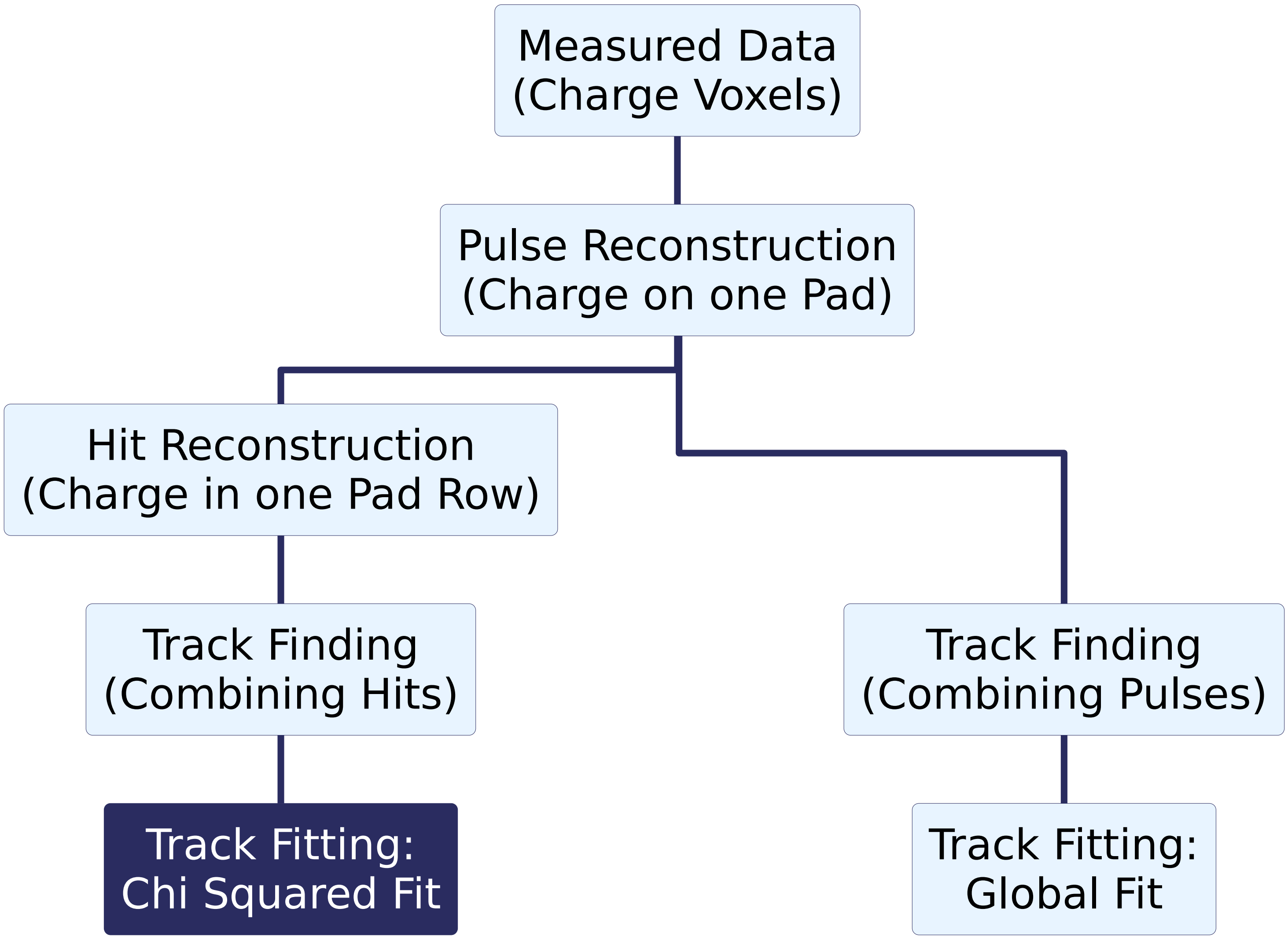
Chi Squared Fit



Least Squares Method

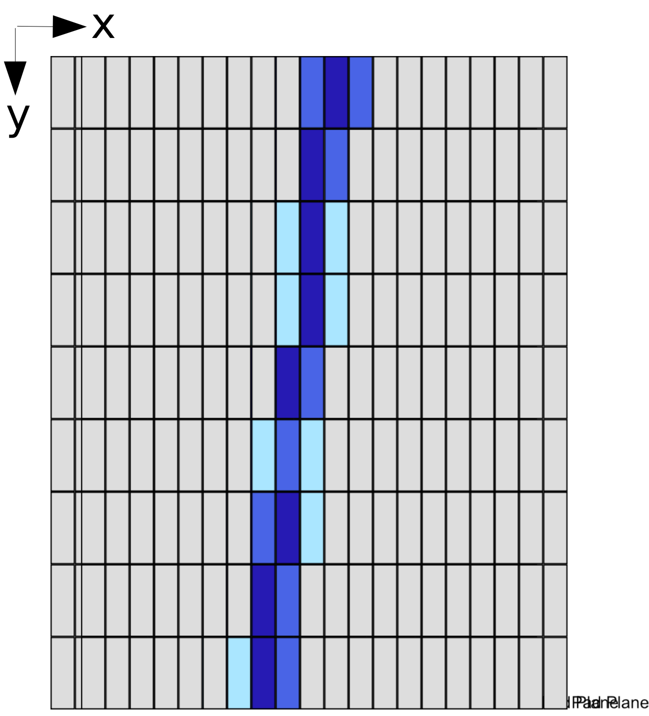
- **N measurements at points y_i** (the measurement points are given by the pad rows)
- **Variables x_i with error σ_i** (the x coordinate of the hit is the measurement)
- **Fit function: $f(y; a_1, a_2, \ldots a_M)$, a_j: parameters to be determined**
- **N > M !** (more measurement points than parameters needed!)
- **For the best values a_j, sum S is a minimum :**

$$S = \sum_{i=1}^{N} \left[\frac{x_{(i)} - f(y_i, a_j)}{\sigma_i} \right]^2 \rightarrow \frac{\partial S}{\partial a_j} = 0, \; j = 1 \ldots M$$

- for S to be a real chi-square, x_i must be Gaussian distributed with mean $f(y_i; a_j)$ and variance σ_i^2

Chi Squared Fit

- Straight Line: $x = f(y) = a\,y + b$ a: SlopeX b: InterceptX



- So, in this case: $S = \sum_i \frac{(x_i - a\,y_i - b)^2}{\sigma_i^2}$ and

$$\frac{\partial S}{\partial a} = -2\sum_i \frac{(x_i - a\,y_i - b)\,y_i}{\sigma_i^2} = 0 \quad \text{and} \quad \frac{\partial S}{\partial b} = -2\sum_i \frac{(x_i - a\,y_i - b)}{\sigma_i^2} = 0$$

with:

$$A = \sum_i \frac{y_i}{\sigma_i^2} \qquad B = \sum_i \frac{1}{\sigma_i^2} \qquad C = \sum_i \frac{x_i}{\sigma_i^2}$$

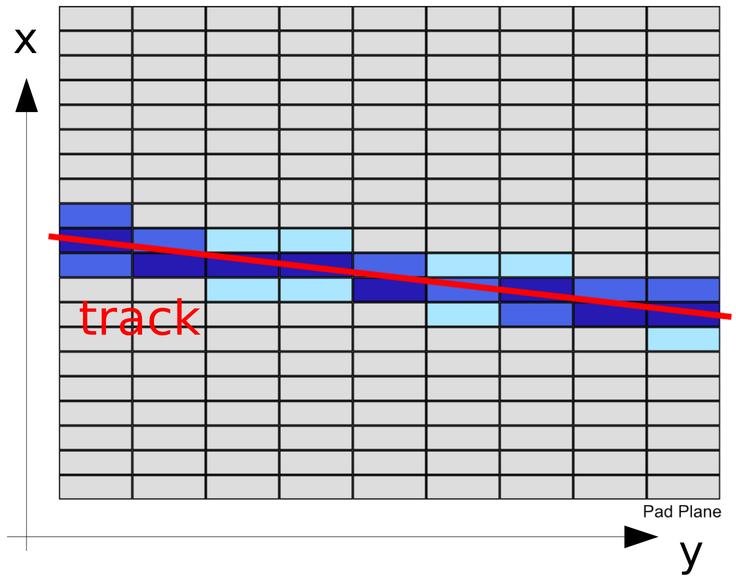
$$D = \sum_i \frac{y_i^2}{\sigma_i^2} \qquad E = \sum_i \frac{x_i\,y_i}{\sigma_i^2} \qquad F = \sum_i \frac{x_i^2}{\sigma_i^2}$$

this results in

$$2(-E + aD + bA) = 0,$$
$$2(-C + aA + bB) = 0$$

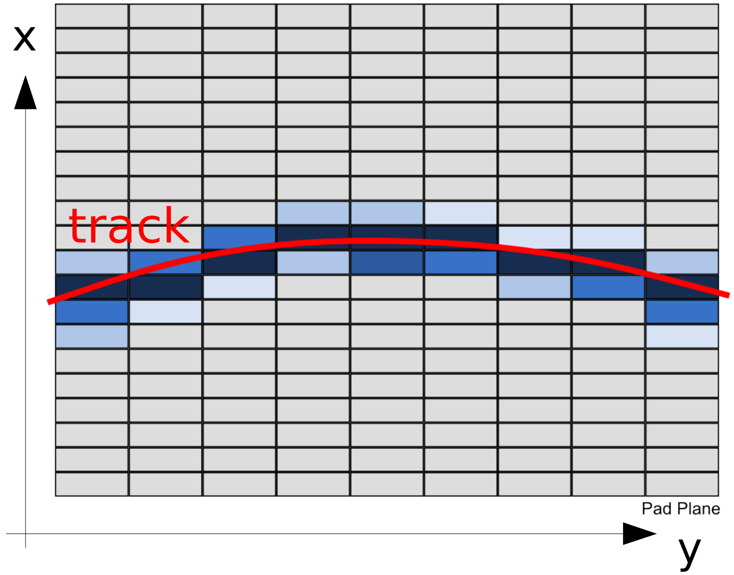
and the parameters a and b are given by:

$$a = \frac{EB - CA}{DB - A^2} \quad \text{and} \quad b = \frac{DC - EA}{DB - A^2}$$



Chi Squared Fit

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



- This leads to $S = \sum_i \frac{\left(x_i - a\,y_i^2 - b\,y_i - c\right)^2}{\sigma_i^2}$

(minimized numerically)

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

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

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

Fast fit method, results can be used for circle fit

Chi Squared Fit

- Circle Fit: $(x-x_0)^2+(y-y_0)^2=R^2$

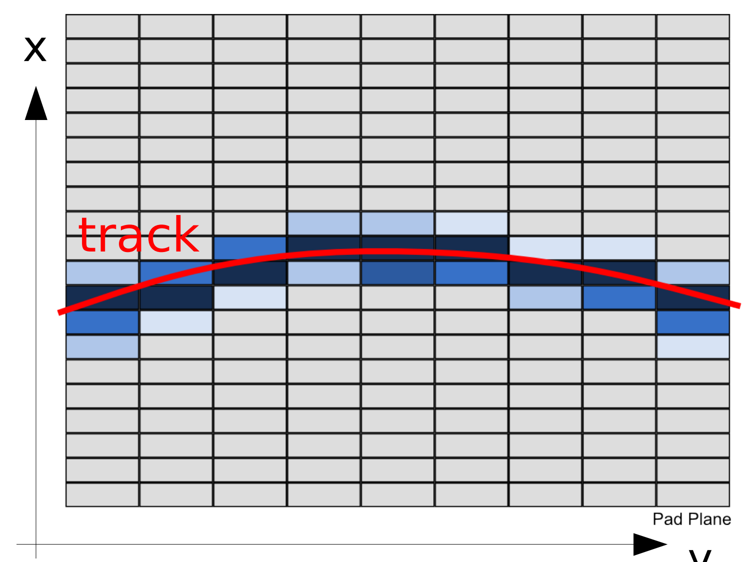
- also in rotated coordinate system, so the function is:

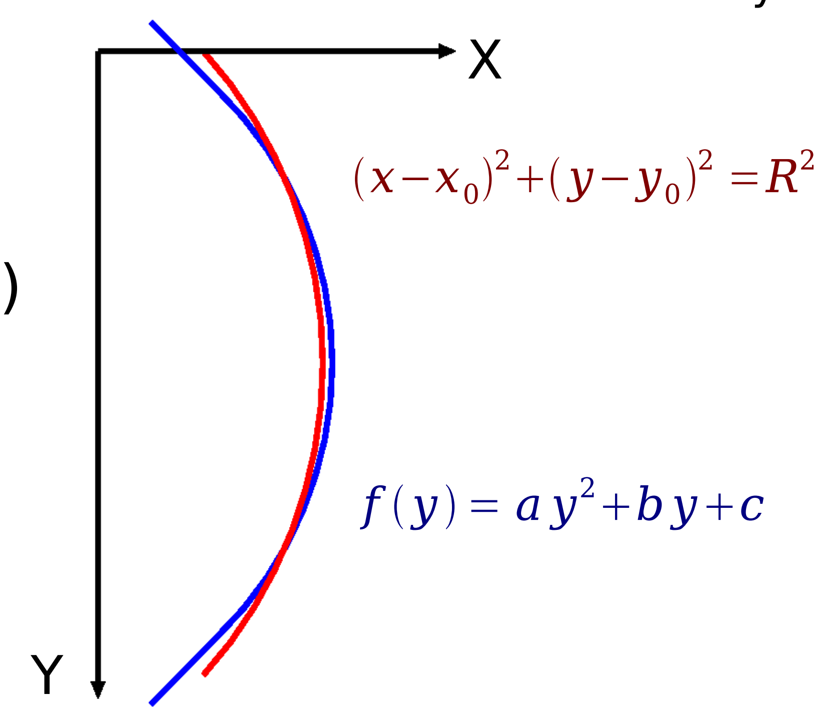
$$x = f(y) = x_0 \pm \sqrt{\frac{1}{C^2}-(y-y_0)^2}$$

- This leads to

$$S = \sum_i \frac{\left(x_i - x_0 \pm \sqrt{\frac{1}{C}-(y-y_0)^2}\right)^2}{\sigma_i^2}$$

(minimized numerically)





Track Object

- Track Object

- Collection of constituents:
  A vector (array) of Hits resp. Pulses belonging to the track

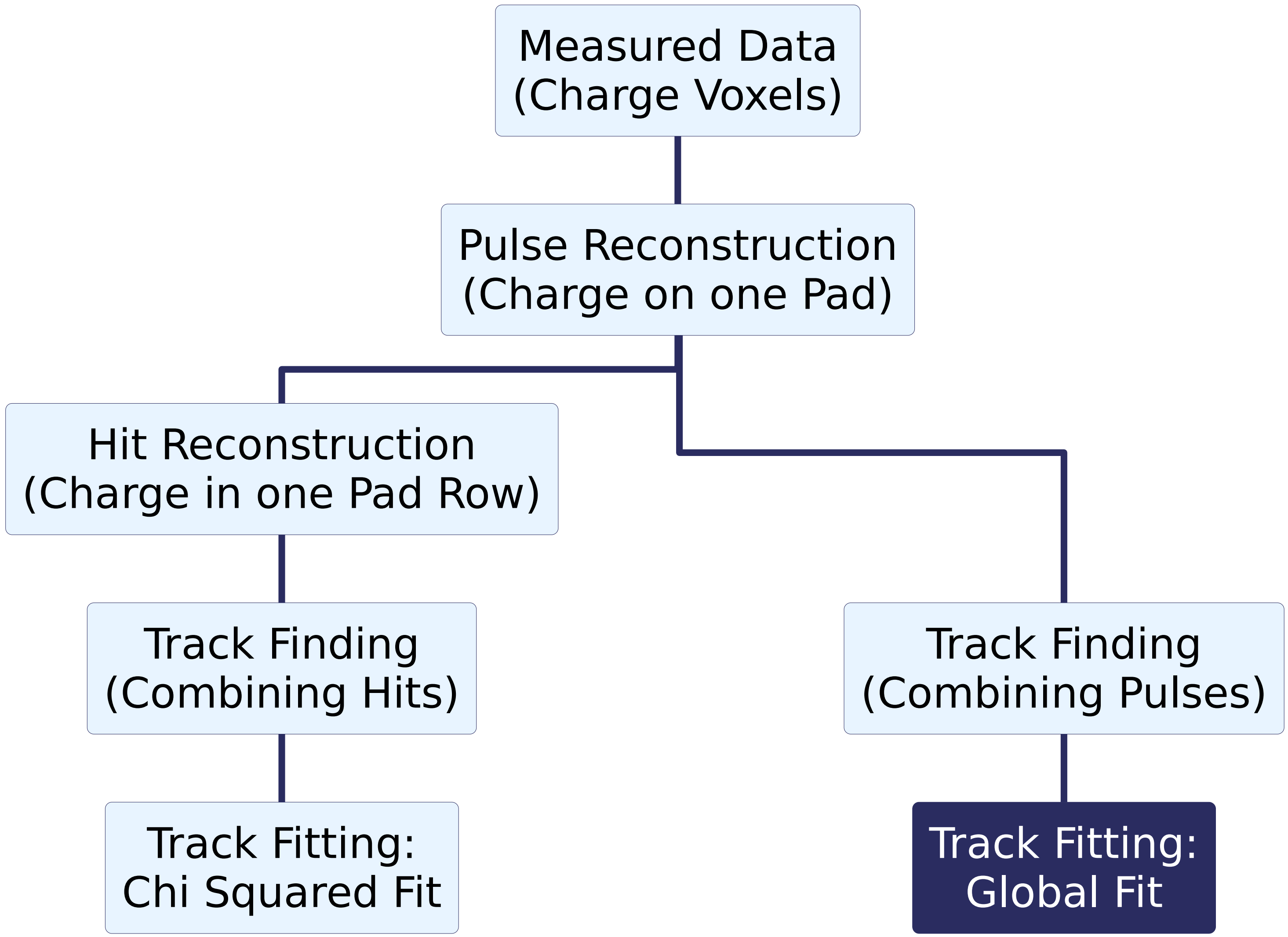
- Track parameters:
  - Intercept (where does it enter the sensitive volume)
  - Slope (angle)
  - Curvature
  - Center of Circle

- Errors of the track parameters

- Chi Squared of the track fit (estimate of the fit quality)

- Optional: Number of parameters, dE/dx,

Global Fit



Measured Data
(Charge Voxels)
Pulse Reconstruction
(Charge on one Pad)
Hit Reconstruction
(Charge in one Pad Row)
Track Finding
(Combining Hits)
Track Finding
(Combining Pulses)
Track Fitting:
Chi Squared Fit
Track Fitting:
Global Fit

Maximum Likelihood Method

- A sample of **n** independent observations $x_1, x_2, \dots x_n$
- Theoretical distribution known: $f(x|a)$,with **a**: Parameter to be estimated
- Calculate the likelihood function:

$$L(a|x) = f(x_1|a) f(x_2|a) \dots f(x_n|a)$$

This can be recognized as the probability for observing the sequence of values $x_1, x_2, \dots x_n$

- Principle: this probability is a maximum for the observed values So the parameter **a** must be such, that ***L*** is a maximum.

- So, **a** can be found by solving: $\frac{dL}{da} = 0$

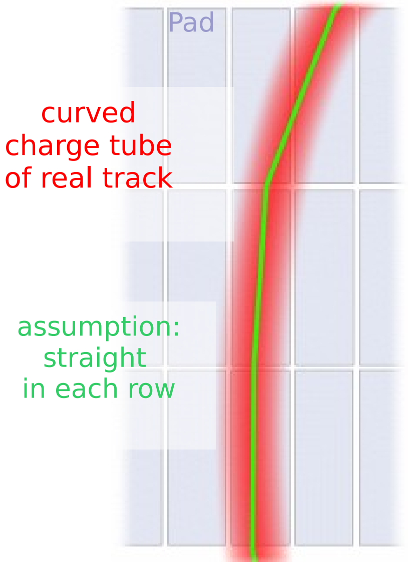
In practice: often easier to maximize the logarithm of ***L***: $\frac{d(\ln(L))}{da} = 0$ since: $\ln(y * z) = \ln(y) + \ln(z)$

this yields results which are equivalent to the above.

Global Fit

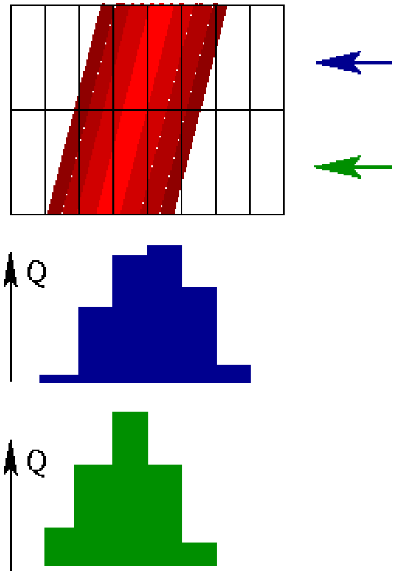
- Assumptions:

- For each row the track can be described by a straight line (height of a pad row much smaller than the radius of the curvature)



- Charge is Gaussian distributed along the track (this is a valid model for the charge deposition)

- Variations of the charge deposition are ignored: assume a constant charge deposition in a row



Global Fit

- Likelihood function describing charge deposition per pad:

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

G : gain factor

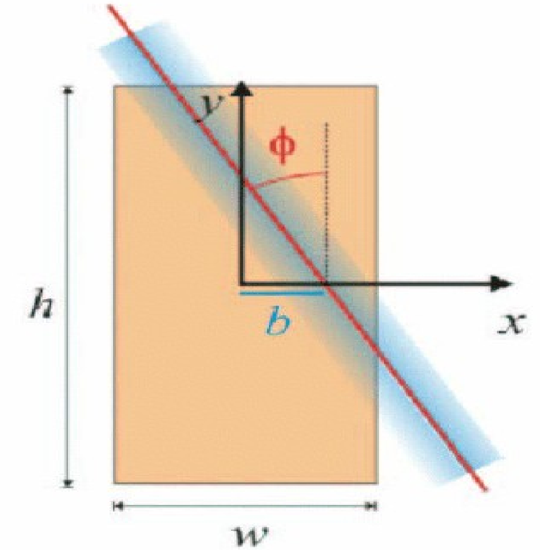
and $p_i = \frac{Q_{exp}}{\sum_{n=1}^{pads/row} Q_{exp}}$ (probability function)

- Logarithm of product of likelihood functions of all pads:

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

width of charge distribution included in fit function as free fitting parameter

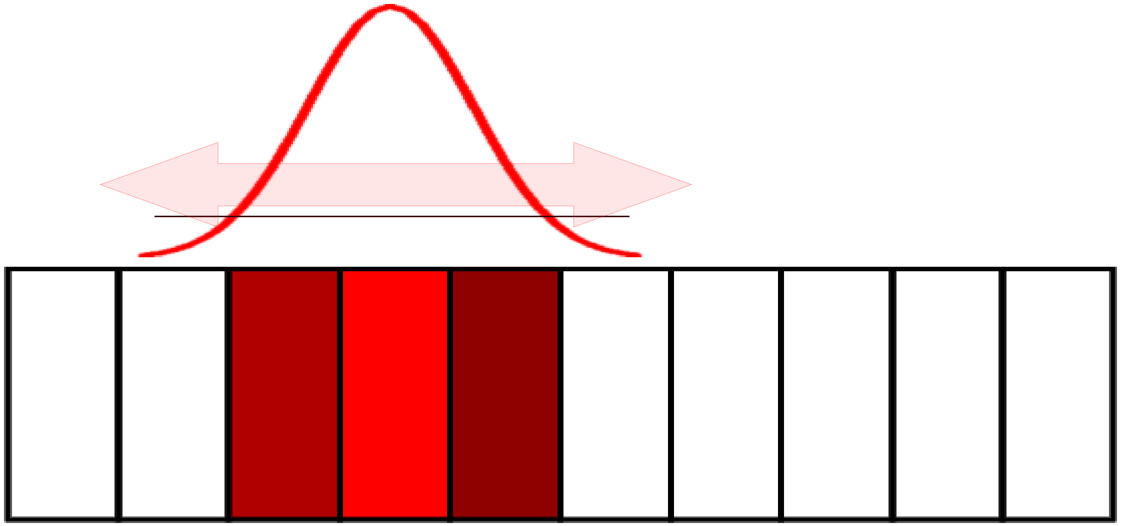
, with $Q_{exp} = \int_{\frac{-h}{2}}^{\frac{h}{2}} dy \int_{\frac{-w}{2}}^{\frac{w}{2}} dx \ \frac{1}{2\pi\,\sigma} e^{\frac{[(x-X_0)\cos(\phi)+y\sin(\phi)]^2}{2\sigma^2}}$



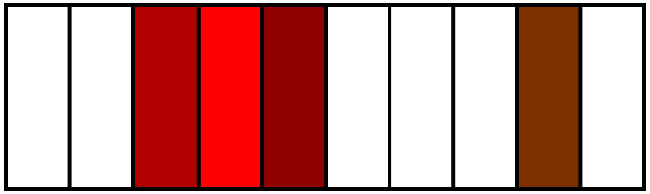
for details see: “TPC Performance in Magnetic Fields with GEM and Pad Readout”, D. Karlen, P. Poffenberger, G. Rosenbaum, 2005

Hit Positions

- Hit is defined as the charge deposition in a row
- The X position of this charge deposition is needed for later resolution calculation, but the Global Fit in general has no hit reconstruction
- To get the position, do a Global Fit in just one line with the width, angle fixed to the result of the track fit:
  this means moving a the charge distribution with fixed width and angle (depending on curvature) along the x axis until it fits best to the deposited charge in this row

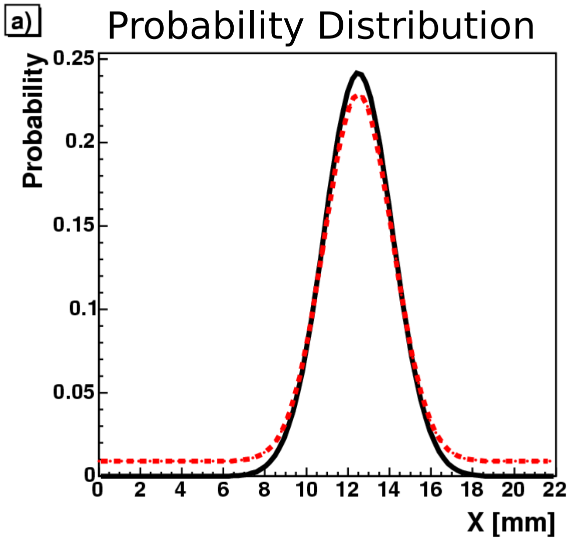


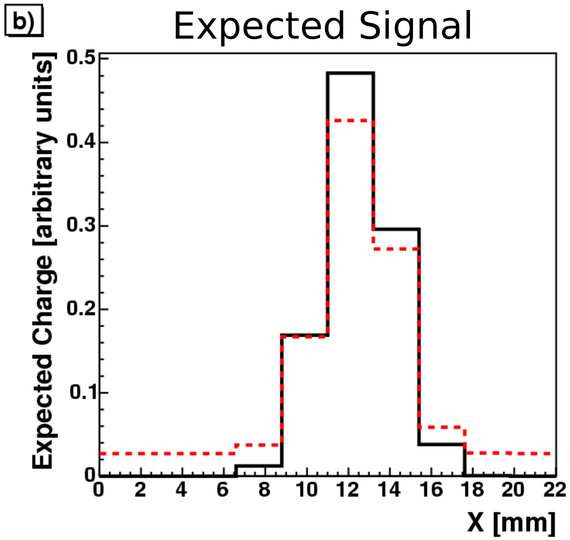
- What to do with noise Pulses? They are not described by the theoretical distribution



- Solution: assign a higher probability for measuring a signal to all pads by introducing a constant offset: noise value N

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





—— without noise value
---- with noise value N=0.01

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

Track Object

- Track Object

- Collection of constituents:
  A vector (array) of Hits resp. Pulses belonging to the track

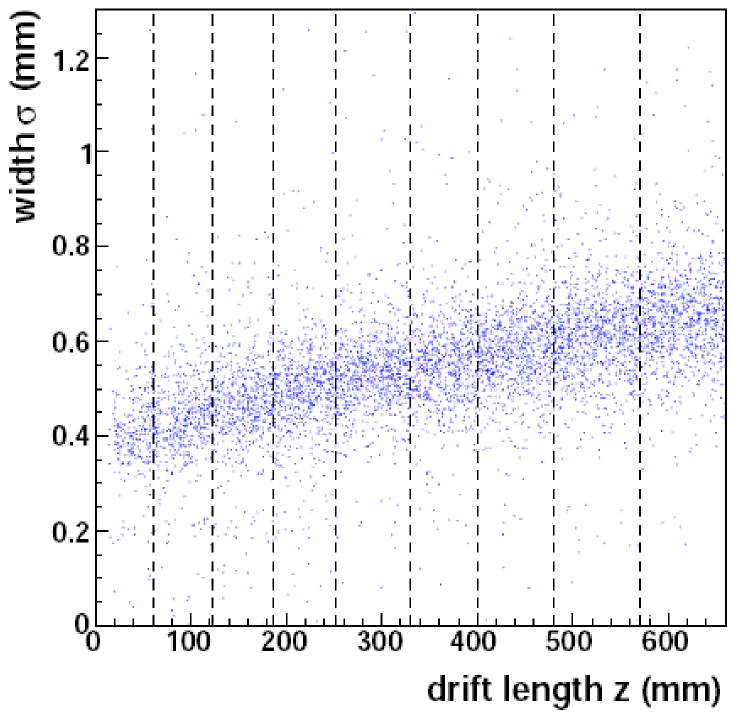
- Track parameters:
  - Intercept (where does it enter the sensitive volume)
  - Slope (angle)
  - Curvature
  - Center of Circle
  - Width of charge distribution

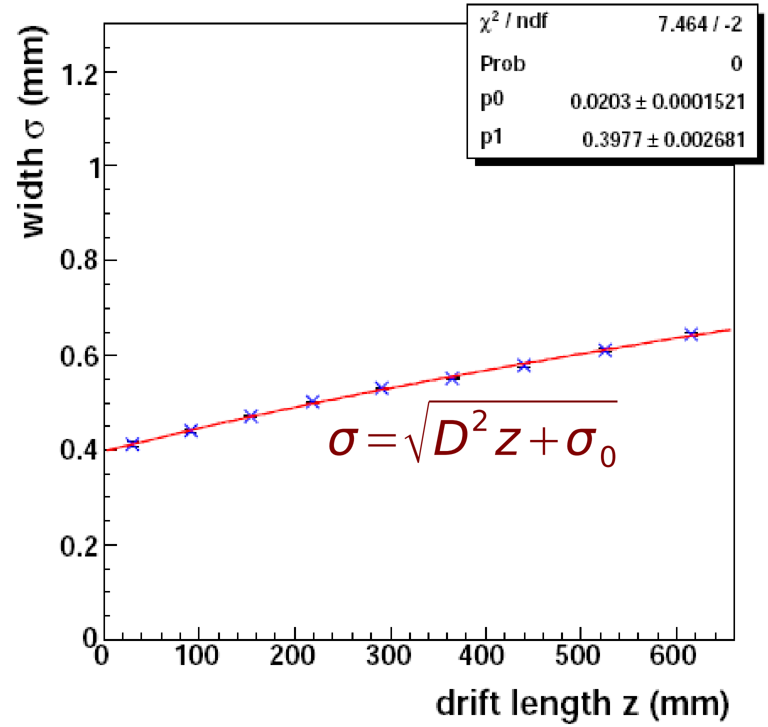
- Errors of the track parameters

- Optional: Number of parameters, dE/dx,

Diffusion Parameters

- From the Global Fit, also the diffusion and defocussing parameters of the setup can be determined

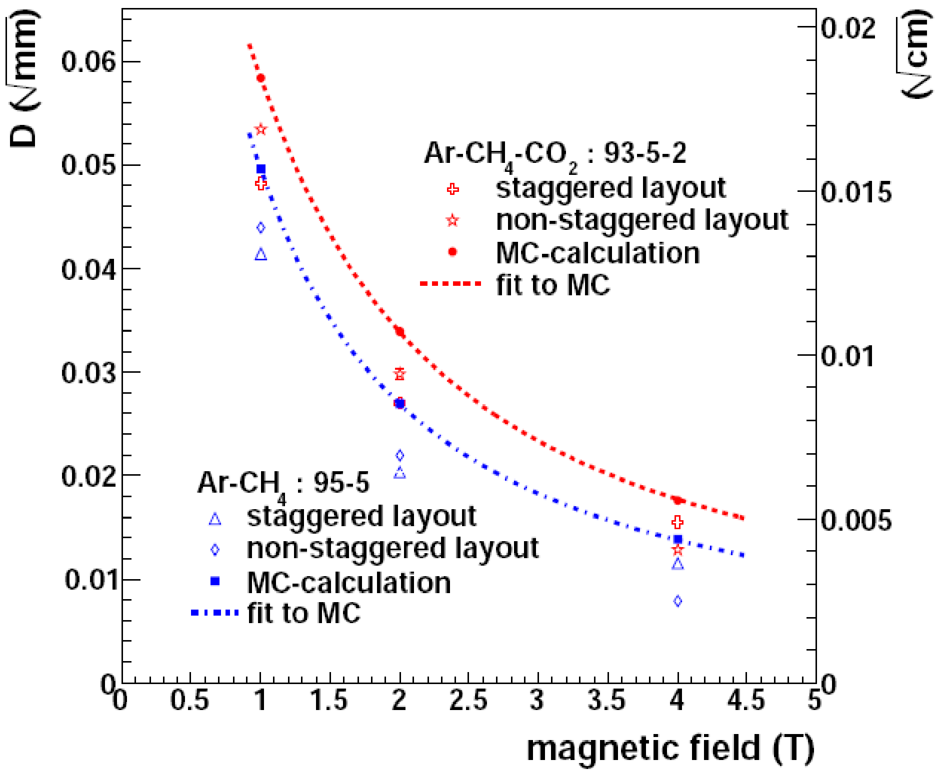


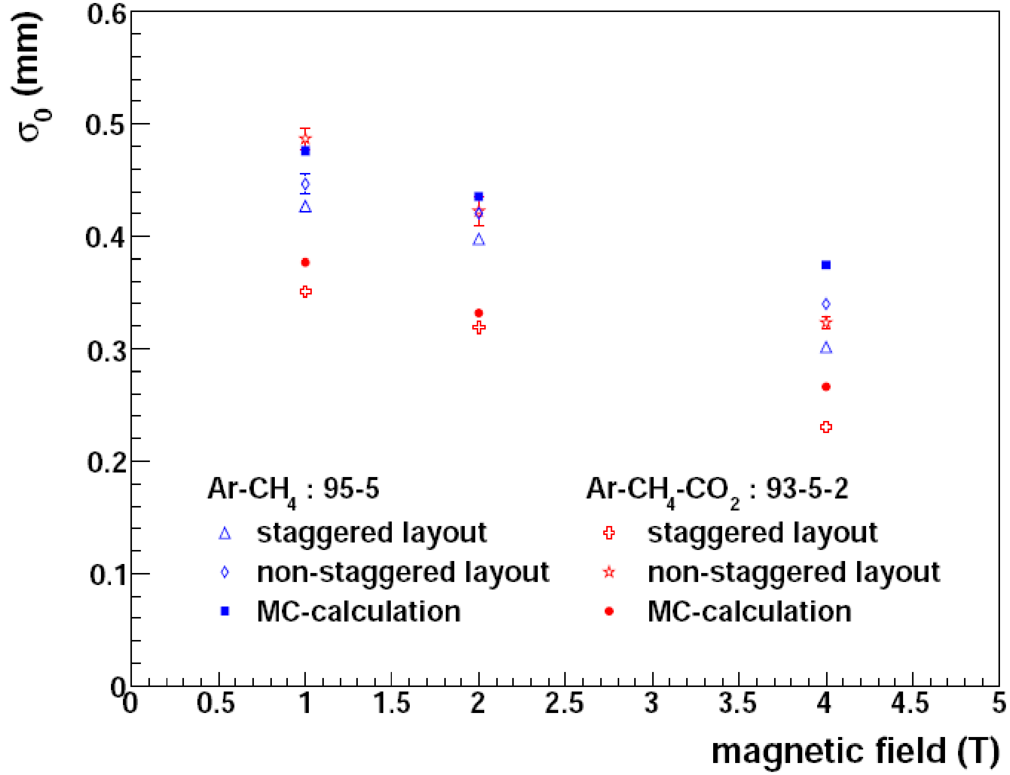


- Fit Gaussian to every interval and get mean values (with errors)
- Width of charge distribution can be described by: $\sigma = \sqrt{D^2 z + \sigma_0}$
- Fit this function with D and σ_0 as free parameters to the mean values of the intervals to get the parameter values

Diffusion Parameters

- Results from simulation compared to results from Global Fit results





- Results from Global Fit are in the right order of magnitude but underestimate the coefficients
  Among other, possible explanation: wrong noise factor (1%), further investigation planned

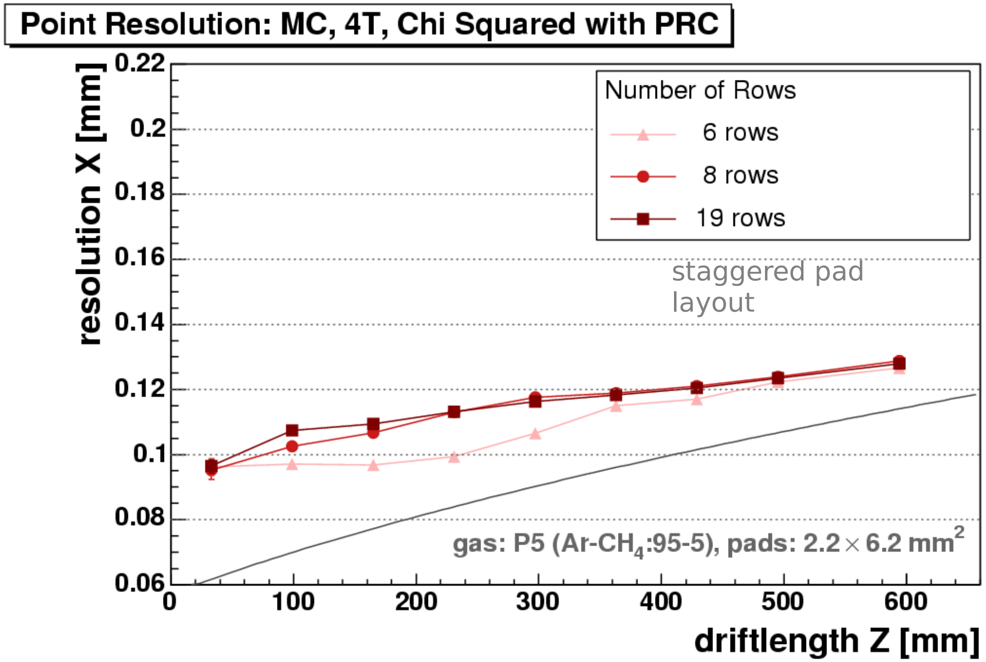
Remarks

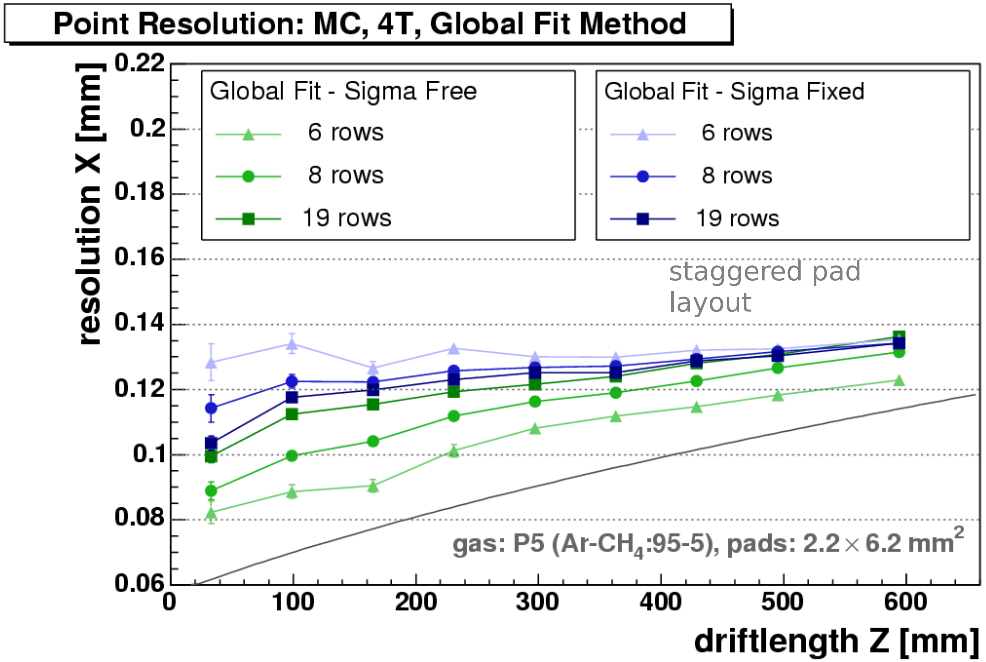
- Global Fit has the advantage that pad response effects are incorporated in the fit function
- Also, missing information (damaged pad) does not affect the fit too much since the term in the sum simply vanishes

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

- Disadvantage: the fit is time consuming with many pad rows: for 6 pad rows Chi Squared Fit and Global Fit need approximately the same time, for 19 rows the Global Fit needs approximately three times longer than the Chi Squared Fit
- If not many pad rows are used, Global Fit can produce too good resolution results
- Width can be fixed during the fit (for a certain Z per row)
- Fit in YZ plane done with Chi Squared straight line fit

Comparison of Fit Methods





- 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

- “Robustness” of the fit methods: influence of damaged pads (dead channels) tested with Monte Carlo simulation for 4T, P5 gas