| TN0010 |
|---|
| **Gaussian Fitting Software** |

Gaussian Fitting

Pyramid products that can read out multistrip position sensing detectors use a software routine that can fit the best Gaussian curve to noisy data. This allows the centroid and width of a transverse beam spot to be determined with much greater precision than the sensor pitch. At typical noise levels, a centroid precision of 0.1 x sensor pitch can be achieved routinely. The fit uses least squares regression to find the best fit to the selected data and is generally more immune to noise than alternatives such as the centre of mass.

The same algorithm can be used for fitting curves that are not true Gaussian but can be considered so if the range of the data is restricted. An example is interpolation of the end of range peak generated by a multi-layer Faraday collector.

Some products (I128, I128S, I6400) execute the algorithm at high speed in their on-board FPGA. Other products (I3200, F3200E) pass the sensor strip data to the PTC DiagnosticG2 software which performs the calculation.

Algorithm

Data for fit

Limit to extend of data for fit / data used for quality measure
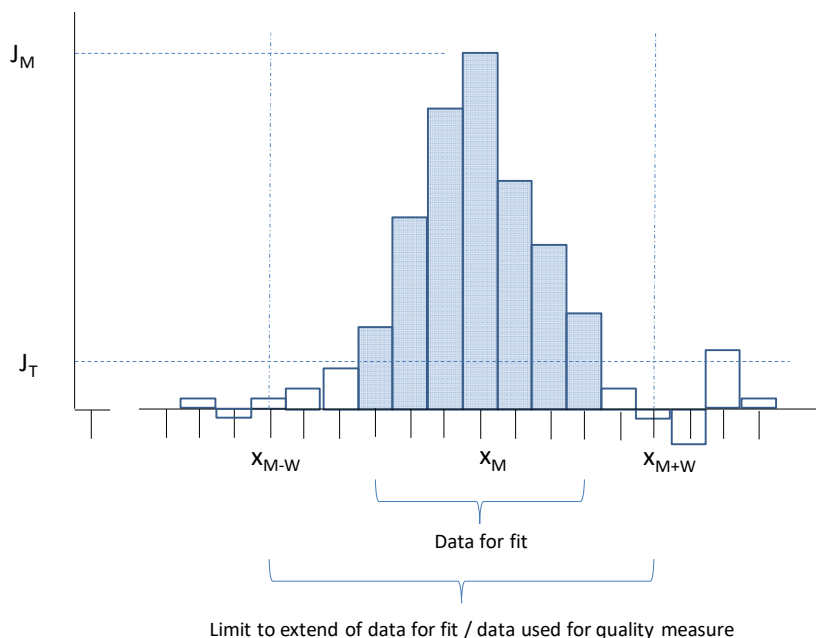
The calculation comprises the following steps in summary:

1. Locate the channel X_M with the maximum signal J_M
2. Select channels either side of the maximum down to a configurable threshold expressed as fraction of the maximum, J_T/J_M. A configurable window caps the maximum distance that the algorithm can step away from

the maximum channel in each direction. This window also sets the span of data that is used to determine the quality of fit.

3. Take logarithms of the selected data and apply weighting.
4. Fit a second order polynominal to obtain the best fit polynominal coefficients.
5. Take exponentials to recover the three coefficients of the Gaussian function that best fits the data.
6. Calculate a chi square quality of fit measure over the points inside the window, expressed as a nominal percentage where 100% is a perfect fit and anything above around 80% is a good fit. This describes how well a Gaussian function models the data.

The detailed implementation includes various measures to maximise accuracy in a 32 bit processor, manage background offsets and handle end effects and wrap around. Note that no iteration is required so the calculation completes in a time known in advance.
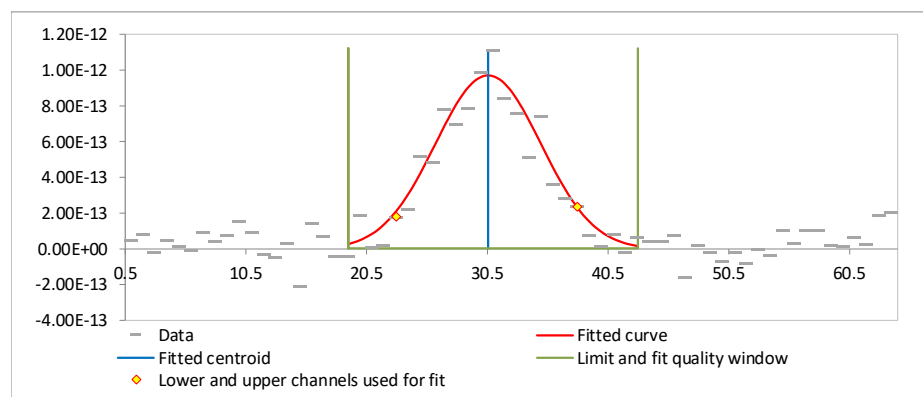
Data
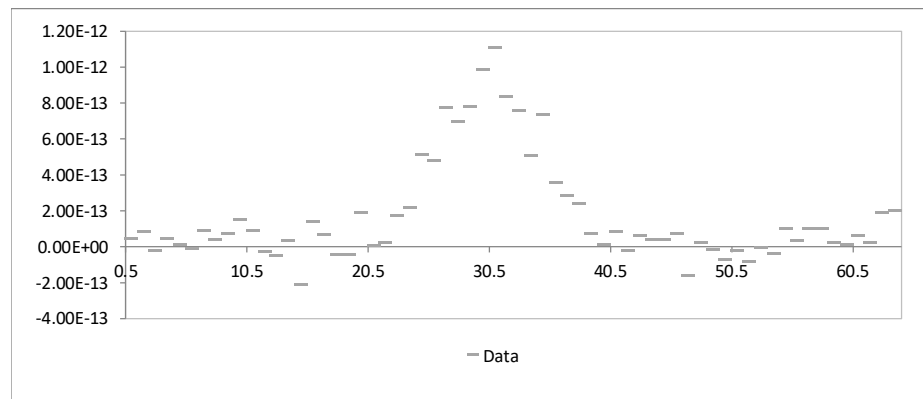
Selecting channels
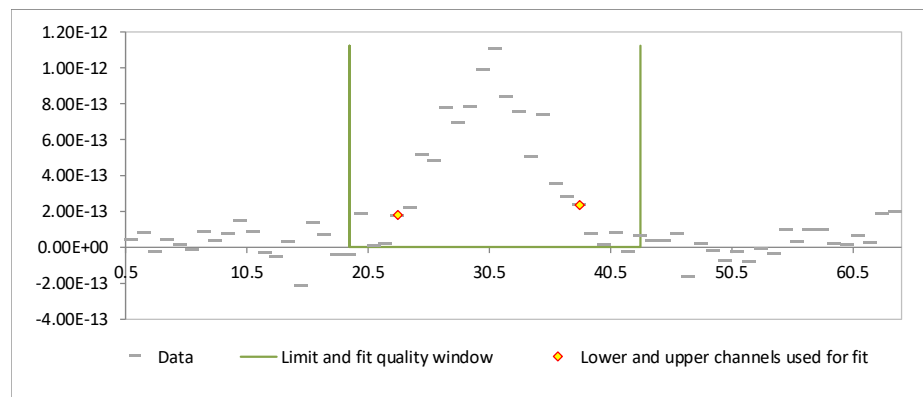
Fit

Control parameters

The control parameters for the fit are exposed in the PTCDiagnosticG2:

Noise threshold: This is the ratio J_T/J_M that defines the range of data channels that will be included in the fit. The algorithm steps out on either side of the maximum until the channel value J is less than J_T. A setting around 5% is typical, although a lower value may be used for narrow peaks with low noise levels.

Window: This is the maximum number of channels M that the algorithm is allowed to step out on both sides of the maximum channel. It should be set approximately to the expected full extent of the peak. A value of 5 for example means that the greatest number of channels that would be included in the fit is 5 + 1 + 5 = 11. The number will be less if the noise threshold is reached beforehand. It also sets the channels that are used in the quality of fit calculation.

Weighting: Weighting should be enabled. This parameter (if exposed to the user) should only be set to 2.

Peak threshold: The fit will not be attempted if the maximum channel value is below this limit.

Sum threshold: The fit will not be attempted if the sum of the selected channel values is below this limit.

The same parameters are included in the xml files that define the behaviour of the Pyramid Scan Dose software used for proton therapy dose delivery.

Handling faulty channels

An advantage of a fitting method over alternatives like centre of mass is that a valid fit will still be obtained if a bad channel is omitted from the fit, provided sufficient number of good channels remain in the fit. There is no requirement that fitted channels are contiguous.
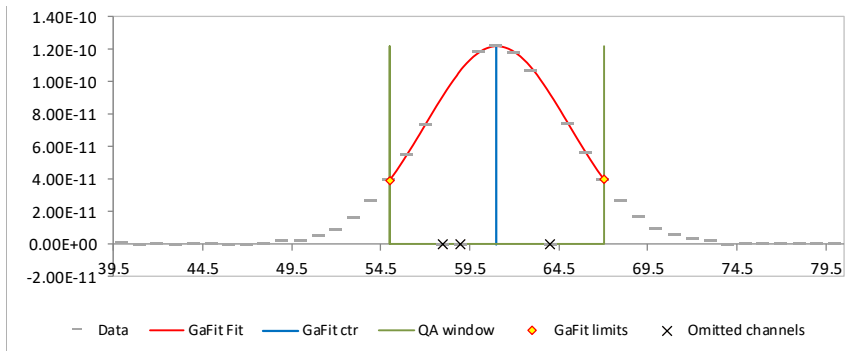
The process is the same as that described earlier, except that the bad channels are not included in the search for the maximum channel, and then they are simply removed from the fitting limits and regression calculations. In this example, channels 58, 59 and 64 are bad, and therefore corrupt the fit to the peak.
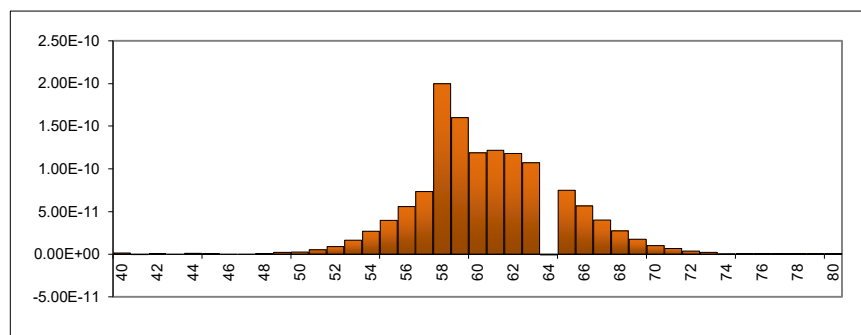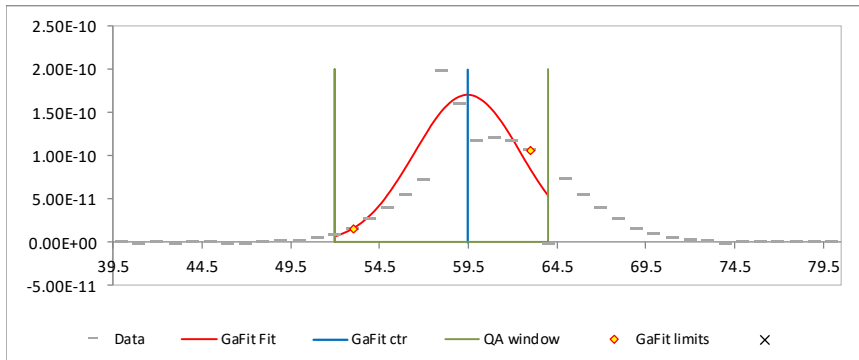
Fit is **Bad**

Omitting those channels allows a valid fit.

Fit is **Good**

Reasons for poor quality or failed fits

The algorithm assumes that a Gaussian curve is a good model for the data. A failed or poor quality fit is invariably because the data does not match that assumption. Nevertheless the fit can often return a good centroid value, especially if the fitting parameters are tuned appropriately. Reasons for bad or failed fits include:
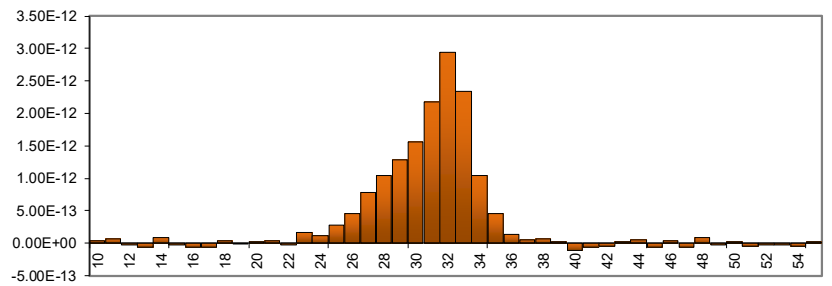
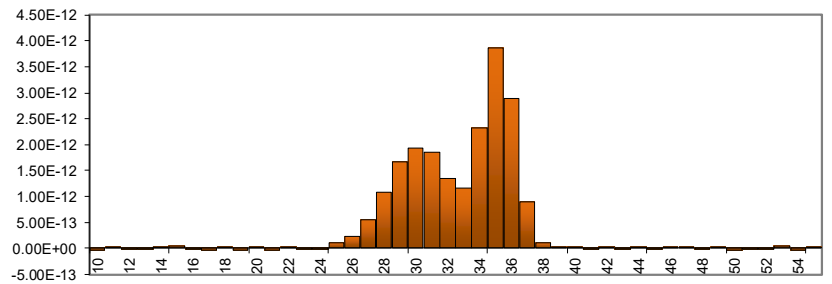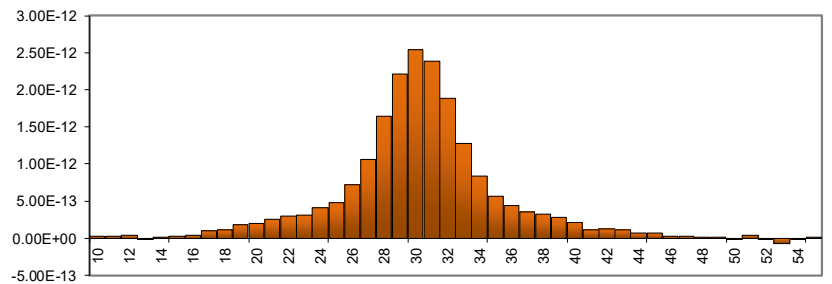| Uncorrected background offset	
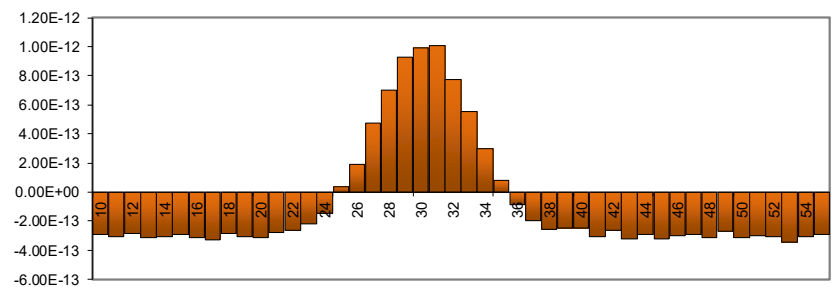
Technical Note

Excess noise
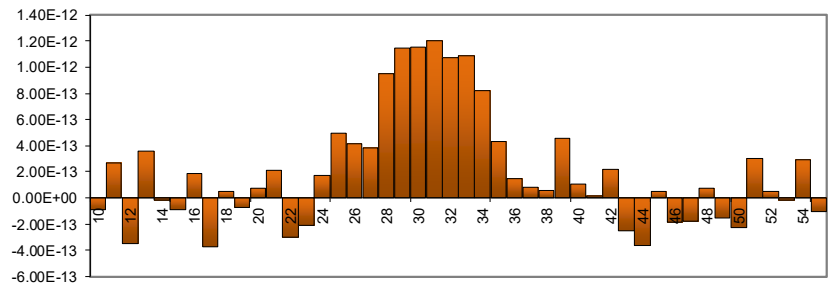
Assymmetric peak

Multiple peaks

Extended halo

Undersampled