474343.