Proceedings of CTD 2020
PROC-CTD2020-27
August 31, 2020

Track Reconstruction on Free-Streaming Data for PANDA at FAIR

Jenny Regina [1]

On behalf of the PANDA Collaboration,
Department of Physics and Astronomy
Uppsala University, Sweden

ABSTRACT

High event rates of up to 20 MHz and continuous detector readout make the event filtering at PANDA a challenging task. In addition, no hardware-based event selection will be possible due to the similarities between signal and background. Therefore PANDA is among a new generation of experiments utilizing a fully software-based event filtering. Track reconstruction will play a key part in the online filtering at PANDA where it will be used together with the event building.

To ensure the quality of the track reconstruction, the existing quality assurance algorithm has been modified to be able to cope with free streaming data. This proceeding addresses a candidate for online track reconstruction algorithms for free streaming data based *e.g.* on a 4D Cellular Automaton procedure. The quality assurance procedure and the results from the tracking at different event rates and level of event mixing are also presented.

PRESENTED AT

Connecting the Dots Workshop (CTD 2020)
April 20-30, 2020

[1] Work supported by Knut and Allice Wallenberg Foundation

1 Introduction

At modern particle physics detectors, advanced trigger systems are divided into several levels. Commonly, the first level is a hardware trigger sorting the detector hits into *events*. At a higher level, the data rates are further reduced by a software-based event filter utilizing *e.g.* tracking or Particle IDentification (PID) information. At the upcoming PANDA (antiProton ANnihilations at DArmstadt) experiment [1] at FAIR (Facility for Antiproton and Ion Research), it will not be possible to use a hardware-based event filtering due to the similarities between signal and background. PANDA will utilize a fully software-based event filtering where *e.g.* online track reconstruction will be used for making trigger decisions. This makes PANDA part of a new generation of experiments employing this strategy, such as CBM [2]. Other experiments [3] actively invest in designing fully software-based trigger systems.

Track reconstruction on hits which are pre-sorted in events is referred to as *event-based* track reconstruction. On the other hand, track reconstruction performed on hits sorted in time without being associated to an event is referred to as *time-based* track reconstruction. The online track reconstruction at PANDA will therefore be time-based since there is no hardware trigger.

2 The PANDA Experiment

PANDA offers unique possibilities to explore QCD in the confinement domain where the relevant degrees of freedom for describing the strong interaction are not yet determined. The four physics pillars of PANDA include nucleon structure investigations, hadrons in nuclei, spectroscopy of charm and exotic states as well as strangeness physics.

Antiprotons, delivered by the HESR (High Energy Storage Ring) will impinge on a stationary proton target. The beam momentum will range between 1.5 GeV/c and 15 GeV/c. At the start setup of PANDA, an initial luminosity of $\mathcal{L} \sim 10^{31}\ cm^{-2}s^{-1}$ is foreseen with an average interaction rate of ~ 2.0 MHz. At later stages of PANDA, this is planned to increase to $\mathcal{L} \sim 2 \cdot 10^{32}\ cm^{-2}s^{-1}$ and an average interaction rate of 20 MHz at 15 GeV/c beam momentum. The earlier stage of PANDA is the focus of this paper. The interaction rate between the antiproton beam and the target will be Poisson distributed. The HESR will be filled to 80% with a 2 000 ns revolution time which leads to an active time of interactions of 1 600 ns and a dead time of 400 ns.

In order to fulfill the physics requirements, the PANDA detector, shown in Fig. 1, is designed to cover an almost full 4π solid angle. It is generally divided into two parts, the barrel shaped target spectrometer which covers the interaction point and the forward spectrometer placed downstream the beam pipe to measure forward going particles. The main distinction between the two parts is the magnetic field; a solenoid field will be present over the barrel spectrometer and a dipole field will be present over the forward spectrometer. Both parts will perform online and offline track reconstruction as well as PID and calorimetry. The central tracking detector at PANDA is the Straw Tube Tracker, which is described in detail Sec. 2.2. The additional detectors mainly dedicated to track reconstruction are: a Micro Vertex Detector based on silicon strip and pixel sensors, planes based on the Gas Electron Multiplier technique and planes consisting of drift tubes in the forward spectrometer.

The official PANDA software, PandaRoot [5] is an extension of FairRoot [6] and is based on the analysis and data processing framework ROOT [7].

2.1 Time Sorted Data

A comparison between the event-based simulation and the time-based simulation is shown in Fig. 2. The first stage, where the Monte Carlo data is generated, is the same in both cases and all data are sorted event-wise. In the event-based simulation, the data are also sorted event-wise in all subsequent stages all the way to the analysis stage. This means that at the reconstruction stage, tracking is done on data already pre-sorted event-wise. However, in the time-based simulation at the digitization stage, the hit data is sorted in time and placed into bursts, *e.g.* 2 000 ns long intervals corresponding to the antiproton beam revolution time. A consequence of this is that the track reconstruction is performed on hits sorted time-wise and events will be

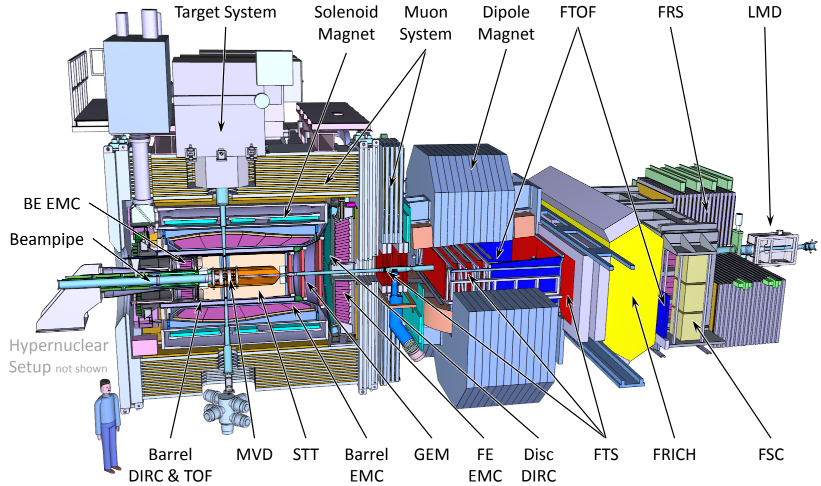


Figure 1: The PANDA detector setup with all sub-detectors highlighted.

mixed as depicted in Fig. 3. The data in the event-based simulation is synchronous from the simulation stage to the analysis stage. In the time-based case, the data is asynchronous with the simulation stage already at the digitization stage and from the reconstruction stage onwards, the data is synchronous.

For the time-based track reconstruction, there are two approaches to adjust the burst intervals in a way so that hits from one track doe not end up at the edges of an interval so the track risk being present in more than one burst of data. The first way is to place the cut between the burst in the HESR dead time where there are no interactions. The other way is to cut the bursts where a gap in the hit data is found. For the results in this proceeding, the first approach is used. No special treatment is given to hits close to the burst boundaries.

2.2 The Straw Tube Tracker

The main tracking detector of PANDA is the Straw Tube Tracker (STT) [4]. It consists of 4 224 closely packed single channel readout drift tubes arranged in hexagonal layers. The tubes are placed in 27 radial layers with 19 layers of tubes placed parallel to the beam pipe for xy-reconstruction and eight central layers with tubes tilted by $\pm 3°$ for z-reconstruction. It will cover a radial distance 15-41.8 cm. The STT comprises of single channel read-out drift tubes consisting of aluminum coated Mylar tubes with a diameter of 10 mm and a gold-plated tungsten-rhenium anode wire in the center. The gas inside the tubes will be Argon-based (90%) with carbon dioxide (10%) as a quench gas. When a particle traverses a tube, electrons in the gas are excited. Due to an applied electric field, the electrons start travelling towards the anode wire knocking out more electrons which in turn do the same, giving rise to an avalanche effect.

The time it takes for an electron to reach the wire is called the *drift time*. The maximum drift time is that of the electrons originating close to the tube wall and is about 250 ns. A circle centered around the anode wire and going through the point of closest approach of the particle trajectory to the wire is called the *isochrone*. The isochrones provide valuable information for the tracking since their radii give the distance between the passage of the high energy particle and the tube wire. This leads to the position resolution in the rϕ-plane of 150 μm. It should also be noted that the isochrone radius can not be measured directly but can be obtained from the pulse shape and a reference time. The resolution of the pulse shape is 1-2 ns on

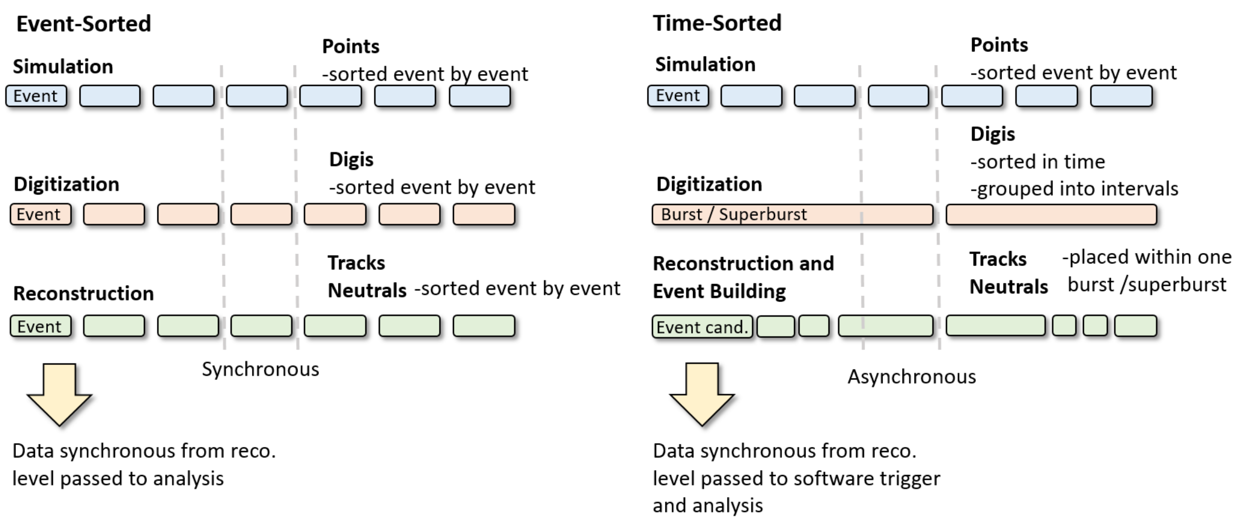


Figure 2: Comparison between the time-sorted data and the event-sorted data in the different simulation stages.

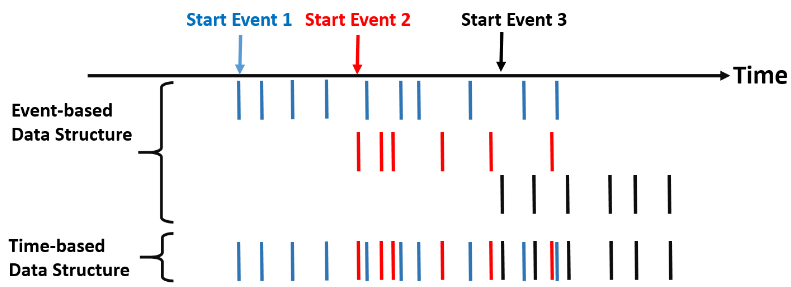


Figure 3: The time structure of the hits for for event-based data and for time-based data.

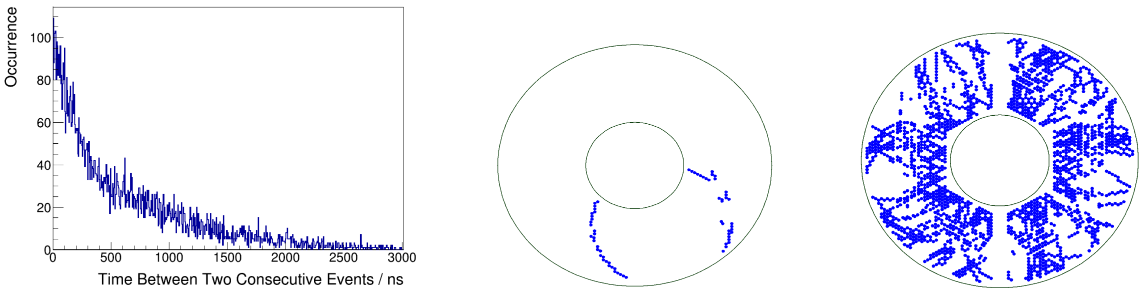
average however, the part of the pulse time currently available in the software is dependent on the isochrone radius.

2.2.1 Event Structure in the Straw Tube Tracker

The time between the start of two consecutive events at 2.0 MHz can be seen in panel a) of Fig. 4. The average time is 500 ns and a peak at smaller time differences is observed. The STT is among the slower tracking detectors as compared to, *e.g.* the Micro Vertex Detector, and there will be an overlap of events in the detector, *i.e.* hits from several events will be present in the detector simultaneously. Panel b) and c) of Fig. 4 illustrates the one event in the STT and 40 events respectively. The latter case is around the average number of events expected within one burst at the maximum event rate of 20.0 MHz so it represents the extreme case. It is clear that overlapping events complicates the track reconstruction.

3 4D Cellular Automaton

A Cellular Automaton [10] based clusterization procedure has been implemented in PandaRoot [8, 9]. It utilizes STT information and the procedure is illustrated in Fig. 5. Panel A), illustrates the trajectories of two particles traversing the STT. The tubes which are hit and give a signal are given an index as displayed in panel B). The state of a tube is defined by the number of active neighbors it has, *i.e.* the number of directly



(a) (b) (c)

Figure 4: Panel (a) shows shows the average time between the start of two consecutive events at 2.0 MHz interaction rate. One example event in the STT (b) and 40 events (c) which is around the average number of events at 20.0 MHz.

adjacent tubes which have also given rise to a signal. Tubes with one or two active neighbors are referred to as unambiguous and marked green in panel B). Tubes with three or more neighbors are referred to as ambiguous and marked red. The unambiguous hit tubes are treated first. The index of each tube is set to the lowest of the indices of its neighbors. This is done iteratively until all hits in one cluster of unambiguous hits have the same index and it is the lowest of all indices in the cluster as shown in panel C). The ambiguous hits are given all indices of their neighbors, panel D), and are assigned to an adjacent cluster of unambiguous hits after a track fit has been performed. The track fit is a Riemann paraboloid fit [11, 12]. Mainly the xy-information is used by the Cellular Automaton but methods for utilizing the z-component also exist, see ref. [13]. One advantage of this algorithm is that it does not assume the interaction point as a constraint so it can be used to reconstruct particles originating from displaced vertices.

The Cellular Automaton is well suited for hit level parallelization and a GPU version of the code exist. A speedup of roughly 100 has been achieved on a GeForce GTX 750 Ti GPU as compared to a i7 3.4 GHz processor.

In order to run the algorithm with event mixing, a 4D version of the algorithm has been developed where time-stamps are utilized in addition to the spatial neighbor relations. The time-stamps are included in a straight forward way by using a cut on the time stamps. If two hits which have been spatially clustered together occur closer to each other in time than this cut they are accepted as neighbors but if they occur further apart in time, they are rejected. The cut value can be varied but the default value is 250 ns which corresponds to the maximum drift time of the electrons in the tubes.

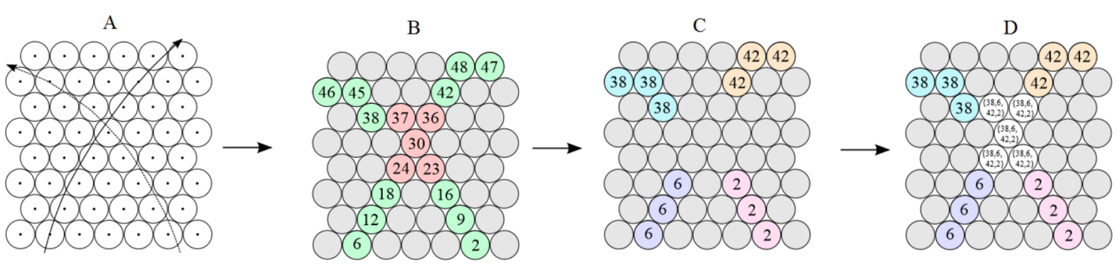


Figure 5: The Cellular Automaton procedure. A description is given in the text.

4 Tracking Quality Assurance

The standard quality assurance for tracking in PandaRoot contains all features needed to evaluate the efficiency and assess the quality of the tracking. It gives high level information such as momentum resolutions and low level information such as the number of true, false and missing hits in each tracking detector. Tracks are divided into six categories according to their hit content, these categories are:

1. **Fully Purely Found**: all hits belonging to the true MC track were found. No impurities from false MC tracks are allowed.

2. **Fully Impurely Found**: all hits belonging to one true MC track were found and impurities are allowed as long as the hits from the true track comprises at least 70% of all hits in the reconstructed track.

3. **Partially Purely Found**: the majority ($> 50\%$) but not all hits in a reconstructed track belong to a true MC track and no impurities of hits from other MC tracks is allowed.

4. **Partially Impurely Found**: impurities are allowed and at least 70% of all hits in the reconstructed track belongs to one true track. Not all hits from the true MC track were found.

5. **Ghost Track**: the reconstructed track does not match a true track. A ghost track appears when the requirement for a single MC track to be reconstructed is not fulfilled or the impurities from false MC tracks exceed 70% of the number of all hits in the reconstructed track.

6. **Clone Track**: one true MC track has been reconstructed more than once. While the *original* reconstructed track is the one with the most hits matching the true track, remaining tracks belonging to the same true track are defined as clones. There is no requirement of shared hits between an original track and its clone.

One track can only belong to one of the categories. The number of tracks in the first four categories make up the number of correctly reconstructed tracks whereas the number of tracks in the last two categories make up the number of incorrectly reconstructed tracks.

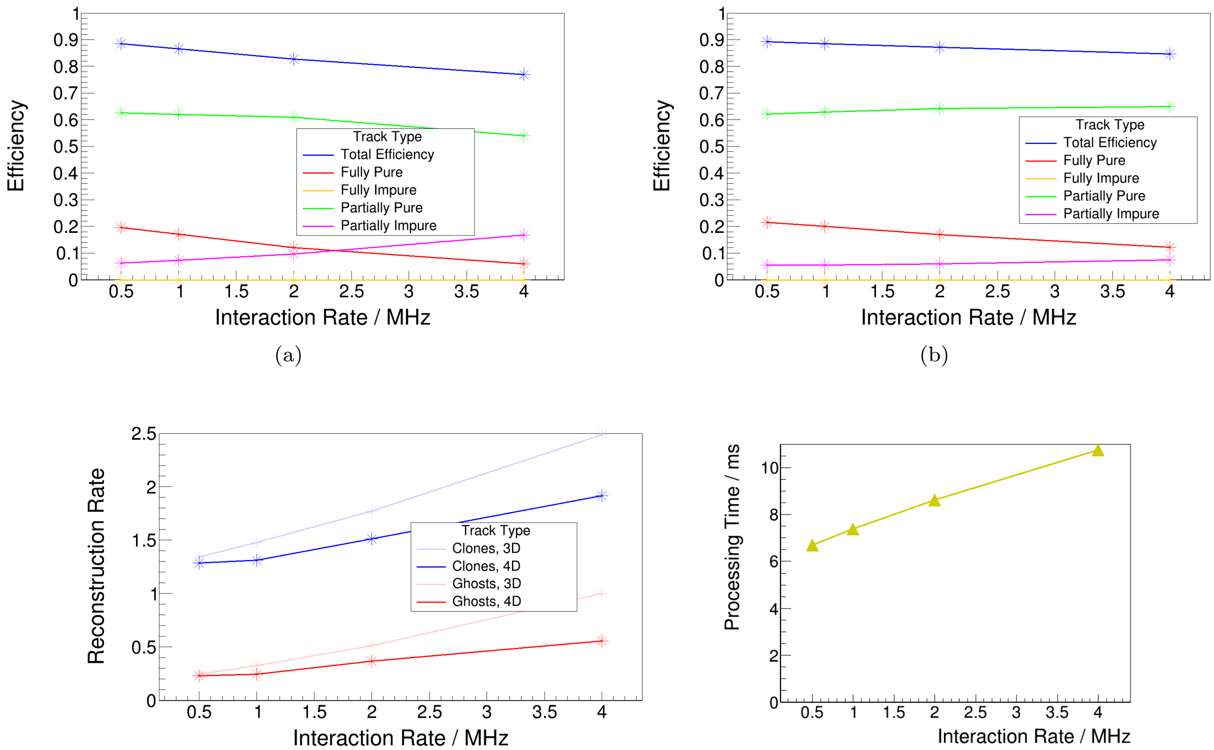
In order to compare the reconstructed track to the relevant track set, an ideal track finder exists in PandaRoot. This track finder gives the tracks which can be reconstructed in a certain sub-detector It is based completely on the ground truth information and the user can decide what hit requirement in the different sub-detectors to use. The resulting tracks are referred to as *ideal tracks* and the realistically reconstructed tracks are compared to the ideally reconstructed tracks for determining the efficiency.

5 Results

Generic background events for $\overline{p}p$ reactions at a beam momentum of 6.2 GeV/c were used for all results in this section. No prior filtering on particle momentum or displaced vertices is performed. The efficiency as a function of the interaction rate is shown in Fig. 6 for both the 3D algorithm (panel a)) and the 4D algorithm (panel b)). The efficiency is calculated as the total number of found tracks over the number of ideal tracks with at least six STT hits. The full efficiency is high and stable for both algorithms over the entire range. For the 3D algorithm the efficiency varies between 89% at 0.5 MHz and 77% at 4.0 MHz. In contrast, the efficiency ranges between 89% at 5.0 MHz and 84% at 4.0 MHz for the 4D algorithm. The efficiencies for all track categories are quite similar at 0.5 MHz for both algorithms. The reason for this is that the level of event mixing is so low that it almost corresponds to the event-based track reconstruction. As the level of event mixing increases, the effect of the timestamp utilization becomes more pronounced.

The clone and ghost rates as functions of the interaction rate are shown in Fig. 6 panel c) for both algorithms. The clone and ghost rates are calculated as the total number of clones or ghosts over the number of ideal tracks with at least six STT hits. The rates are both quite high, especially the clone rate. Both increases with increasing interaction rate due to the increasing level of event mixing. At 0.5 MHz interaction rate, the 3D and 4D algorithm yield similar results. The higher the interaction rate, the more the results diverge from each other and the timestamp utilization greatly helps suppressing both the ghost and clone rate. At 4.0 MHz, the ghost rate is reduced by almost 50% when using timestamps. Since both the clone and ghost rates are high, developments of clone merging and ghost suppression techniques are currently ongoing.

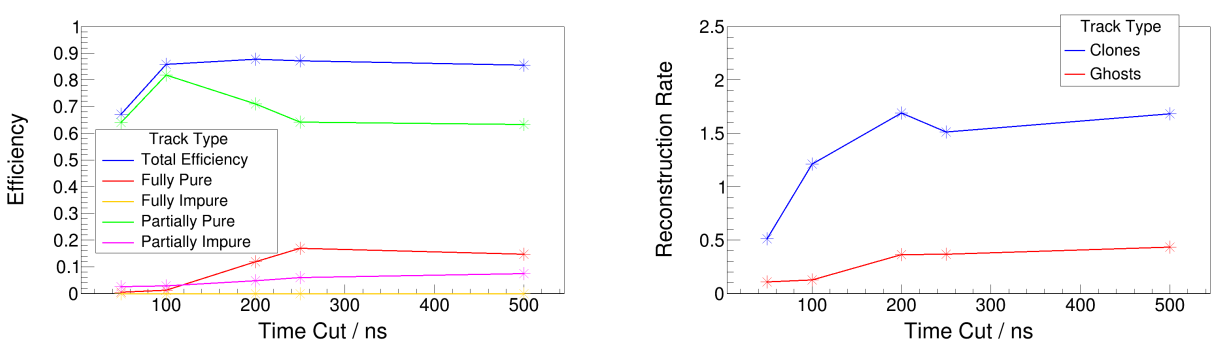
The performance of the part of the tracking obtaining the neighborhood relations was tested and the results can be seen in panel d) of Fig. 6. Due to the increasing number of hits being processed in each data burst with increasing interaction rates, the processing time also increases.



(a) (b)

(c) (d)

Figure 6: Efficiency vs. the interaction rate for the 3D algorithm (a) and the 4D algorithm (b) as well as the ghost and clone rate vs. the interaction rate for both algorithms (c). The different track categories are included. The processing time per data burst is shown in panel d).



(a) (b)

Figure 7: Efficiency vs. the time-cut (a) and the fake rate vs. the time-cut (b). The different track categories are included.

The effect of varying the time-cut was also investigated, the results can be seen in Fig. 7 for the efficiency (panel a)) and the clone and ghost rate (panel b)). The time-cut is varied between 50 ns and 500 ns. When tightening the time-cut, the total efficiency decreases. The efficiency is stable in the region between 100 and 500 ns time cut and the sub-categories of correctly found tracks are stable within the range 250 to 500 ns time cut. The number of fully purely found tracks rapidly decreases for time cuts lower than 250 ns whereas the number of partially pure tracks increases. This indicates that true hits are rejected even at time cuts slightly lower than 250 ns. The most dramatic effect of tightening the time-cut can be seen in the clone category (blue line, panel b)). A clone rate around 1.5 is observed at larger values of the time-cut and this is reduced to almost 0.5 at a 50 ns time-cut.

6 Conclusions

A 4D Cellular Automaton based hit clusterization procedure is available for PANDA. This track finder can be used for tracking particles originating from displaced vertices. The hit clusterization can assume time-based hit data and first results are promising. The efficiency is high, 84-89%, over the range of interaction rates of 0.5-4.0 MHz. The time-stamp utilization helps to suppress the clone and ghost rate as well as keeping the efficiency stable over the relevant range of interaction rates. During the online data taking at PANDA, this might be run together with a primary track finder.

References

[1] PANDA Collaboration, *Strong Interaction Studies with Antiprotons*, Physics Performance Report for PANDA, (2009) https://arxiv.org/abs/0903.3905v1

[2] I. Kisel for the CBM Collaboration, *Event Topology Reconstruction in the CBM Experiment*, IOP Conf. Series: Journal of Physics: Conf. Series **1070** (2018) 012015 Conference Contribution, 34th Winter Workshop on Nuclear Dynamics

[3] The LHCb Collaboration, *Design and performance of the LHCb trigger and full real-time reconstruction in Run 2 of the LHC*, JINST **14** (2019) P04013

[4] PANDA Collaboration, *Strong Interaction Studies with Antiprotons*, Technical Design Report for the: PANDA Straw Tube Tracker, (2012) https://arxiv.org/abs/1205.5441v2

[5] S. Spataro and the PANDA collaboration, *The PandaRoot framework for simulation, reconstruction and analysis*, J.Phys.: Conf. Ser. **331** (2011) 032031.

[6] M. Al-Turany et al. *The FairRoot Framework*, J.Phys.: Conf. Ser. **396.2** (2012) 022048.

[7] A. Naumann, *ROOT — A C++ framework for petabyte data storage, statistical analysis and visualization*, Computer Physics Communications **180** (2009) 2499–2512. doi:https://doi.org/10.1016/j.cpc.2009.08.005

[8] J. Schumann, *Entwicklung eines schnellen Algorithmus zur Suche von Teilchenspuren im "Straw Tube Tracker" des PANDA-Detectors*, Bachelor Thesis, Fachhochschule Aachen, Campus Jülich, Jülich (2013)

[9] J. Schumann, *Beschleunigung eins Spurfindealgorithmus fur den Straw Tube Tracker des PANDA-Detektors durch Parallelisierung mit CUDA C*, Master Thesis, Fachhochschule Aachen, Campus Jülich, Jülich (2015)

[10] S. Wolfram, *Statistical mechanics of cellular automata* Rev. Mod. Phys. **55** (1983), 601–644. doi: 10.1103/RevModPhys.55.601. url: https://link.aps.org/doi/10.1103/RevModPhys.55.601.

[11] A. Strandlie, J. Wroldsen, R. Frühwirth and B. Lillekjendlie, *Particle tracks fitted on the Riemann sphere*, Computer Physics Communications **131** (2000) 95-108.

[12] A. Strandlie, J. Wroldsen and R. Frühwirth, *Treatment ofmultiple scattering with the generalized Riemann sphere track fit* Nuclear Instruments and Methods in Physics Research A **488** (2002) 332–341.

[13] W. Ikegami Andersson, *Exploring the Merits and Challenges of Hyperon Physics: with PANDA at FAIR*, Disertation, Uppsala Universitet, Uppsala (2020)