Journal of Physics: Conference Series **396** (2012) 022042 doi:10.1088/1742-6596/396/2/022042

Pattern recognition in a high-rate GEM-TPC

J Rauch[1] for the GEM-TPC Collaboration

[1] Physik-Department E18, Technische Universität München, James-Franck-Strasse 1, 85748 Garching, DE

E-mail: j.rauch@tum.de

Abstract. A pattern recognition software for a continuously operating high-rate Time Projection Chamber with Gas Electron Multiplier amplification (GEM-TPC) has been designed and tested. Space points are delivered by a track-independent clustering algorithm. A true 3-dimensional track follower combines them to helical tracks, without constraints on the vertex position. Fast helix fits, based on a conformal mapping on the Riemann sphere, are the basis for deciding whether points belong to one track. To assess the performance of the algorithm in a high-rate environment, $p\bar{p}$ interactions at a rate of $2 \times 10^7\,\mathrm{s}^{-1}$, the maximum rate foreseen for $\overline{\mathrm{P}}$ANDA, have been simulated. The pattern recognition is capable of finding different kinds of track topologies with high efficiency and provides excellent seed values for track fitting or online event selection. The feasibility of event deconvolution has been demonstrated: Different techniques to retain the tracks from an event with known time from other tracks in the TPC are presented in this paper.

1. A GEM-TPC in a high-rate environment

A Time Projection Chamber (TPC) [1] is a well-known and established device for three-dimensional track reconstruction, combining a drift chamber and a Multi Wire Proportional Chamber (MWPC). Charged particles traversing the detector gas create electron ion pairs, which are separated by an electric drift field. The ionization electrons drift towards the MWPC, where gas amplification occurs and signals are induced on a segmented pad plane. In combination with a drift time measurement, a three-dimensional image of the track can be obtained. Since a TPC is a gaseous detector, it is light-weight and has a very low material budget. It can measure a charged particle's momentum by determining the curvature and dip angle of its helical path in a magnetic field. Moreover, the specific energy loss dE/dx can be measured and used for particle identification.

The ions created during amplification in the MWPC have to be prevented from drifting back into the TPC, since they would heavily distort the drift field and alter the paths of the ionization electrons. With a gating grid sitting above the MWPC wires, ion backflow factors ϵ (number of ions drifting back into the TPC per incoming ionization electron) of around 1.4 can be reached [2]. However, this grid requires a triggered operation, which poses an inherent limit of $\mathcal{O}(1\,\mathrm{kHz})$ to the rates at which the detector can be read out.

On the other hand, interaction rates in high energy physics experiments are ever growing: Within the luminosity upgrade of LHC, ALICE [2] plans an upgrade with Pb Pb interactions at a rate of 50 kHz, and $\overline{\mathrm{P}}$ANDA [3] aims for $\bar{p}p$ interactions at astounding 20 MHz. In order to take full advantage of a TPC in these *high-rate environments*, the rate restrictions due to the triggered gating grid have to be overcome.

Published under licence by IOP Publishing Ltd

One solution is to use a stack of Gas Electron Multiplier (GEM) foils [4] as an amplification stage. GEMs provide an intrinsic ion backflow suppression and therefore open the possibility to operate a TPC without a gating grid and at high rates. Ion backflow factors of $\epsilon = 4$ can be reached with GEMs. Depending on event rates and track densities, drift distortions in the range of a few cm can occur (cf. Sec. 3).

Regardless of that, high interaction rates bring up another issue: If they become higher than the inverse drift times of the ionization electrons, which are $\mathcal{O}(100\,\mu s)$, events start to overlap and tracks from many events will be present in the TPC at the same time. This is called *event mixing*. Since there are no more "gaps" between single events now, the next logical step is to abandon the trigger entirely and read out the GEM-TPC continuously. It now acts as a *continuous tracking pipeline*.

This new paradigm has considerable consequences and requires a new approach of track and event reconstruction. There is no more clear assignment of tracks to events. In other words, the track time t_0, the time when the particle passed the detector, is unknown a priori. The z-coordinate of a position in the TPC can be calculated from the drift time t_{drift} and the drift velocity v_{drift} of the ionization electrons:

$$z = t_{\text{drift}} \cdot v_{\text{drift}}. \tag{1}$$

The drift time is

$$t_{\text{drift}} = t_{\text{meas}} - t_0, \tag{2}$$

where t_{meas} is the time when the ionization electrons arrive at the readout. The lack of knowledge of t_0 translates into a z-offset of the tracks, proportional to t_0. This is illustrated in figures 13 and 16.

For this reason, the Pattern Recognition (PR) must, at the first stage, work *standalone* in the TPC, and it has to be robust against drift distortions, since they are z-dependent (cf. Fig. 7) and the correction of them requires the absolute z-positions (or t_0) of the tracks. But first and foremost, the PR has to cope with very high track densities and various track topologies. A three-dimensional track follower (cf. Sec. 2.2) has been developed its performance with regard to this requirements has been evaluated (cf. Sec. 3).

The next essential step in the reconstruction chain is to obtain a track time t_0 for each track. This can be achieved using several techniques described in Sec. 4. Finally, drift distortion corrections can be applied and TPC tracks can be fitted and merged with hits or tracks from other detectors. All of this is presented in Sec. 5, together with a first study on event selection.

2. Pattern recognition algorithms

The task of the PR in the field of particle tracking is to decide which detector hits originate from one physical track and group them together. In the GEM-TPC, a clustering algorithm delivers space points. As charged particles traversing a magnetic field describe helical tracks (assuming a homogeneous magnetic field and small material effects), the *pattern* to be recognized are helices. In addition, the PR should sort the hits along the track properly and provide good estimates of the track parameters.

2.1. Cluster finding

In the first step, before PR, the signals from the individual readout-pads are combined into points in space, so-called *clusters*. This is done in a *three-dimensional* and *track independent* algorithm (cf. Fig. 1). Since there is no explicit or implicit assumption on the track direction, as it is the case when e.g. clustering only signals from one pad-row, it works equally well for all kinds of track topologies like low- and high-p_t tracks, tracks curling in the magnetic field, secondary tracks etc.

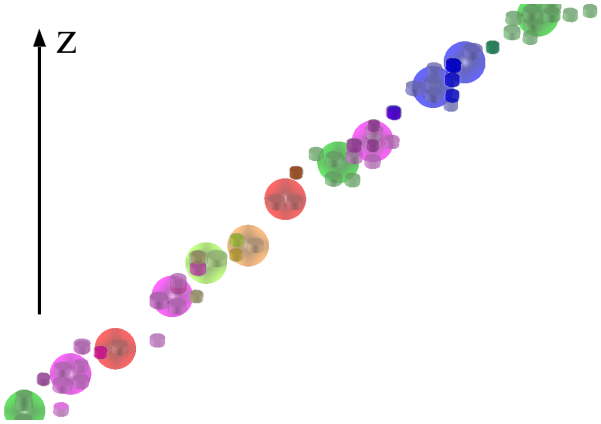


Figure 1. TPC cluster finding algorithm at work. *Signals* (shown as cylinders) are combined into larger *clusters* (shown as spheres). The cylinder radii are proportional to the signal amplitudes in a logarithmic scaling. The drift direction is from top to bottom, against z-direction. Signals split between adjacent clusters appear in darker colors.

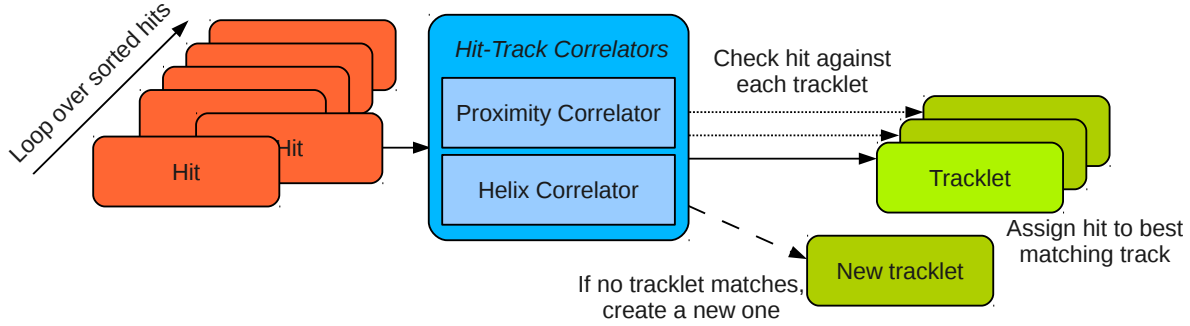


Figure 2. Scheme of the track building process. See text for explanations.

In essence, the algorithm groups signals around local maxima. The signals are sorted by decreasing amplitude and are then checked against all up to then existing clusters. If a signal is close enough to a cluster, it will be assigned to it. Signals can also be split in equal parts between adjacent clusters. If no matching cluster is found, a new one is started with the signal. The cluster position is then defined by the Center of Gravity (CoG) of its signals.

2.2. The Riemann track follower

The *Riemann Pattern Recognition* is a track following algorithm, that associates three-dimensional space points to tracks by using different correlations between hits and tracks: Proximity of a hit to hits already assigned to the track, and proximity of a hit to the helix made up by the hits in the track.

During PR of a complex event as shown in Fig. 13, helix parameters have to be estimated $\mathcal{O}(10^5)$ times. Therefore it is important to have a fast method to fit a helix to a set of hits. This is obtained via an extended Riemann fit [5].

2.2.1. Track building Figure 2 illustrates the *track building* process: To begin with, the clusters are presorted by z, radius R or angle ϕ. The idea is that the PR goes from areas of low track density, where tracks can be separated by their distance in space, to areas of high track density. The very first *tracklet* (i.e. a possibly incomplete set of clusters) is built and contains only one

hit at this time, then the algorithm loops through the presorted clusters. Each hit is checked against each existing tracklet. If one or several matching criteria (*"hit-track correlators"*) are fulfilled, the hit may then be assigned to the best matching tracklet.

If a hit-track correlator is applicable, it delivers a *matching quality*. Two correlators are applied: The *Proximity Correlator* checks proximity in space, by finding the nearest cluster in the tracklet. It is always applicable, and the matching quality is the distance of the two cluster positions. The *Helix Correlator* calculates the distance of the cluster to the helix that has been fitted to the tracklet. Newton's method is employed to find the minimum distance. If the tracklet has not been fitted, the correlator is not applicable. The matching quality is the distance to the helix.

If the matching quality is smaller than a definable cut (*proximity-* and *helix-cut*), the tracklet *survives* the correlator.

Dynamic scaling of the cuts These two cuts are dynamically scaled, depending on the number of hits n of the so far existing tracklets. With better track quality q, the helix-cut c_{helix} is narrowed, which gives better track separation power in areas of high cluster density, whereas the proximity-cut c_{prox} is opened, which makes the process less prone to track splitting. The effective cuts are defined by the following functions:

$$c_{\text{prox,eff}} = (1 + 3q)\, c_{\text{prox}}, \tag{3}$$

$$c_{\text{helix,eff}} = (2 - q)\, c_{\text{helix}}, \tag{4}$$

with

$$q = \begin{cases} \frac{n}{40} & n < 40 \\ 1 & n \geq 40. \end{cases} \tag{5}$$

These scaling functions are not dictated by any physical laws, however, they turned out to be effective and contribute to the good performance of the PR shown in Sec. 3.

Hit assignment The number of applicable *and* survived correlators defines the *correlation level*. The hit is added to the tracklet that reaches the highest level. If there is more than one tracklet, the best matching quality decides to which tracklet the hit is assigned. If no correlator is survived, a new tracklet is started with the hit.

To avoid following the wrong track in an area where two or more tracks are crossing, clusters which match well to more than one tracklet are excluded.

After adding a hit to a tracklet, a fast helix fit is performed if the tracklet contains more than a minimum number of hits, and the helix parameters are updated correspondingly.

2.2.2. Fast helix fits on the Riemann sphere In the TPC, the magnetic- and drift-field are parallel and very homogeneous. The readout plane is perpendicular to the latter two. Additionally, the material budget in the TPC is very low. Therefore, charged particles traversing the detector describe helical tracks in good approximation. Projected on the pad plane, the helices reduce to circles or circular arcs.

Therefore, a helix fit can be performed in two steps: A fast circle fit is obtained via a plane fit of the hits projected onto the Riemann sphere, and a straight line fit of the hit angles on the circle vs. their z-positions yields the dip angle of the helix.

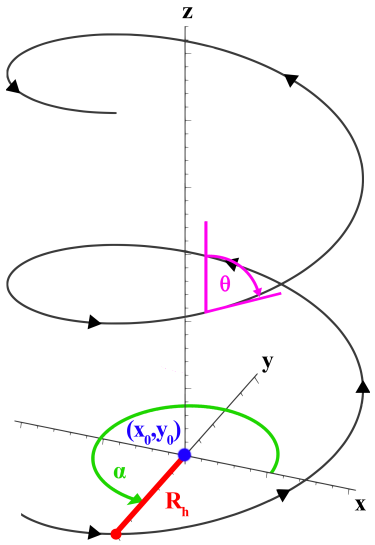


Figure 3. A helical track in the TPC can be described by the center of the helix on the pad-plane (x_0, y_0), the radius of the helix R_h, and slope m and offset t (cf. Eq. 6). Each point on the helix is defined by its angle α. The dip angle θ is the angle between a line parallel to the z-axis and a tangent on the helix.

Helix parametrization A helix can be described with 5 parameters (cf. Fig. 3): The center of the helix on the pad-plane $\mathbf{c}_h = (x_0, y_0)$, the radius of the helix R_h, and finally slope m and offset t.

For each point on the helix, an angle α can be defined. It is the angle against the x-axis and can become larger than 2π for curling tracks. There is a linear relation between angle α and z-position of a point:

$$z = m\alpha + t. \tag{6}$$

The dip angle θ can be calculated as follows:

$$\theta = \arctan\left(-\frac{m}{R_h}\right) + \frac{\pi}{2}. \tag{7}$$

The Riemann transformation The Riemann Transformation is a *stereographic projection*. It maps a plane (usually the complex plane $\mathbb{R}^2$) onto a sphere. This so-called Riemann sphere sits on top of the origin of the complex plane and has a diameter of 1. A point on the plane is connected to the north pole with a straight line, and the intersection of this line with the surface of the sphere defines the projection of the point onto the sphere. The projection is smooth, bijective and conformal (conserving angles).

The projection of a point on the plane $\mathbf{x_p} = (x, y) = (R\cos\phi,\, R\sin\phi)$ to a point on the sphere $\mathbf{x}_\mathrm{s} = (x_\mathrm{s}, y_\mathrm{s}, z_\mathrm{s})$ obeys the following equations. R and ϕ are the polar coordinates of the point.

$$\begin{aligned} x_\mathrm{s} &= R \cdot \cos\phi \,/\, (1 + R^2), \\ y_\mathrm{s} &= R \cdot \sin\phi \,/\, (1 + R^2), \\ z_\mathrm{s} &= R^2 \,/\, (1 + R^2). \end{aligned} \tag{8}$$

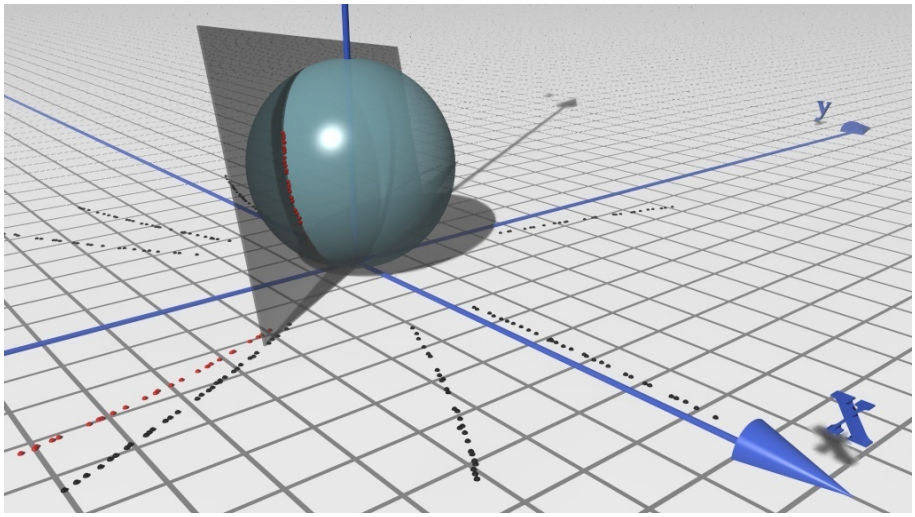


Figure 4. Stereographic projection of a track to the Riemann Sphere.

Circles and straight lines on the plane map to circles on the sphere, which again define planes in space (cf. Fig. 4). Consequently, the nonlinear task of circle-fitting on the plane can be reduced to the linear task of plane-fitting on the Riemann sphere. This can be used for fast fitting of trajectories of charged tracks [6]. In our application, the complex plane $\mathbb{R}^2$ is therefore replaced by the readout plane of the TPC. To maximize the surface of the projection of the pad-plane onto the sphere, the radius of the TPC clusters is divided by a scaling factor s before the actual projection.

Plane fit on the Riemann sphere A linear least-squares plane fit is performed by building a matrix $\mathbf{C}$ from the residua $\mathbf{d}$ of the hits to the average hit position $\mathbf{a}$ in the Riemann space:

$$\mathbf{C} = \frac{\sum_{\text{Hits}} w_{\text{Hit}}\, \mathbf{d}_{\text{Hit}}\, \mathbf{d}_{\text{Hit}}^{\mathbf{T}}}{\sum_{\text{Hits}} w_{\text{Hit}}} \tag{9}$$

with

$$\mathbf{d}_{\text{Hit}} = \mathbf{x_s} - \mathbf{a}, \tag{10}$$

$$\mathbf{a} = \frac{\sum_{\text{Hits}} w_{\text{Hit}} \mathbf{x_s}}{\sum_{\text{Hits}} w_{\text{Hit}}} \tag{11}$$

with weighting factor

$$w_{\text{Hit}} = \frac{1}{\sqrt{\sigma_x^2 + \sigma_y^2}} \tag{12}$$

and cluster error

$$\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z). \tag{13}$$

$\mathbf{C}$ has three eigenvectors which are the three principal axes of a covariance ellipsoid of the hits around $\mathbf{a}$. The shortest axis (which is the eigenvector $\mathbf{u}_{\text{j}}$ to the smallest eigenvalue λ_{j} of $\mathbf{C}$) is a

vector perpendicular to the plane the hits lie on. To get the normal vector $\mathbf{n}$, $\mathbf{u}_\mathrm{j}$ still has to be normalized:

$$\mathbf{C} \cdot \mathbf{u_i} = \lambda_i \, \mathbf{u_i} \tag{14}$$

$$\mathbf{n} = \frac{1}{u_j} \mathbf{u_j} \mid \lambda_j < \lambda_{k \neq j}. \tag{15}$$

The origin of the plane is $\mathbf{a}$. Any point $\mathbf{r}$ on the plane then satisfies the equation:

$$(\mathbf{r} - \mathbf{a}) \cdot \mathbf{n} = 0. \tag{16}$$

Calculation of the circle parameters This plane fit, projected back onto the pad-plane, delivers a circle, which has no constraints in radius R_h and position of its center $\mathbf{c}_h$. These two quantities are calculated from the normal vector of the plane $\mathbf{n} = (n_x, n_y, n_z)$, the distance of the plane to the origin d and the scaling factor s:

$$R_h = 0.5 \, |r_2 - r_1|, \tag{17}$$

$$\mathbf{c}_h = 0.5 \, (r_1 + r_2) \, \frac{1}{\sqrt{n_x^2 + n_y^2}} \begin{pmatrix} n_x \\ n_y \\ 0 \end{pmatrix}, \tag{18}$$

with

$$a_{1,2} = -0.5 \sqrt{1 - n_z^2} \pm 0.5 \sqrt{1 - n_z^2 - 4d^2 - 4dn_z}, \tag{19}$$

$$x_{1,2} = -dn_x + a_{1,2} \, n_z, \tag{20}$$

$$z_{1,2} = -a_{1,2} \, n_x - dn_x, \tag{21}$$

$$r_{1,2} = \mathrm{signum} \, (x_{1,2}) \; s \sqrt{\frac{z_{1,2}}{1 - z_{1,2}}}. \tag{22}$$

Straight line fit and hit sorting In a second step, slope m and offset t of the helix (cf. Eq. 6) are fitted and the clusters are sorted along the tracklet. The position of each cluster on the helix can be defined by its angle α (cf. Fig. 3). To determine α_i for every cluster in the tracklet, they are sorted by their z-coordinates. The angle α_1 of the first hit is simply the angle between the vector pointing from the center of the helix to the hit and the x-axis. Now α_i of each subsequent hit is defined by the angle α_{i-1} of the previous hit plus the angle between the hit and the previous hit with respect to the center of the helix. This method grants consistent angles also for curling tracks, because α is not restricted to $[0, 2\pi]$.

Now a straight line fit of the hit angles α versus the z-positions of the clusters is performed, which delivers the helix parameters m and t. With the latter two, the dip angle θ can be calculated according to Eq. 7.

After that, the hits are sorted again, now by their angle. Steep tracklets ($\theta < 30°$ or $\theta > 140°$) are left sorted by their z-coordinates. This technique makes sure that the hits are well sorted along the tracklet, which is important for track fitting.

2.2.3. Track merging In the process of track building, the actual tracks might not be found as a whole. The assumption of helical tracks holds only approximately, as the momentum of a particle changes along its path due to energy loss. Therefore, especially curlers are likely not to be recognized as one single tracklet in the first place. Moreover, outliers, noise hits or hits from other tracks wrongly assigned to a tracklet can alter the helix parameters so much, that the next

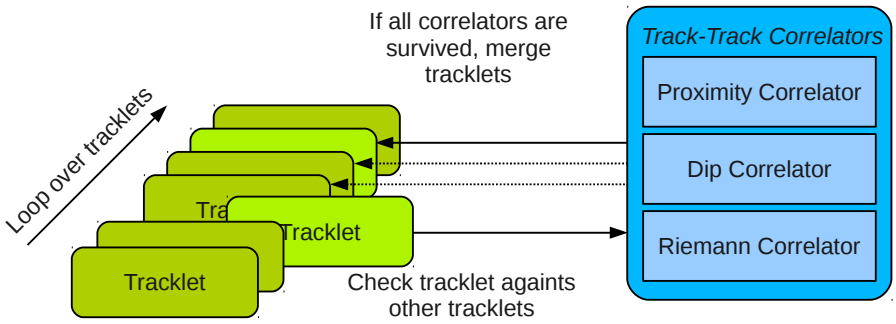


Figure 5. Scheme of the track merging process. See text for explanations.

correct hit does not survive all correlators any more. Also in track crossing areas, splitting can occur. Finally, particles with low energy loss can lead to fragmentary tracks, or the track may exit and re-enter the chamber one or more times. In the real chamber, dead channels or chips can also cause gaps. If the PR is sectorized (cf. Sec. 2.2.4), it is necessary to recombine tracks crossing the sector borders.

Therefore, a second level *tracklet merging* is performed. Similar to the track building process, the tracklets are presorted along z, and then compared to each other. Again, there are several *track-track correlators* which all (in this point the merging is different to the track building) have to be applicable and survived.

The *Proximity Correlator* compares the position of the first and last hits of the two tracklets. If the smallest distance of any combination of first and last hits is smaller than the *proximity-cut*, the correlator is survived.

The *Dip Correlator* compares the dip angles θ of the two tracklets if both tracklets are fitted. But not only has the absolute difference of the dip angles θ to be smaller than the *angle-cut*, also the relative z-positions of the tracklets have to match. Thus, the distance of the two helices at a the starting point of the tracklet with less hits is calculated. It has to be smaller than the *helix-cut*. If only one of the tracklets is fitted, the helix-distances of all hits of the unfitted tracklet are calculated and compared with the helix-cut. If none of the tracklets is fitted, the correlator is not applicable and the tracklets cannot be merged.

Finally, the tracklets have to pass the *Helix Correlator*. For tracklets with few hits, the helix fit might not be very accurate, and for straight tracks, some parameters of the helix (i.e. radius, center) are not well defined. Thus it is not reasonable to directly compare the helix parameters. Instead, a new track is created temporarily, containing the hits of both tracklets. A helix fit is performed and a *RMS-cut* on the Root Mean Square (RMS) of the distance of the hits to the helix is applied. If the two tracklets together do not have enough hits to be fitted, this correlator is not applicable.

2.2.4. Sectorization Since every cluster has to be checked against every track, the process of track building requires computational power of $\mathcal{O}(n_\mathrm{c} \cdot n_\mathrm{t})$, where n_c is the number of clusters and n_t is the number of tracks, respectively. In order to retain good performance also for large numbers of clusters and tracks, the TPC volume is split into sectors. Track building and merging is done for each sector separately, and only then global merging is performed.

2.2.5. Multistep approach The pattern reconstruction efficiency for different track topologies depends strongly on the presorting of the clusters mentioned in Sec. 2.2.1. Performance is best for tracks in sorting direction (i.e. very steep tracks and very low momentum curlers for sorting along z, tracks with high transverse momentum for radial sorting, curling tracks for angular

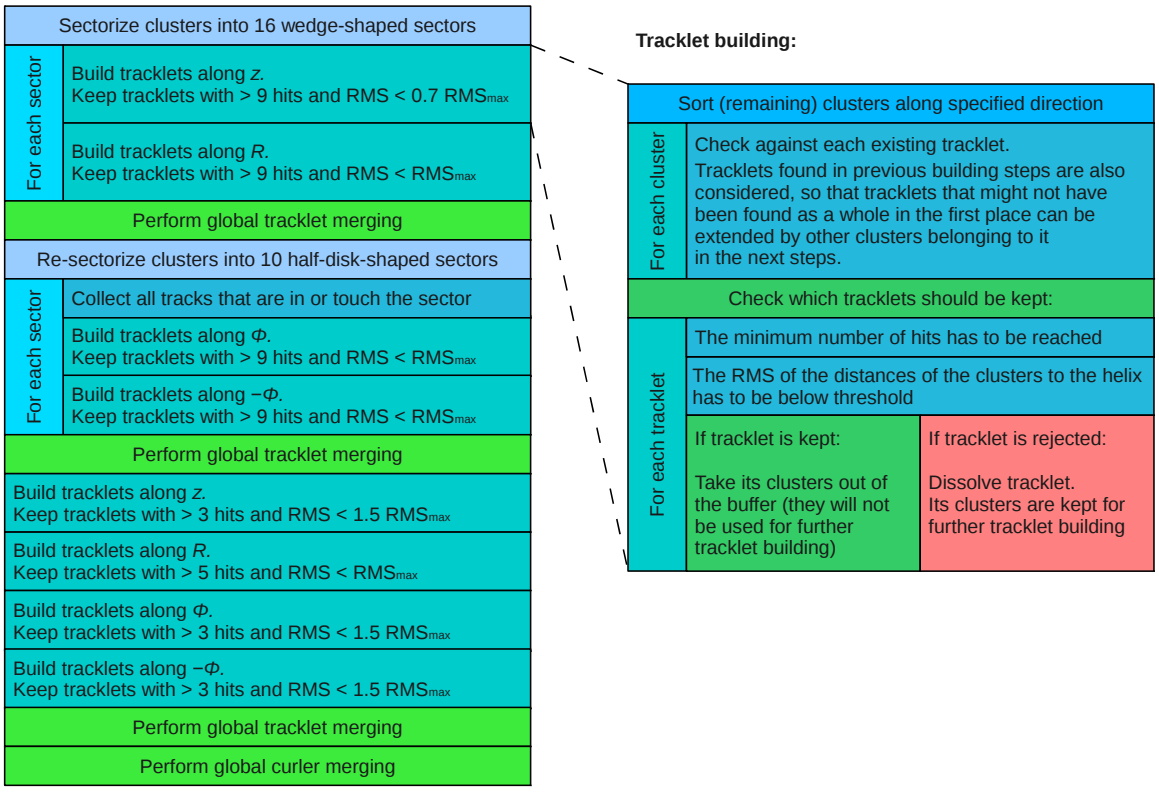


Figure 6. Scheme of the multistep approach (left) and tracklet building (right).

sorting). Thus, it is advantageous to run the PR more than once, and use a different presorting in each step. Tracks that reach certain quality criteria (i.e. a minimum number of hits and an RMS of distances of the hits to the helix smaller than a certain cut) are kept, the remaining clusters are sorted again, and the procedure is repeated. The full scheme is illustrated in Fig. 6. This approach yields high efficiencies for all kinds of track topologies (cf. Sec. 3), and has been used for the studies presented in Sec. 3.3 and 3.4.

For a fully mixed event like the one shown in Fig. 13, the computing time needed for PR with this approach is less than 4 min on a standard 3.1 GHz office PC, corresponding to $\mathcal{O}(50)$ ms/track.

2.2.6. Estimation of track-parameters For track fitting, *seed values* of the track parameters (i.e. momentum p, position $\mathbf{x} = (x, y, z)$ and direction $\mathbf{a}$) have to be provided. They can be calculated from the helix prefit and the magnetic field $\mathbf{B} = B\hat{\mathbf{e}}_\mathbf{z}$. The track parameters can be

Scaling factor s	24.6
Minimum number of hits for a helix fit	3
Cuts for the Hit-Track Correlators	
Proximity-cut	1.9 cm
Stretch of proximity-cut in z	1.6
Helix-cut	0.2 cm
Cuts for the Track-Track Correlators	
Proximity-cut	15 cm
Angle-cut	0.2 rad
Helix-cut	0.5 cm
RMS-cut	0.3 cm
$\mathrm{RMS}_{\mathrm{max}}$ for the multistep-approach	0.15 cm

Table 1. PR parameters used for the $\overline{\mathrm{P}}$ANDA-TPC simulation studies.

calculated at any hit angle α:

$$\mathbf{x} = \begin{pmatrix} x_0 + R_h \cdot \cos\alpha \\ y_0 + R_h \cdot \sin\alpha \\ m\alpha + t \end{pmatrix}, \tag{23}$$

$$\mathbf{a} = \frac{1}{\sqrt{1 + \left(\frac{1}{\tan\theta}\right)^2}} \begin{pmatrix} (y - y_0)/R_h \\ (x_0 - x)/R_h \\ 1/\tan\theta \end{pmatrix}, \tag{24}$$

$$p = \left| \frac{R_h}{\sin\theta} \cdot 3 \times 10^{-4} B \right|. \tag{25}$$

For R_h given in cm and B in kG, the output of the last equation is the momentum p in GeV c^{-1}.

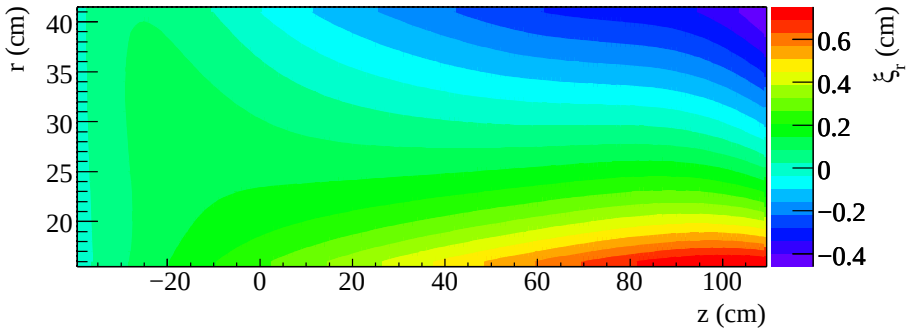
3. Performance studies

To asses the performance of the PR algorithm, several Monte Carlo (MC) performance studies have been conducted for the $\overline{\mathrm{P}}$ANDA TPC. $\overline{\mathrm{P}}$ANDA is a fixed target experiment with a $\bar{p}$ beam impinging on a hydrogen target. The GEM-TPC (the central tracker), is a cylindrical detector with an outer (inner) diameter of 84 cm (30 cm). $p\bar{p}$-interactions have been simulated with a Dual Parton Model (DPM) event generator [7], and a physics event generator has been used to simulate one particular η_c decay channel. The *primary* particles produced by these event generators have been transported through the detector and a 2 T solenoid magnetic field with GEANT [8]. During propagation, *secondary* particles like delta electrons, products from further decays etc. may be created. Energy loss in the TPC gas (Ne/CO_2 in a 90/10 mixture) has been calculated with the ALICE model [9]. The detector response has been obtained by a microscopic simulation of electron drift (400 V cm^{-1} drift field), GEM amplification (gain $G = 4000$), pad response, signal shaping and -sampling.

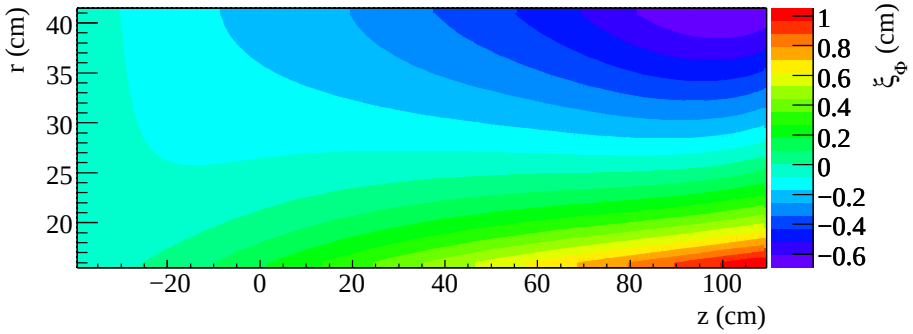
Moreover, distortions from space-charge effects due to ion backflow ($\epsilon = 4$) from the GEM amplification have been simulated, resulting in a drift distortion map (cf. Fig. 7). This map gives distortions in radial and azimuthal directions, depending on radius and z-coordinate of an ionization electron.

3.1. Momentum resolution of the fast helix fit

5000 π^+ particles with a fixed momentum of 1.0 GeV c^{-1} and a dip angle of $\theta = 60°$ have been simulated and processed with the PR algorithms. From the helix fit, the momentum can



(a) Total drift distortions in radial direction ξ_r (cm)



(b) Total drift distortions in azimuthal direction ξ_ϕ (cm)

Figure 7. Azimuthal and radial drift distortions of ionization electrons [10]. The graphs show the deviations from a straight line drift, experienced by an electron which starts its drift at (r, z) in the TPC volume.

be calculated according to Eq. 25. The momentum resolution is around 3 % and very well comparable to the momentum resolution obtained by a Kalman fit performed with GENFIT [11], as can be seen in Fig. 8.

3.2. Figures of merit

PR efficiencies for single tracks and mixed events, as well as event deconvolution capabilities have been studied. The following definitions have been employed:

$$\textit{Purity} \text{ of a track} = \frac{\text{number of “correct” clusters in track}}{\text{total number of clusters in the track}}, \tag{26}$$

$$\textit{Completeness} \text{ of a track} = \frac{\text{number of “correct” clusters in track}}{\text{total number of clusters from the MC track}}, \tag{27}$$

$$\text{PR } \textit{efficiency} = \frac{\text{found tracks}}{\text{number of tracks in TPC}}. \tag{28}$$

Here, the number of tracks in TPC is the number of MC tracks with at least one cluster in the TPC volume. The number of “correct” clusters in track is number of clusters in the track originating from the associated MC track. Finally, a track is “found” if it has a purity and completeness $> 50\,\%$. Therefore, a track can only be “found” once, even if it is split by the PR.

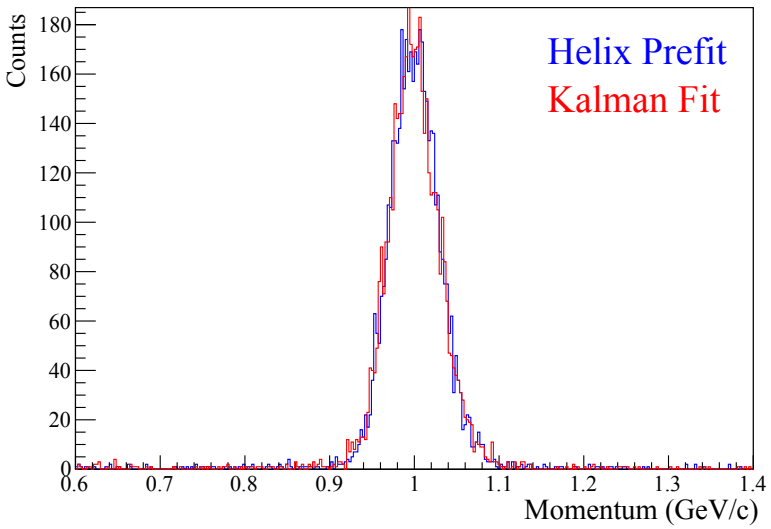


Figure 8. Momentum distribution before (in blue, track parameters obtained from the PR helix fit) and after a Kalman fit with 3 iterations (in red).

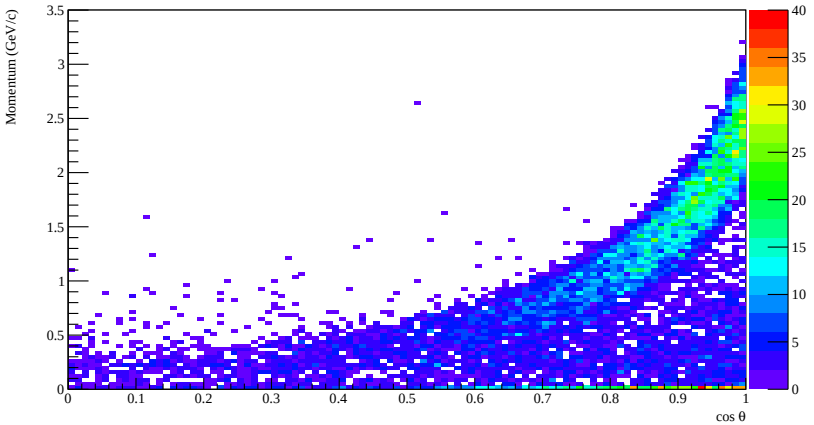


Figure 9. Phase space distribution (momentum vs. cosine of dip angle θ) of K tracks from η_c decays (cf. Eq. 29). The tracks are boosted into forward direction.

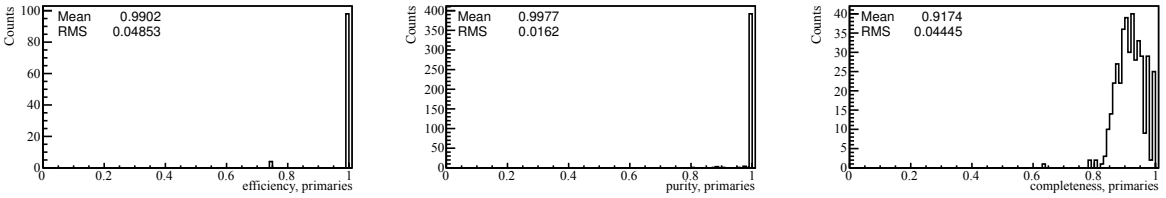
3.3. Pattern recognition performance for single events

One of the $\overline{\text{P}}$ANDA benchmark channels [3] has been simulated with a physics event generator:

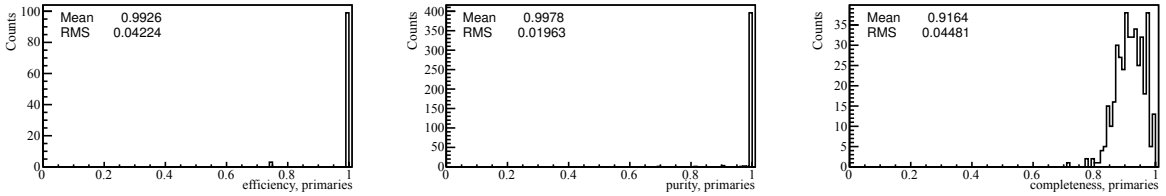
$$\eta_c \to \phi\phi \to K^+K^-K^+K^- \tag{29}$$

A pseudo-scalar η_c meson decays into two ϕ mesons, each of which decays into a K^+K^- pair. The phase space distribution of tracks from this decay is shown in Fig. 9.

PR performance for these pure *physics events* has been analyzed. As it can be seen in Fig. 10, these low density events do not really challenge the PR. Efficiencies and purities are above 99 %, and the track completeness almost never goes below 80 %. The performance is virtually unaffected by drift distortions, noticeable effects are less than 0.1 %.



(a) 0 % drift distortions.



(b) 100 % drift distortions.

Figure 10. PR efficiency, purity and completeness for single $\eta_c \to \phi\phi \to K^+K^-K^+K^-$ physics events without and with drift distortions.

3.4. Pattern recognition performance for mixed events

For the following study, one η_c *physics event* (Eq. 29) generated at $t_0 = 0$ has been mixed with 2000 *background events*, corresponding to an interaction rate of $2 \times 10^7\,\mathrm{s}^{-1}$. The background events are $\bar{p}p$-interactions simulated with the DPM event generator at a beam momentum of $4.73\,\mathrm{GeV}/c$. Momentum and phase-space distributions for these events can be seen in Fig. 11.

Approximately 4000 tracks are present in the chamber, corresponding to a maximum average track density of $0.5\,\mathrm{cm/cm}^3$ (i.e. track-length per volume; cf. Fig. 12). Fig. 13 gives an impression of the situation the PR has to cope with.

The performance without and with drift distortions has been compared. The figures of merit defined in Sec. 3 have been evaluated. They are shown for all tracks and again separated for primary and secondary tracks.

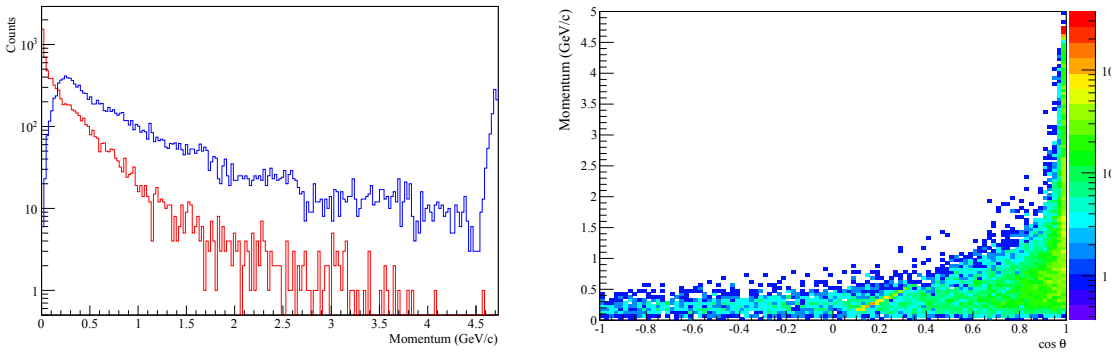
As it can bee seen in the Figures 14 and 15, PR performance is virtually unaffected by drift distortions, noticeable effects are typically less than 1 %. The overall efficiency is around 77 %. Track purities are also high, about 94 %. Track completeness is around 70 %. Here, the situation for primary tracks is fine, with an average completeness of 78 %. For secondary tracks, the situation is worse, with an average completeness around 54 %. However, one has to keep in mind that 30 % of the tracks have momenta lower than 50 MeV (cf. Fig. 11(a)).

4. Event deconvolution techniques

$\overline{\mathrm{P}}$ANDA is designed as a high-rate experiment without specialized trigger hardware. Instead, each sub-detector has to acquire and process signals autonomously. The event selection then takes place in compute nodes, where the information of all detector systems is available. Each sub-system can contribute to a trigger decision [3]. Whenever any sub-detector has found an interesting event signature at a certain time t_0, all other detectors have to provide their hits and tracks in a time window around t_0.

Therefore, a track time has to be estimated for each track in the TPC. This can be done by using several techniques, some of which are described in the following sections.

On the other hand, the TPC could also contribute to a trigger decision and e.g. find displaced secondary vertices or decays in the TPC-volume.



(a) Momentum distribution of primary (blue) and secondary (red) tracks.

(b) Phase space distribution of primary tracks.

Figure 11. Momentum and phase space distributions of tracks from DPM events at a beam momentum of 4.73 GeV/c. At this momentum, a peak from elastically scattered protons shows up. The acceptance for target tracks goes from $\theta_{\text{min}} = 8°$ to $\theta_{\text{max}} = 160°$.

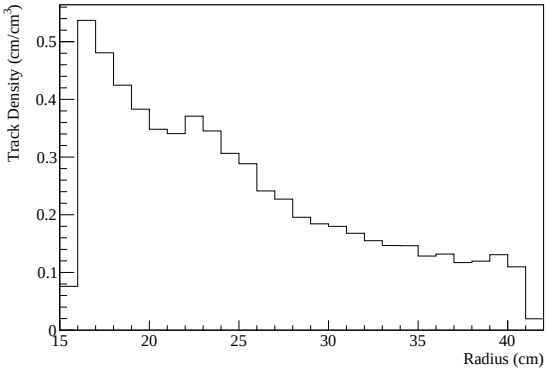


Figure 12. Track density vs. radius for DPM events at an interaction rate of $2 \times 10^7\,\text{s}^{-1}$.

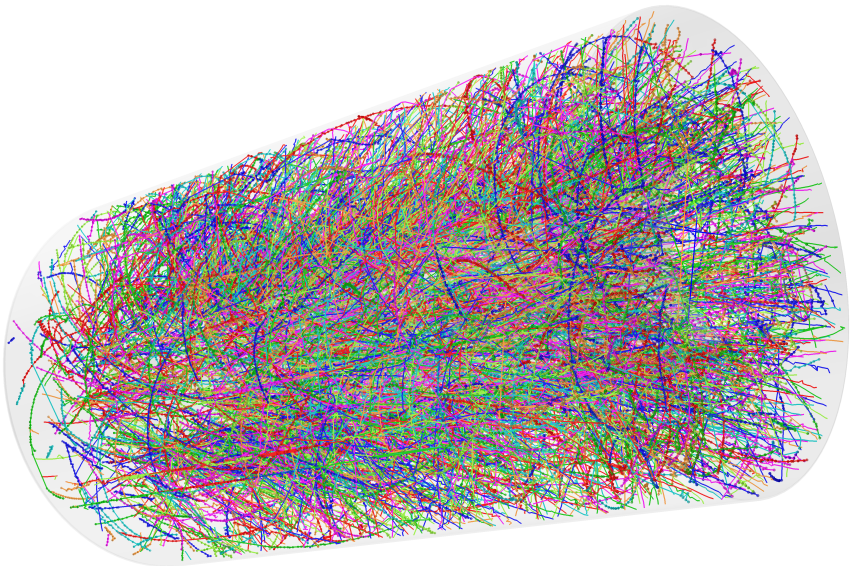


Figure 13. One drift frame in the $\overline{\text{P}}$ANDA TPC at an interaction rate of $2 \times 10^7\,\text{s}^{-1}$ after PR.

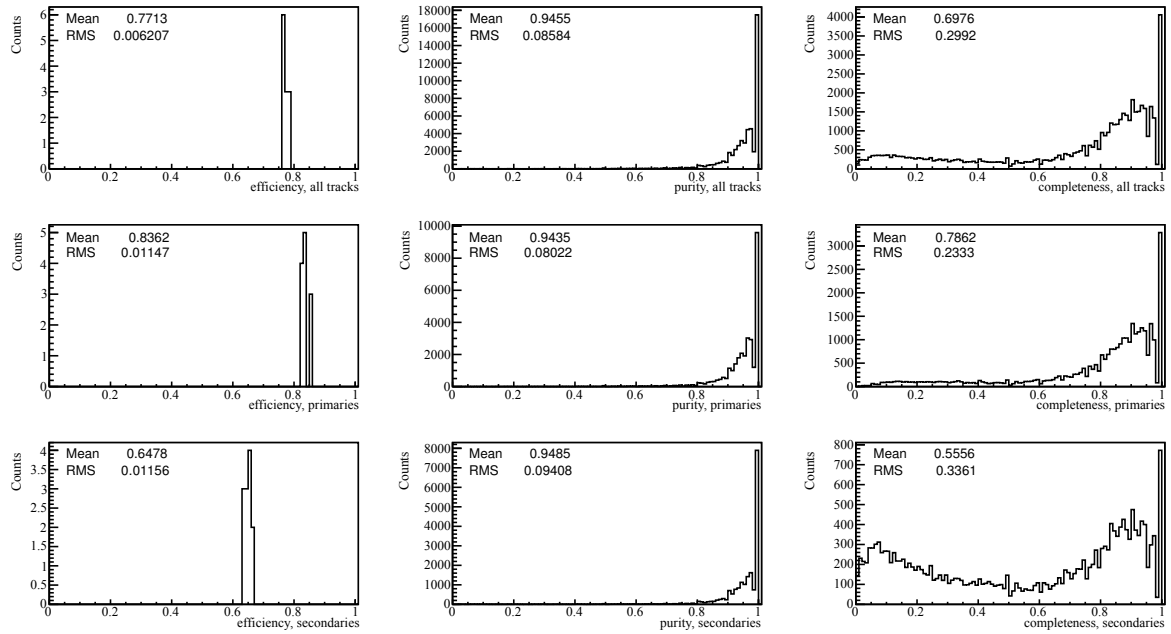


Figure 14. PR efficiency, purity and completeness at an interaction rate of $2 \times 10^7\,\mathrm{s}^{-1}$ without drift distortions.

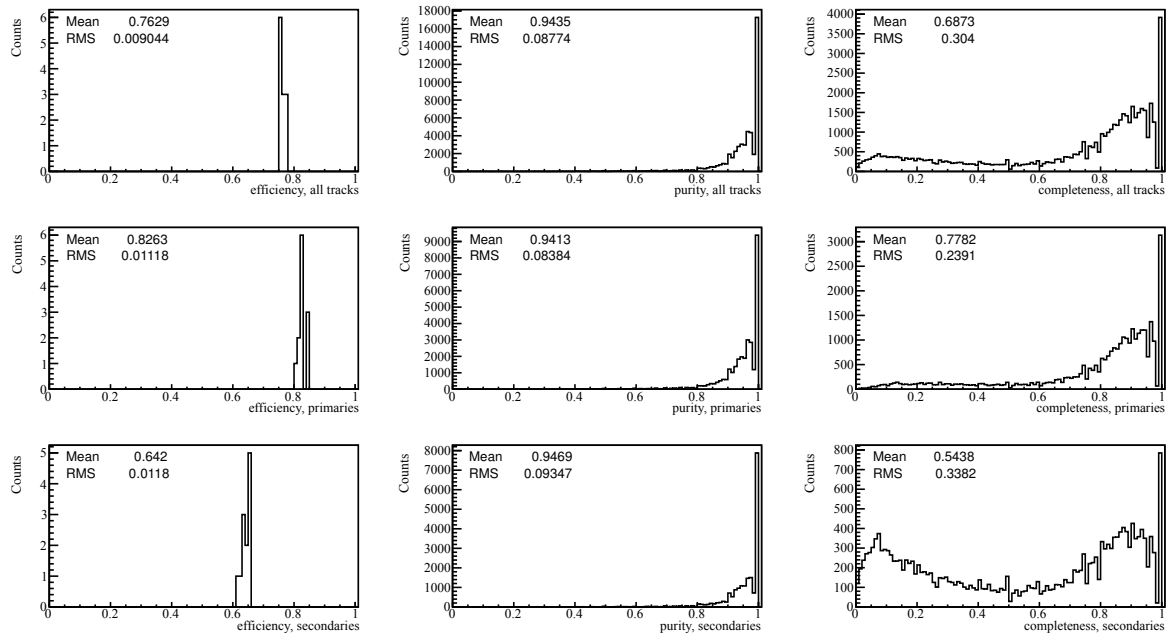


Figure 15. PR efficiency, purity and completeness at an interaction rate of $2 \times 10^7\,\mathrm{s}^{-1}$ with drift distortions.

For the event deconvolution study presented in Sec. 5, this general approach has been rephrased a bit: One physics event has been mixed with background events like described in Sec. 3.4. Tracks which might belong to the physics event with known t_0 are selected from all tracks present in one drift frame.

4.1. Target pointing
If the tracks from the physics event come from the primary vertex (the Interaction Point (IP)) in good approximation, *target pointing* can be used to separate physics from background tracks (cf. Fig. 16): For each track found in the TPC, the Point of Closest Approach (POCA) of the corresponding helix to the IP is calculated, and a cut on the distance of the POCA to the IP is applied. Primary tracks from earlier or later events are shifted in z and will not come close to the origin of the physics tracks. Thus, they can be rejected. Of course, this can only work reliably for primary tracks, because the POCAs of secondary tracks to the IP will be more or less randomly distributed.

The accuracy of the POCA calculation depends on space charge distortions, track-length measured in the TPC and extrapolation length. With this method, the amount of background tracks can be reduced by around 85 %. But that also means that still $\mathcal{O}(10^2)$ of the background tracks cannot be rejected without loosing physics tracks. Therefore additional methods of background reduction have to be applied.

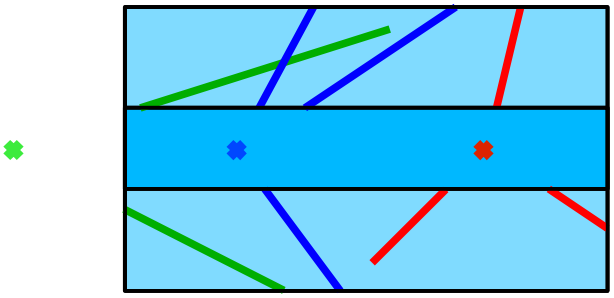
In a more general approach, the track time can be estimated with this technique: The track has to be shifted in z until the distance of its POCA to the IP becomes minimal. From this estimate of the real z-position, t_0 can be calculated according to equations 1 and 2.

4.2. End-cap penetration
A track that seems to end in the middle of the chamber (like the upper bold green track piece in Fig. 16) probably exited the TPC through the drift cathode or GEM-stack. This can be used to estimate t_0 of the track, again by shifting it in z until its end is at either cathode- or anode plate. For split tracks not properly merged in the PR, this method will yield wrong results.

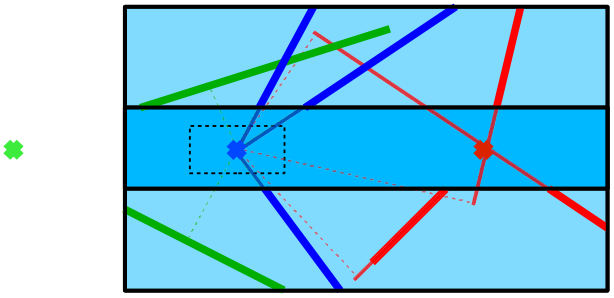
4.3. Correlation with other tracking detectors
So far, only the TPC has been considered. However, another powerful tool to reject background-tracks is to correlate TPC tracks with information from other detectors. In the case of $\overline{\text{P}}$ANDA, the Micro Vertex Detector (MVD) and the forward GEM-trackers can be used.

For the studies presented in Sec. 5.1, the following approach has been employed (cf. Fig. 17): All remaining tracks have been fitted with GENFIT and extrapolated to the MVD. Because of the very good time resolution of the MVD of around 5 ns, there will be no MVD hits for background-tracks. However, for the physics-tracks, where the track time t_0 is known, there will also be hits in the MVD at this time. By requiring a minimum number of MVD hits within a certain road width around the extrapolated track, the proper physics-tracks can be selected. The remaining tracks are then refitted with the added MVD hits, which also improve the resolution of the tracks.

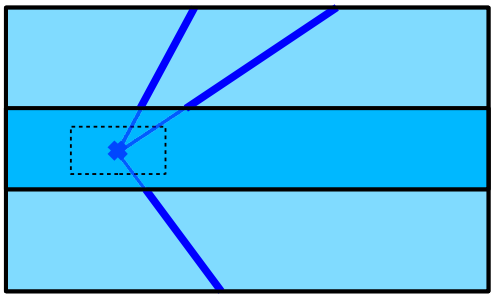
5. Event deconvolution performance
After the PR has processed the whole drift frame, we now want to find the needle in a haystack: The η_c physics tracks have to be selected from thousands of background tracks. Therefore we apply a target pointing scheme like explained in Sec. 4.1: The POCA of all tracks to the IP is calculated, and a variable target-cut in z as well as a fixed cut in r (5 cm) on the distance of the POCA to the IP is applied. The POCA is calculated from the helix prefit of the PR, no distortion corrections are applied, material effects etc. are not taken into account. Finally, the



(a) Tracks from different events (shown in green, blue and red) are present in one drift frame.



(b) For every track, the POCA to the IP is calculated.



(c) Only tracks with their POCA to the IP being in a fiducial volume around the IP survive.

Figure 16. Scheme of target pointing for selecting tracks coming from the IP. As a showcase, tracks from three events present in one drift frame are shown.

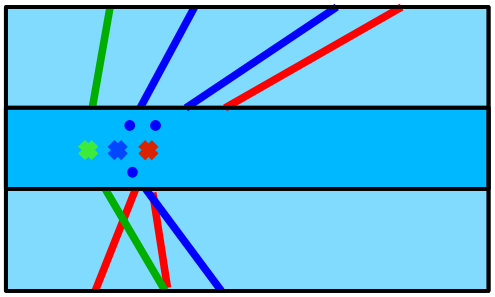
physics and background tracks that survive the cuts are counted. Again, we have to define some figures of merit to assess the event deconvolution performance:

$$\text{Event deconvolution purity} = \frac{\text{survived physics tracks}}{\text{survived physics tracks} + \text{survived background tracks}}, \quad (30)$$

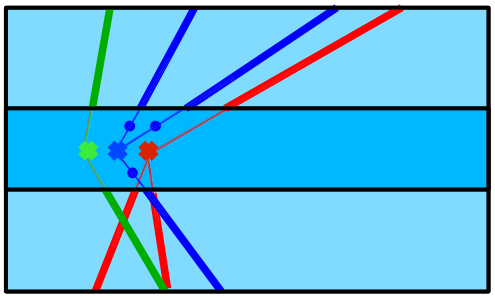
$$\text{Single track efficiency} = \frac{\text{survived physics tracks}}{\text{expected physics tracks}}, \quad (31)$$

$$\text{Full event efficiency} = \frac{\text{events with all physics tracks survived}}{\text{generated events}}. \quad (32)$$

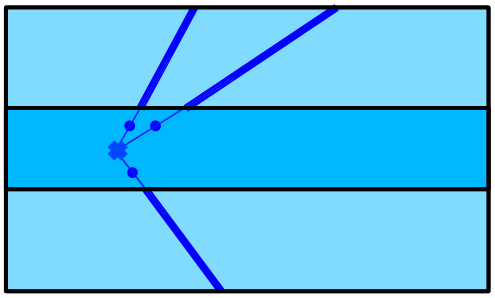
In this study, the geometrical acceptance of the TPC is canceled on MC level. Only events where all 4 generated K tracks span a sufficient track-length of a few cm in the TPC volume are taken into account. That means that 100 % full event efficiency could be achieved, assuming perfect PR and event deconvolution.



(a) Tracks from different events (shown in green, blue and red), that survived the target pointing are present in one drift frame.



(b) Every track is extrapolated to the MVD.



(c) Only tracks with a minimum number of MVD hits within a certain road width around the extrapolated track survive.

Figure 17. Correlation of TPC tracks with the MVD. As a showcase, tracks from three events present in one drift frame are shown. Points measured with the MVD are depicted as blue dots.

The results of the event convolution for different distortion settings as a function of the target cut are shown in Fig. 18. For no or small drift distortions, a target cut of 4 cm is already sufficient to obtain a full event efficiency of around 70 % with drift distortions.

It can be seen that the full event efficiency is virtually unaffected by small distortions, whereas at full distortions, more tracks are lost. In this case, opening the target cut does not restore the efficiency reached without distortions, due to the fixed cut in r.

For a target cut of 2 cm, the event deconvolution purity is around 1.5 %. This means that for one physics event with all 4 tracks retained, still around 270 background-tracks survive the target cut.

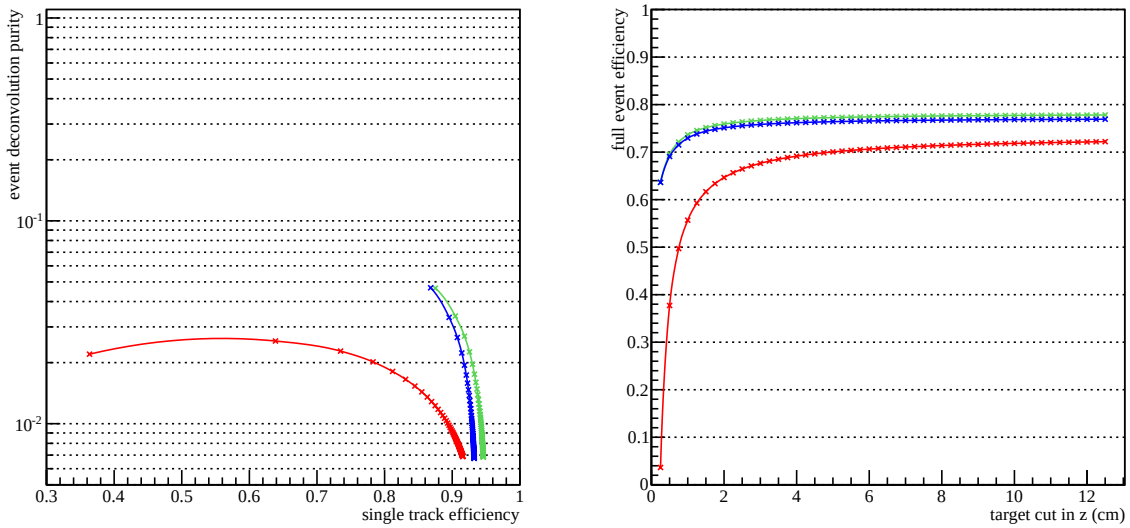


Figure 18. Event deconvolution performance for $\eta_c \rightarrow \phi\phi \rightarrow K^+K^-K^+K^-$ physics events mixed with 2000 background events with 0 % (green), 20 % (blue) and 100 % (red) drift distortions. The points are connected with a smoothed line to guide the eye.

5.1. Correlation with the micro vertex detector and event selection

At this point, after first event deconvolution via target pointing, it is assumed that the drift distortions due to space charge can be fully corrected. Therefore, the following study has only been performed for undistorted mixed events at an interaction rate of $2 \times 10^7\,\mathrm{s}^{-1}$.

The remaining tracks are correlated with the MVD (cf. Sec. 4.3) to further reduce the amount of background tracks. At least two MVD hits within a road width of 3 mm around the extrapolated track, are required.

Events with at least four charged tracks are selected and the invariant masses of all oppositely charged pairs are calculated. Figure 19 shows the invariant mass distribution, exhibiting a clear ϕ peak. Selecting further events with at least two ϕ candidates in a mass window of $\pm 30\,\mathrm{MeV}$ around the nominal mass of $1019.445\,\mathrm{MeV}\,c^{-2}$ results in the invariant mass distribution shown in Fig. 20. Here, the η_c mass of $2980\,\mathrm{MeV}\,c^{-2}$ has already been subtracted.

The spectrum exhibits a clear peak, well centered around 0, with almost no background. This is remarkable, and it should be mentioned that e.g. no vertex-fits or other constraints than described have been employed.

6. Conclusions

The continuous operation of a GEM-TPC in a high-rate environment requires specialized reconstruction algorithms. PR has to be performed in the TPC alone and has to be robust against drift distortions. High track densities pose an additional challenge. True three-dimensional clustering and PR algorithms have been developed, which stand up to these requirements and are capable of finding various kinds of track topologies with high efficiency.

Continuous, triggerless readout of a GEM-TPC at high rates leads to event mixing. A priori, the track times are unknown. Therefore, a high level physics selection has to be performed and TPC tracks within a time window around t_0 have to be selected and grouped together. It has been demonstrated that, despite of tracks from around 2000 different events being present in the TPC at the same time, tracks belonging to a particular event can be filtered out with good efficiency.

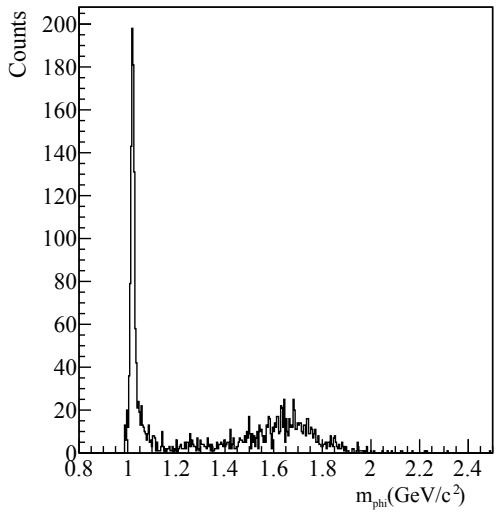


Figure 19. Invariant mass distribution of oppositely charged track pairs which survive a target-cut of 3.5 cm and have at least two additional hits in the MVD within a road width of 3 mm.

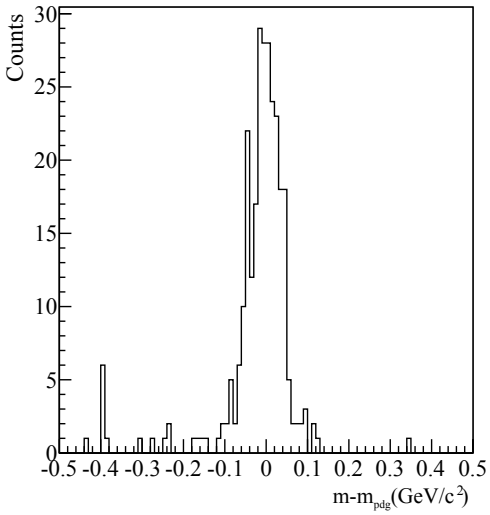


Figure 20. Invariant mass distribution of events with two ϕ candidates. The mass of the η_c has already been subtracted.

References

[1] Nygren D R and Marx J N 1978 *Phys. Today* **31N10** 46–53
[2] Alme J *et al.* 2010 *NIMA* **622** 316–367
[3] $\overline{\text{P}}$ANDA Collaboration 2005 $\overline{\text{P}}$ANDA*: Technical Progress Report* (GSI FAIR)
[4] Sauli F 1997 *Nucl. Inst. Methods* **A386** 531–534
[5] Frühwirth R, Strandlie A and Waltenberger W 2002 *NIMA* **490** 366 – 378
[6] Strandlie A, Wroldsen J, Frühwirth R and Lillekjendlie B 2000 *Comput. Phys. Commun.* **131** 95 – 108
[7] Capella A, Sukhatme U, Tan C and Tran Thanh Van J 1994 *Physics Reports* **236** 225 – 329
[8] CERN 1993 *CERN Program Library Long Write-up* **W5013**
[9] ALICE Collaboration 2000 *ALICE-DOC-2003-011*
[10] Böhmer F V *et al.* 2009 *IEEE Nuclear Science Symposium Conference Record* 2273–2279
[11] Höppner C, Neubert S, Ketzer B and Paul S 2010 *NIMA* **620** 518–525