Statistical Interpretation Problems

Group 14

J. Calcutt, S. Henry, K. Mulder, C. Sarasty

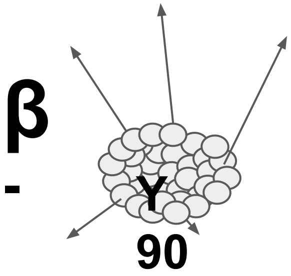
Introduction

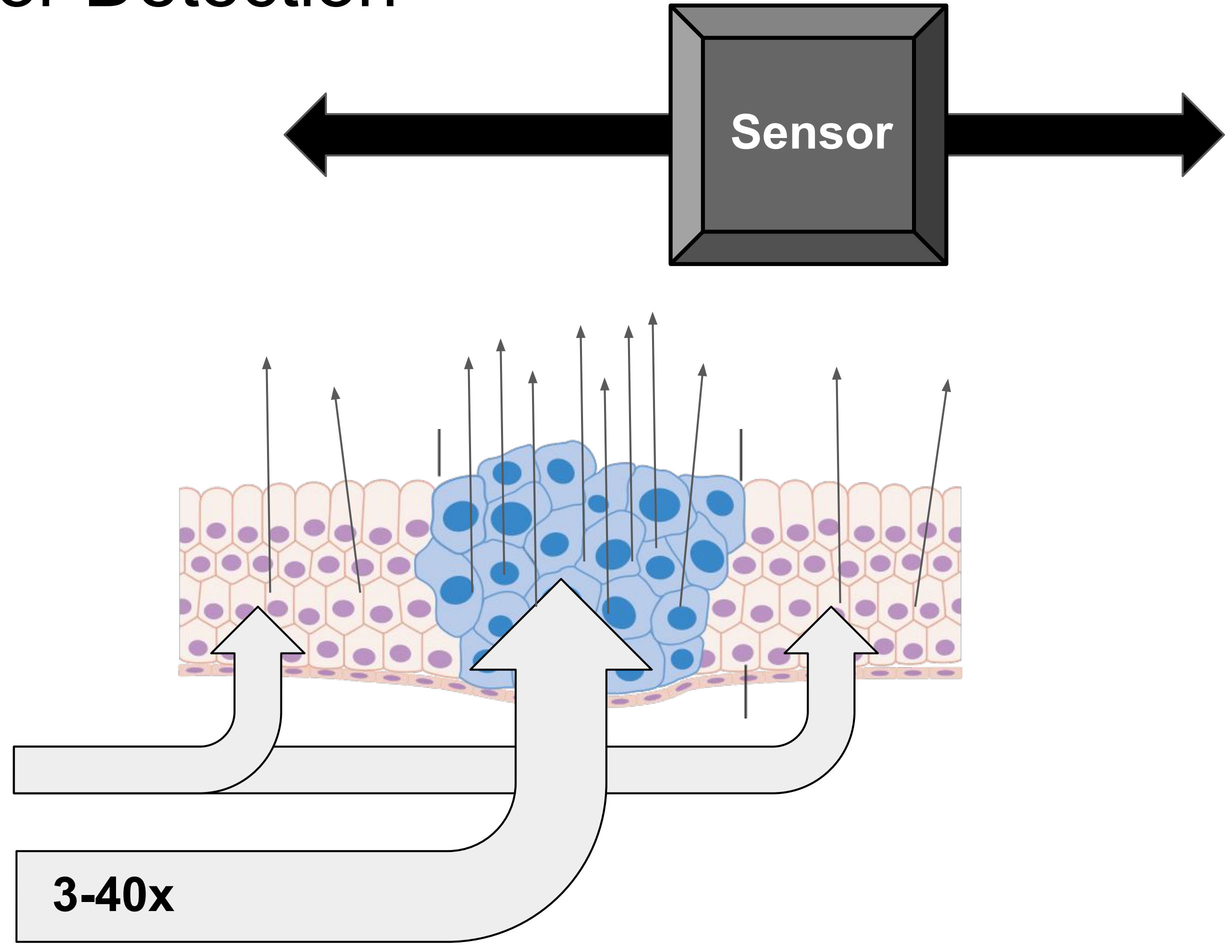
Our exercise is to point out any problems with this paper's statistical interpretations

- L. Alunni Solestizi et al., 2018 JINST 13 P07003
- Instrumentation paper regarding a sensor used for the detection of tumors with the use of radiopharmaceuticals
- https://iopscience.iop.org/article/10.1088/1748-0221/13/07/P07003
- Spoiler: there's loads

Background - Tumor Detection

Yttrium 90





MT9V011 Sensor

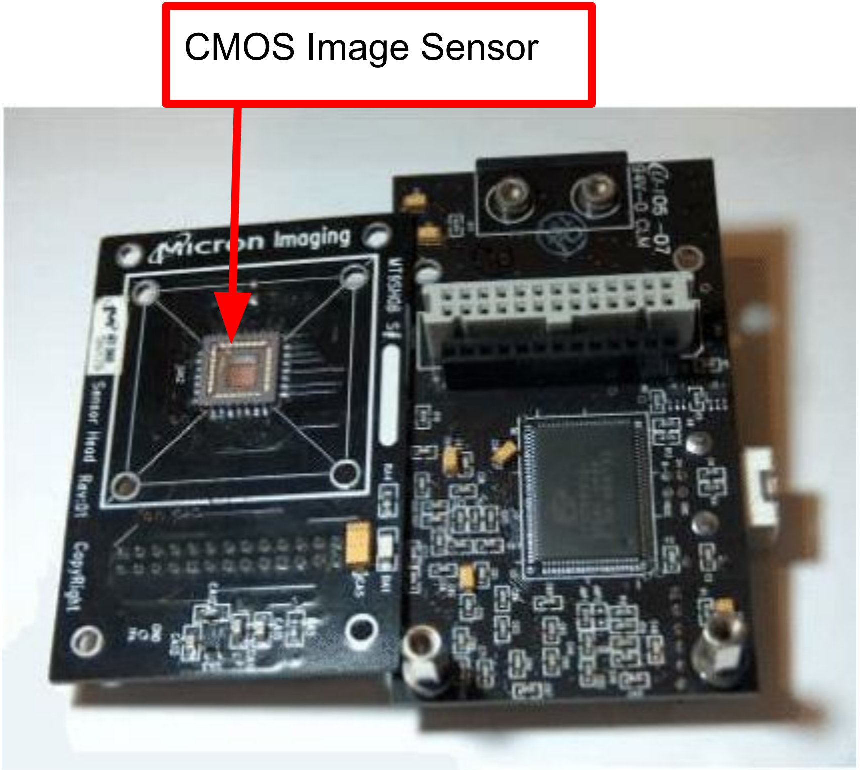
CMOS image sensor

- Ionization activates transistors in pixels

640 x 480 (307,200) Pixels

The group had 2 sensors

- One with a filter to protect the sensor
- One without to “detect lower energy electrons” → Damaged electronics from **9 years of irradiation**



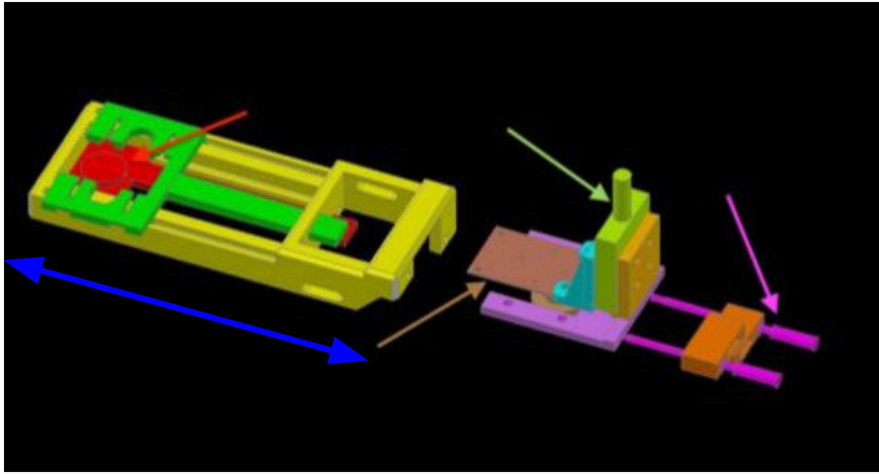
Experimental Setup

Red: Radioactive Source

- Various collimators placed over source

Brown: Sensor

- 2 heights above source (2,8mm)



The sensor slides laterally (blue direction) to collect data at different positions -- controlled by 2 screws (purple)

The authors make an attempt to determine the location of the source with the maximum count number for various

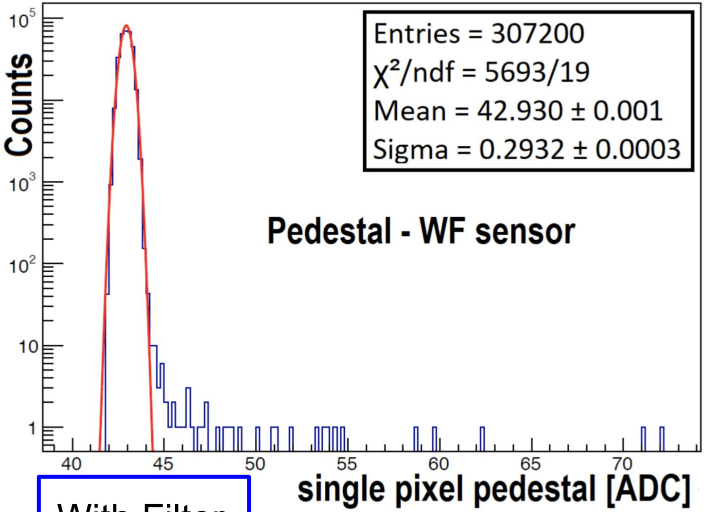
Electronics “Calibration”

They took data with the sensor in a dark box and without a source

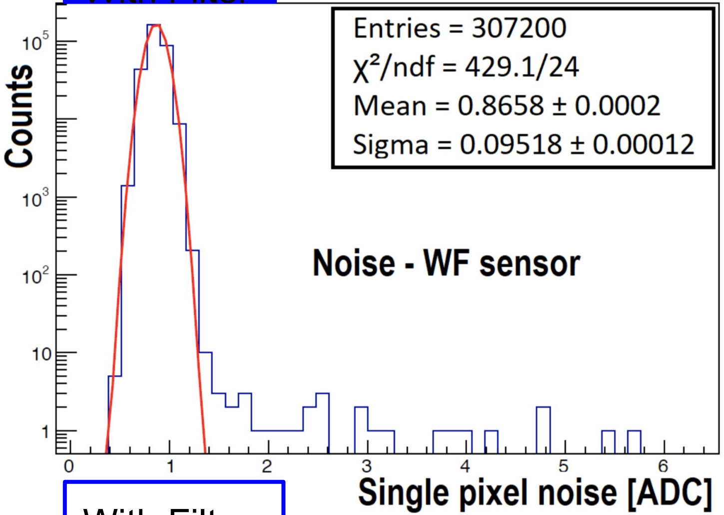
“Analyzed” the pedestal and noise for each channel

- Both With Filter (WF) and #NoFilter (NF)

For WF: χ^2/ndof are huge → Bad Fits



With Filter



With Filter

Entries = 307200
χ^2/ndf = 5693/19
Mean = 42.930 ± 0.001
Sigma = 0.2932 ± 0.0003

Entries = 307200
χ^2/ndf = 429.1/24
Mean = 0.8658 ± 0.0002
Sigma = 0.09518 ± 0.00012

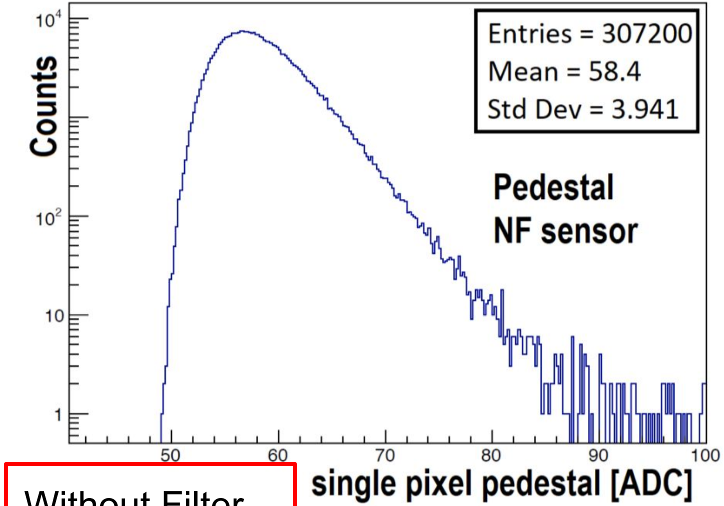
Electronics "Calibration"

They took data with the sensor in a dark box and without a source

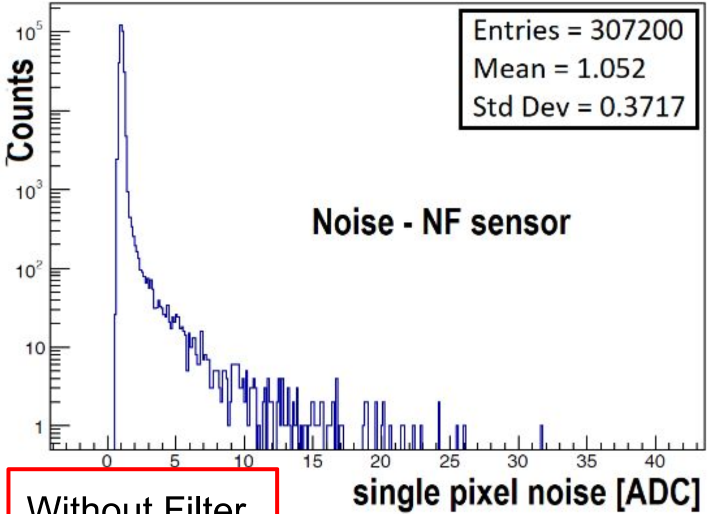
"Analyzed" the pedestal and noise for each channel

- Both With Filter (WF) and #NoFilter (NF)

Make no attempt to see if/how noise and pedestal values correlate (i.e. make a plot of pedestal vs. noise)



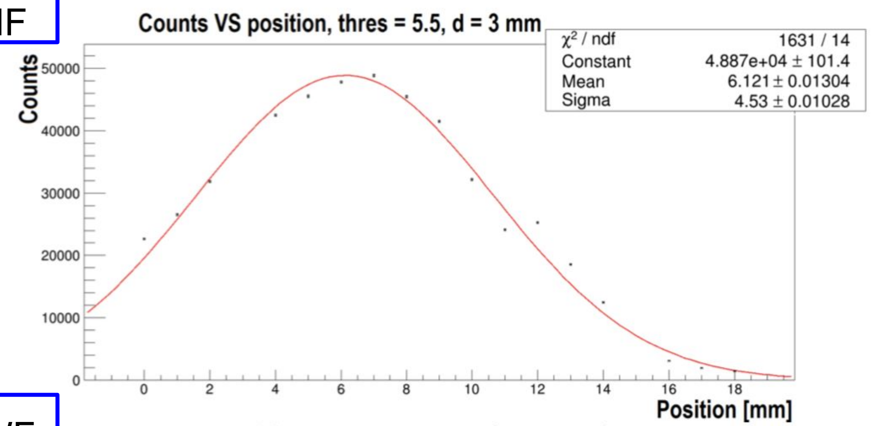
Without Filter



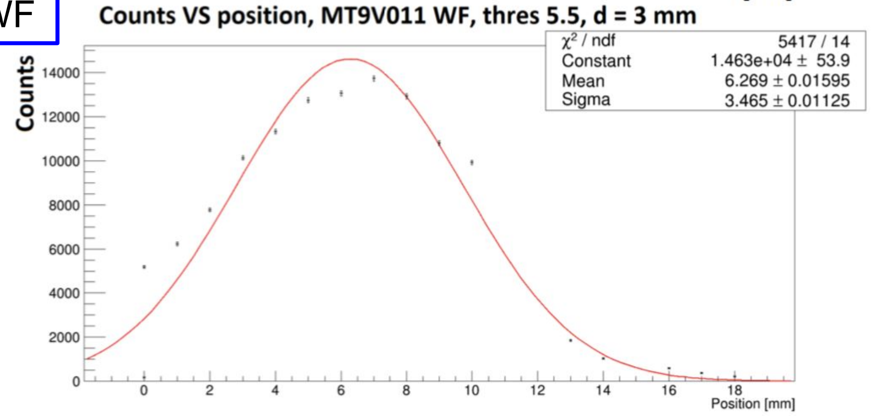
Without Filter

Gaussian Fit & Statistical Errors

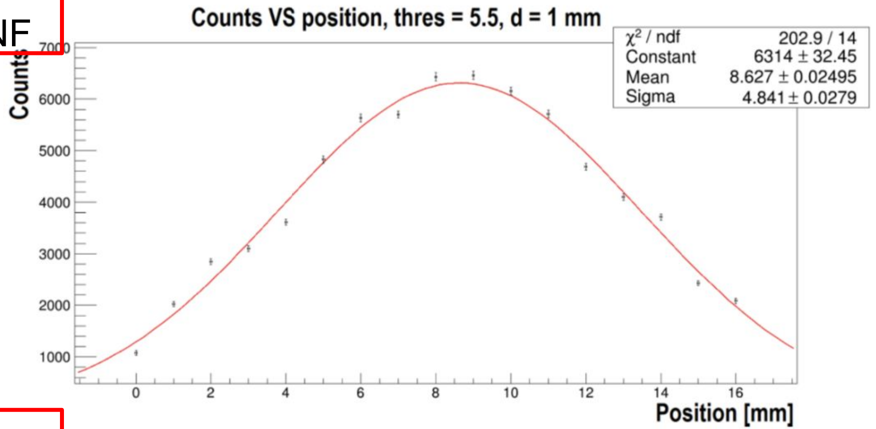
NF



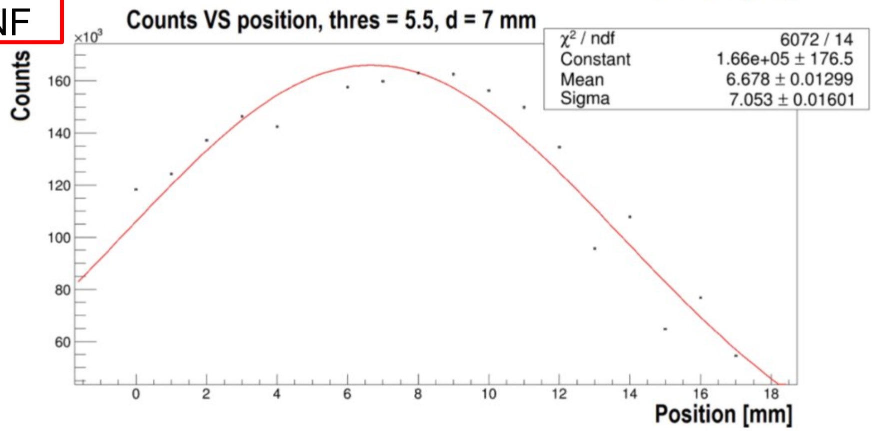
WF



NF



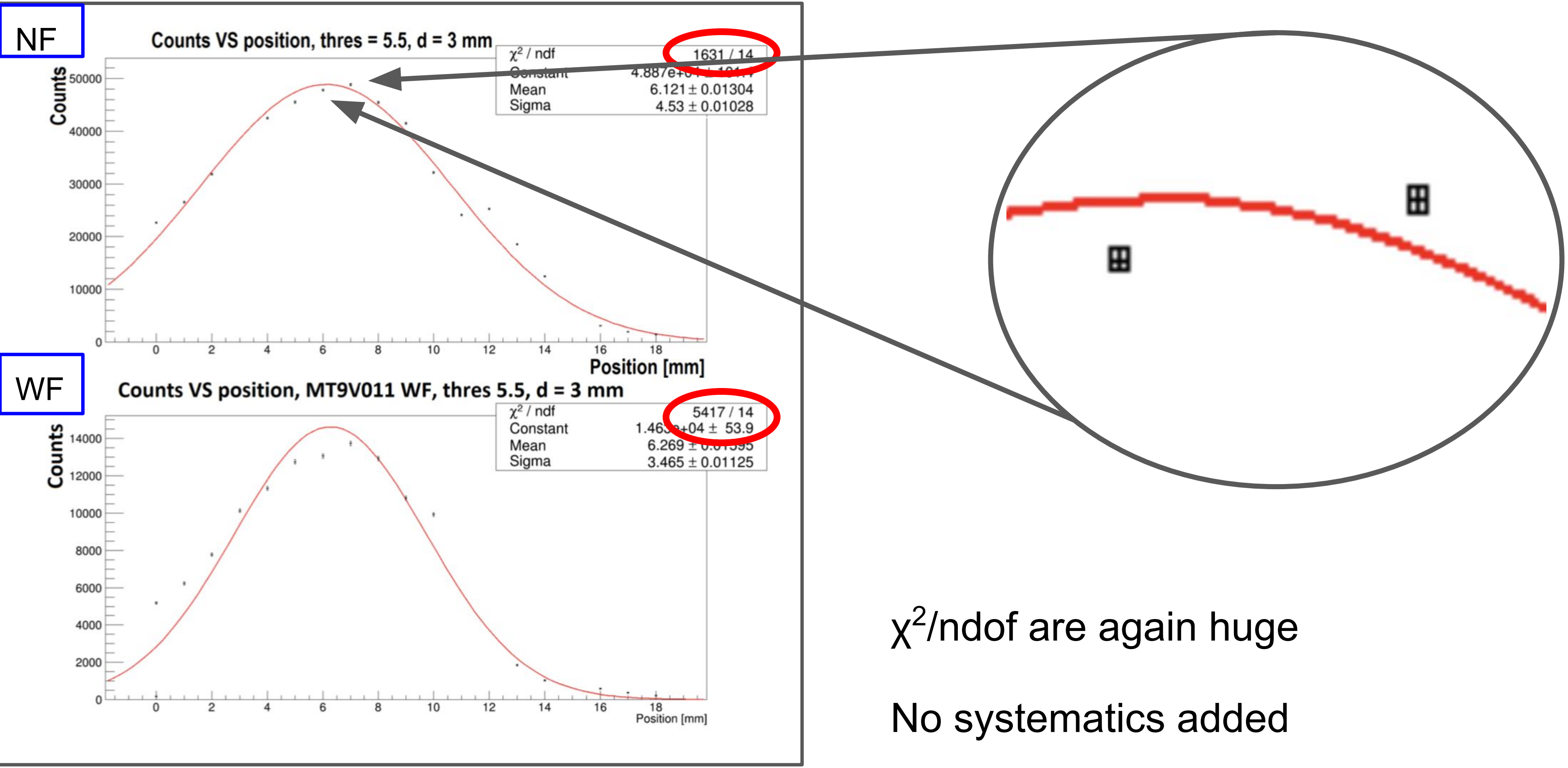
NF



Source-sensor distance 8mm, collimator 3mm diameter

Source-sensor distance 8mm, collimator 1mm (above), 7mm (below)

Gaussian Fit & Statistical Errors



χ^2/ndof are again huge

No systematics added

Source-sensor distance 8mm, collimator 3mm diameter

Gaussian Fit & Statistical Errors

The values of the maxima (in cps) and σ resulting from the gaussian fits are reported in table 1 for a source-sensor distance of 2 and 8 mm, associating their statistical errors, and for each collimator. The statistical error on σ is dominated by the uncertainty on the relative source-sensor position. The last row of the table refers to the acquisition using the non-collimated source with the sensor positioned centrally with respect to the source.

Measurement uncertainties

No explanation of where the uncertainties come from

	$h = 2\,\text{mm}$	
Collim. diameter [mm]	Counts peak [cps]	σ [mm]
1	87 ± 2	1.97 ± 0.05
3	689 ± 11	2.16 ± 0.04
5	1512 ± 18	2.55 ± 0.03
7	2239 ± 20	3.07 ± 0.03
	$h = 8\,\text{mm}$	
1	10.52 ± 0.05	4.84 ± 0.03
3	81.5 ± 0.2	4.53 ± 0.03
5	101.8 ± 0.4	6.15 ± 0.03
7	276.7 ± 0.3	7.05 ± 0.03
non-collim.	442 ± 0.9	—

χ^2 / ndf	6072 / 14
Constant	$1.66e+05 \pm 176.5$
Mean	6.678 ± 0.01299
Sigma	7.053 ± 0.01601

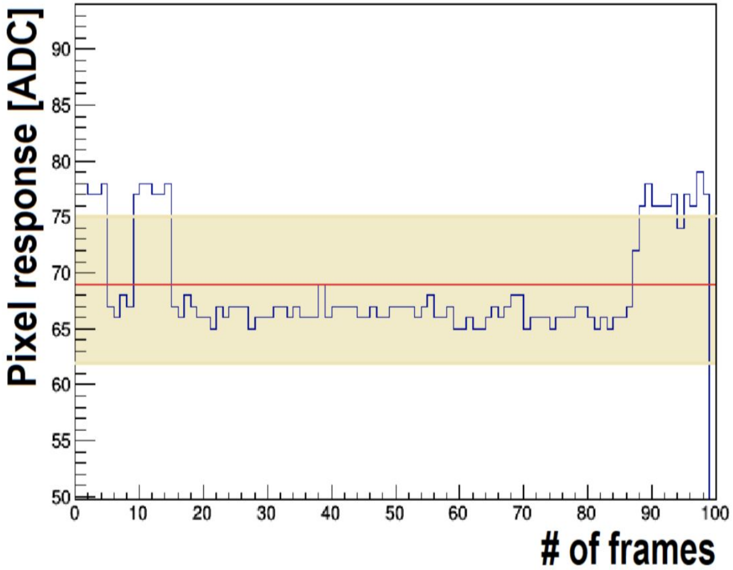
Conclusion

There are plenty questionable things within the paper

- Bad fits
- Virtually no treatment of uncertainty
- Obtuse procedure
- Lots of hanging threads and dead ends

There is a lot more that we could not fit in within the time constraints

My Favorite Line in the Paper



critical pixels in 100 frames, and all of them have shown an anomalous response. An example of such a pixel is reported in figure 9. One can observe the two-value response of this pixel, probably due to some electronic instability. The response goes out the range of 5.5 σ around the pedestal (coloured

Thanks

Backup: Bad pixel removal method

Calculate the average times a pixel goes over the set threshold of 5.5 sigma (average size of the noise in the pixel). Call this value lambda.

Get the number of times an individual pixel goes over threshold. (n)

Arbitrarily chosen:
If lambda is greater than 2, pixel is bad if n>lambda +7sqrt(lambda)
If lambda is less than or equal to 2, pixel is bad if the value 1-p is lower than epsilon (which they don't give).

P is given by the summation of the the poisson distribution.