LC-TOOL-2016-001

http://www-flc.desy.de/lcnotes

A track finding method for a TPC with pixelated readout based on local road search with macro pixels

Claus Kleinwort, DESY, Germany

September 29, 2016

Abstract

This document describes a local track finding algorithm using a road search with *macro pixels* developed for a prototype TPC with pixelated readout.

1 Introduction

A track finding method based on local road search with macro pixels has been developed especially for the data taken by the Large Prototype TPC (LPTPC) [1] with *GridPix* readout [2] at the DESY test beam in March 2015 and implemented as a MarlinTPC [3] processor.

In this setup the drift direction perpendicular to the readout plane is along the Z-coordinate. This is instrumented with 160 GridPix chips grouped into 20 *octoboards* in three mechanical modules .

For each GridPix chip the pixel address space is rearranged into larger bins to reduce the combinatorics. Each such bin (in the XY-plane) can contain several *macro pixel* hits separated in the Z-direction. All macro pixel hit pairs with certain distance and content are used to define straight line search roads. Roads compatible in longitudinal and transverse direction with the expectation from a track are accepted as track segments.

Track segments are combined into larger objects in an hierarchical way to reduce combinatorics and extrapolation. The combination is performed first for all chips in an octoboard, next for all octoboards in a module and finally between all modules.

In section 2 the basic concepts are presented, section 3 describes the sequence of processing steps, section 4 shows some performance tests and section 5 explains the steering parameters of the TIMEPIXLOCALROADSEARCHPROCESSOR.

2 Concepts

2.1 Macro Pixel hit

A macro pixel hit is a collection of at least two pixel hits contained in a larger voxel and characterised by the first two moments (center $\mathbf{c}$ and variance $\mathbf{V}$) of the contributing pixel positions.

2.2 Macro pixel bin

The 256×256 pixels in a GridPix chip are grouped in $n_b \times n_b$ bins of size $s_{mp} \times s_{mp}$ (e.g. $n_b = 8, s_{mp} = 32$). If n_b is no divisor of 256 the remaining pixels are attributed to the border bins $(0, n_b - 1)$. The pixel hits contained in a bin are arranged into macro pixel hits.

2.3 Macro pixel road

A macro pixel road is a straight line in three dimensions defined by a seeding pair of macro pixel hits containing all macro pixel hits within a maximum distance. It is characterised by the longitudinal and transversal variance of the contributing macro pixel hits with respect to this line. The combinatorics scales with the square of the number of seeding macro pixel hits.

2.4 Macro pixel segment

A macro pixel segment is a collection of pixel hits from several macro pixel hits described by a straight line or helix fit.

2.5 Local neighbourhood

The combination of macro pixel segments to tracks is performed exploring the hierarchical structure of the readout (chips in octoboards in modules) in the corresponding local neighbourhood to reduce combinatorics and extrapolation (uncertainties) and to account for different levels of systematics like the mounting precision. This hierarchy should be part of the geometry description.

2.6 Equivalence classes

Equivalence classes are used to combine macro pixel segments to tracks. For a set of segments a *matching* relation is defined for pairs of elements. An equivalence class is the subset of all directly or indirectly related elements. The original set is decomposed in disjunct subsets.

3 Sequence

3.1 Input

Input are all pixel hits with the Z-position determined from the drift time inside the physical detector volume. Used are the chip, row and column number, the measured spatial position (in millimeter) and its (diagonal) covariance matrix. The measurement errors in X- and Y-directions are expected to be equal (σ_{xy}). The errors can be adjusted by the TPCHIT_FIXERRORS_FORTRACKING MarlinTPC processor. A constant magnetic field is assumed to describe the tracks by a helix or straight line.

3.2 Segment finding for a GridPix chip

The following steps are performed independently for each GridPix chip (with an example in figure 1).

3.2.1 Creation of macro pixels

All pixel hits are grouped according to row and column number into macro pixels bins of size s_{mp}. These are subdivided into macro pixel hits according to the Z-coordinate.

The pixel hits are first sorted in Z-direction and using a sliding window (e.g. of size $\pm 2 \cdot \sigma_Z$) the core of a cluster is looked for. The window with the maximum number of unused pixel hits inside is selected and all pixel hits in a region (e.g. $\pm 3.5 \cdot \sigma_Z$) around its center of gravity define a cluster and are marked as used. From clusters with at least two pixel hits (in different rows and columns to suppress noisy rows or columns) a macro pixel hit is created. This procedure is iterated until no pixel hits are left.

3.2.2 Selection of seeding pairs

For the seeding only the macro pixel bins containing not too many macro pixels hits are selected to limit the combinatorics. All pairs of unused macro pixel hits with some minimum number of pixel hits in both of them and some minimum distance are considered as seeding pairs for macro pixel roads and are ordered according to distance. The integer distance is defined as the maximum of the absolute differences in row and column numbers of the macro pixel bins. This definition is independent of the orientation of the track on the chip. The Euclidian distance would give preference to tracks along the chips diagonal.

3.2.3 Searching with macro pixel roads

From the list of seeding pairs macro pixel roads are build in decreasing order of their distance d_{pair}. For each pair the road direction $\mathbf{e}_{xy} = (e_x, e_y, 0)$ in the XY-plane can be derived from the two macro pixel hit (center) positions $\mathbf{c}_1$ and $\mathbf{c}_2$:

$$(\Delta x, \Delta y, \Delta z) = \mathbf{c}_2 - \mathbf{c}_1, \quad \Delta s = \sqrt{\Delta x^2 + \Delta y^2}, \quad e_x = \frac{\Delta x}{\Delta s}, \; e_y = \frac{\Delta y}{\Delta s} \tag{1}$$

For a macro pixel hit at position $\mathbf{c}$ the arc-length along the road (s_{arc}) and the predicted position ($\mathbf{c}_{\text{road}}$) on the road is:

$$s_{\text{arc}} = (\mathbf{c} - \mathbf{c}_1) \cdot \mathbf{e}_{xy} \tag{2}$$

$$\mathbf{c}_{\text{road}} = (1 - \zeta) \cdot \mathbf{c}_1 + \zeta \cdot \mathbf{c}_2, \quad \zeta = s_{\text{arc}}/\Delta s \tag{3}$$

The distance to the road in XY (r_{xy}) and Z (r_z) can be calculated using the measurement directions $\mathbf{m}_{xy} = (e_y, -e_x, 0)$ and $\mathbf{m}_z = (0, 0, 1)$:

$$r_k = (\mathbf{c} - \mathbf{c}_{\text{road}}) \cdot \mathbf{m}_k, \; k = xy, z \tag{4}$$

For the matching with the road the variance $\mathbf{V}$ inside the macro pixel hit is projected onto the measurement directions and added (yielding the average squared single pixel hit residual $\langle r^2_{\text{pix}} \rangle$):

$$v^2_{\text{k}} = r^2_k + \mathbf{m}_k \cdot \mathbf{V} \cdot \mathbf{m}^t_k \tag{5}$$

The macro pixel hit is assigned to the road if the squared pulls v^2_{xy}/σ^2_{xy} and v^2_z/σ^2_z are below some cut values. To the χ^2 of the road $n_{\text{pix}} \cdot (v^2_{xy}/\sigma^2_{xy} + v^2_z/\sigma^2_z)$ is added with n_{pix} as the number of pixel hits inside the macro pixel hit.

Roads fulfilling the following requirements are called valid:

Content The number macro pixels (n_{mpix}) is at least three.

Longitudinal shape For a uniform distribution of macro pixel hits along the track the variance of the arc-length normalised to the track length should be $1/12$. For roads seeded by hits from different tracks in the same chip this value can be larger up to $1/4$. Together with the maximal arc-length gap it is limited to some maximal value.

Transversal shape The χ^2 normalised with the total number of pixel hits in the road ($\chi^2_{\text{norm}} = \chi^2/(n_{\text{pix,tot}} - 4)$) does not exceeds a certain value.

If a valid road contains all macro pixel hits on the chip the road search is stopped immediately. After the first valid road has been found the road search will stop after trying the seeding pairs with the same distance or one smaller than that by one unit ($d_{\text{pair}} \geq d_{\text{pair,first}} - 1$) to be able to find better rows while allowing for some inefficiency. If more than one valid roads have been found the one with the smallest normalised χ^2 divided by the number of macro pixels ($\chi^2_{\text{norm}}/n_{\text{mpix}}$) is selected to create a (level zero) macro pixel segment. The corresponding macro pixel hits are marked as used. If there are still sufficient unused ones the procedure restarts with a new selection of seeding pairs (3.2.2).

3.2.4 Segment fitting

Assuming a constant magnetic field in Z-direction the pixel hits of the macro pixel segments are fitted with a circle [4] in the XY- and a straight line in the ZS-projection using the average of the positions as reference point. The five track parameters $\mathbf{p}$ with covariance matrix $\mathbf{V}_p$ defined there are the curvature and for both projections the directions at and distances to the point of closest approach in XY. For zero magnetic field straight lines are fitted in both projections.

3.3 Segment combination

Macro pixel segments are combined into larger segments in an hierarchical way on several levels:

1. For all chips in an octoboard.
2. For all octoboards in a (mechanical) module.
3. Between all (neighbouring) modules.

3.3.1 Segment matching

For each target hierarchy structure (single octoboard, single module, multiple modules) the list of segments in its subcomponents is created. For each segment pair from this list the overlap and distance is estimated from the segment reference points (average pixel hit positions) $\mathbf{r}_i$ and track length l_i in the XY-plane:

$$(\Delta x, \Delta y) = \mathbf{r}_2 - \mathbf{r}_1, \quad \Delta s = \sqrt{\Delta x^2 + \Delta y^2}, \quad d_{\text{rel}} = \frac{\Delta s}{l_1 + l_2} \tag{6}$$

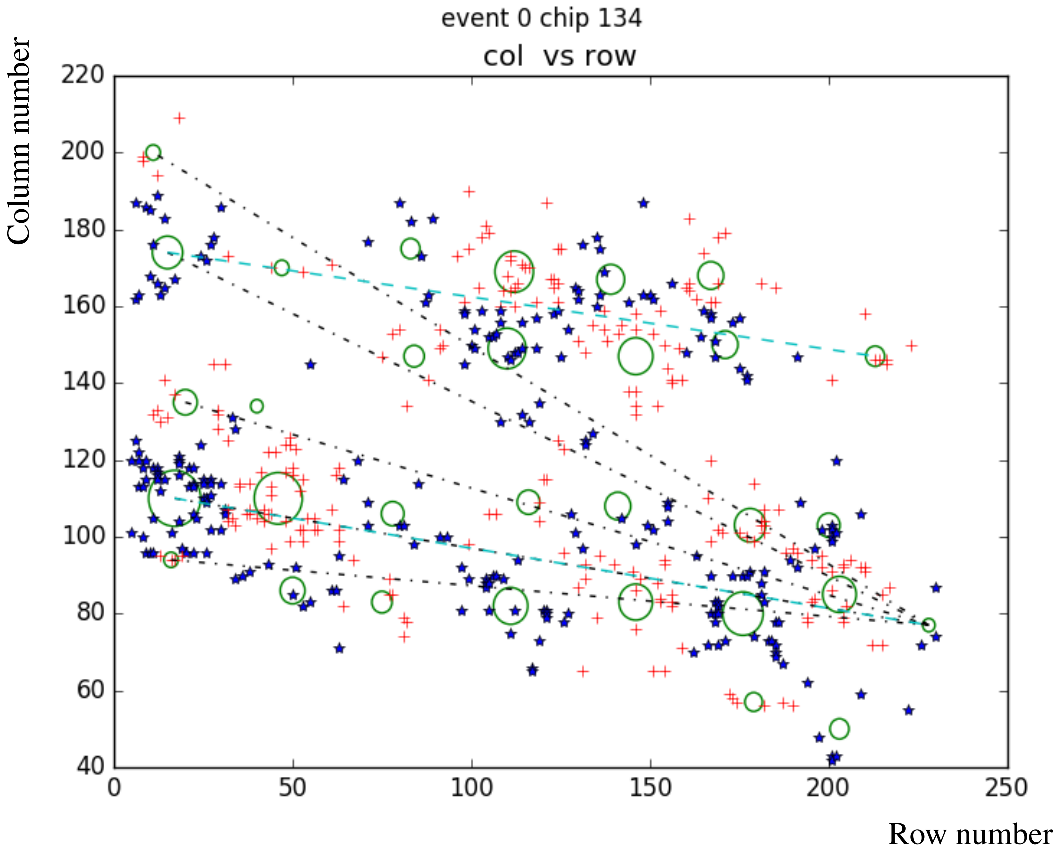

Figure 1: GridPix chip with two tracks. The stars and crosses show the single pixel hit positions (row, column) in a checkerboard pattern indicating the macro pixel bins. The circles represent the macro pixel hits which the area proportional to the number of pixel hits inside, the dashed-dotted lines as examples the seeding macro pixel hit pairs with maximal distance (for the first iteration) and the dashed lines the seeds for accepted macro pixel segments.

The relative distance d_{rel} must have a minimum value (0.25) to avoid too much overlap and maximum value to avoid too large gaps and too much extrapolation. The relative midpoint is used as new reference point:

$$\mathbf{r}_{\text{mid}} = \frac{l_2}{l_1 + l_2} \cdot \mathbf{r}_1 + \frac{l_1}{l_1 + l_2} \cdot \mathbf{r}_2 \tag{7}$$

For both segments the (n_{par}) parameters and the covariance matrices are propagated to this midpoint and compared there:

$$\chi^2_{\text{norm}} = (\mathbf{p}_{\text{mid},1} - \mathbf{p}_{\text{mid},2}) \cdot (\mathbf{V}_{p,\text{mid},1} + \mathbf{V}_{p,\text{mid},2})^{-1} \cdot (\mathbf{p}_{\text{mid},1} - \mathbf{p}_{\text{mid},2})^t / n_{\text{par}} \tag{8}$$

Pairs of segments with χ^2_{norm} smaller than some cut value are added to the set of matching segments.

3.3.2 Segment merging

Segments are merged from the equivalence classes of matching segments. For each class the list of segments is replaced by a single segment with the combined list of pixel hits and the track parameters calculated from a new fit at the new reference point. The level of the new segment corresponds to the hierarchy level (1–3) of the combination.

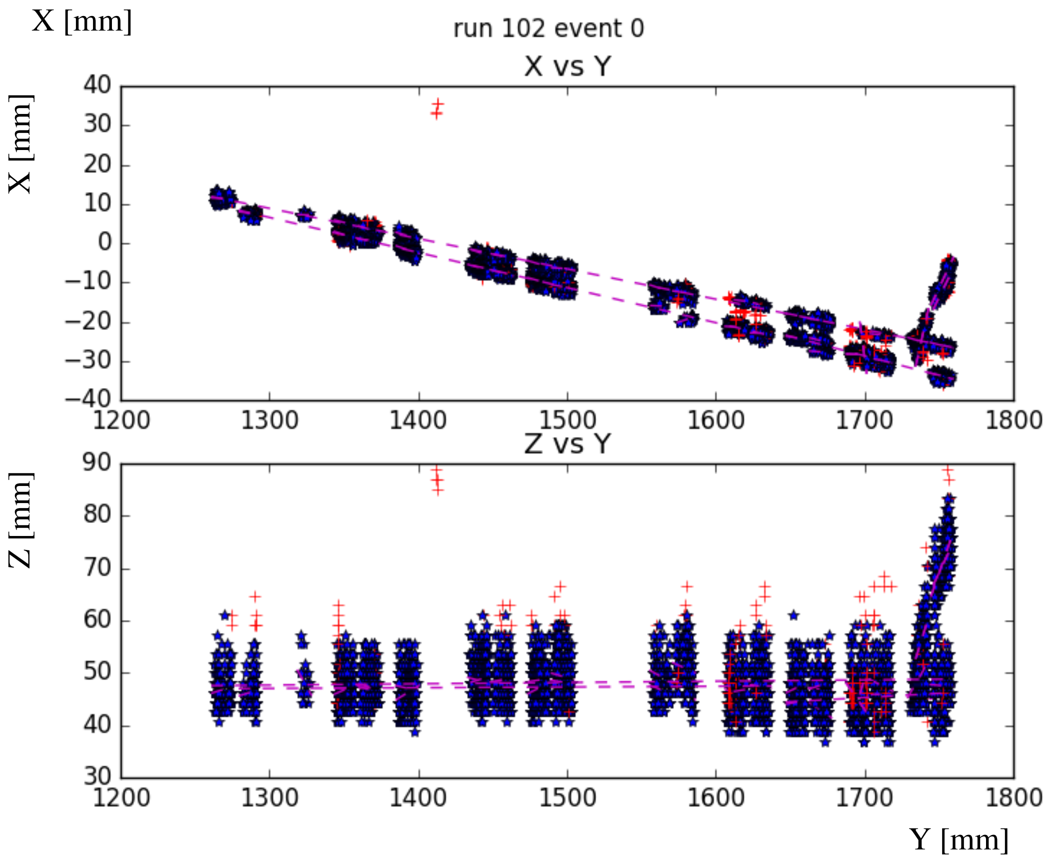

Figure 2: The complete event from figure 1 after segment combination. The stars indicate pixel hits associated to a track (segment), the crosses unused pixel hits and the dashed lines the combined segments.

3.4 Output

From the fit results and the list of pixel hits a TPC track is constructed for each segment with a minimum level (as illustrated in figure 2). At the first hit as reference point the parameters and covariance matrix are converted to the LCIO parametrisation [5].

4 Performance

The performance of this method for the LPTPC has been checked with GridPix data from the DESY test beam campaign in March 2015. Unless mentioned otherwise the default settings described in section 5 are applied.

4.1 Data used

For several runs at different Z-positions and with or without magnetic field data files with a few hundred events have been provided by A. Shirazi from the University of Siegen. From all runs the first hundred events have been used for some basic checks of the method as summarized in table 1. With run 102 some more detailed studies have been performed. The other runs still suffer from noise, e. g. a significant fraction of the pixel hits have a measured Z-position outside the LPTPC volume. For all runs rough estimates of the XY- and Z-resolution have been obtained from the single chip track fit residuals. In XY the resolution varies between 0.2 and 1.6 mm.

run number	62	67	71	80	84	88	102
drift distance [mm]	560	280	12	400	200	5	47
magnetic field [T]	0.	0.	0.	1.	1.	1.	0.
XY-resolution [mm]	1.6	1.2	0.4	0.6	0.4	0.2	0.7
Z-resolution [mm]	4.0	3.5	3.0	3.0	3.0	3.0	3.0
all pixel hits / event	24680	25050	24490	16790	16000	14630	5020
accepted hits / event	13850	14110	13300	10140	6518	5265	4874
pixel / macro pixel	6.4	7.2	11.4	13.6	15.3	12.1	11.1
macro pixel / chip	18.3	16.0	9.8	12.8	11.3	7.9	14.1
int. track length [mm]	542.5	608.4	507.0	503.1	468.5	416.2	651.2
cpu time [msec]	9.6	9.8	7.3	8.0	6.3	6.0	5.7

Table 1: Summary table for different data runs. The different transverse hit densities (XY-resolutions) are reflected in different number of pixel hits per macro pixel hit and macro pixel hits per chip. The integrated track length per event is about 500 mm. The required cpu time varies with the number of (accepted) pixel hits.

4.1.1 Shortcomings of the geometry description

The geometry description in the corresponding GEAR [6] file describes only a flat list of GridPix chips without any hierarchy structure in a local coordinate system with the origin in the center module. This leads to two problems:

1. The track fitting involves a twofold ambiguity in terms of the track direction. This is resolved by assuming the tracks are coming from the origin of the coordinate system. This implies now different track directions (by π) for the two outer modules with the origin in the center module complicating the segment combination. To circumvent this the segment finding, fitting and merging is performed in a different system (e.g. the global LPTPC system) according to an offset specified in the steering. There all modules should be on the same side of the origin.

2. The missing hierarchy structure defining the *local neighbourhood* for the segment combination has to be deduced from the hit data itself. Each pixel hit contains its chip number. The octobaord number is assumed to be the chip number divided by eight and the module number to be linear in the octoboard number with offset and slope defined in the steering. This may be wrong for future setups and the code of this processor would have to be adjusted.

4.2 Influence of bin size

The size and with this the number of macro pixel bins per chip has been changed ($n_b = 6, 8$ (default), $12, 16$) for run 102 as compiled in table 2. The performance as indicated by the integrated track length and fraction of used pixel hits per event varies only weakly. The required cpu time changes only moderately with the number of bins.

bin size	number of bins	pixel per macro pixel	macro pixel per chip	int. track length [mm]	used pixel fraction [%]	cpu time [msec]
16	16	6.5	27.4	674.5	83.1	8.0
21	12	7.8	22.6	671.5	87.5	7.0
32	8	11.6	15.4	673.5	90.5	5.7
42	6	15.9	11.7	667.4	90.5	5.5

Table 2: Summary table for variation of macro pixel bin size. Compared are the averages for the number of pixel hits per macro pixel hit and macro pixel hits per chip for all chips and per event the integrated track length, fraction of used pixel hits and the cpu time on a DESY workgroup server. For comparison the LCIOOUTPUTPROCESSOR takes 134 msec. The reproducibility of the timing measurement is about 0.5 msec.

4.3 Monitoring plots

With the data from run 102 monitoring plots have been created. Figure 3 displays the number of pixel hits per macro pixel hit and the number of macro pixel hits per GridPix chip. Figure 4 illustrates the fraction of macro pixel roads tried for segment finding for a GridPix chip and the average number of macro pixel hits as a function of the number of segments found. Figure 5 presents the distributions of the (final) segment level and the track length.

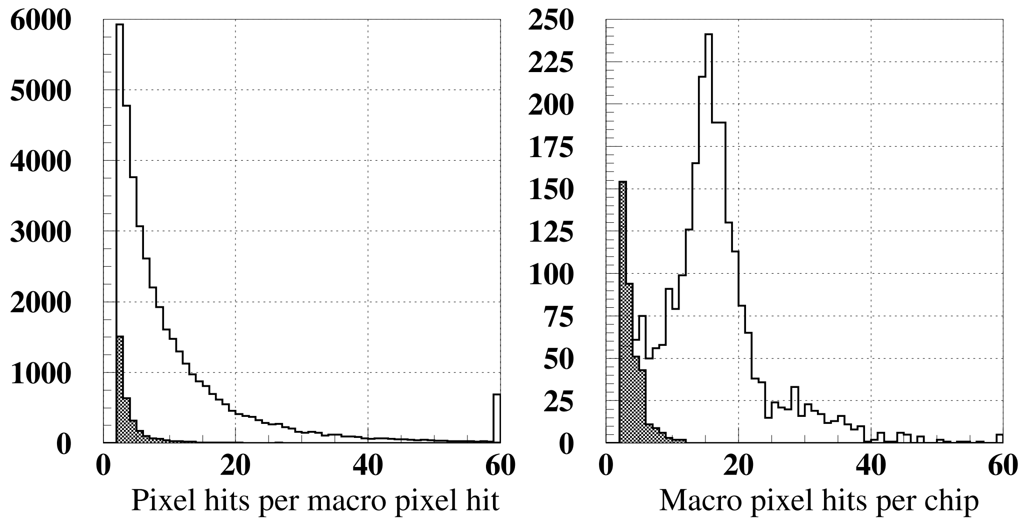

Figure 3: The left plot shows the number of pixel hits per macro pixel hit with a mean of 11.6 and the right one the number of macro pixel hits per chip with a mean of 15.4. The shaded histograms are for the cases of unused macro pixels hits or chips without any segment found.

5 Steering parameters

The steering parameters with the defaults indicated in parentheses are:

InputHits ("TPCHits"): The name of the input collection of TPC hits .

OutputTracks ("RoadSearchTracks"): The name of the output collection with the found tracks.

BFieldScaleFactor (1.0): Scale factor for the magnetic field (map), allows to switch off the magnetic field (section 3.2.4).

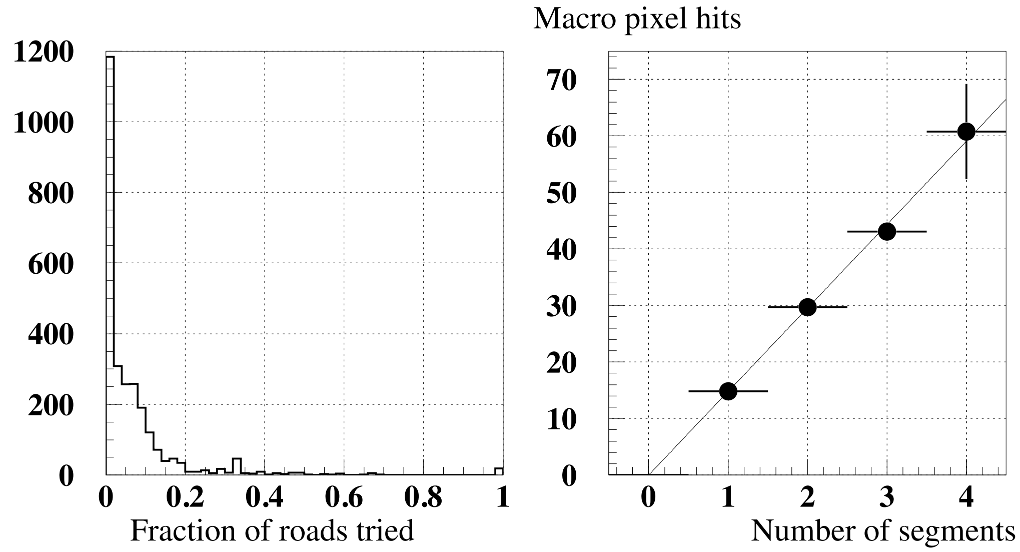

Figure 4: The left plot shows the fraction of the $n_{\rm mpix} \cdot (n_{\rm mpix} - 1)/2$ possible roads tried for segment finding. The average is about 7.1%. The right plot shows the average number of macro pixel hits per chip (for chips with segments) as function of the number of segments found for the chip. The offset is 0.003 and the slope 14.8.

MinTrackLevel (1): Minimum segment level for track output (section 3.4).

MacroPixelSize (32): Macro pixel (bin) size (section 3.2.1).

MinZmeas (0.) Minimum (accepted) measured Z value (section 3.1).

MaxZmeas (600.) Maximum (accepted) measured Z value (section 3.1).

SigmaZ (3.0): Resolution in Z for clustering (section 3.2.1).

Xoffset (0.): Offset in x of (local) coordinate system (section 4.1.1, item 1).

Yoffset (1500.): Offset in y of (local) coordinate system (section 4.1.1, item 1).

MaxMult (3): Maximum multiplicity for road seeding macro pixel bins (section 3.2.3).

MinPixels (3): Minimum number of pixels hit for road seeding macro pixel hits (section 3.2.3).

MinDist (2): Minimum distance for road seeding macro pixel bins (section 3.2.3).

MaxPull2XY (15.0): Maximal squared pull to road in XY (section 3.2.3).

MaxPull2Z (15.0): Maximal squared pull to road in Z (section 3.2.3).

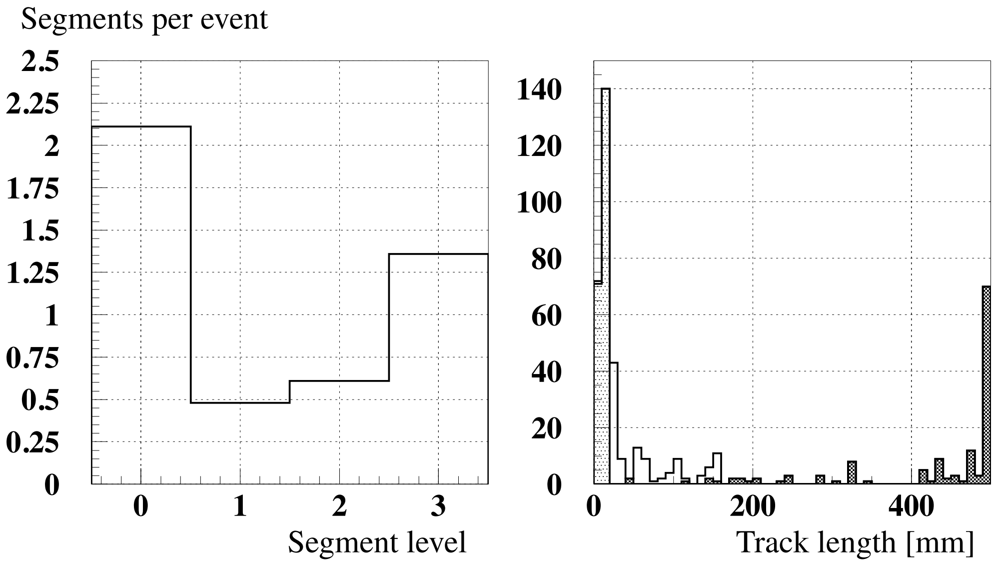

Figure 5: The left plot shows the number of segments found per event as function of the final segment level (0: single chip, 1: single octoboard, 2: single module, 3: multiple modules). The right plot shows the track length l_{xy} in the XY-plane per segment. The dotted bins on the left are from remaining single-chip-segments ($\langle l_{xy} \rangle = 10.4$) and the shaded ones from multi-module-segments ($\langle l_{xy} \rangle = 428.$) .

MaxChi2 (5.0): Maximal χ^2/ndf for road (section 3.2.3).

MaxSvar (0.15): Maximal relative arc-length variance for road (section 3.2.3).

MaxGap (4.): Maximal (arc-length) gap for road (section 3.2.3).

OctoboardOffset (8): Octoboard number offset (in module=(octo+offset)/scale) (section 4.1.1, item 2).

OctoboardScale (12): Octoboard number scale (in module=(octo+offset)/scale) (section 4.1.1, item 2).

Chi2CutChip (30.): Maximal χ^2/ndf for chip segment matching (section 3.3.1).

Chi2CutOcto (30.): Maximal χ^2/ndf for octoboard segment matching (section 3.3.1).

Chi2CutMod (50.): Maximal χ^2/ndf for module segment matching (section 3.3.1).

DistCut (2.): Maximum relative (to track length sum) distance (of centers) for segment matching (section 3.3.1).

ReferencePointAtPca (false) Flag for using the point of closest approach (PCA) as track reference point instead of position of first hit (section 3.4).

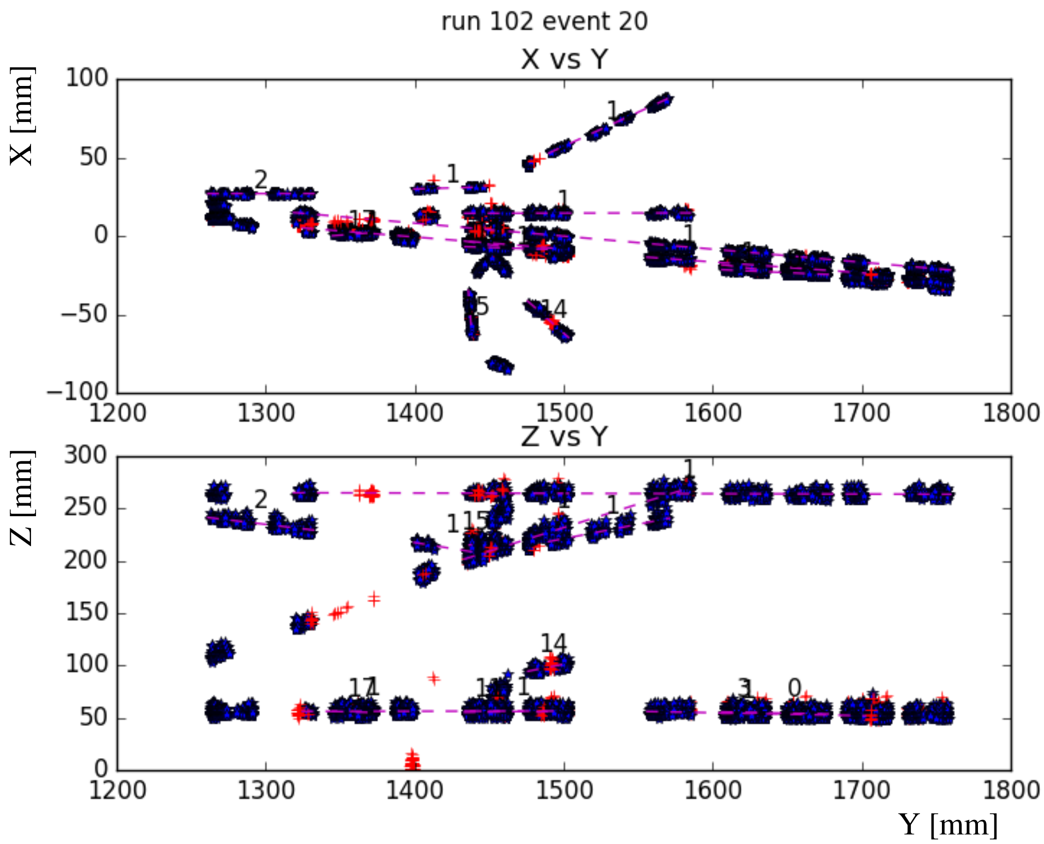

Figure 6: Busy event with tracks in various directions.

6 Summary

A simple and fast local track finding method for GridPix data based on local road search with macro pixels for curved or straight tracks has been presented. It relies on reasonable pixel hit efficiency and estimates of the measurement errors. The method tries to avoid any directional bias (as demonstrated in figure 6). An implementation in MarlinTPC is available as the TIMEPIXLOCALROADSEARCHPROCESSOR.

References

[1] P. Schade, J. Kaminski on behalf of the LCTPC collaboration, A large TPC prototype for a linear collider detector, Proceedings of the 12th International Vienna Conference on Instrumentation, *Nucl. Instr. and Methods A*, 628:128, 2011.

[2] J. Kaminski, Development of a TPC readout based on GridPix, Talk given at the 7. Detector Workshop of the Helmholtz Alliance, Göttingen, March 2014, http://indico.desy.de/conferenceDisplay.py?confId=9389

[3] MarlinTPC homepage http://ilcsoft.desy.de/portal/software_packages/marlintpc/

[4] V. Karimäki, Effective circle fitting for particle trajectories, *Nucl. Instr. and Methods A*, 305:187, 1991.

[5] T. Krämer, Track Parameters in LCIO, LC-DET-2006-004, http://www-flc.desy.de/lcnotes

[6] GEAR homepage http://ilcsoft.desy.de/portal/software_packages/gear/