Reconstruction Methods at DESY

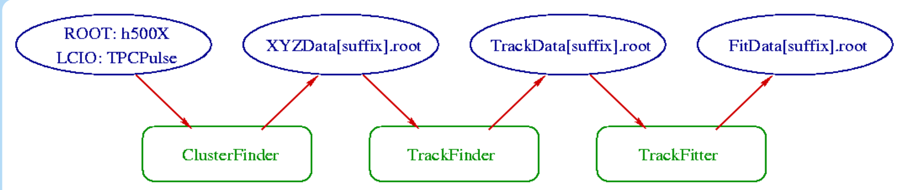
Ralf Diener

- **Outline:**
  - **Pad Response Correction: Performance?**
  - **Number of Rows: Resolution and Stability?**
  - **Double Track Resolution:**
    Status of ongoing work

bmb+f - Förderschwerpunkt
Elementarteilchenphysik
Großgeräte der physikalischen Grundlagenforschung

MultiFit

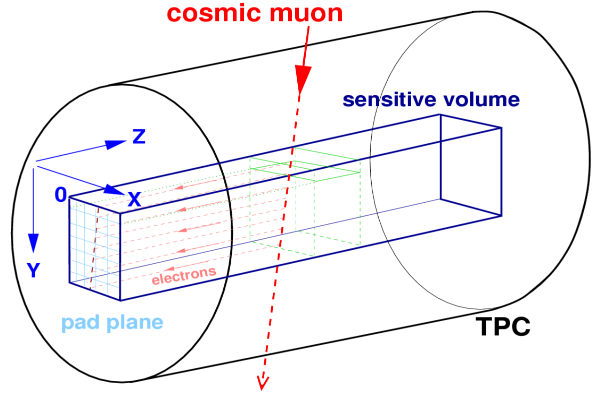
- **Reconstr. Software MultiFit**
  - 3 Step process: Hit Reconstruction
    → Track Finding
    → Track Fitting



MultiFit Module

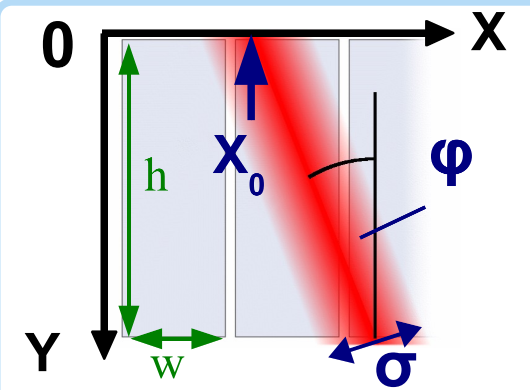
- 2 Track Fit Methods implemented:
  (both for straight line and circular arc track hypothesis)
  - Chi Squared Method:
    fits track hyphothesis to reconstructed hits
    ← Pad Response Correction (PRC) implemented
    in hit reconstruction
  - Global Fit Method(*):
    fits track hyphothesis to measured pulses (signals on the pads) → built-in PRC

(*): *"TPC Performance in Magnetic Fields with GEM and Pad Readout"*
D. Karlen, P. Poffenberger, G. Rosenbaum Nucl.Instrum.Meth. A555 (2005) 80-92



Koordinaten-System

- Fit results: Intercept X_0, Slope X,
  Circ. Arc: Curvature,
  Global Fit: Width σ (can be fixed per track and row during fit depending on z; calculated from D and σ_0)

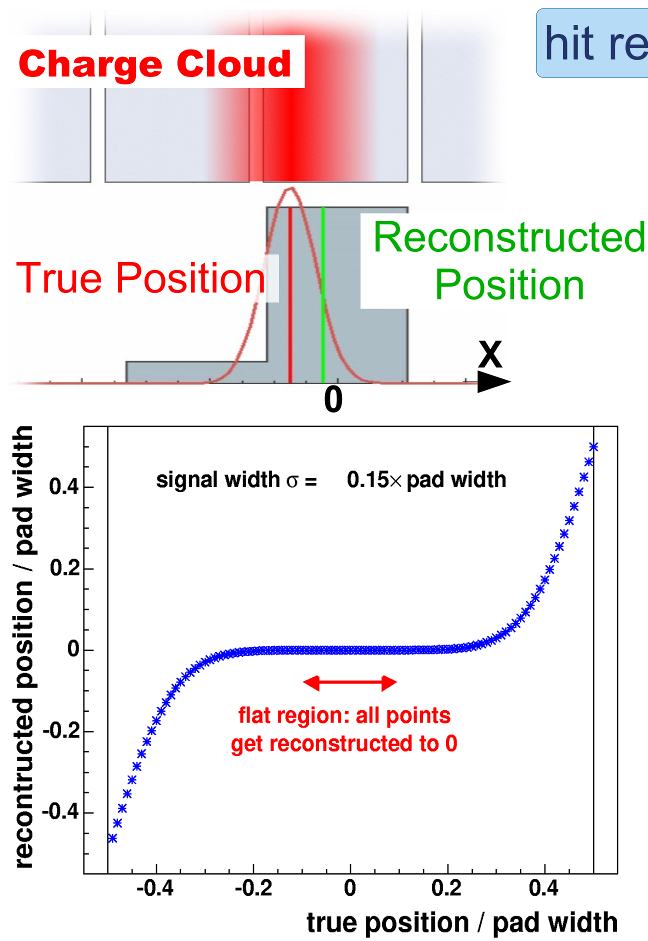


Spurparameter

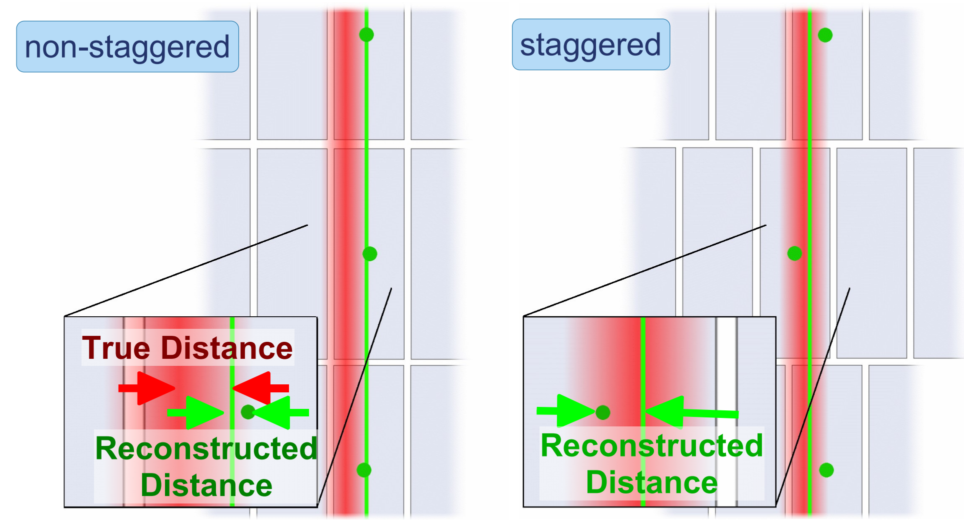
Pad Response Function

- Chi Squared fit works on hit coordinates → reliable hit reconstruction necessary!
- Pad Response Function (PRF):
  not enough charge sharing → Center of Gravity method reconstructs hit towards the pad with the highest signal

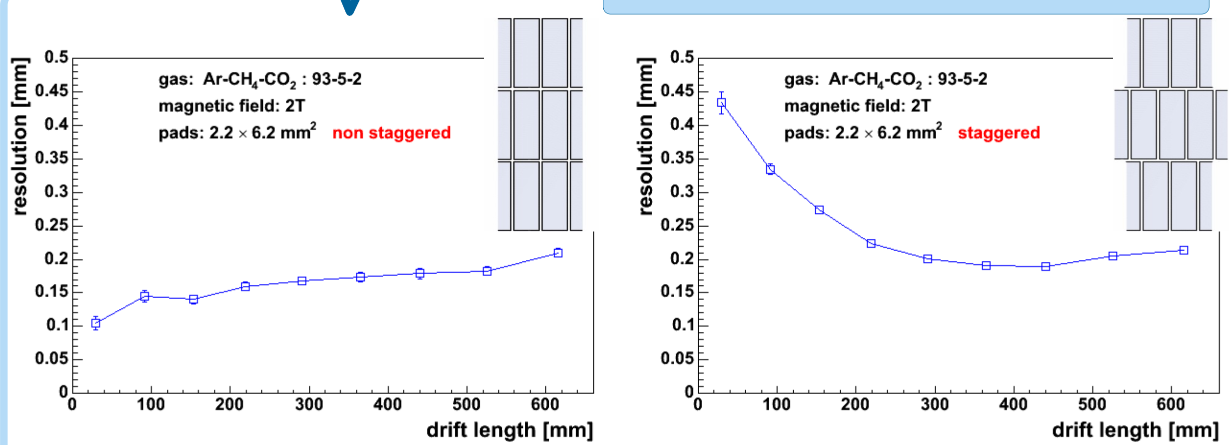
hit reconstruction



effects of the pad response



influence on the resolution



PRC Implementation in MultiFit

- Pad Response Function (Gaussian charge cloud)

$$Q_{pad}(y) = \int_{-\infty}^{+\infty} \left(\Theta\left(\psi - \frac{\Delta}{2}\right) * \Theta\left(-\psi + \frac{\Delta}{2}\right) \right) \times \left(\frac{Q_{max}}{\sqrt{2\pi}\,\sigma_S} * \exp\left[\frac{-(y-\psi)^2}{2\sigma_S^2} \right] \right) d\psi$$

- Pad Response Correction

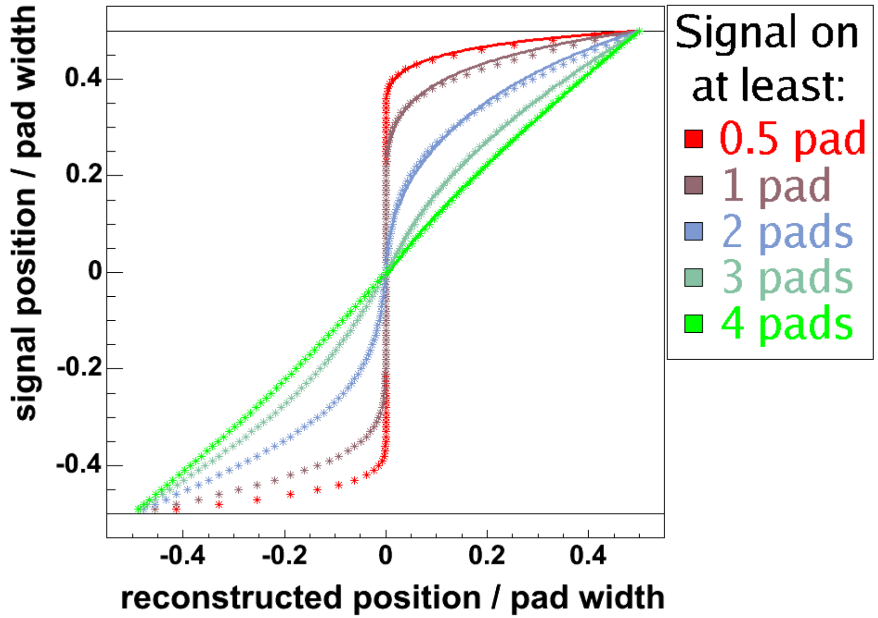
$$F_{noflat} = P_1 x + P_2 \sqrt{x} + \left(\frac{1-P_1}{2} - \frac{P_2}{\sqrt{2}} \right) \cdot \sqrt[3]{2x}$$

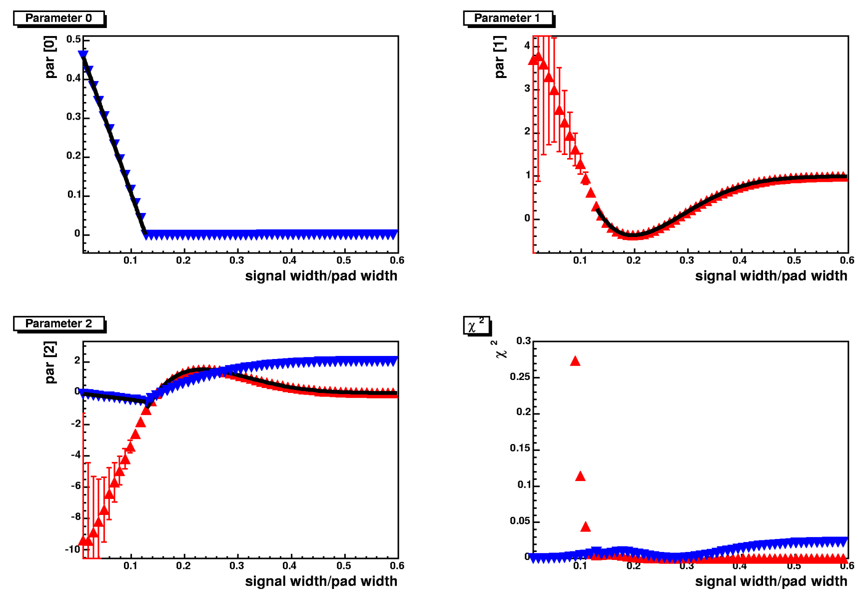
$$F_{flat} = P_0 x + P_2 \sqrt{x} + \left(\frac{1-2P_0}{2} - \frac{P_2}{\sqrt{2}} \right) \cdot \sqrt[3]{2x}$$

- Parameters: dependent on width σ

$$P_0 = a_{01}\left((1-\sigma)/(a_{00}) \right)$$
$$P_1 = a_{15}\sigma^5 + a_{14}\sigma^4 + a_{13}\sigma^3 + a_{12}\sigma^2 + a_{11}\sigma + a_{10}$$
$$P_{2,flat} = a_{26}\sigma$$
$$P_{2,noflat} = a_{25}\sigma^5 + a_{24}\sigma^4 + a_{23}\sigma^3 + a_{22}\sigma^2 + a_{21}\sigma + a_{20}$$

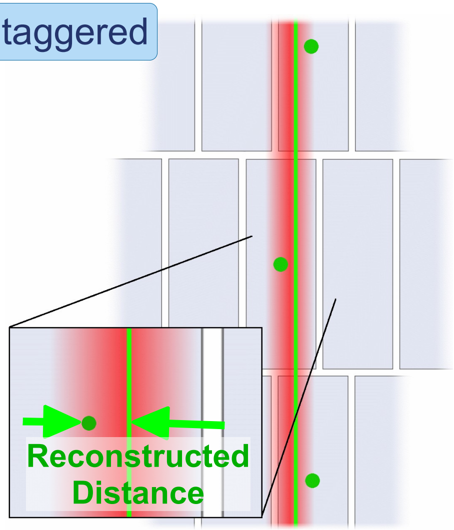
needed input for reconstruction: diffusion / defocussing coefficients



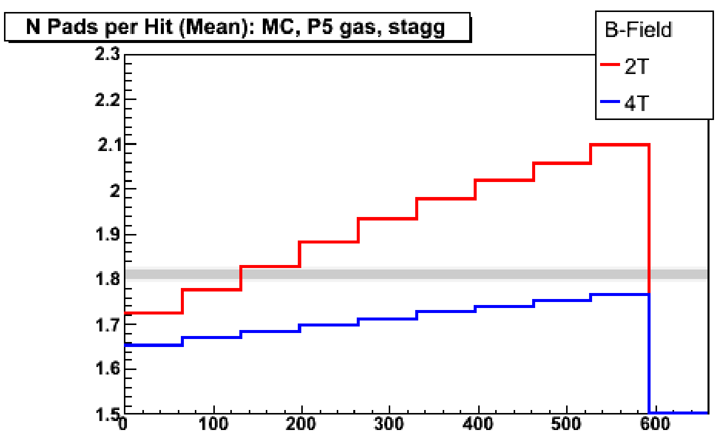


Effect of PRC: Hit Position (Monte Carlo)

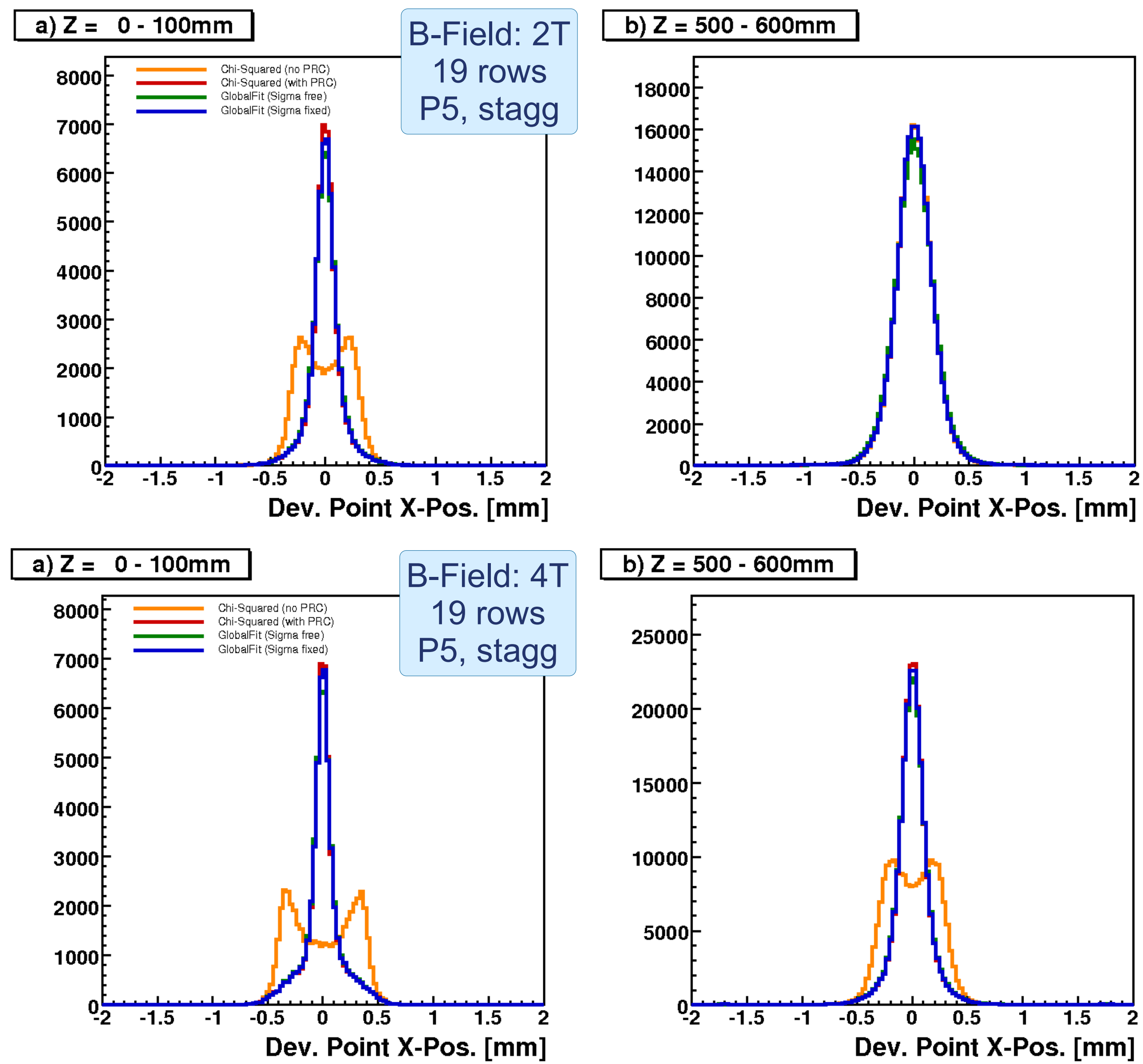
staggered



Hit Width

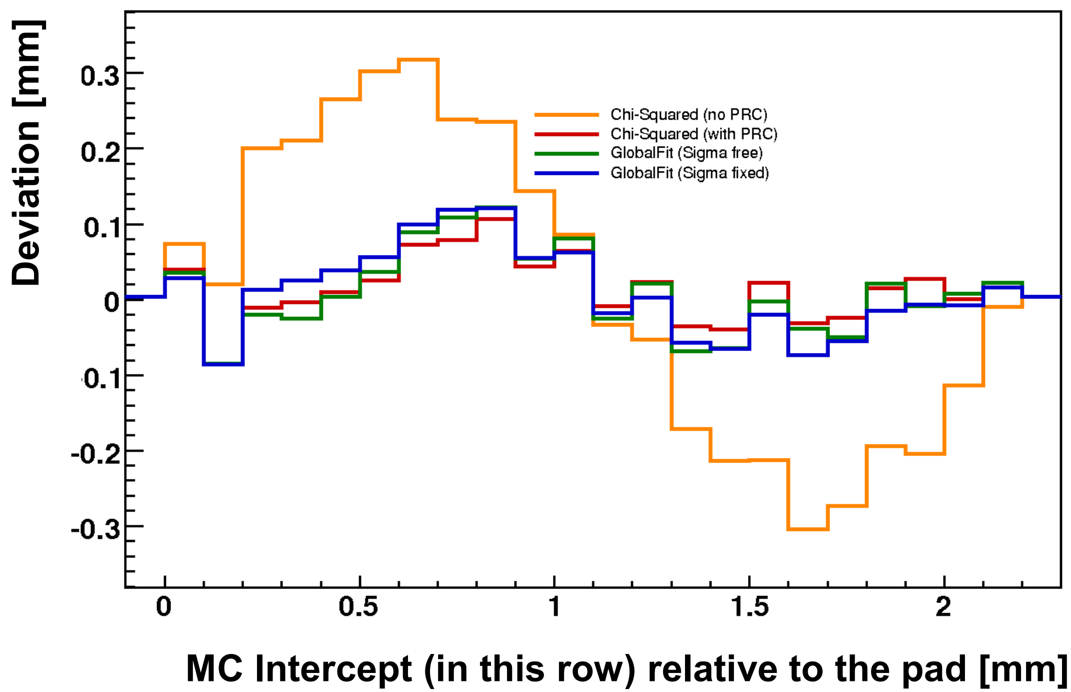


deviation: hit position ↔ MC track



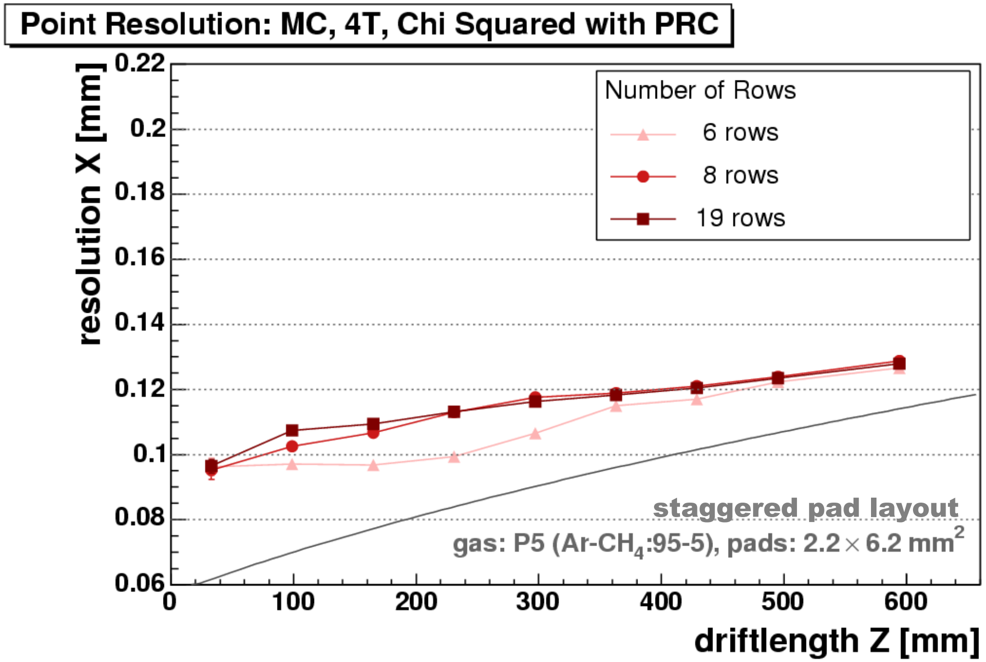
Distance of Hits from the MC Track

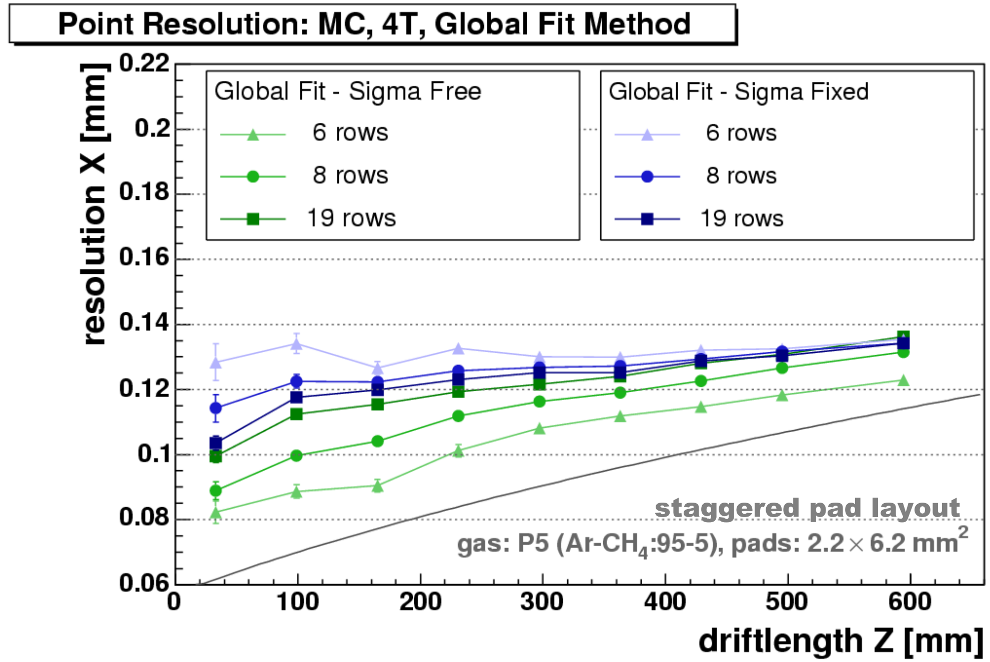
- Mean distance of the reconstructed hits from the Monte Carlo track truth in dependence of the position of the track relative to the pad :



- The track angle is not taken into account (would alter the charge distribution on the pads)
- The correction happens during the hit reconstruction (where the track parameters are unknown) → How to take it into account?:
  - Iterative process: PRC – Track Fit – PRC (... more than once?)

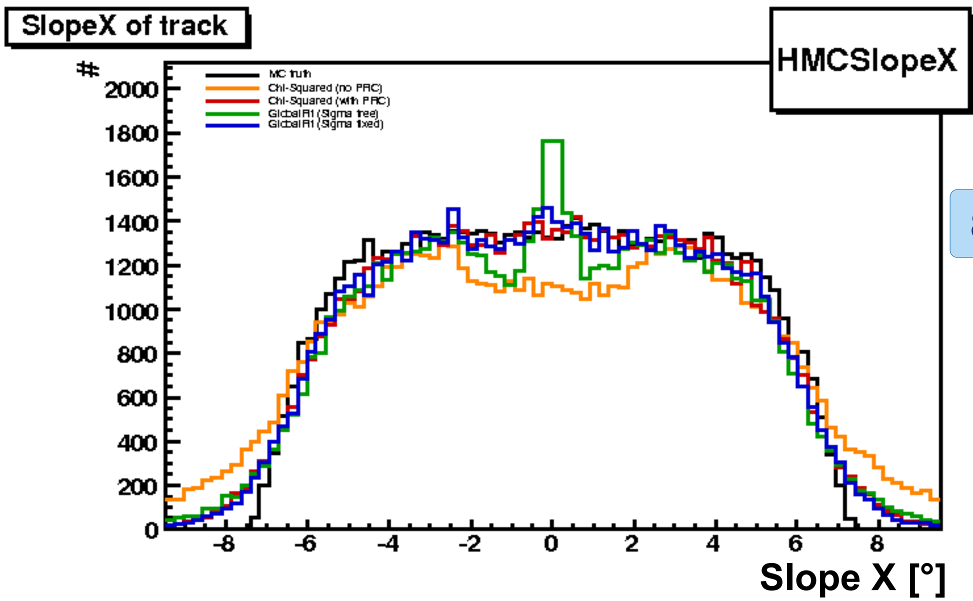
Influence of the Number of Rows

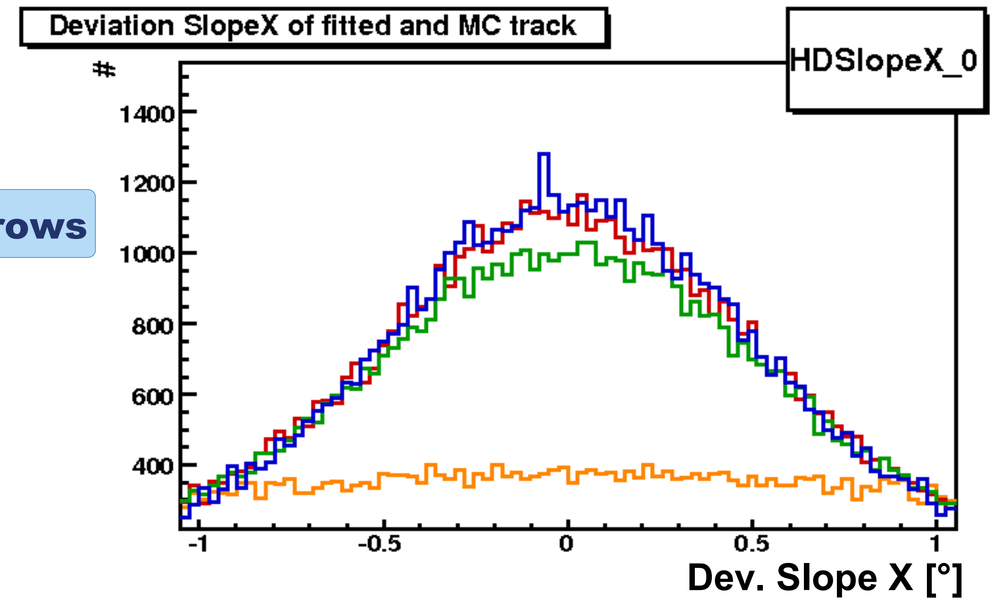




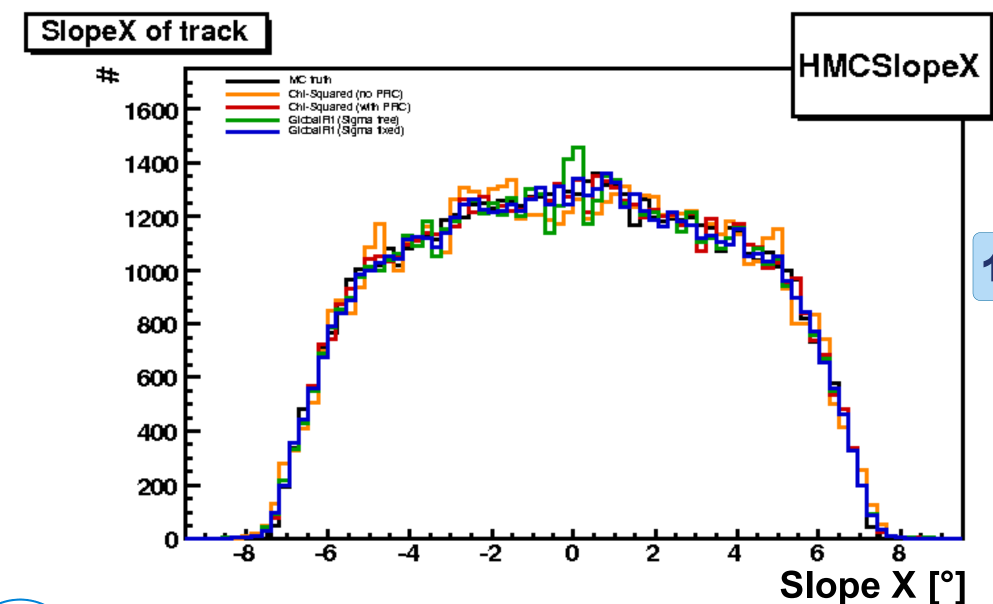
- Chi Squared Method:
  - 6 rows in comparison too good
  - 8 rows already reasonable
  - 19 rows results show expected shape and are comparable with Global Fit results for 19 rows
- Global Fit with free σ:
  - 6 rows unreasonably good
  - 8 and 19 rows tend to more reasonable results
- Global Fit with fixed σ:
  - results conservative and scale with increasing number of rows
- Both flavors comparable at 19 rows

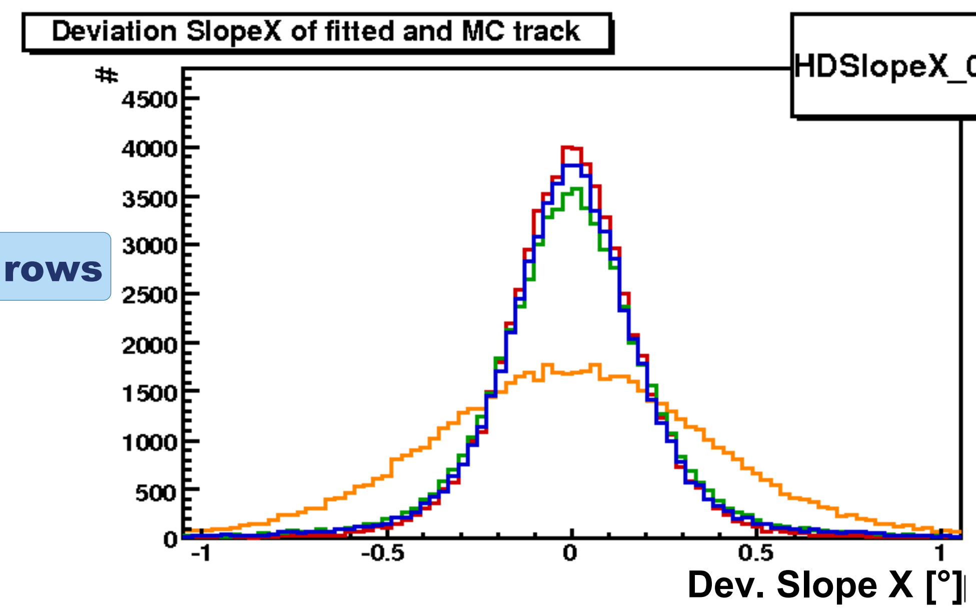
Number of Rows: Track Fit – Slope (Angle)





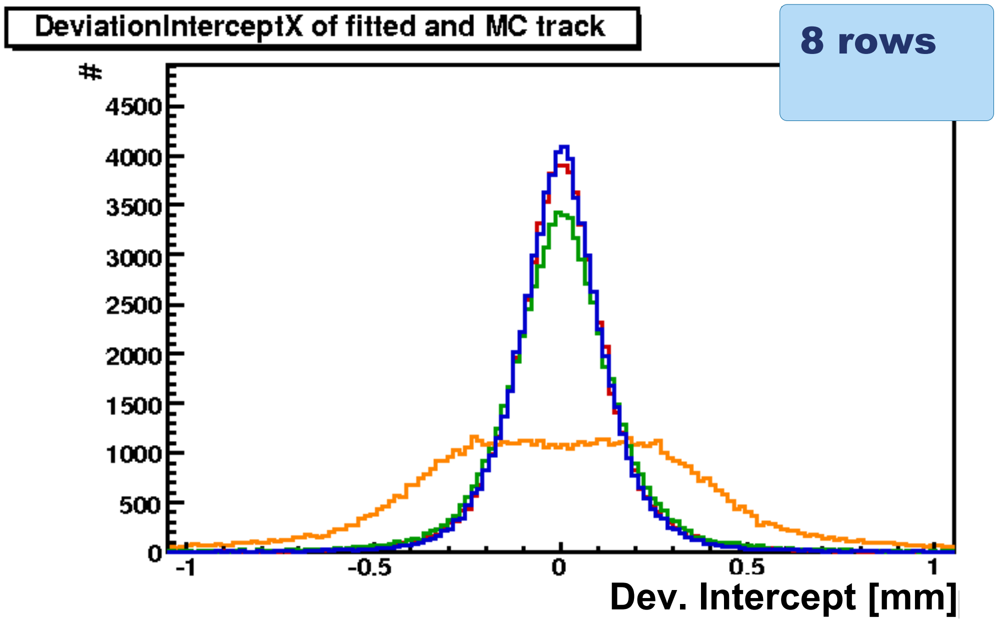
8 rows

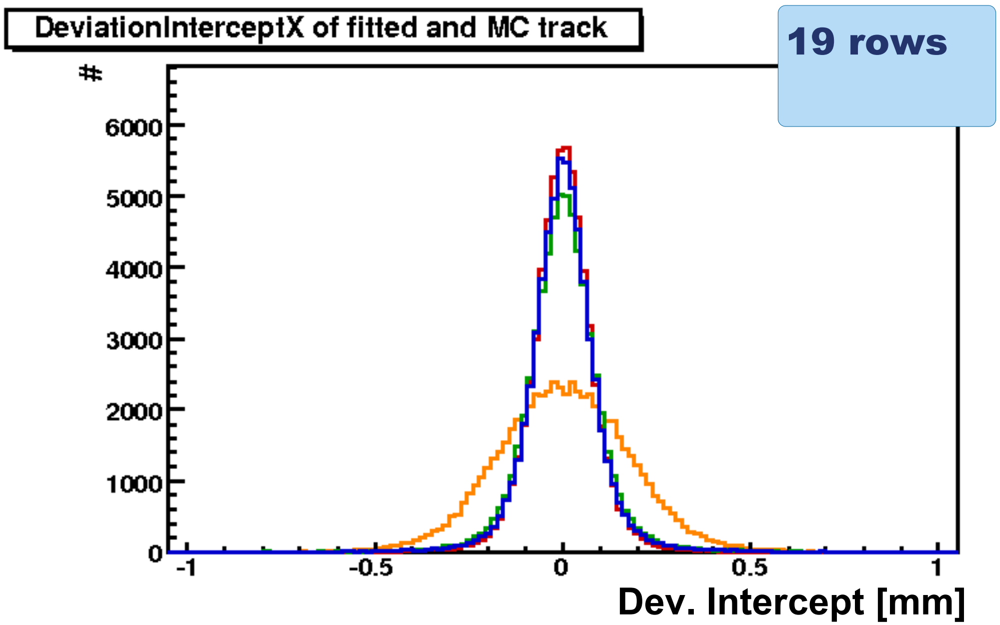




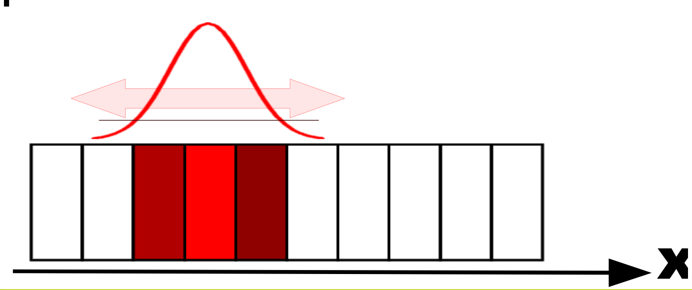
19 rows

Number of Rows: Track Fit - Intercept

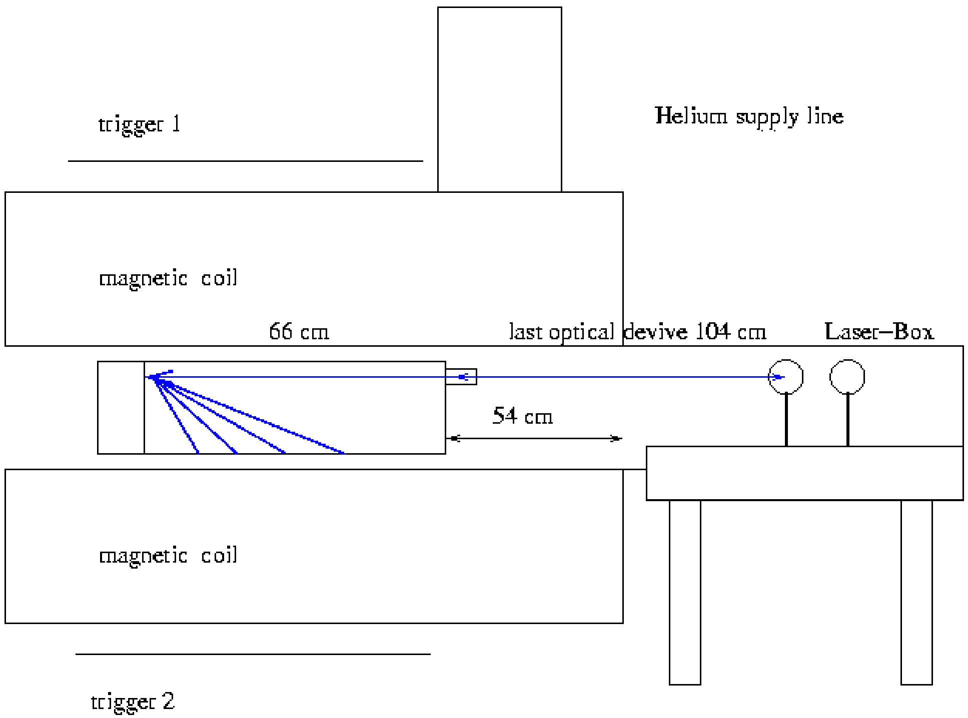




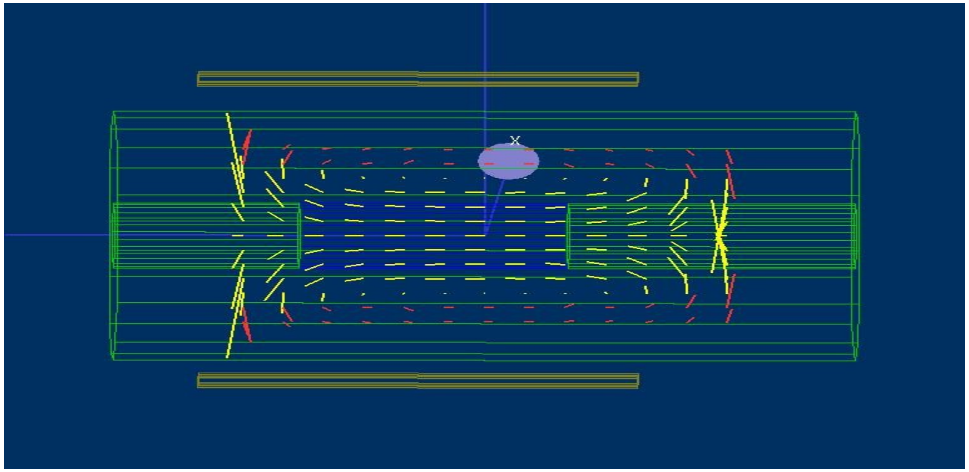
- Increased number of rows make fit methods more stable (as expected)
- Influence on the resolution probably dominated by the hit position reconstruction; difference between methods:
  - Chi Squared: hit reconstruction independent of the track fit
  - Global Method: hit reconstruction influenced by quality of track fit

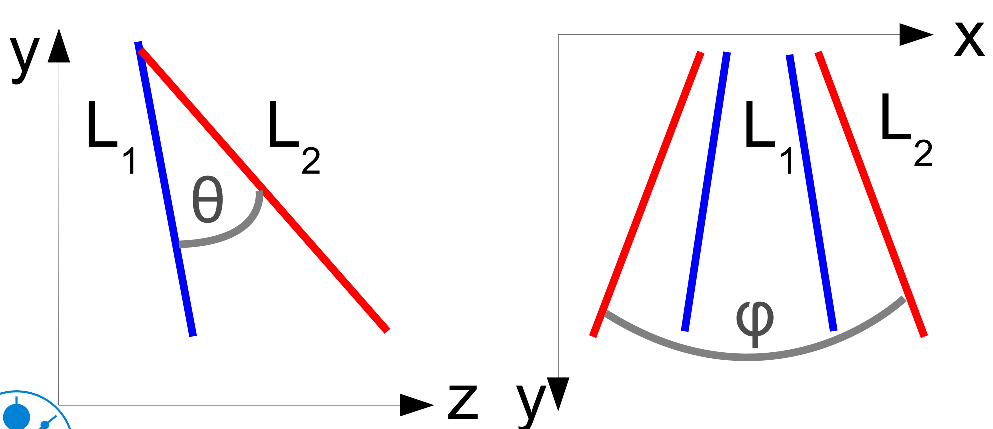


Double Track Resolution



Simulation of the Magnet for 1 T (P. Krstonosic)

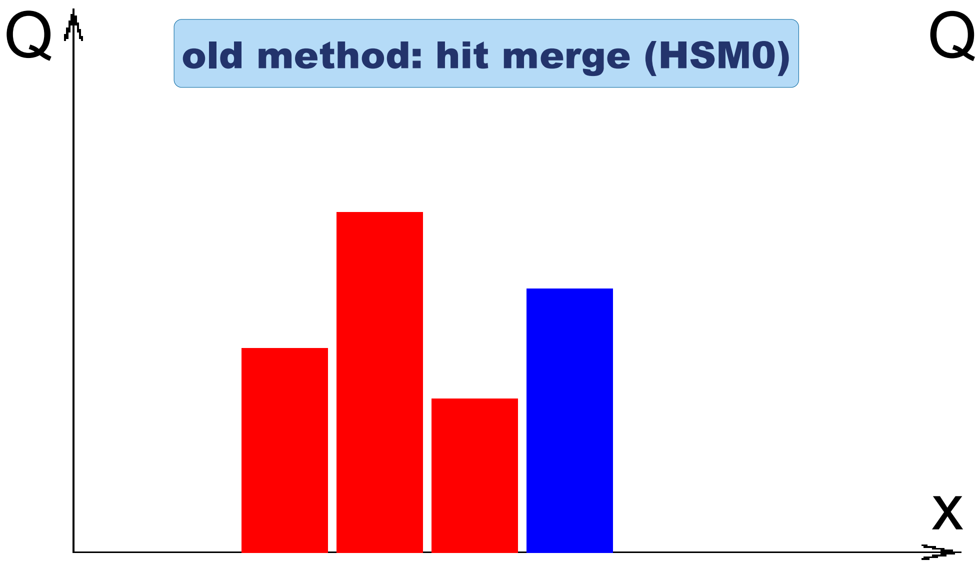


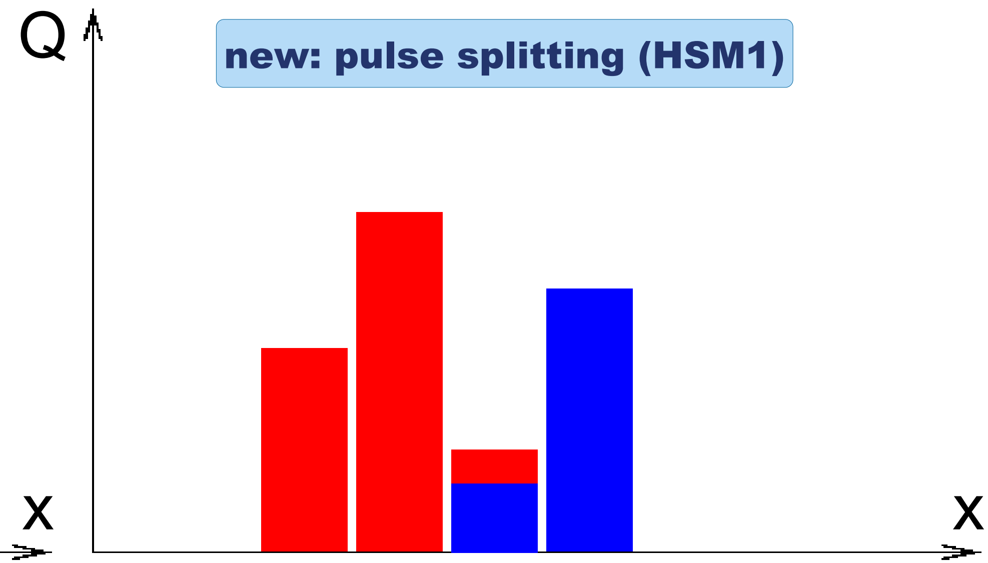


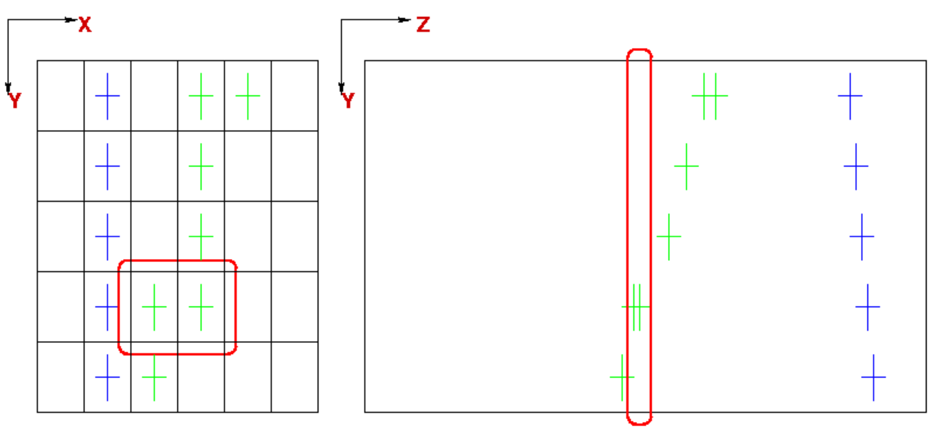
$\Delta x \approx \alpha \cdot L$

different inclination of the θ-angle lead to different separation in the sensitive volume

Double-Track Reconstruction: Hit Separation



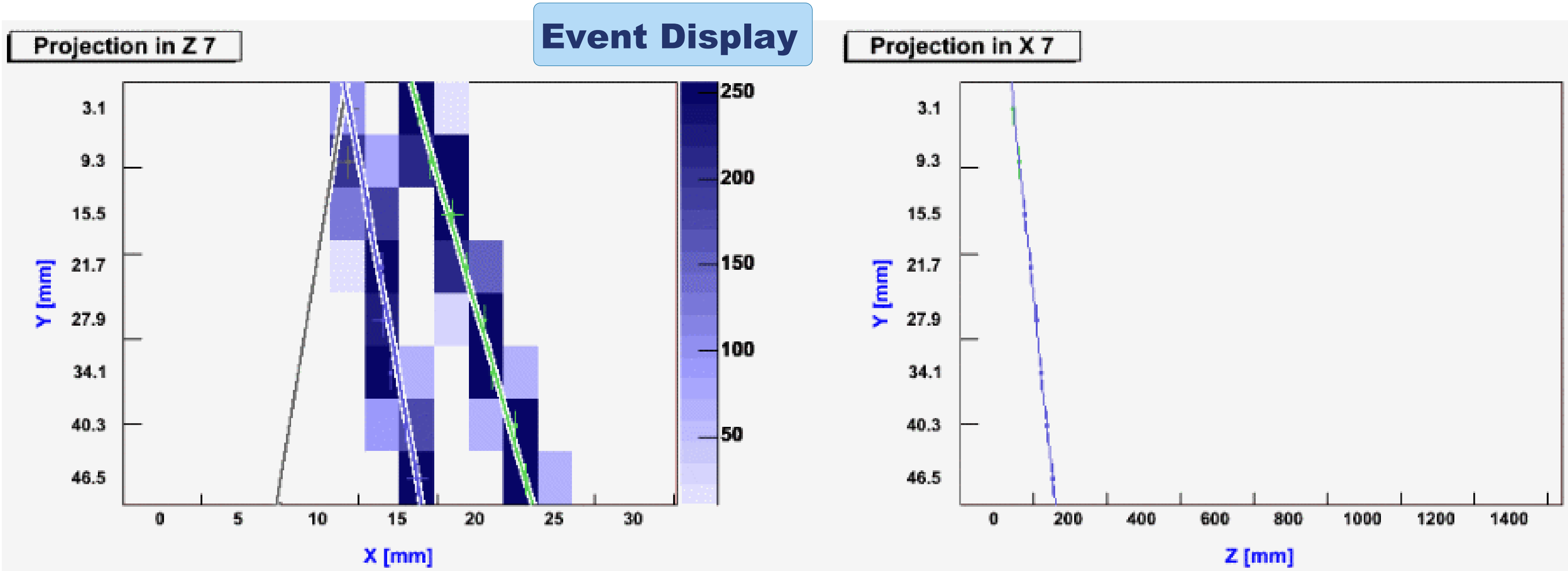


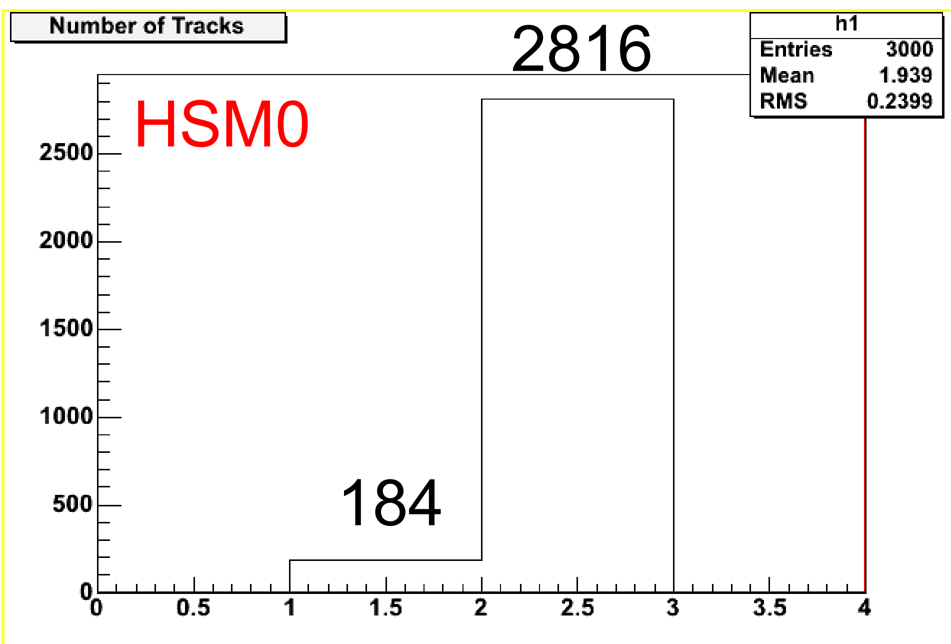


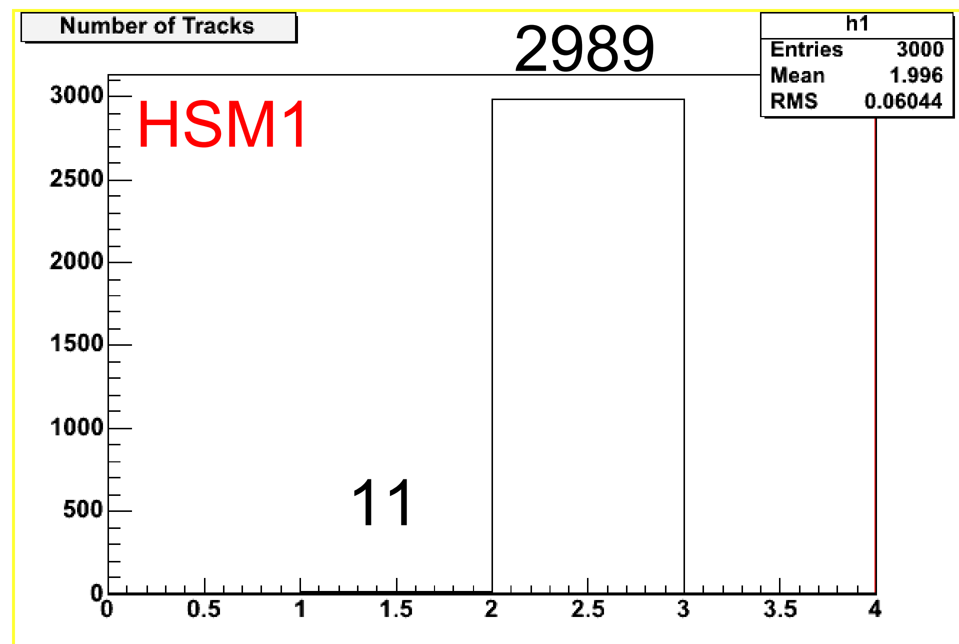
- Find the first local minimum by calculating the derivative of the pulses
- Calculate a charge depending weight
  $x = Q_1 / (Q_1+Q_2)$
  resp. $1\text{-}x = Q_2 / (Q_1+Q_2)$, default value $x = 1$
- Produce two Pulse collections: first store pulses before minimum + minimum, second erase pulses before minimum (to repeat the method)

Performance Hit Separation: Monte Carlo

Event Display

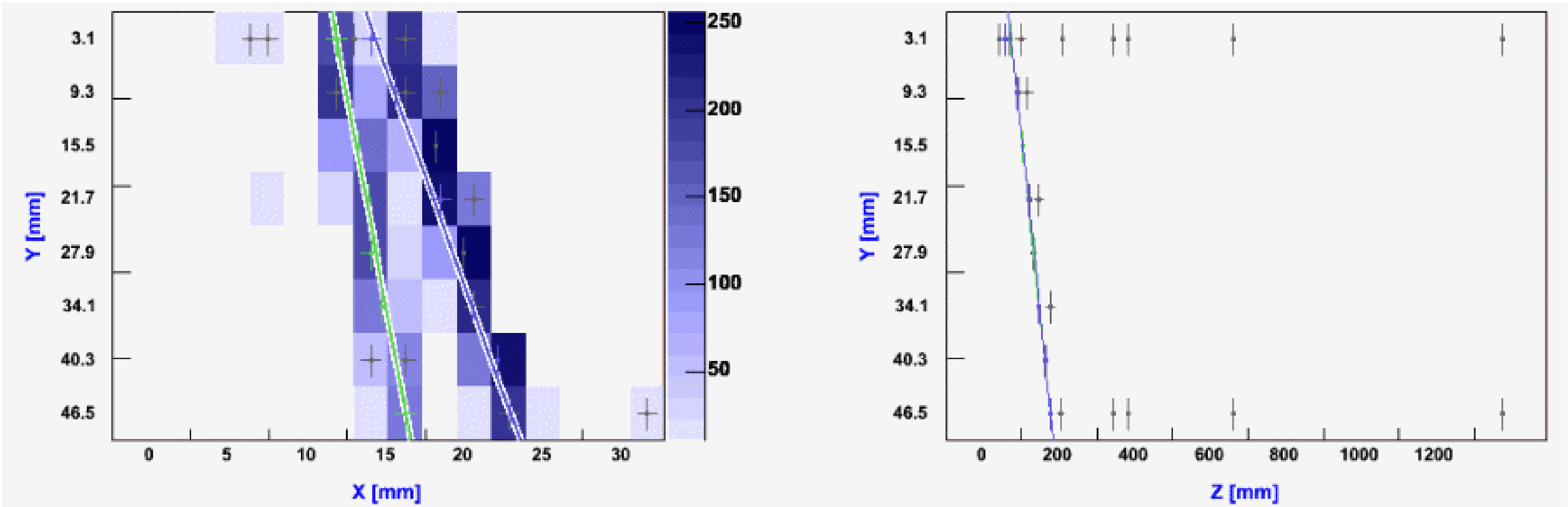


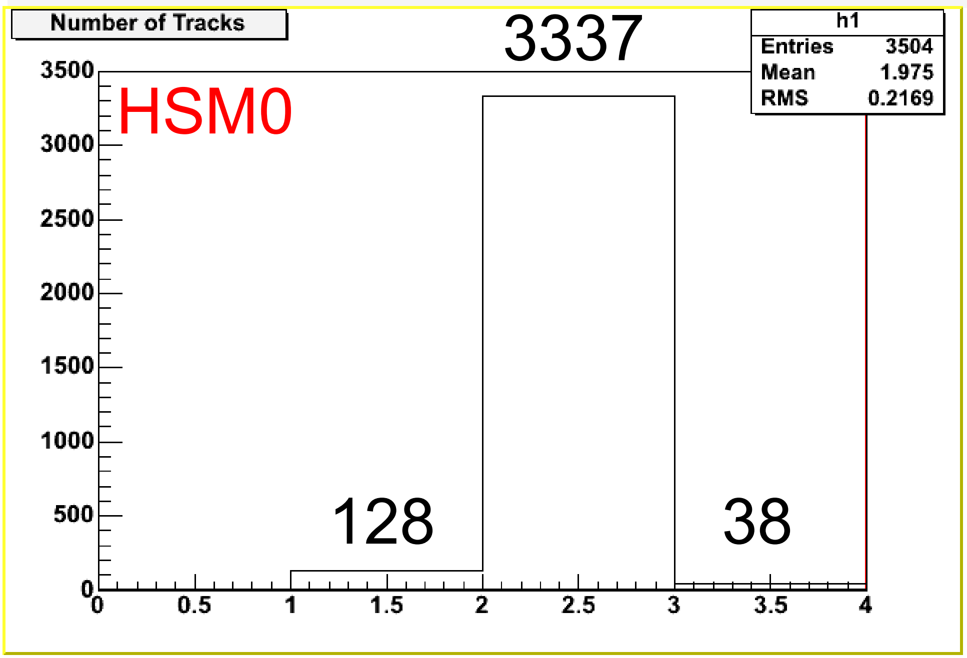


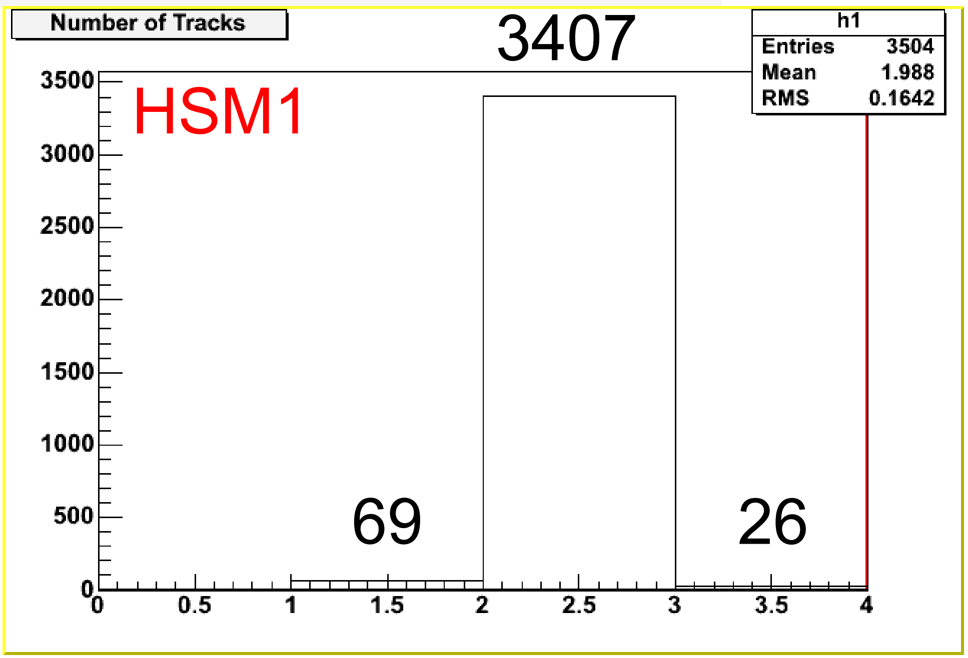


Performance Hit Separation: Measurement

Event Display

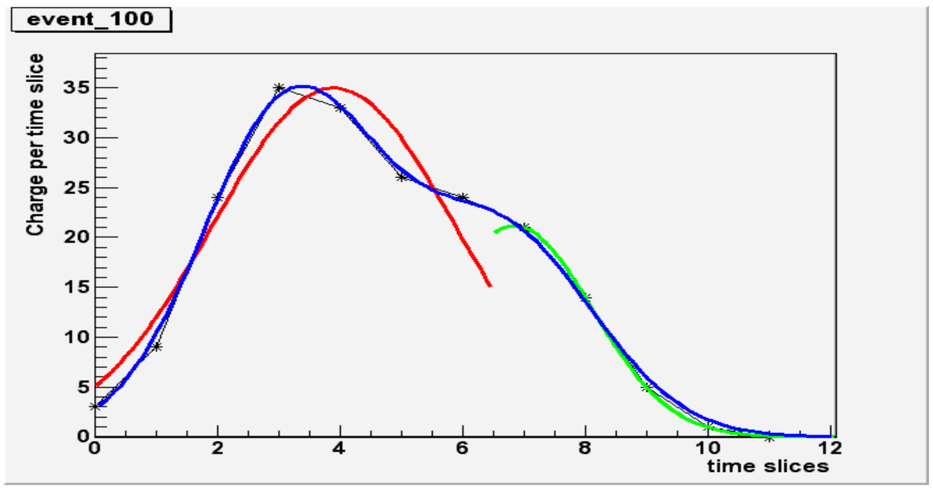


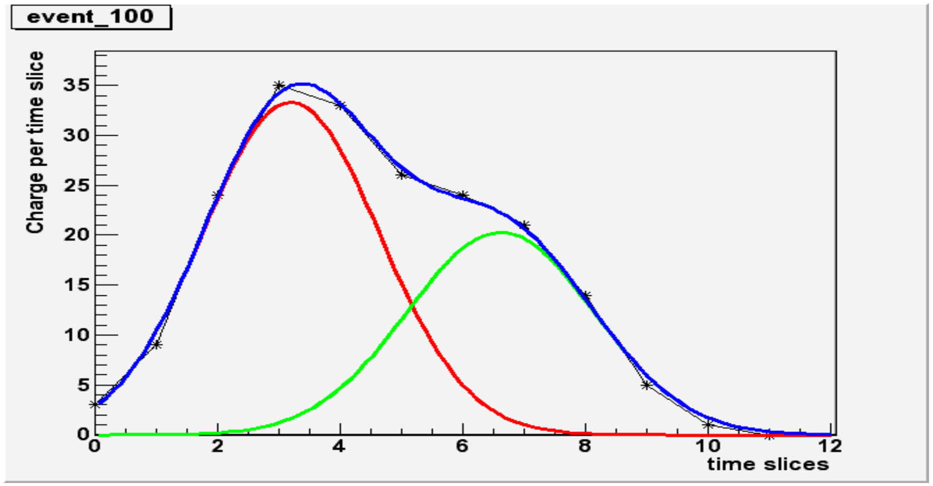




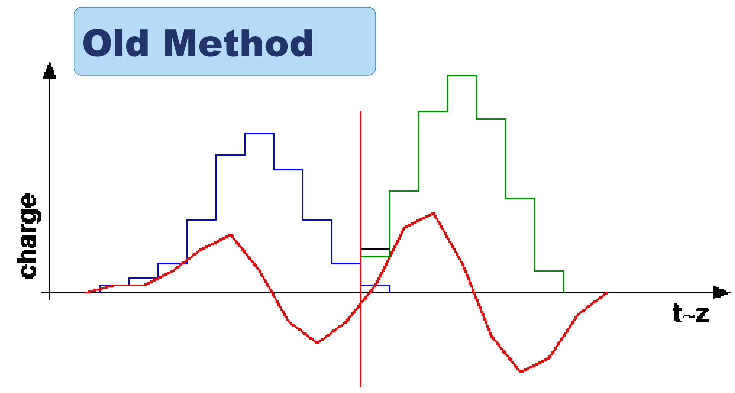
Double Track Separation: Status

- Separation of pulses in z: implementation of new method based on double-Gaussian fit in progress





- Fit two single gauss functions from → and ← to get start values for the mean$_{1,2}$ and const$_{1,2}$ ($\sigma_{1,2}$ is fixed value)
- In a second step fit a double Gaussian function with the start parameters from the first step
- Calculate from the mean$_{1,2}$ values of this function the z-separation



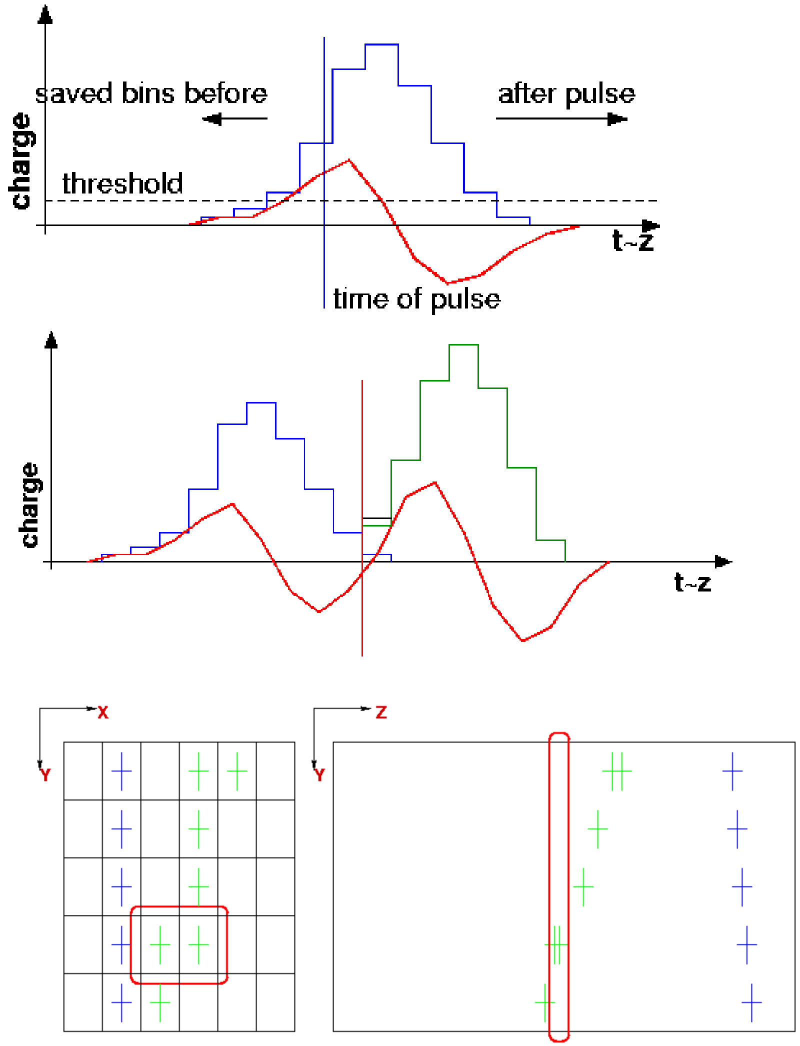
Double Track Separation: Status

- More detailed Studies of hit separation method
- Implementation of new methods?
- Global Fit Method: modifications in MultiFit:
  - Number of tracks from TrackFinder
    → use appropriate likelihood function for fit (done)
  - Implementation of re-calculation of hit x-positions in progress
- Goals:
  - Find value up to which tracks can be separated
    and efficiency of reconstruction
  - Study influence of 2 tracks nearby
    on the track fit and single point resolution

APPENDIX

MultiFit: Hit Reconstruction

- Find pulses in raw data:
  - detect pulses by threshold
  - time: inflexion point of rising slope
- Separation of pulses:
  - Change in slope
    (ignore variations in the order of noise)
- Combine pulses to hits:
  - start with biggest pulse
  - use recursive method in a time window
    - add the pulse if it is smaller
    - take care of damaged pads
  - calculate hit coordinates
    - x: center of gravity (charge)
    - y: center of the row
    - z: error weighted mean of time of pulses



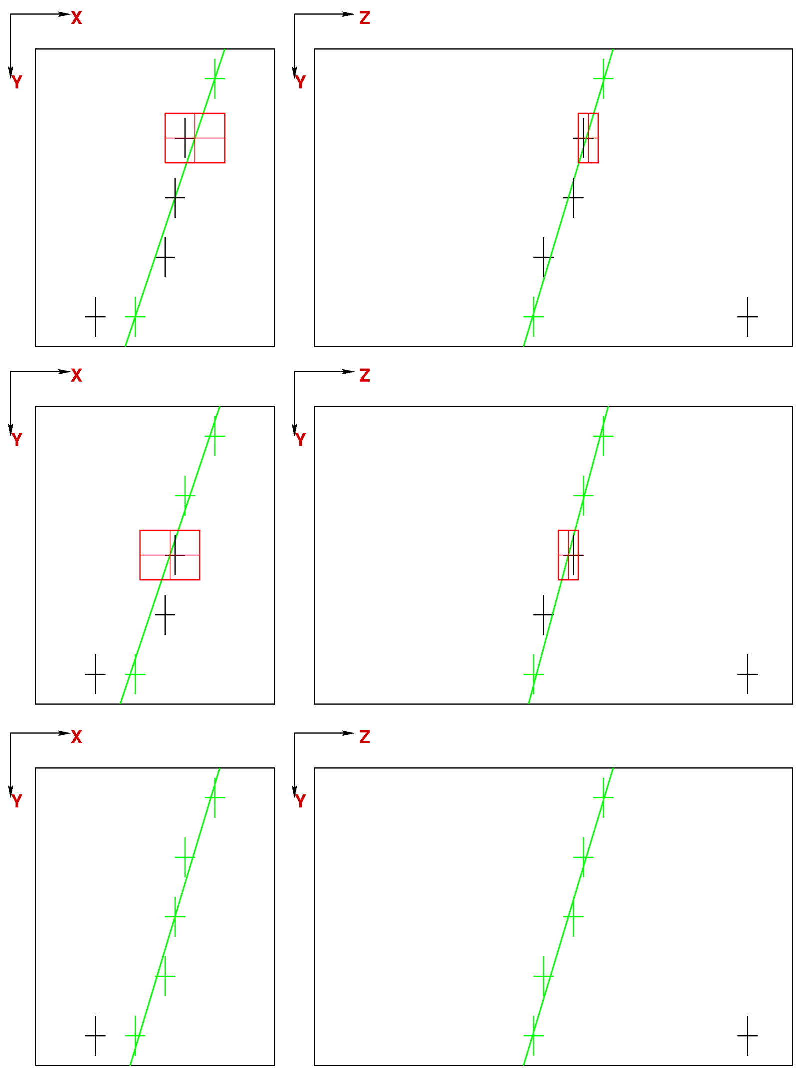
MultiFit: Track Finder

1) First track hypothesis from two points -> fit straight track search in a time window for a hit in the next row

2) After adding the hit:
   - re-fit the track with new hit
   - repeat this procedure in the next row...

3) ... until reaching the last row.

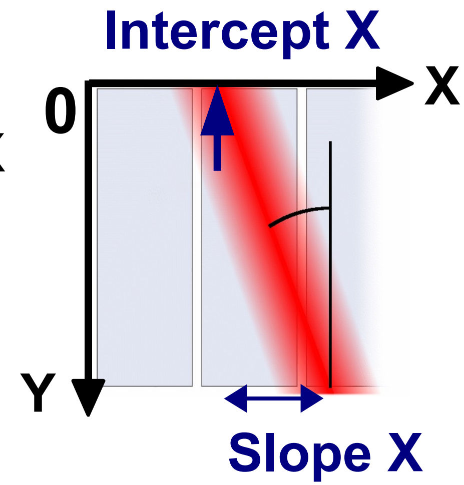
To avoid false tracks:
- only small gaps
- minimal number of hits



Track Fitting: Chi Squared Method

- **Straight line** $x = f(y) = a\,y + b$

  a: SlopeX
  b: InterceptX



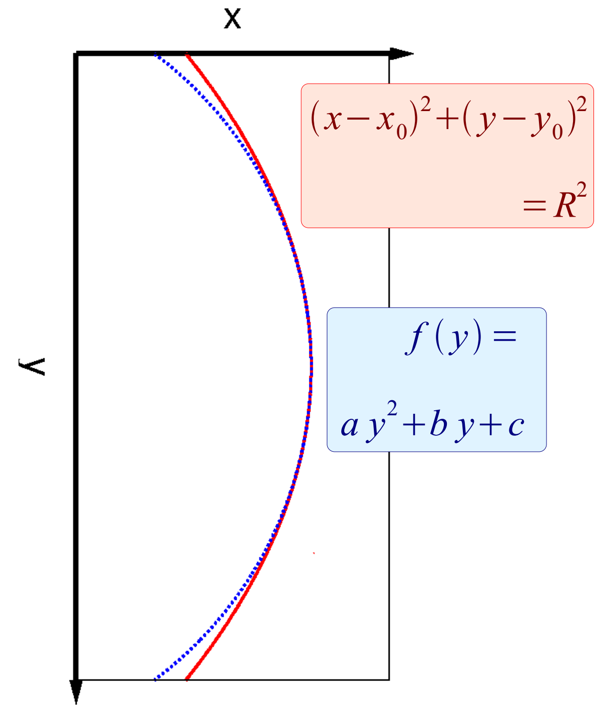
- **2nd degree polynomial:** $x = f(y) = a\,y^2 + b\,y + c$
  - rotated coordinate system

$$Radius\ \ R = \frac{a}{2}\ \ ,\ \ Curvature\ \ C = \frac{1}{R}$$

$$Center\ (x_{0,}\,y_0) \rightarrow solve\,equation\,system:$$

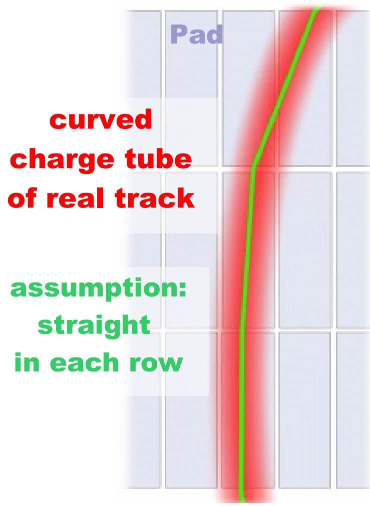
$$(x - x_0)^2 + (y - y_0)^2 = R^2 \quad for\ 2\ points\ (x_{1,}\,y_1),(x_{2,}\,y_2)$$

- **Circular arc:** $(x - x_0)^2 + (y - y_0)^2 = R^2$
  - rotated coordinate system
  - initialized with results from polynomial method
  - Fit function: $x = f(y) = x_0 \pm \sqrt{\frac{1}{C^2} - (y - y_0)^2}$

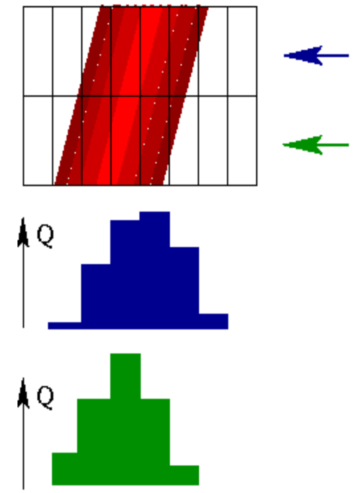


Global Fit Method: Basics

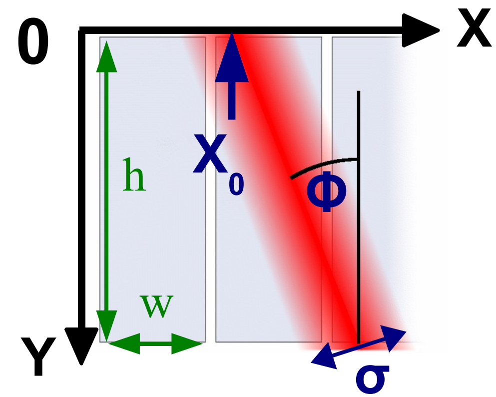
- Assumptions:
  - In each row the track can be described by a straight line



  - XY track fit uses a Gaussian model for charge cloud



- Three (four) parameter fit:
  - Intercept X_0 (x at y=0)
  - Azimuthal angle ϕ
  - Width of the charge cloud σ
    (can be fixed in MultiFit: calculated dependent on drift length per track and per row from diffusion and defocussing coefficient)
    → more stable fit (one parameter less)
    → quite stable against small disturbances (dead channels)
  - Curvature **C** (in case of curved track hypothesis)



Global Fit Method: Principle

- Likelihood function describing charge deposition per pad:

$L_i = p_i^{n_i}$, with $n_i = \frac{N_i}{G}$: number of primary e$^-$, N_i

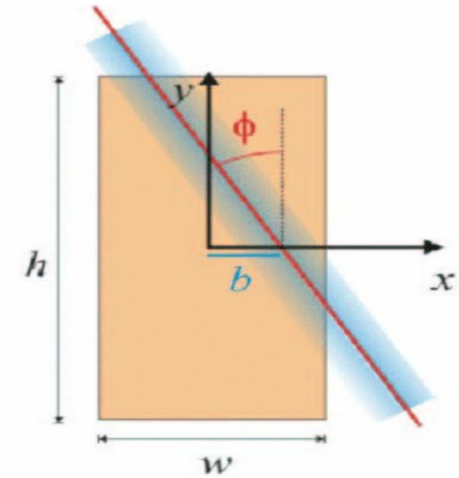
G: gain factor

and $p_i = \frac{Q_{\exp}}{\sum_{n=1}^{pads/row} Q_{\exp}}$ (probability function)

- Product of likelihood functions of all pads:

$$\ln L = \sum_{Pad} Q_{measured} \ln \left[\frac{Q_{expected}}{\sum_{Row} Q_{expected}} \right]$$

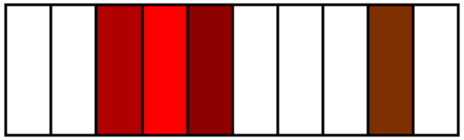
, with $Q_{\exp} = \int_{\frac{-h}{2}}^{\frac{h}{2}} dy \int_{\frac{-w}{2}}^{\frac{w}{2}} dx \, \frac{1}{2\pi\sigma} e^{\frac{[(x-X_0)\cos(\phi)+y\sin(\phi)]^2}{2\sigma^2}}$



- In MultiFit implementation: hit positions will be recalculated after the track fit by fitting the likelihood function with a fixed width to the pulse data of one row

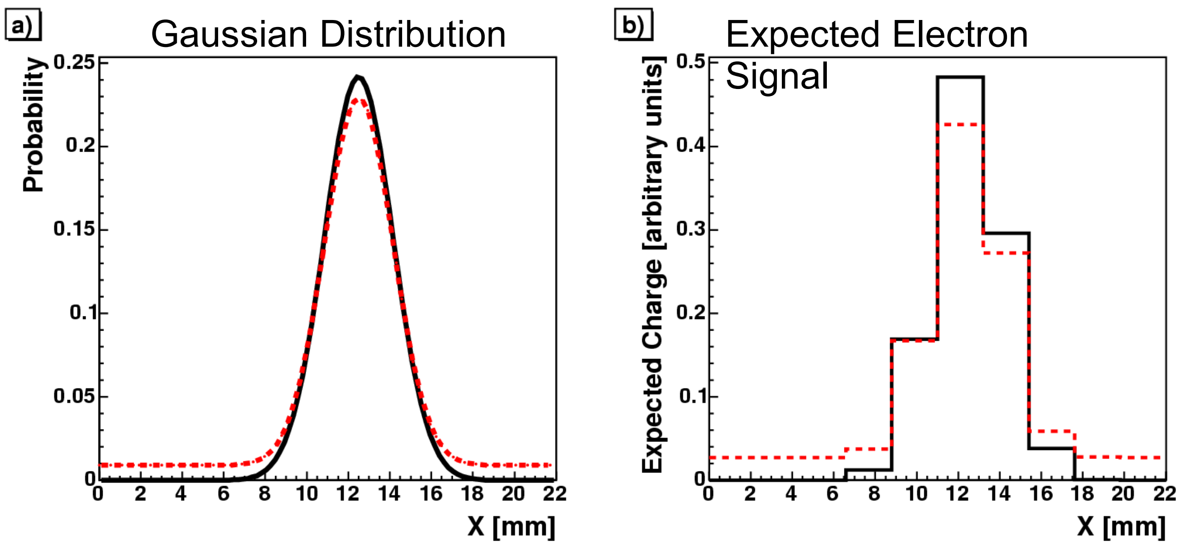
Global Fit Method: Noise Value

- In original, Canadian implementation (JTPC): no clustering → problems with noise pulses



- To make fit more robust, assign a higher probability for measuring a signal to all pads by introducing a constant offset: noise value N

$$p_i \rightarrow \frac{p_i + N}{1 + N \cdot n_{row}} \qquad \rightarrow \qquad \ln L = \sum_{Pad} Q_{measured} \, \ln\left[\left(\frac{Q_{expected}}{\sum_{Row} Q_{expected}} + N\right) / (1+N)\right]$$



Example: pad row with 10 pads, pitch: 2.2mm

— without noise value

---- with noise value N=0.01

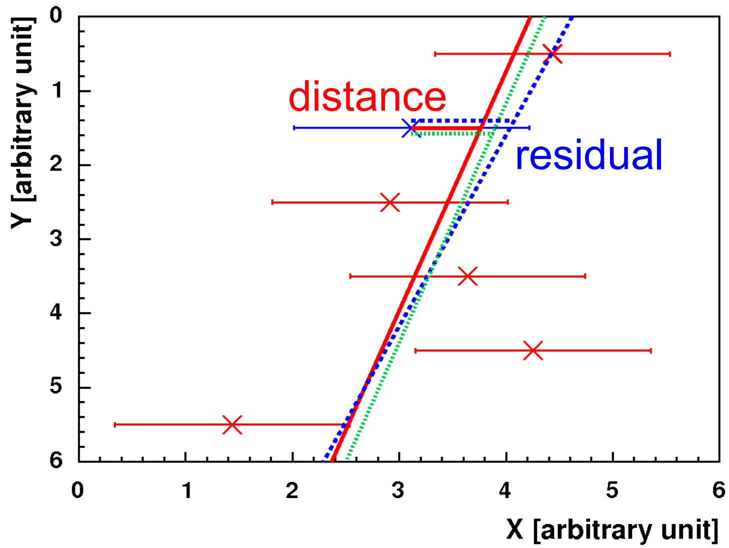
- In JTPC: N=0.01; studies with MultiFit indicated that this is a good value for our implementation too, although we should filter out noise in the ClusterFinder

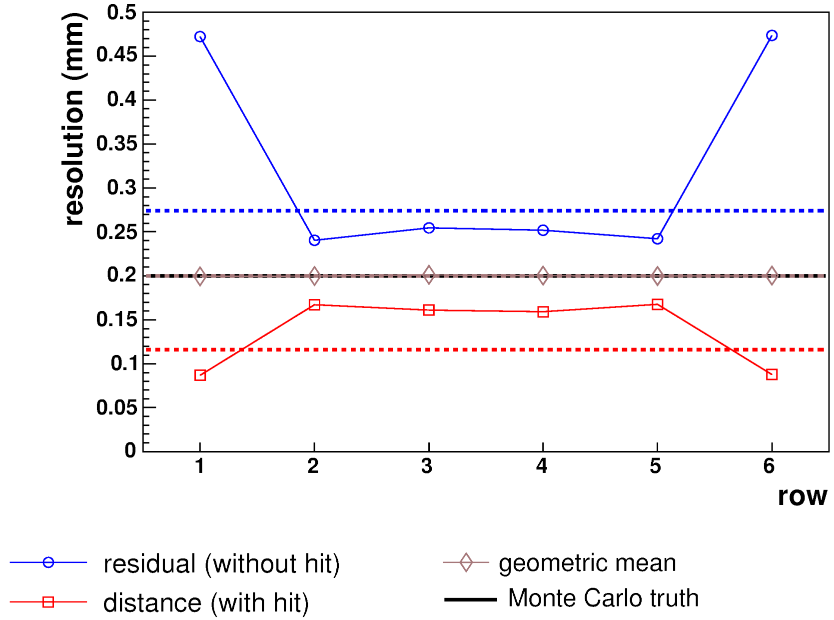
Point Resolution: Geometric Mean Method

- True track position not known
  → calculate two residuals
  - once for track fit including the point (denoted "*distance*")
  - once for track fit without the point (denoted "*residual*")
- Determine the width of both distributions by Gaussian fit
- Resolution calculated from geometric mean of both values:

$$\sigma = \sqrt{\sigma_{with} \cdot \sigma_{without}}$$

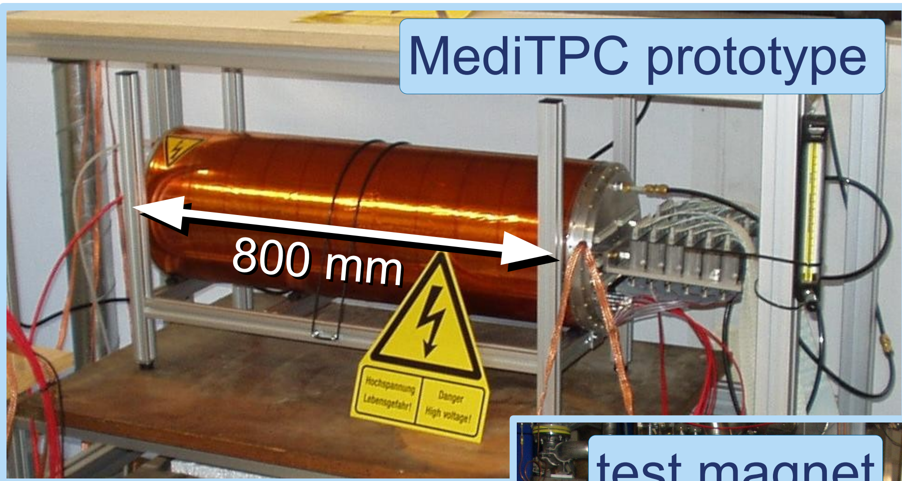
- Proven for
  - straight tracks : analytically
  - curved tracks : MC studies





Measurement Setup and Data Sets

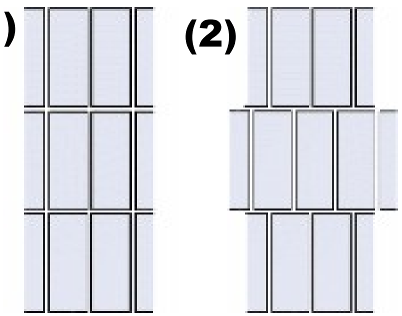
- Length: 800 mm , Ø : 270 mm
- Sensitive volume:
  666.0 x 49.6 x 52.8 mm^3



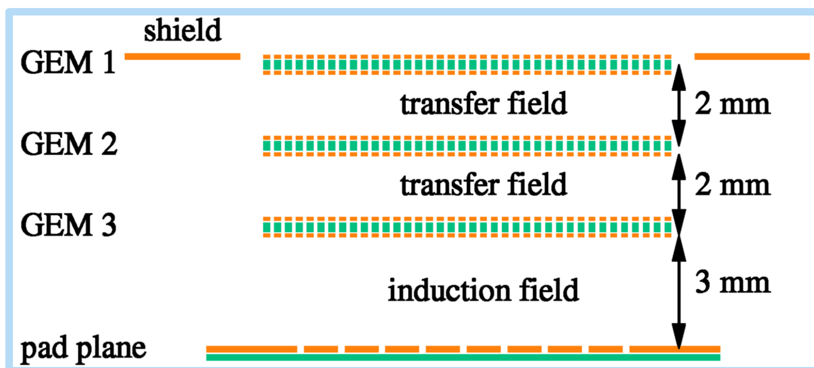
- Magnetic field up to 5.25 T (deviation<7%)
- Data Sets for 0, 1, 2 and 4 T

- Studies with cosmic muons
- Gas: TDR $Ar:CH_4:CO_2$ 93:5:2
  P5 $Ar:CH_4$ 95:5
- Pad layouts:
  - non-staggered (1)
  - staggered (2)
  24 columns, 8 rows
  pitch: 2.2 x 6.2 mm^2

(1) (2)



- Triple GEM amplification setup:
  - Transfer fields : 1500 V/cm
    Induction field : 3000 V/cm
  - About 320 - 335 V per GEM



Monte Carlo Simulation - Principle

- Muon generator: cosmic muons with realistic angular and energy spectra
- Detector and trigger geometry taken into account

→ only straight tracks!

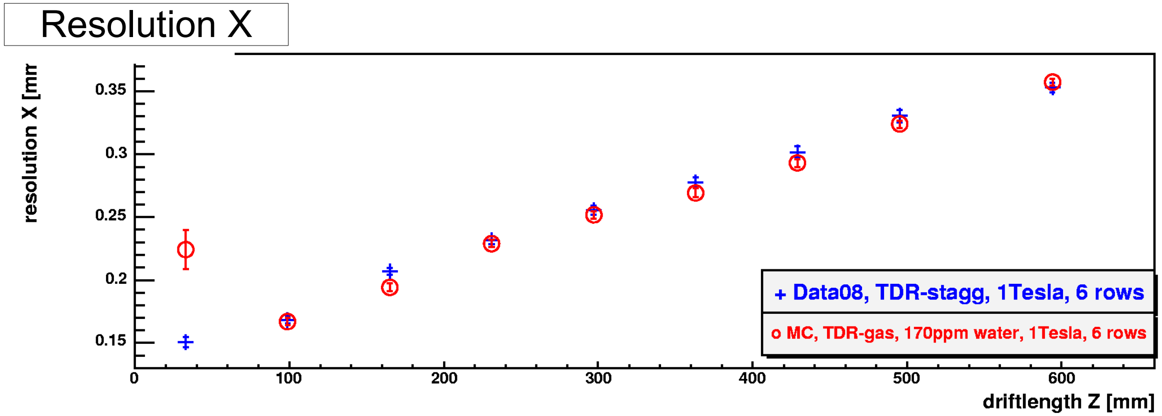
Gas properties: mixture, pressure, water content etc.

E- & B-field

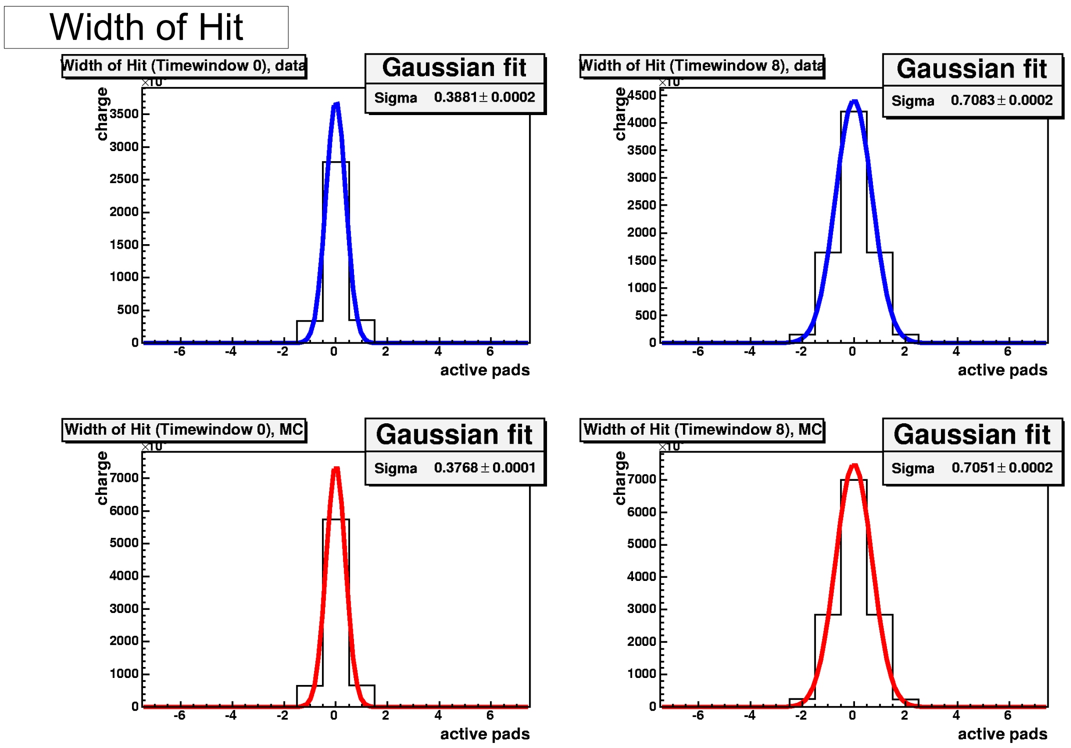
- Primary ionization simulated with HEED → 3D e^- distribution
- Drift:
  - Velocity and diffusion parameters from GARFIELD
  - Gaussian position smearing: $x' = x + N_{random} \cdot C_D \cdot \sqrt{L}$
- GEM amplification:
  - Electrons forced into the nearest hole
  - Amplification with effective gain (smearing: Polya distributed)
  - Drift between GEMs and last GEM to pad plane as above
- Collection on the pad plane and readout simulation

Monte Carlo Simulation - Performance

- X Resolution:
  good agreement between simulated and measured data (except first bin)



- Width of Hits:
  good agreement between simulated and measured data



Point Resolution Studies: Introductory Remarks

- Cuts:
  - Angle: $\varphi <$ 0.1 rad (5.73°)
    $\Theta <$ ~0.44 rad (25.0°)
  - Exclude outer columns:
    only hits taken into account with (*nearly*) complete charge measured
  - Minimum of 6 hits per track
- Gas mixtures: TDR (Ar-CH4-CO2: 93-5-2)
  P5 (Ar-CH4: 95-5)

diffusion coefficient D
defocussing constant σ_0

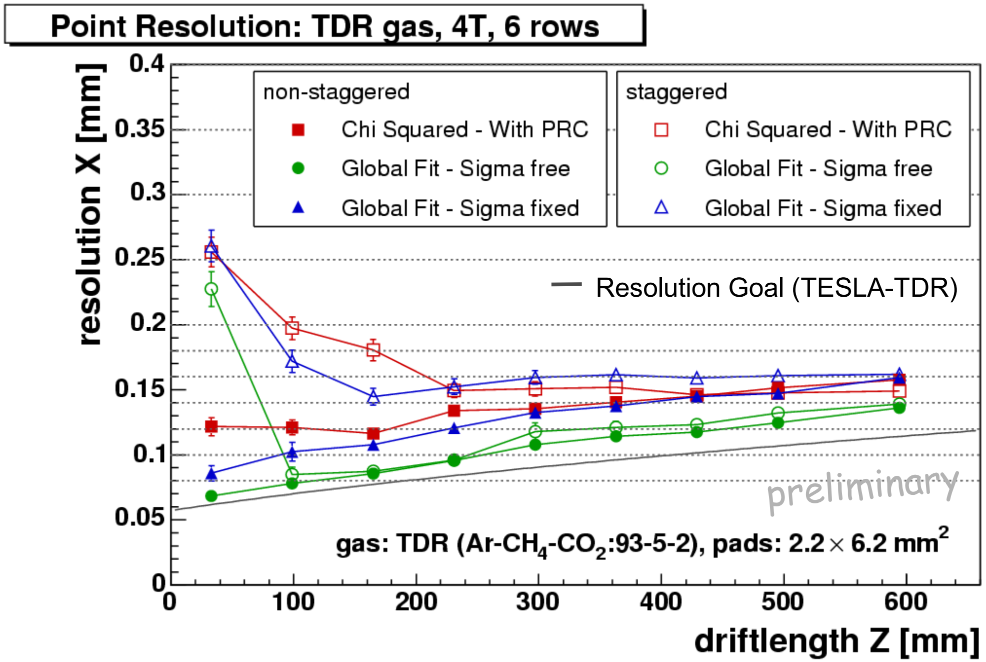
$$\sigma = \sqrt{D\,z + \sigma_0}$$

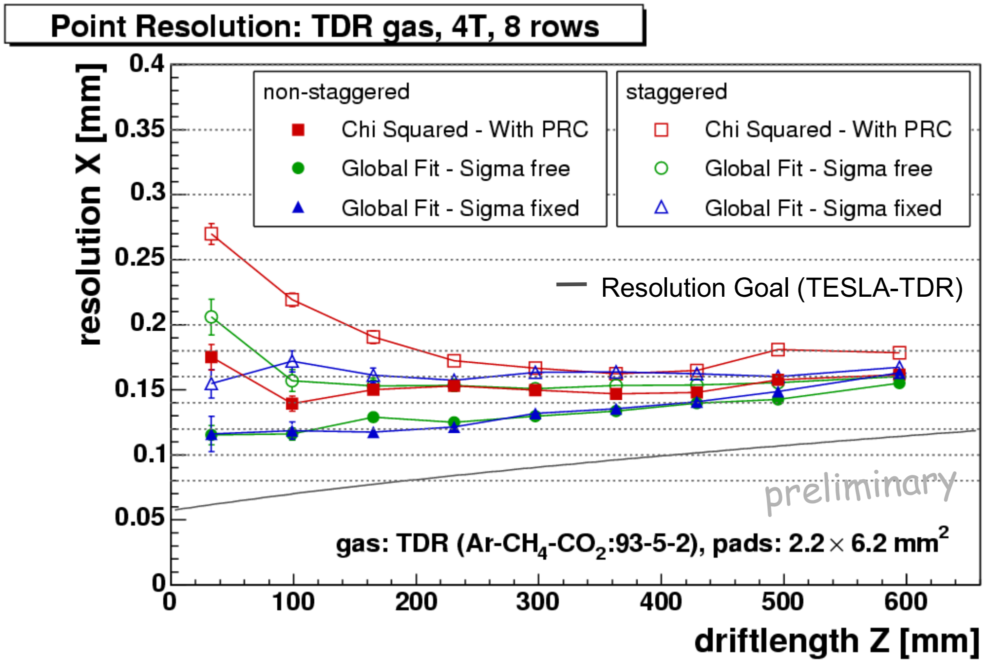
derived from GARFIELD7 simulation (0ppm water content)

	P5		TDR	
B (T)	D (mm) 10^{-4}	σ_0 (mm^2)	D (mm) 10^{-4}	σ_0 (mm^2)
0	571	0,288	202	0,180
1	24,05	0,227	34,1	0,142
2	7,24	0,190	11,5	0,110
4	1,92	0,140	3,00	0,070

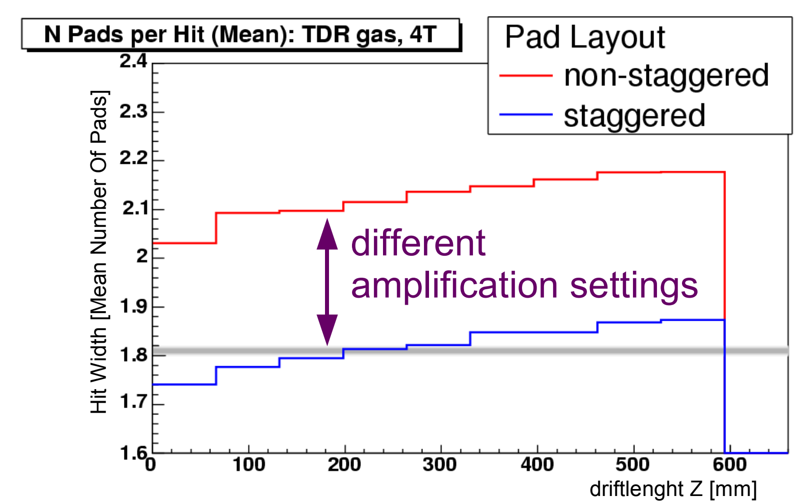
- Problem of measured data: top and bottom row (#1 and #8) show crosstalk with the surrounding shield →
  - resolution calculated with all 8 rows too pessimistic (contains not perfect hits)
  - resolution calculated with only inner 6 rows too optimistic (relation between fit parameters and data points too small)
- Both values will be presented
- 8 rows deliver more conservative results (upper limit)

Point Resolution Results: TDR gas



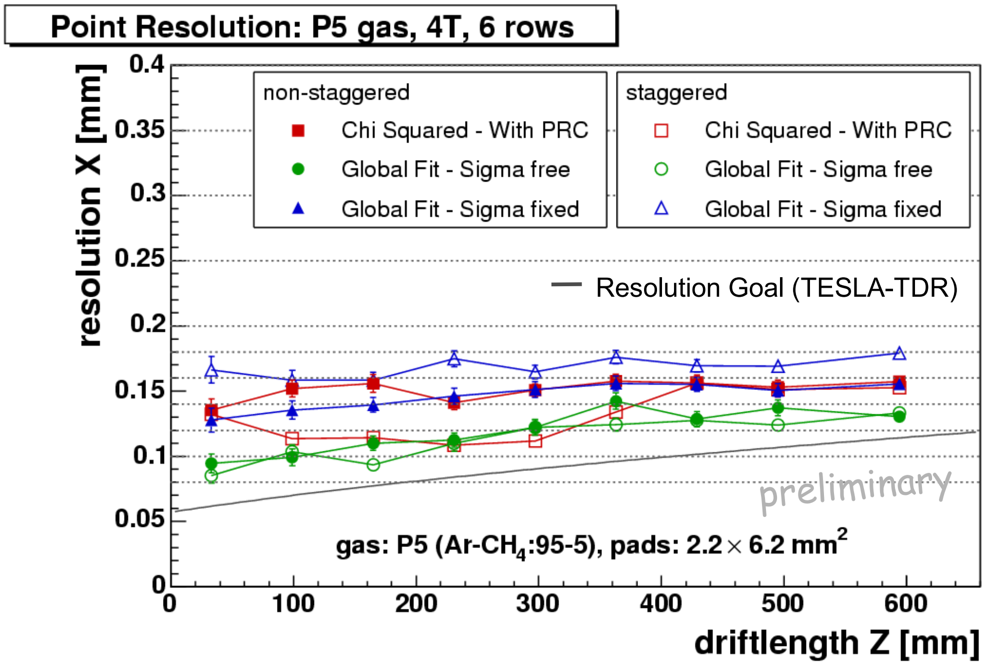


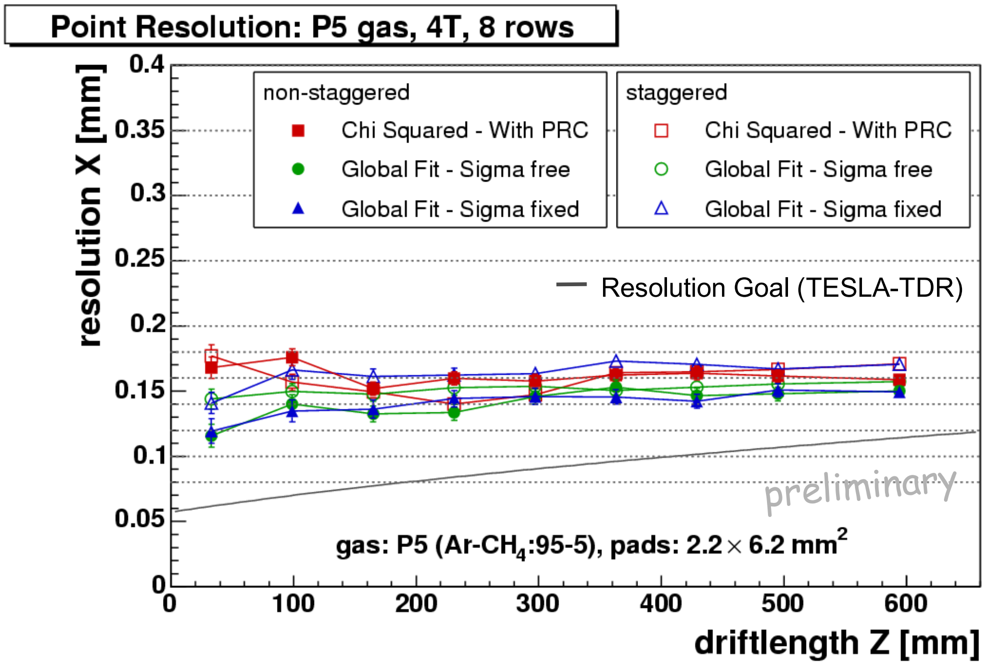
- Deviation between non-staggered and staggered results ← charge sharing too small
- Especially at short drift distances: results from staggered layout affected by charge sharing limit
- Results for 6 rows unreasonably good esp. Global Fit with free σ
- Resolution: ~ 120-180 μm (Z = 0-660 mm)



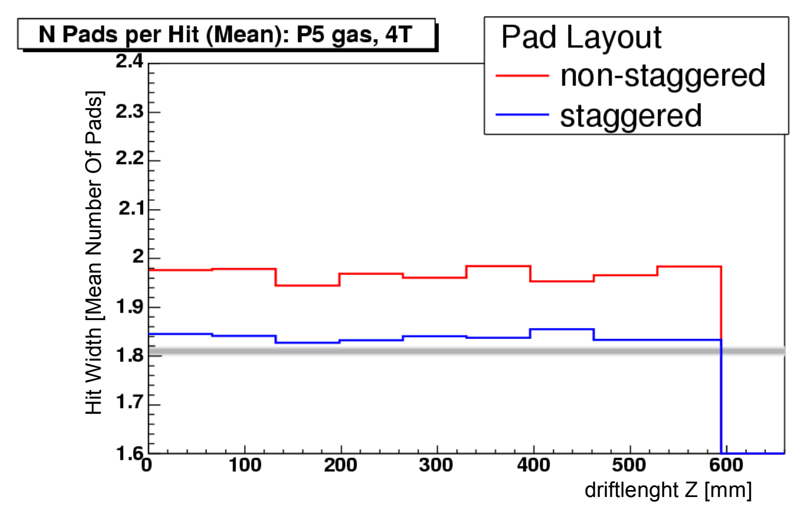
extracted from measurement runs

Point Resolution Results: P5 gas



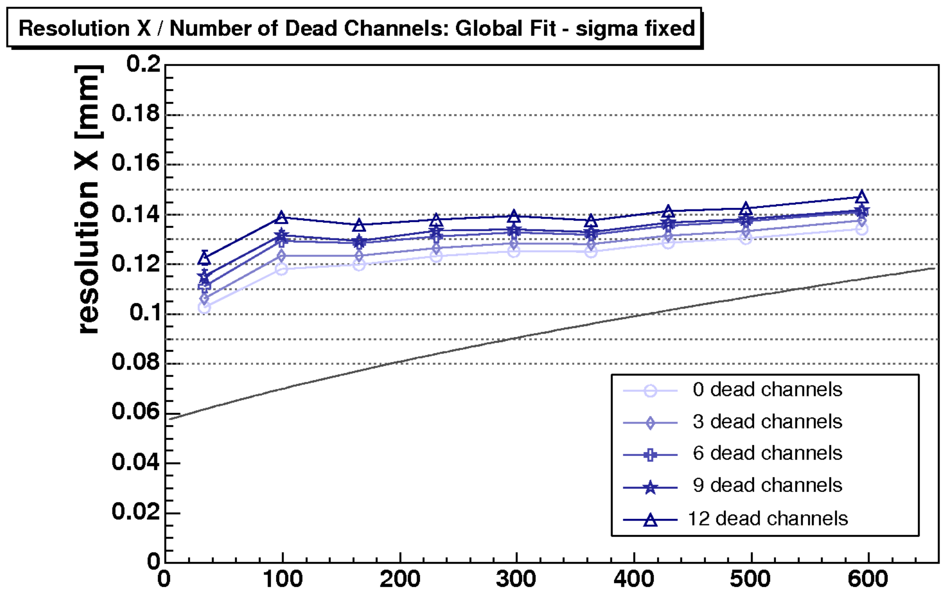


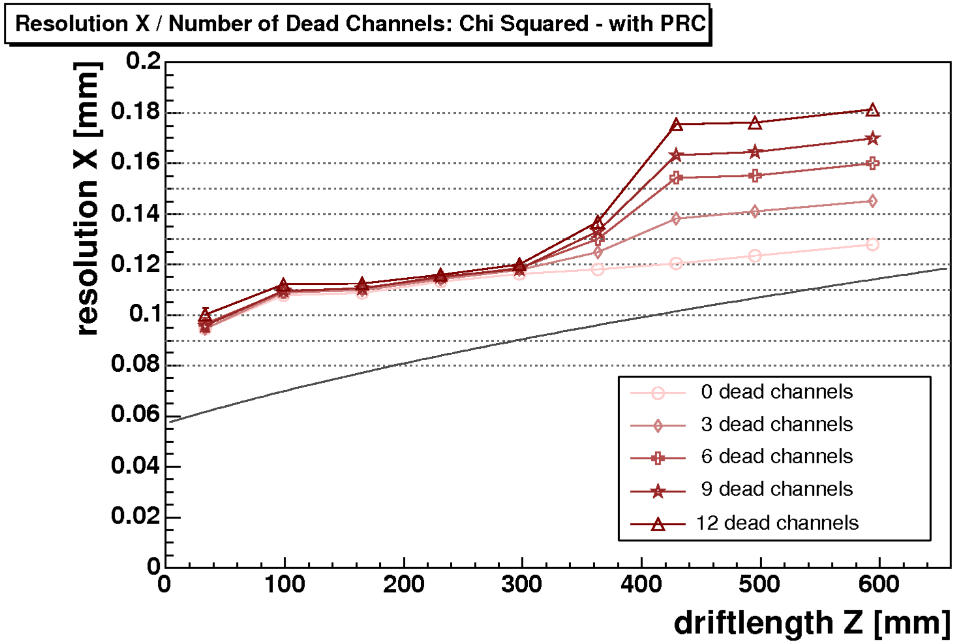
- Again deviation between non-staggered and staggered results, but here smaller ← charge sharing too small
- Some results from staggered layout also increase at short drift distances, but much less (no big drift dependence of width)
- Results for 6 rows a bit better than for 8, but spread of results smaller for 8 rows
- Resolution: ~ 120-170 μm (Z = 0-660 mm)

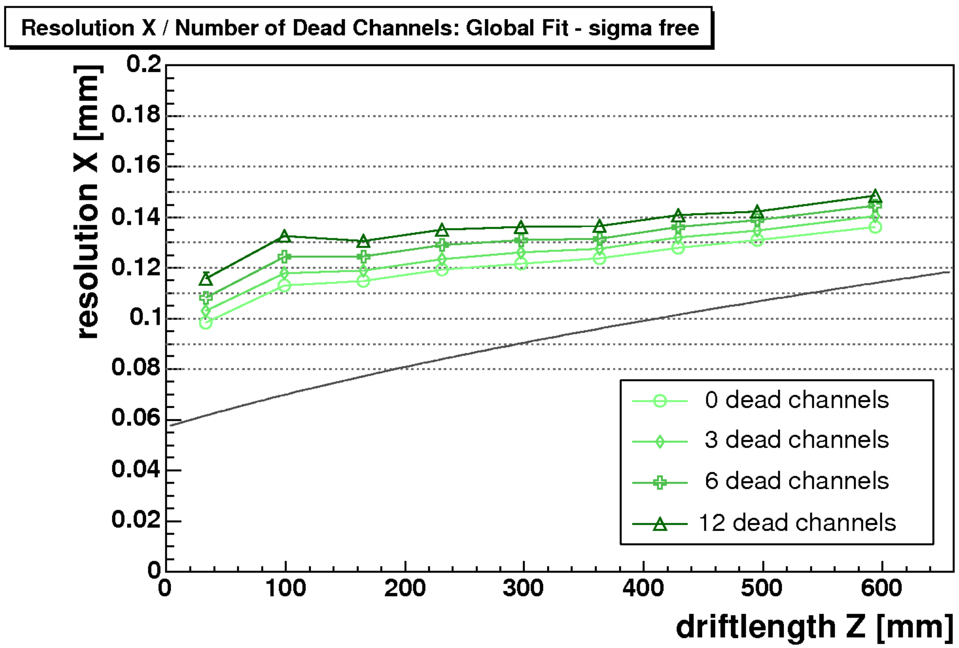


extracted from measurement runs

Influence of the Dead Channels

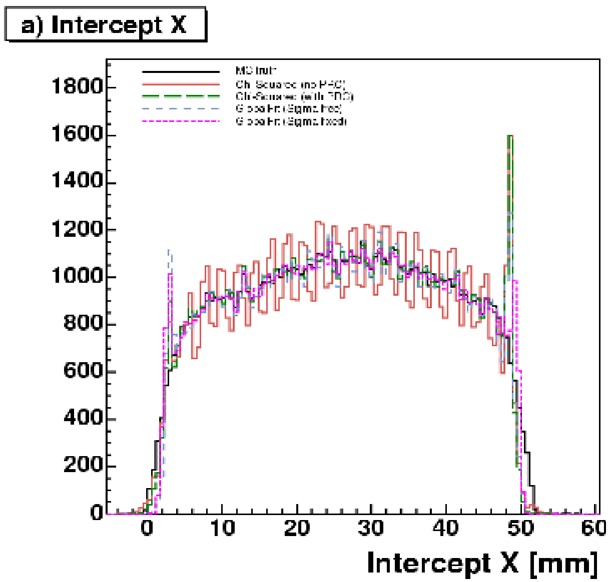


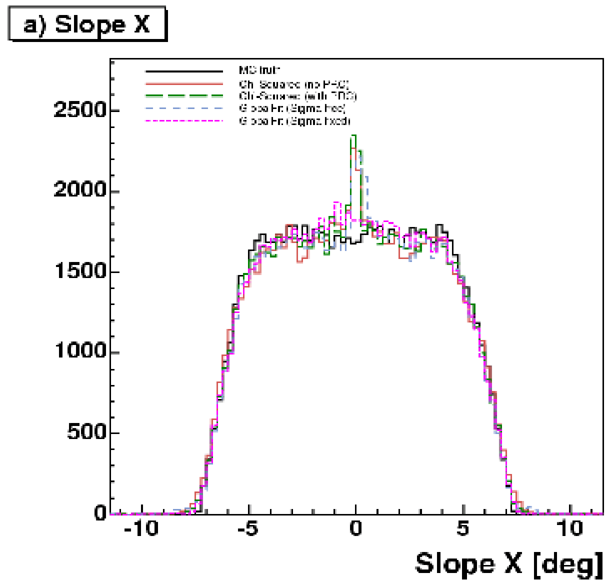




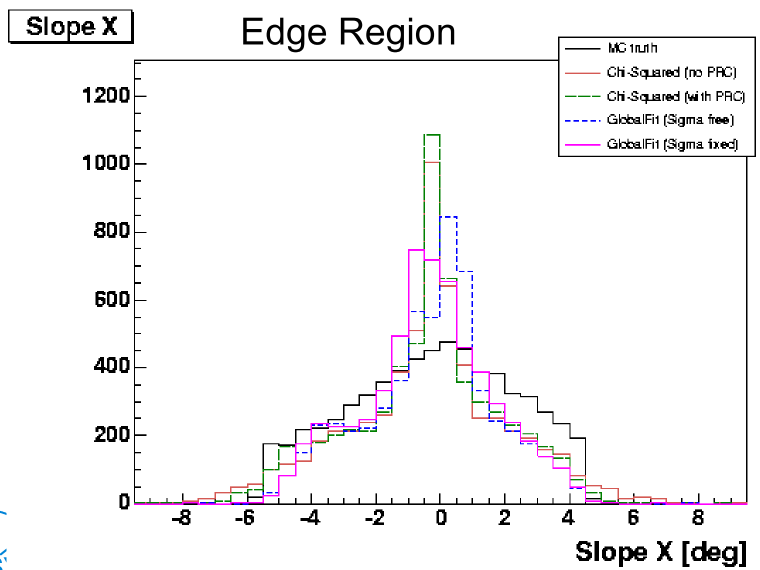
Edge Effects

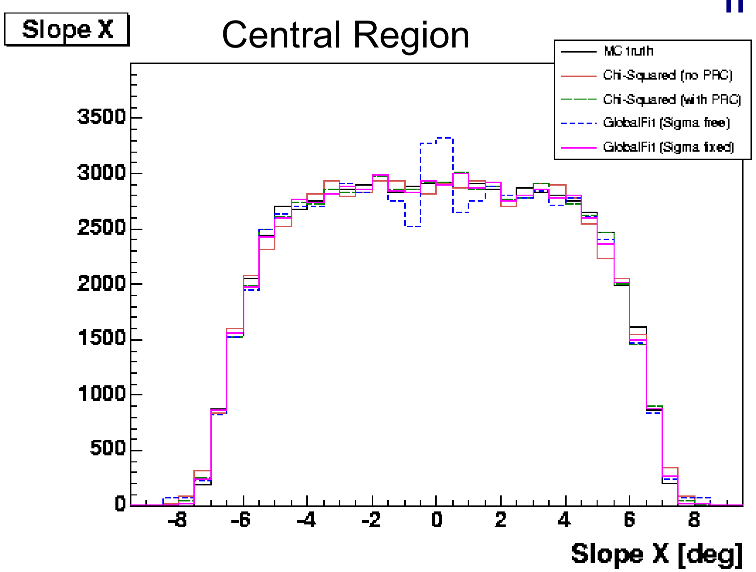
- Interesting due to nearly straight tracks and edges of readout structures





- Edge: 2 outer pad columns on each side
- Not whole charge information measured → tracks get reconstructed at angle ~0° and therefore at same Intercept
  → Edges excluded in analysis





Monte Carlo Data:
4T, P5,
staggered pad layout
8 rows