# Pattern Recognition in a High Rate GEM-TPC

#### Johannes Rauch on behalf of the GEM-TPC Collaboration

Physik Department E18 Technische Universität München Germany

> CHEP 2012 May 24, 2012





Bundesministerium für Bildung und Forschung











### ... At High-Rates

Technische Universität München







• High rates: rate > 
$$\frac{1}{t_{\text{drift}}}$$
,  $t_{\text{drift}} = \mathcal{O}(100 \, \mu \text{s})$ 

- Overlapping events!
- ullet ightarrow TPC acts as continuous analog tracking pipeline







- High rates: rate >  $\frac{1}{t_{drift}}$ ,  $t_{drift} = \mathcal{O}(100 \, \mu s)$
- → Overlapping events!
- $\bullet \ \rightarrow$  TPC acts as continuous analog tracking pipeline







• High rates: rate > 
$$\frac{1}{t_{\text{drift}}}$$
,  $t_{\text{drift}} = \mathcal{O}(100 \, \mu \text{s})$ 

- $\bullet \ \rightarrow \text{Overlapping events!}$
- ullet ightarrow TPC acts as continuous analog tracking pipeline







- High rates: rate >  $\frac{1}{t_{\text{duiff}}}$ ,  $t_{\text{drift}} = \mathcal{O}(100 \, \mu \text{s})$
- → Overlapping events!
- $\rightarrow$  TPC acts as continuous analog tracking pipeline







• High rates: rate > 
$$\frac{1}{t_{\text{drift}}}$$
,  $t_{\text{drift}} = \mathcal{O}(100 \, \mu \text{s})$ 

- Overlapping events!
- ullet ightarrow TPC acts as continuous analog tracking pipeline





### Applications and Requirements



#### Examples:

- Panda-TPC:  $2 \times 10^7 \, \mathrm{s}^{-1} \, \bar{p} p$
- ALICE-upgrade: 50 kHz Pb Pb

Pattern recognition requirements:

- Fast processing (feasible for online reconstruction)
- Robust against drift distortions ( $\epsilon \approx$  4 with GEM amplification)
- Efficient for all kinds of track topologies



### Applications and Requirements



Examples:

- Panda-TPC:  $2 \times 10^7 \, \mathrm{s}^{-1} \, \bar{p}p$
- ALICE-upgrade: 50 kHz Pb Pb

Pattern recognition requirements:

- Fast processing (feasible for online reconstruction)
- Robust against drift distortions ( $\epsilon \approx 4$  with GEM amplification)
- Efficient for all kinds of track topologies











# Signals Combined to Clusters







# Signals Combined to Clusters











6	( )	$\boldsymbol{\Delta}$
~~	$\sim$	LΛ
N	$\sim$	Ň
て	(7	9

Pattern Recognition













- Combines locally adjacent signals to *clusters*.
- Three-dimensional and track independent









- Clusters (hits) are presorted by *z*, radius or angle.
- The first tracklet is started with the first hit.
- Clusters are checked against tracklets with *hit-track correlators*.
- Cuts are dynamically scaled.







- Split tracks due to sectorization, re-entering tracks etc.
- → Tracklet merging
- Tracklets are checked against each other with track-track correlators





#### Circles on the plane

- $\rightarrow$  circles on the sphere
- nonlinear circle-fitting
  - No constrainte on circle
- No constraints on circle parameters!



• Additional straight line fit in  $(\phi, z) \rightarrow$  fast helix fit





- Circles on the plane  $\rightarrow$  circles on the sphere
  - $\rightarrow$  planes in space
- nonlinear circle-fitting
   → linear, fast plane-fitting
- No constraints on circle parameters!



• Additional straight line fit in  $(\phi, z) \rightarrow$  fast helix fit





- Circles on the plane
   → circles on the sphere
  - ightarrow planes in space
- nonlinear circle-fitting
   → linear, fast plane-fitting
- No constraints on circle parameters!



• Additional straight line fit in  $(\phi, z) \rightarrow$  fast helix fit





- Circles on the plane
   → circles on the sphere
  - $\rightarrow$  planes in space
- nonlinear circle-fitting
   → linear, fast plane-fitting
- No constraints on circle parameters!



• Additional straight line fit in  $(\phi, z) \rightarrow$  fast helix fit





- Circles on the plane
   → circles on the sphere

  - ightarrow planes in space
- nonlinear circle-fitting
   → linear, fast plane-fitting
- No constraints on circle parameters!



• Additional straight line fit in  $(\phi, z) \rightarrow$  fast helix fit





- Circles on the plane
   → circles on the sphere
  - → circles on the spher
  - ightarrow planes in space
- nonlinear circle-fitting
   → linear, fast plane-fitting
- No constraints on circle parameters!



• Additional straight line fit in  $(\phi, z) \rightarrow$  fast helix fit





# Sectorization and Multistep Approach









# Sectorization and Multistep Approach





# Pattern Recognition Performance





# PR Performance Studies



- Monte Carlo study for the PANDA-TPC
- $2 \times 10^7 \ \bar{p}p$  interactions per second (background)
- $\rightarrow$  In one drift frame:

One  $\eta_c \rightarrow \phi \phi \rightarrow K^+ K^- K^+ K^-$  event mixed with 2000 background-events

• PR performance without and with drift distortions ( $\epsilon = 4$ )





# PR Performance Studies



- Monte Carlo study for the PANDA-TPC
- $2 \times 10^7 \ \bar{p}p$  interactions per second (background)
- → In one drift frame:
   One η<sub>c</sub> → φφ → K<sup>+</sup>K<sup>-</sup>K<sup>+</sup>K<sup>-</sup> event mixed with 2000 background-events
- PR performance without and with drift distortions ( $\epsilon = 4$ )





# • Purity = number of "correct" clusters in track total number of clusters in the track

• Completeness = number of "correct" clusters in track total number of clusters from the physical track

• A track is "found" if it has a purity and completeness > 50 %

• PR efficiency =  $\frac{\text{number of found tracks}}{\text{number of generated tracks in TPC}}$ 





- Purity = number of "correct" clusters in track total number of clusters in the track
- Completeness = number of "correct" clusters in track total number of clusters from the physical track
- A track is "found" if it has a purity and completeness > 50 %
- PR efficiency =  $\frac{\text{number of found tracks}}{\text{number of generated tracks in TPC}}$



- Purity = number of "correct" clusters in track total number of clusters in the track
- Completeness = number of "correct" clusters in track total number of clusters from the physical track
- $\bullet\,$  A track is "found" if it has a purity and completeness  $>50\,\%\,$
- PR efficiency =  $\frac{\text{number of found tracks}}{\text{number of generated tracks in TPC}}$



- Purity = number of "correct" clusters in track total number of clusters in the track
- Completeness = number of "correct" clusters in track total number of clusters from the physical track
- A track is "found" if it has a purity and completeness > 50 %
- PR efficiency =  $\frac{\text{number of found tracks}}{\text{number of generated tracks in TPC}}$

Single Events,  $\eta_c \rightarrow \phi \phi \rightarrow K^+ K^- K^+ K^-$ 



Efficiency

Purity

#### Completeness





Technische Universität München


### 2000 Mixed Events $\bar{p}p$ , no Distortions





### 2000 Mixed Events $\bar{p}p$ , with Distortions







### **Computing Time Requirements**



- Pattern recognition computing time on a standard 3.1 GHz office PC, one thread, fully mixed event  $\approx$  3:40 minutes
- $\approx$  4000 tracks
- $\rightarrow$  computing time per track  $\approx 55 \ \mu s$

## Event Deconvolution Performance $\eta_c \rightarrow \phi \phi \rightarrow K^+ K^- K^+ K^-$







• Tracks from a physics event ( $\eta_c$ ) are mixed with  $\bar{p}p$  background events in the TPC.



Technische Universität München



- The point of closest approach (POCA) of each track to the interaction point (IP) is calculated.
- This calculation is done with the helix parameters from the PR.
- No material effects or distortion corrections are taken into account.







- A fiducial volume around the interaction point is made up.
- Tracks with their POCAs to the IP outside the fiducial volume are rejected.



Technische Universität Müncher



- Tracks with their POCA close enough to the IP survive.
- With this technique, the amount of background tracks can be reduced by almost one order of magnitude.
- In this step it is more important to retain all physics tracks than to reject all background tracks.



### Correlation with the MVD and Event Selection

- For the remaining tracks, (ideal) drift distortion corrections are applied.
- The tracks are fitted and correlated with the PANDA-MVD.
- 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.
- The invariant masses of all oppositely charged pairs are calculated.
  → A clear φ peak shows up (m<sub>φ</sub> = (1019.445 ± 0.020) MeV c<sup>-1</sup>).



Technische Universität Müncher





### Correlation with the MVD and Event Selection

- Events with at least two  $\phi$  candidates in a mass window of  $\pm 30~{\rm MeV}~c^{-1}$  around the nominal mass are selected.
- The invariant masses of two  $\phi$  candidates are calculated.

 $\rightarrow$  A clear  $\eta_c$  peak shows up  $(m_{\eta_c} = (547.853 \pm 0.024) \, {\rm MeV} \, c^{-1}$  has been subtracted).





Conclusion



- 3-dimensional clustering and pattern-recognition algorithms for a high-rate GEM-TPC
- Efficient at high track densities
- Finds all kinds of track topologies.
- Robust against drift distortions.
- Excellent seed values for event-deconvolution and track-fitting
- Event-deconvolution feasible
- The algorithms presented are in use for the reconstruction of data taken with the GEM-TPC prototype installed in the FOPI spectrometer at GSI, Germany





Conclusion



- 3-dimensional clustering and pattern-recognition algorithms for a high-rate GEM-TPC
- Efficient at high track densities
- Finds all kinds of track topologies.
- Robust against drift distortions.
- Excellent seed values for event-deconvolution and track-fitting
- Event-deconvolution feasible
- The algorithms presented are in use for the reconstruction of data taken with the GEM-TPC prototype installed in the FOPI spectrometer at GSI, Germany



## Thanks for your attention!













Counts

#### **DPM Momenta**



Primary tracks, Secondary tracks.

ПП

Technische Universität München





#### Track topology:

- E || B field configuration (2T solenoid B field)
  - $\longrightarrow$  Helical tracks inside the TPC volume



- Two independent projections
- 3 (circle) + 2 (sine curve) = 5 parameters





Track topology:

● E || B field configuration (2T solenoid B field)
 → Helical tracks inside the TPC volume



- Two independent projections
- 3 (circle) + 2 (sine curve) = 5 parameters

$\bigotimes$	The	Detection	Strategy
--------------	-----	-----------	----------





## 🖇 The Hough Transform

Task:

- Inside some data-set, find a pattern with a given **parametrization**  $P(p_1, ..., p_N)$ General concept:
  - Transform each data point into the parameter ("Hough"-) space {p<sub>i</sub>}
  - Points actually lying on the pattern looked for will create a maximum
  - Transform search for a pattern into a search for a maximum



Technische Universität Müncher



H. Li, M. A. Lavin, R. J. Le Master: "Fast Hough Transform: ...", 1986



- ... is not trivial:
  - Conventional histogramming methods fail (memory consumption!)

#### Instead: trade memory consumption for processing load

- Subdivide the parameter space into subvolumes "nodes"
- Perform an iterative tree search
- Example of 5 simulated tracks:





H. Li, M. A. Lavin, R. J. Le Master: "Fast Hough Transform: ...", 1986



- ... is not trivial:
  - Conventional histogramming methods fail (memory consumption!)

#### Instead: trade memory consumption for processing load

- Subdivide the parameter space into subvolumes "nodes"
- Perform an iterative tree search
- Example of 5 simulated tracks:





H. Li, M. A. Lavin, R. J. Le Master: "Fast Hough Transform: ...", 1986



- ... is not trivial:
  - Conventional histogramming methods fail (memory consumption!)

#### Instead: trade memory consumption for processing load

- Subdivide the parameter space into subvolumes "nodes"
- Perform an iterative tree search
- Example of 5 simulated tracks:





H. Li, M. A. Lavin, R. J. Le Master: "Fast Hough Transform: ...", 1986



- ... is not trivial:
  - Conventional histogramming methods fail (memory consumption!)

#### Instead: trade memory consumption for processing load

- Subdivide the parameter space into subvolumes "nodes"
- Perform an iterative tree search
- Example of 5 simulated tracks:





H. Li, M. A. Lavin, R. J. Le Master: "Fast Hough Transform: ...", 1986



- ... is not trivial:
  - Conventional histogramming methods fail (memory consumption!)

#### Instead: trade memory consumption for processing load

- Subdivide the parameter space into subvolumes "nodes"
- Perform an iterative tree search
- Example of 5 simulated tracks:





H. Li, M. A. Lavin, R. J. Le Master: "Fast Hough Transform: ...", 1986



- ... is not trivial:
  - Conventional histogramming methods fail (memory consumption!)

#### Instead: trade memory consumption for processing load

- Subdivide the parameter space into subvolumes "nodes"
- Perform an iterative tree search
- Example of 5 simulated tracks:







### Implementation on a GPU using CUDA<sup>TM</sup>



- Massively parallel implementation
- Each node (parameter sub-volume) is connected to one thread
- $\mathcal{O}(10000)$  threads run in parallel  $\rightarrow$  hide memory latencies
- Optimal hardware mapping to the problem: 32 threads in a *warp*, each sub-volume has  $2^{DIM} = 32$  daughter volumes (DIM = 5)



Example event (right half passed to pattern recognition)

## The GEM-TPC

### The Time Projection Chamber (TPC)



- A TPC is a gaseous detector for three-dimensional tracking of charged particles.
- It combines a drift chamber and a two-dimensional measurement device on the anode side of the drift volume.
- Multi Wire Proportional Chambers (MWPCs) or Gas Electron Multipliers (GEMs) are used for amplification.

### Gas Amplification: MWPCs vs. GEMs

#### MWPCs

- No intrinsic ion-backflow suppression.
- Have to be operated in gated mode to avoid space charge buildup.
- Gating limits the maximum event rate that can be read out to  $\mathcal{O}(100 \, \mathrm{s}^{-1})$ .

#### GEMs

- Intrinsic ion-backflow suppression.
- TPC can be operated and read out *continuously*.
- Suited for high event rates and track densities.











### **Technical Drawings**

Technische Universität München





Technische Universität München



- A voltage-difference of (300 to 400) V is applied between the Cu layers.
  → Large electric fields in the holes of O(50 kV cm<sup>-1</sup>).
- Incoming primary electrons are guided into the holes, where gas amplification occurs.
- Most of the ions flow back onto the top of the GEM, and the electrons are extracted.



### Pulse Shape Analysis (PSA)



- The PSA combines consecutive samples from a readout-pad to *signals*.
- Signals are defined by their time and amplitude.
- The PSA finds local maxima.
- The signal amplitude is determined by the amplitude of the maximum sample.
- The signal time is defined by the time of the maximum sample minus the rise time of the signal-shaper.









- All signals are sorted by decreasing amplitude.
- The algorithm loops over the signals, each signal is checked against all clusters.
- A cluster matches a signal if:
  - The sinal lies on a pad or is an immediate neighbor to a pad that is already in the cluster.
  - The signal is close enough in *z*-direction, e.g. it lies within a certain *time slice* around the *z*-component of the center of gravity of the cluster.
- If no matching cluster is found, a new cluster is created.
- If exactly one cluster matches, the signal is assigned to it.
- If more than one cluster matches, the signal is split between the clusters.





- The cluster amplitude is the sum of the signal amplitudes.
- The cluster position is the center of gravity of the signal positions.
- The error of the custer position is estimated from the amplitude weighted standard deviation of the signal positions.



# Event Mixing and Event Deconvolution at High Interaction Rates




# Event Mixing in a High-Rate Environment



- In PANDA, event rates of up to 2 × 10<sup>7</sup> p
  p-interactions per second (i.e. the *full rate*) are foreseen.
- At this high rates, tracks from several thousands of events overlap or mix in the TPC.
- The event time of these tracks is not known a priori!





• The TPC is perfectly suited for a high-rate experiment without specialized trigger hardware.

۲

- The TPC could for example find displaced secondary vertices or decays in the TPC-volume.
- If an event has been found by other detectors, TPC tracks that belong to that event have to be filtered out from the background.
- Two techniques were used to retain the tracks from an event with known event time:
  - Target pointing.
  - Correlations with other detectors.





#### Correlations with other Detectors



- To further reject background tracks, TPC tracks can be correlated with information from other detectors.
- In the case of PANDA, the Micro Vertex Detector (MVD) and the forward GEM-trackers can be used.





### Correlation with the MVD





- Tracks from different events are mixed in the TPC.
- The MVD has a time resolution of around 5 ns.

 $\rightarrow$  For the tracks, where the event time  $t_0$  is known, there will also be hits in the MVD at this time.



### Correlation with the MVD





• All tracks in the TPC are extrapolated into the MVD.



### Correlation with the MVD





Only tracks with a minimum number of MVD hits within a certain road-width around the TPC tracks survive.



## Space Charge Distortions



 Ions from primary ionizations and ion-backflow from the GEMs create a space charge.

 $\rightarrow$  Position-dependent distortions of the drift paths of the primary electrons.





# Space Charge Distortions

- Positions and reconstructed momenta for undistorted (black), distorted (red) and corrected (green) tracks.
- Distortions can be corrected.
- However, distortion correction requires knowledge of z-position of the track.
   → Correction can only be done after event deconvolution.

 $\rightarrow$  The PR has to work for distorted tracks!



Technische Universität Müncher