GRAEME WINTER
CCP4 / DIAMOND WORKSHOP 2022

DATA COLLECTION STRATEGIES FOR PHOTON COUNTING DETECTORS

OVERVIEW

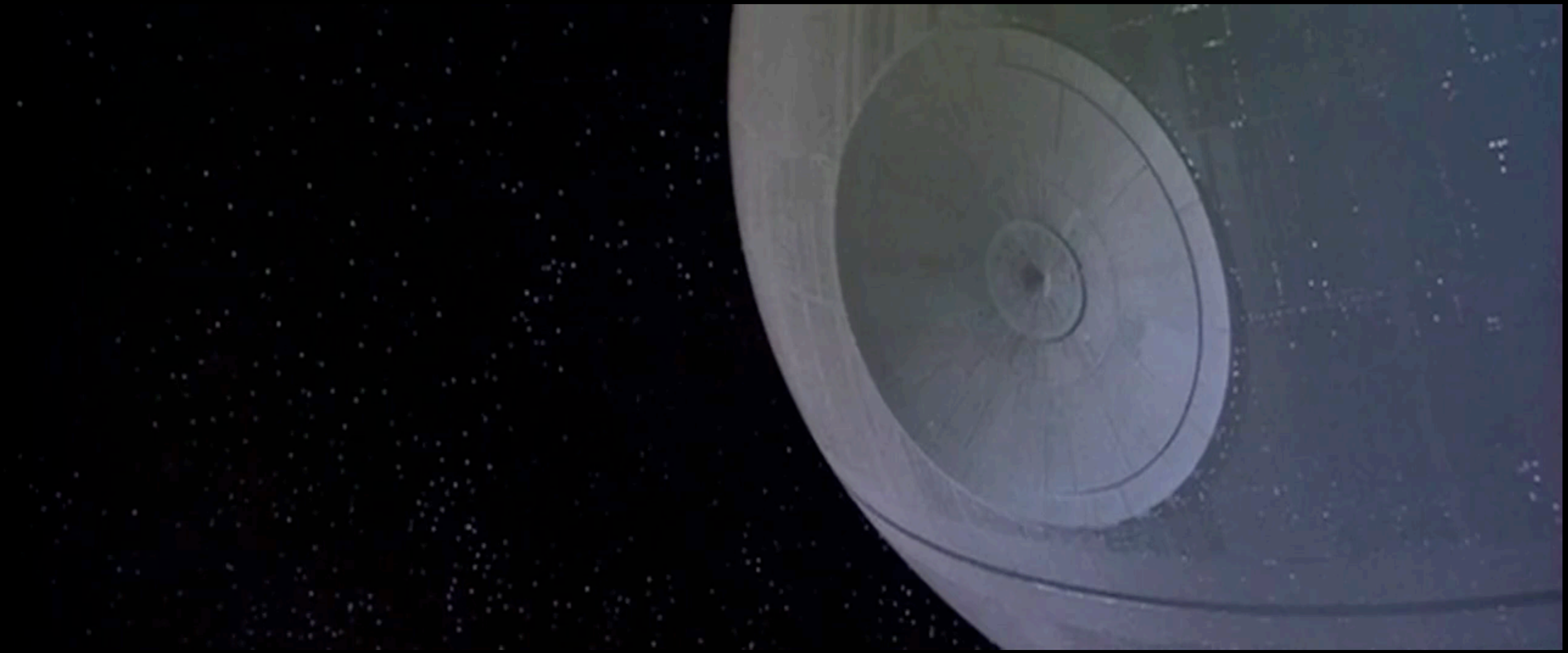
- Before we start advice: if you remember two things…
- Data collection strategies
- Detectors
- Photon counting detectors
- Diffraction data collection advice
- Tactical suggestions
- Conclusions

IF YOU ARE GOING TO REMEMBER ONE THING…

Ask your beamline scientist / local contact for advice

“What would you suggest?”

AND IF YOU CAN REMEMBER TWO THINGS



DON'T BE DARTH

DATA COLLECTION STRATEGIES

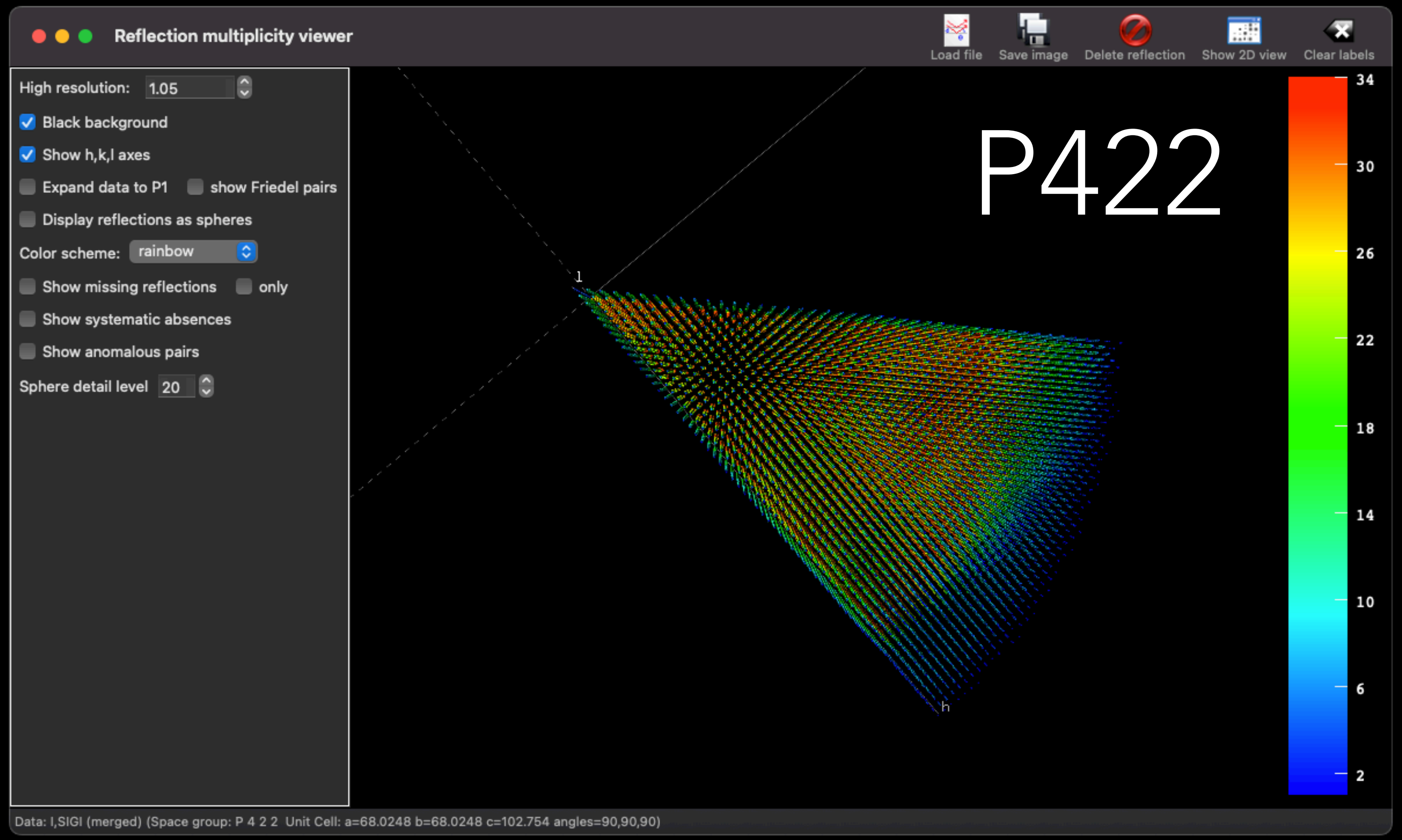
GOALS FOR DATA COLLECTION

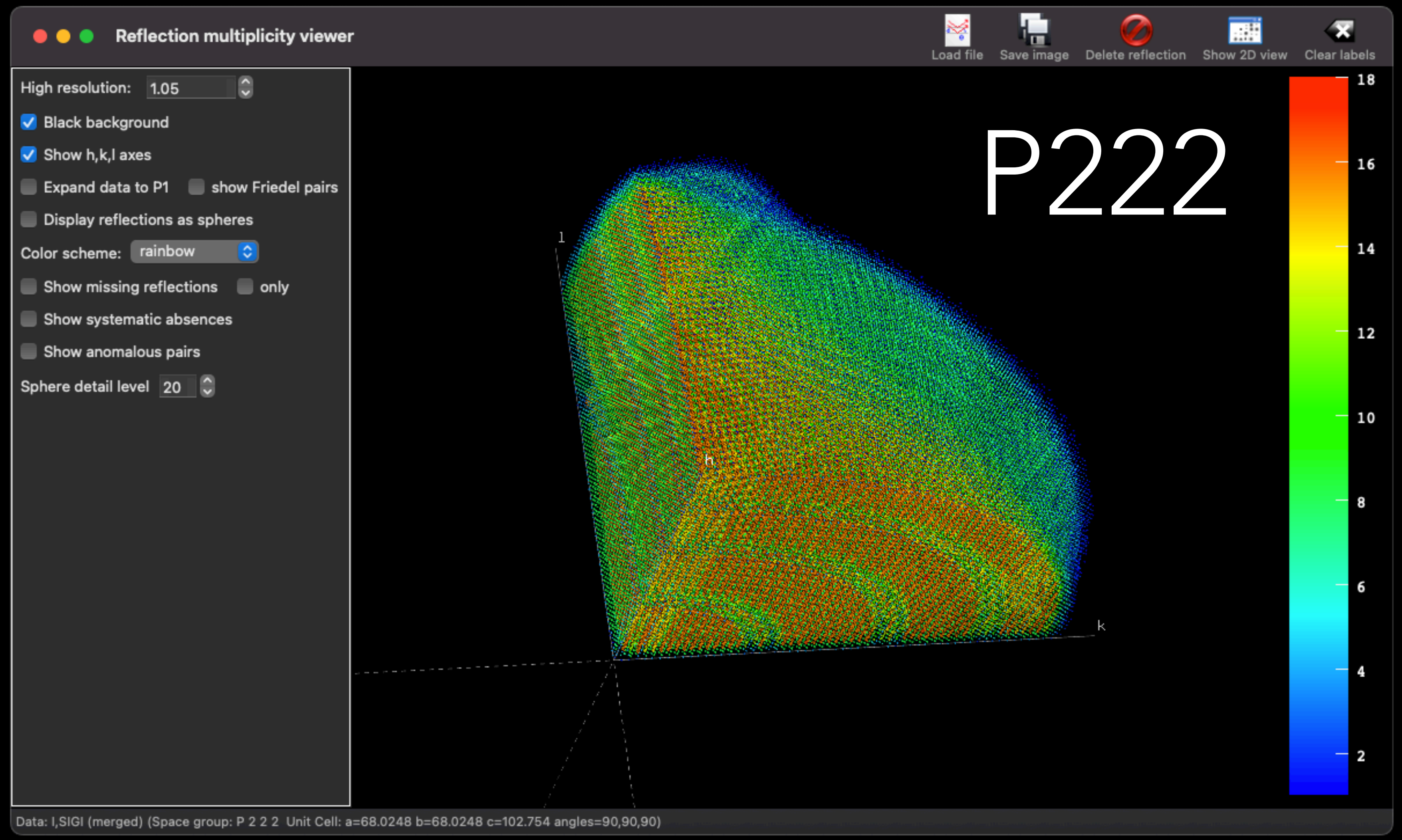
- Measure all the unique reflections
- Measure enough copies that you can assess how well they agree / determine symmetry etc.
- Maximise the “quality” of the signal you record
- Minimise the radiation damage (quantitive discussion of this will follow in subsequent lectures)

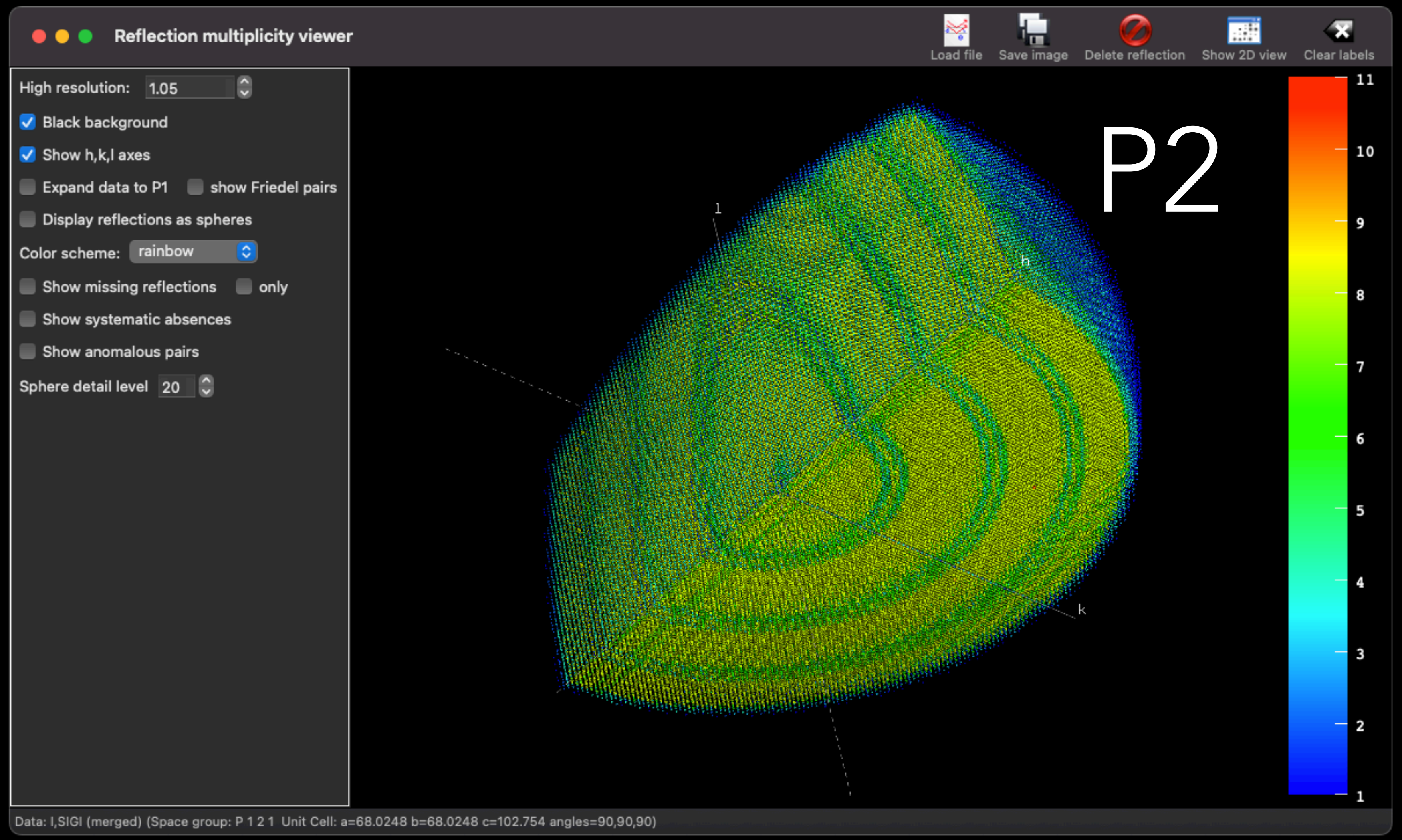
UNIQUE REFLECTIONS: GEOMETRIC STRATEGY

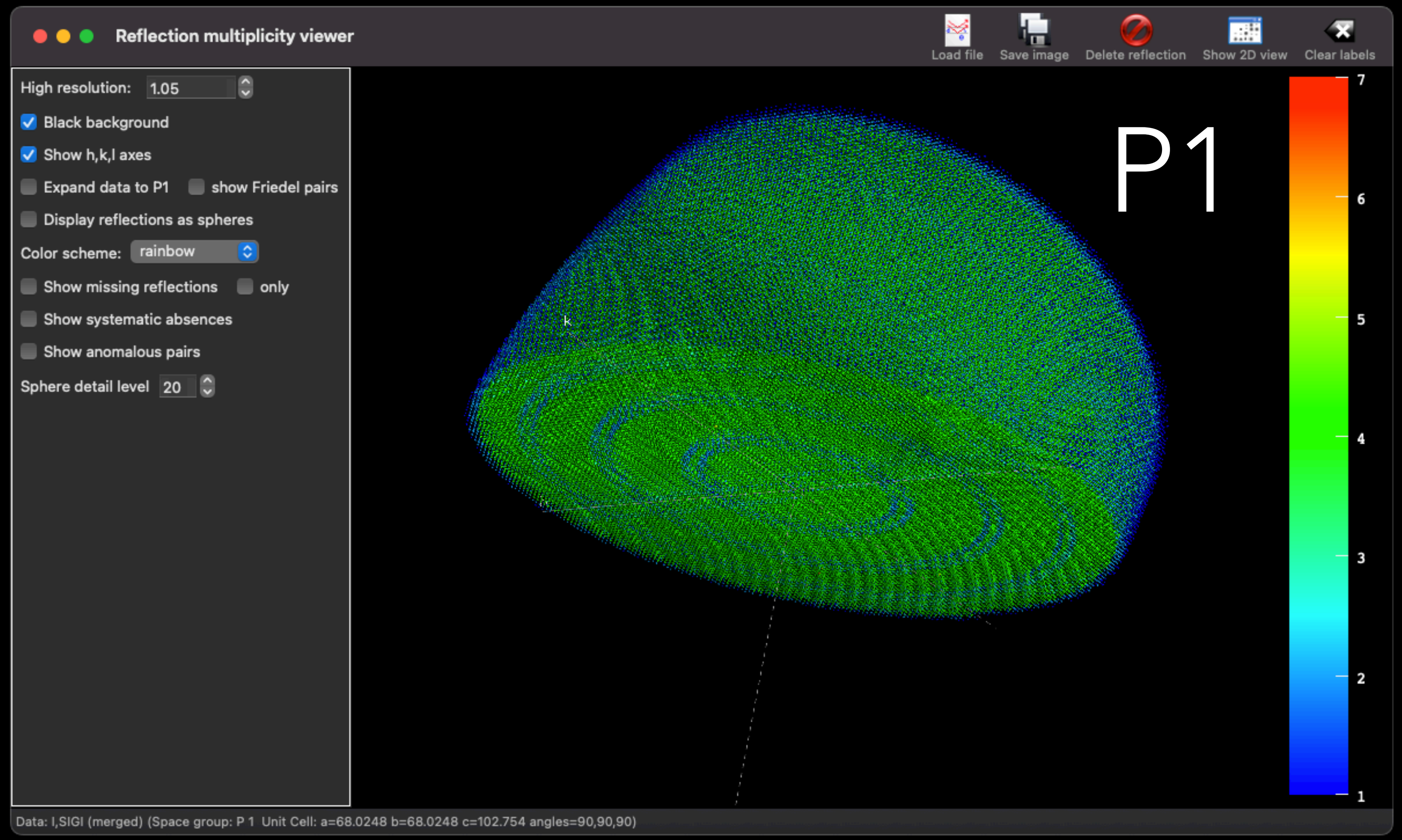
- Asymmetric unit: whole volume of reciprocal space “divided by” the crystal symmetry - you want all of this ideally
- For high symmetry crystals can get away with a relatively small rotation to get complete data
- For low symmetry data may need much more
- More measurements of reflections also improves the quality of the average

ASYMMETRIC UNITS

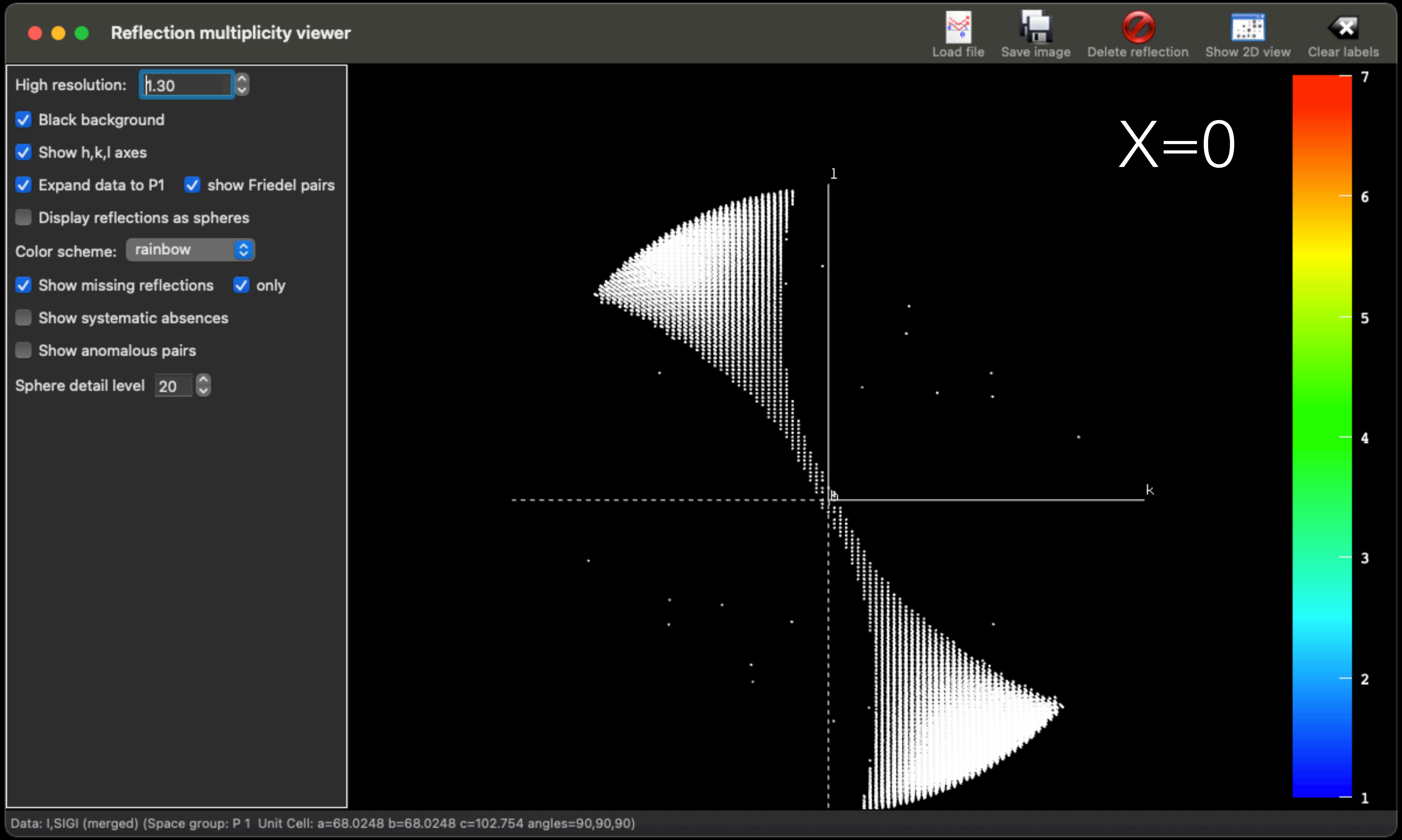




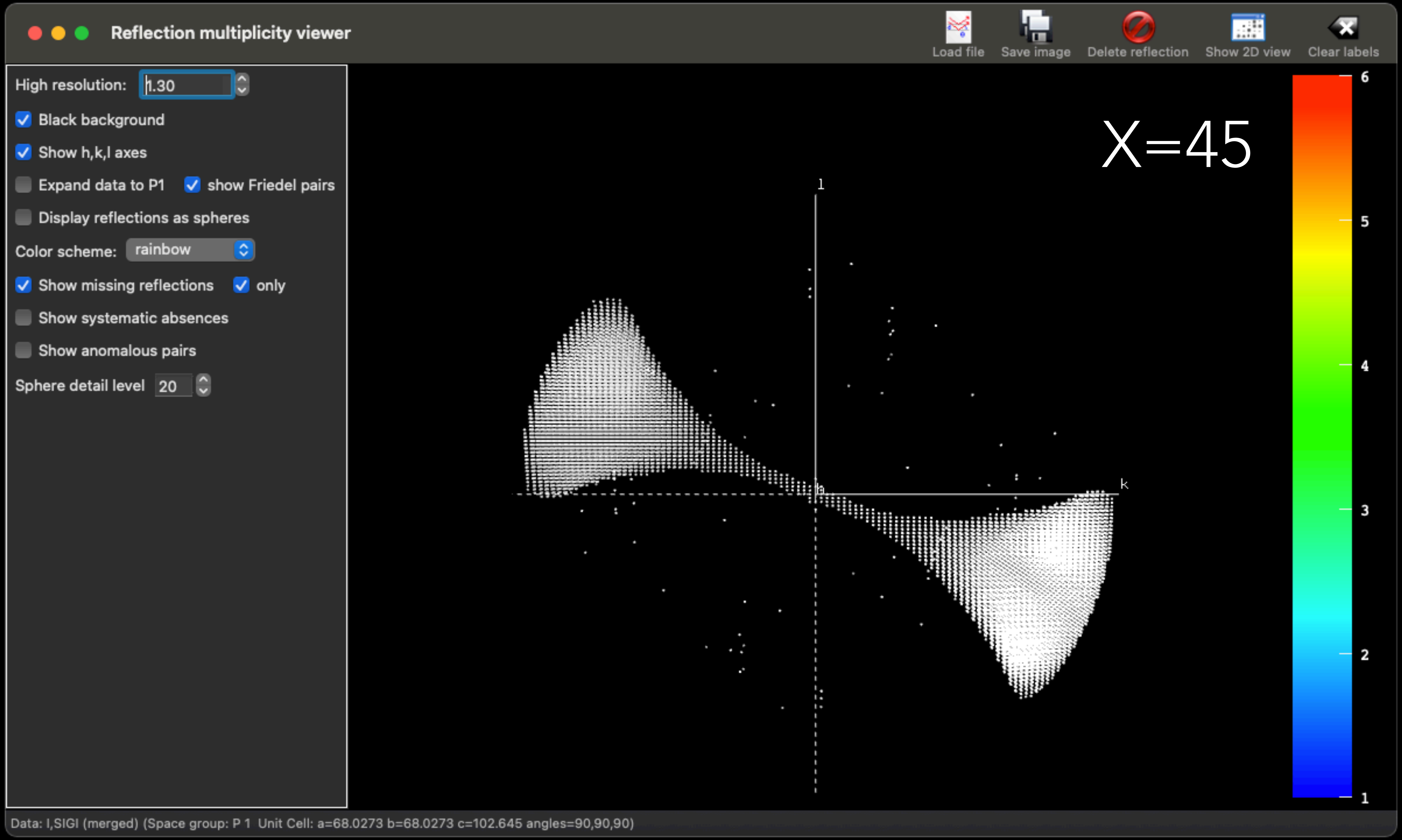




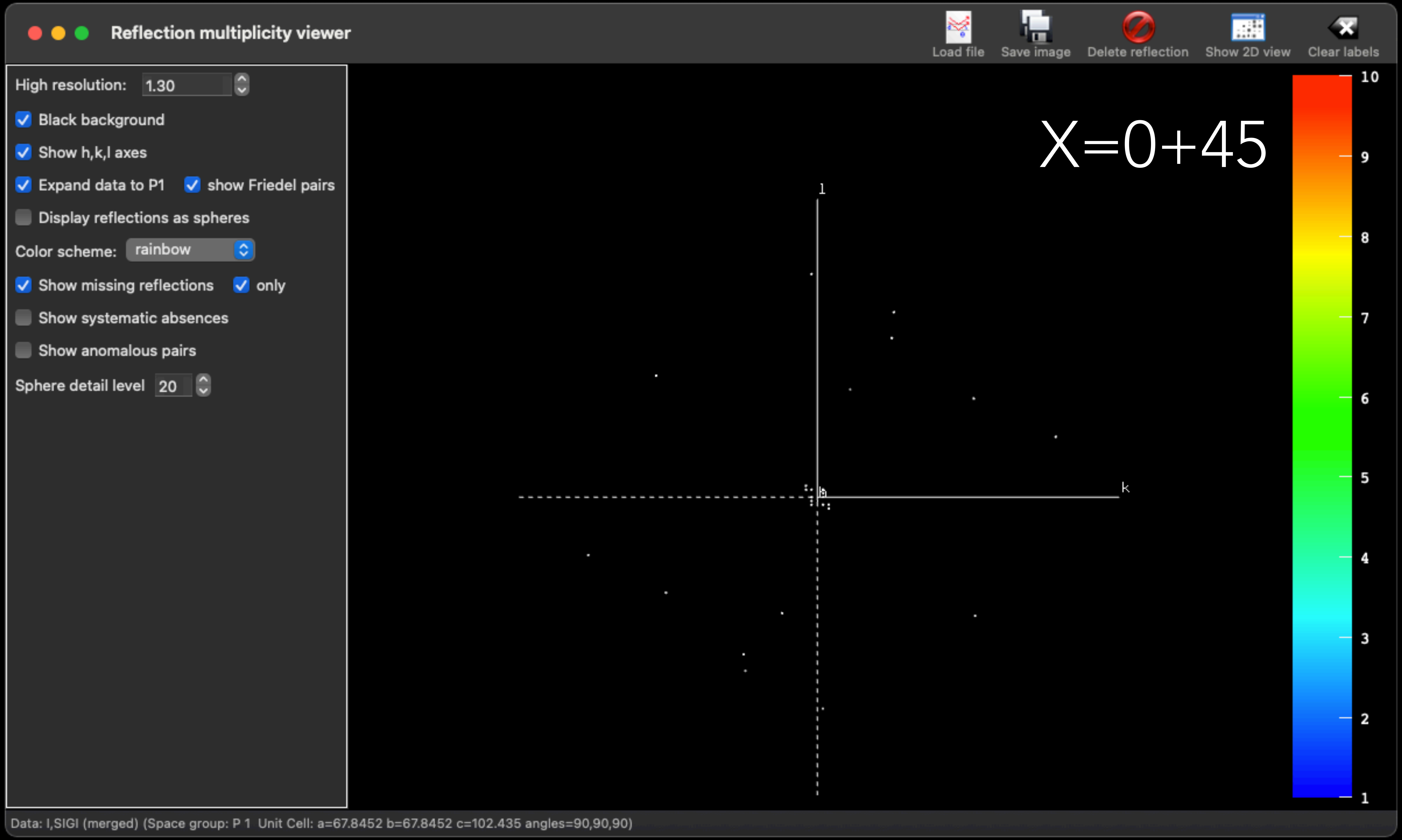
BLIND REGIONS - ISSUE IN LOW SYMMETRY



BLIND REGIONS - ISSUE IN LOW SYMMETRY



BLIND REGIONS - MULTIPLE ORIENTATIONS



RULES OF THUMB - ASSUMING PIXEL ARRAY DETECTOR

- Unless you have a specific reason not to, collect no less than 360° of data
- If you can, collect with more than one sample orientation w.r.t. beam
- Merging multiple data sets is easy and has ancillary benefits - better coverage of the absorption surface, truly independent measurements of reflections etc.
- Cost of more data with a lower dose is simply in processing time - and how long do you spend getting your samples

LEGITIMATE REASONS FOR < 360°

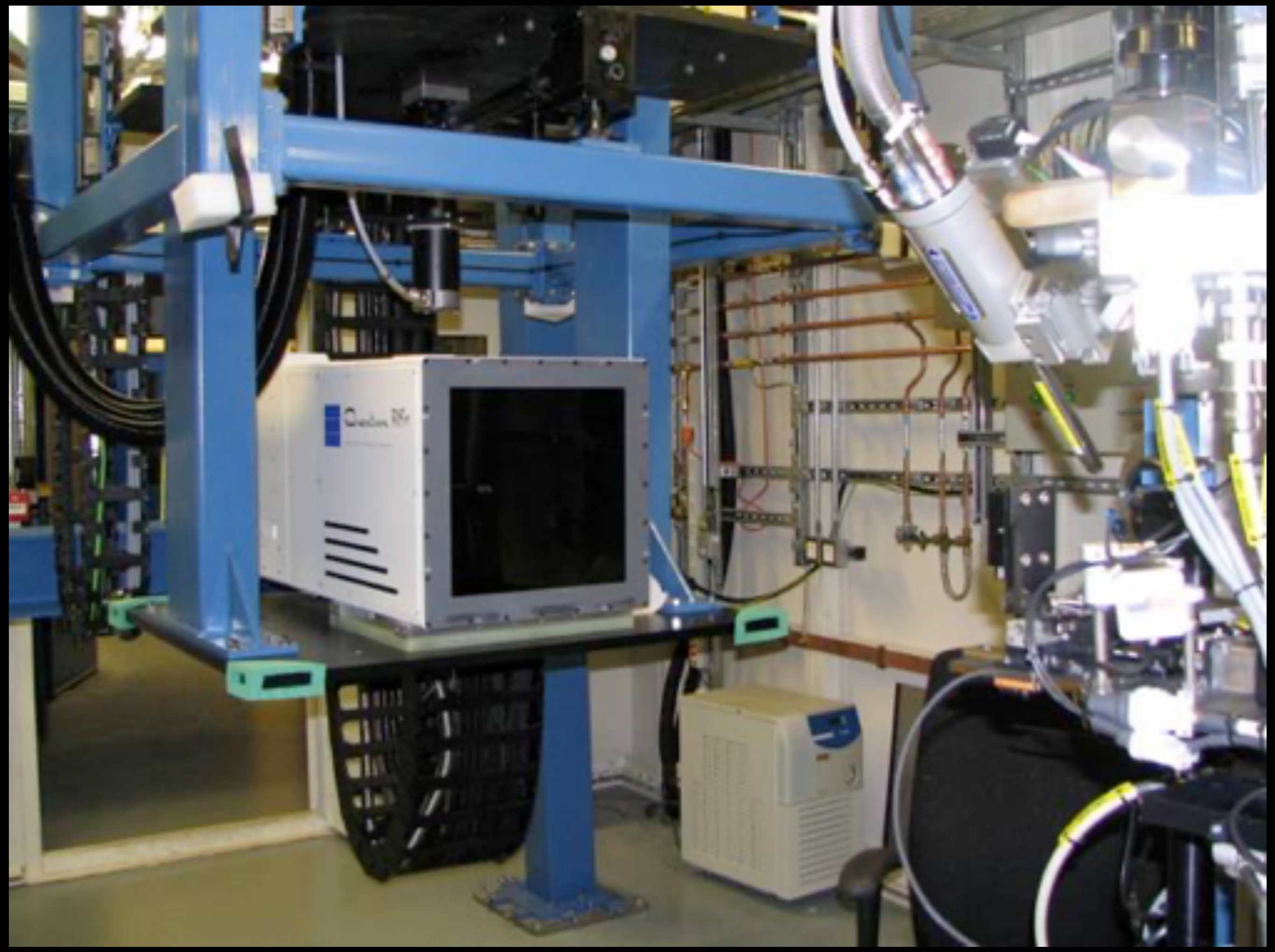
- Sample mounting precludes full rotation - in situ / many samples on mesh
- I know precisely what I am doing for this fragment screen - 120 / 180° is known to be enough

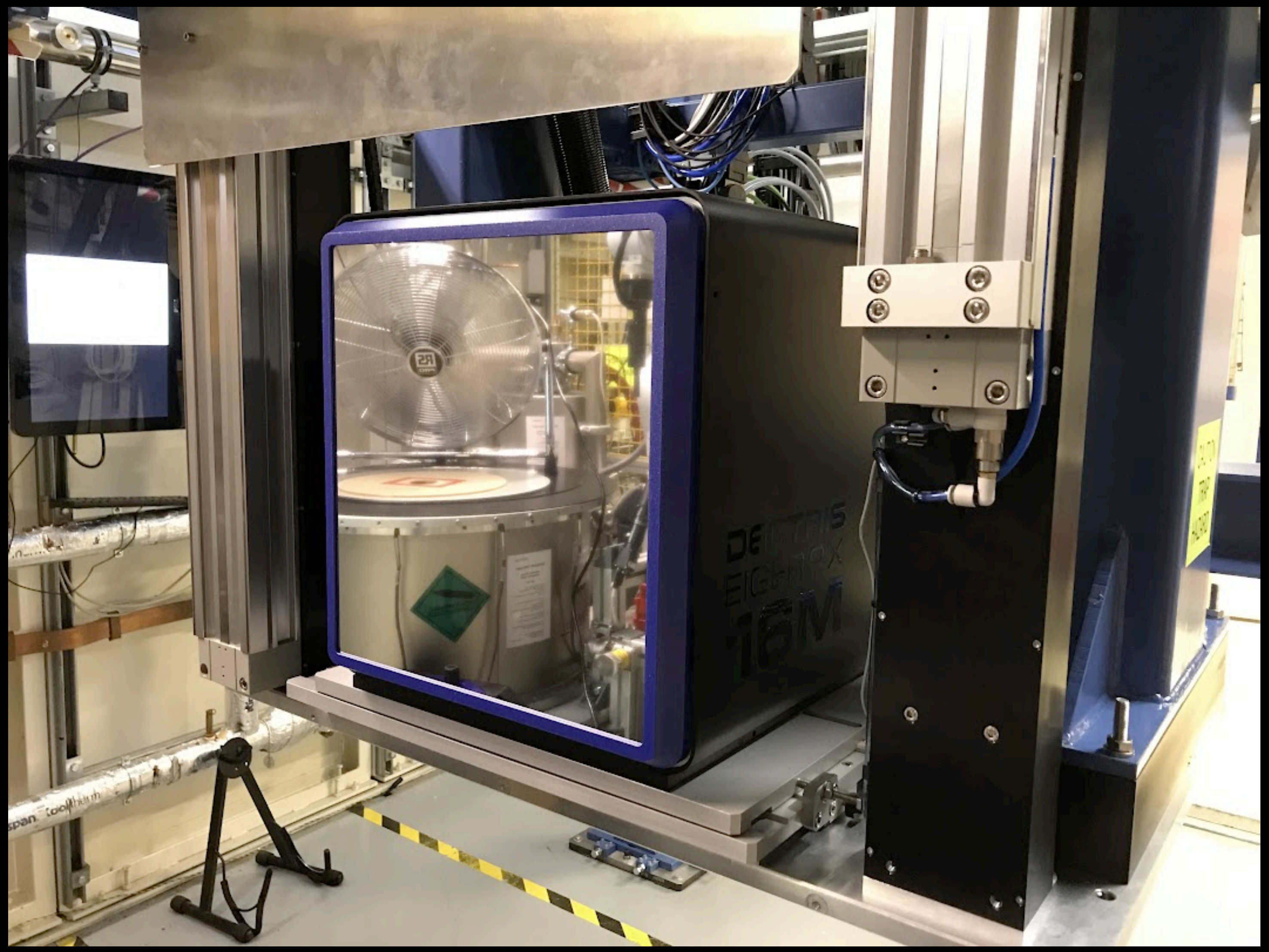
DETECTORS

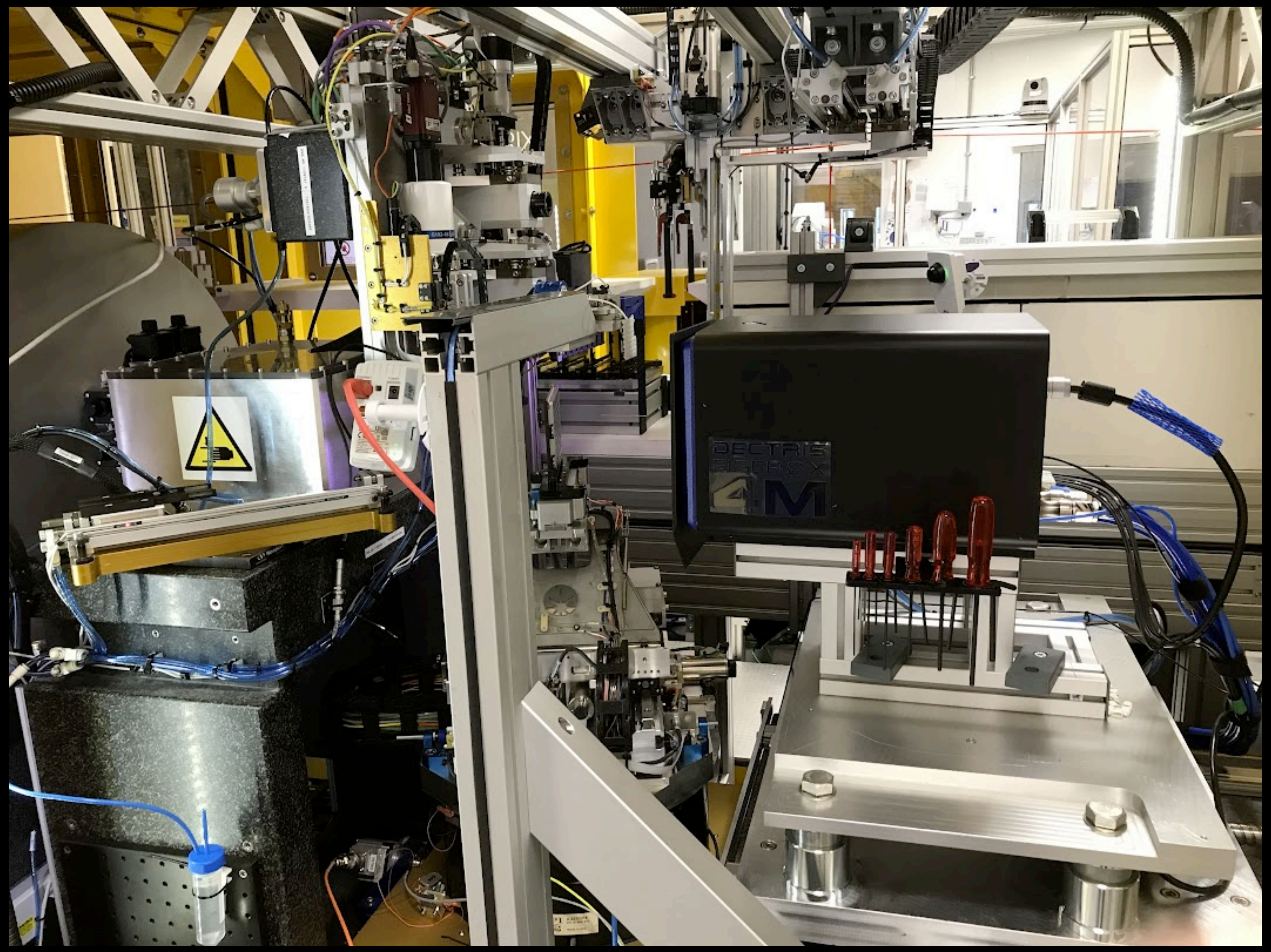
DETECTORS

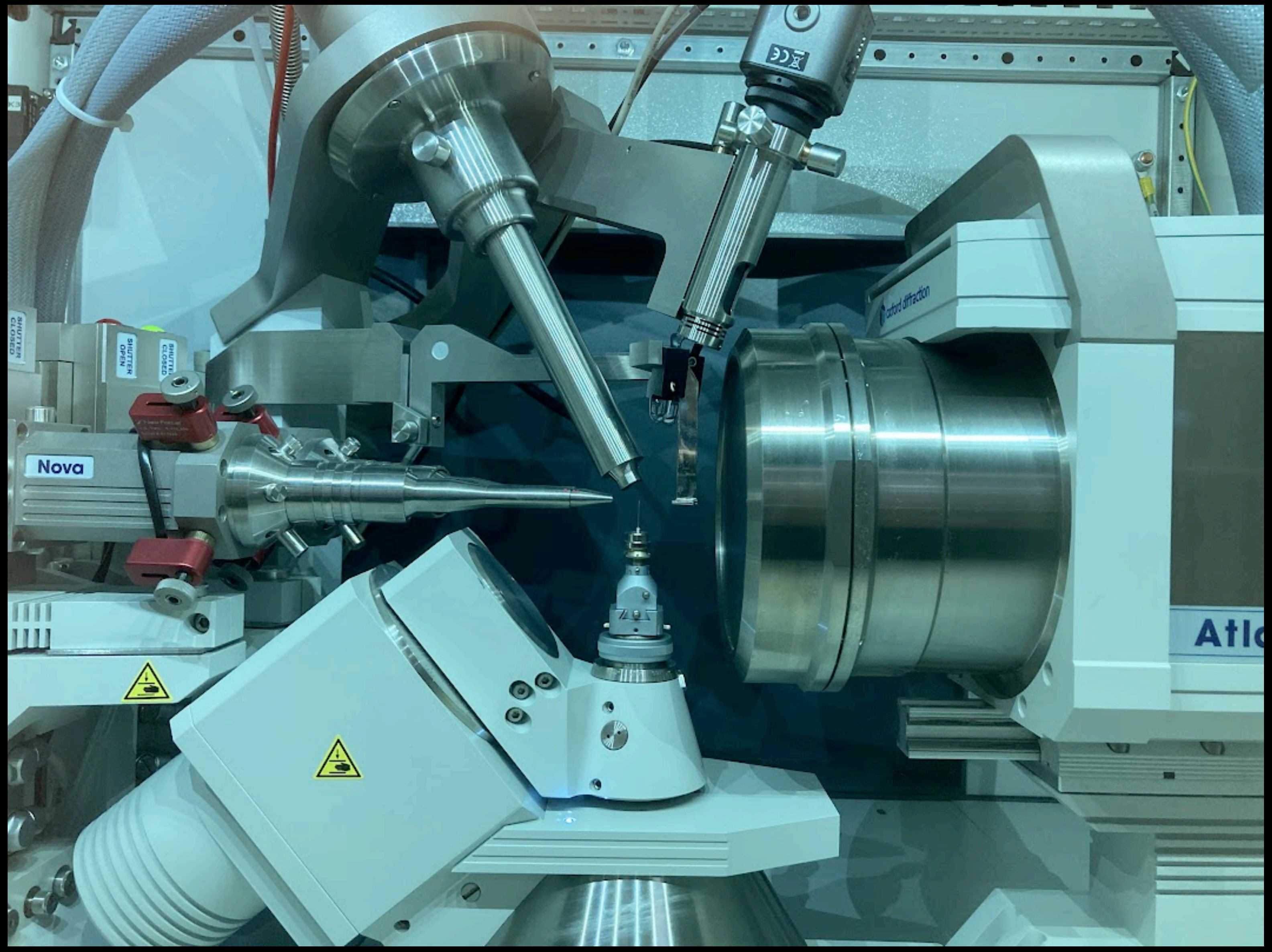
- Convert X-ray photons into some sort of “intensity value”
- Started off with photographic film (data processing involved chemicals, darkroom, film scanner)
- Most likely to encounter CCD or photon counting pixel array detectors
- No detector is perfect but they are very good - knowing the properties will help you to collect good data

DETECTORS





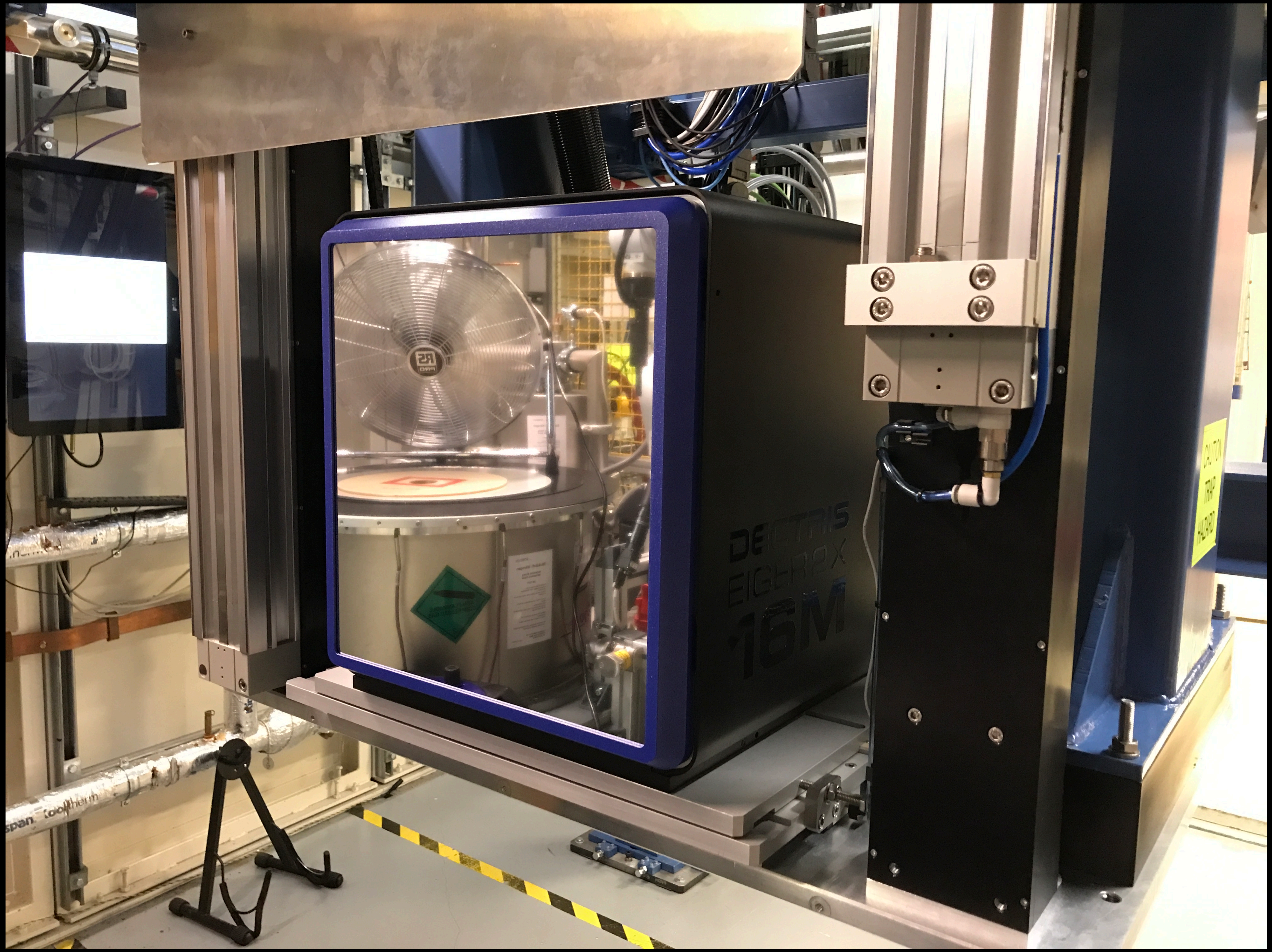




COMMON DETECTOR TYPES

- Integrating:
  - Not worried about photons / second, total photons / pixel problem
  - Read-out noise, dark current
  - Potentially slow (CCD - *not* Jungfrau) - 0.5s / frame read out time
  - Variable gain (counts / photon)
- Photon counting:
  - Limited photons / second rate
  - No read-out noise, dark current
  - Substantially faster - allows shutterless collection
  - Nuances relating to charge sharing etc.

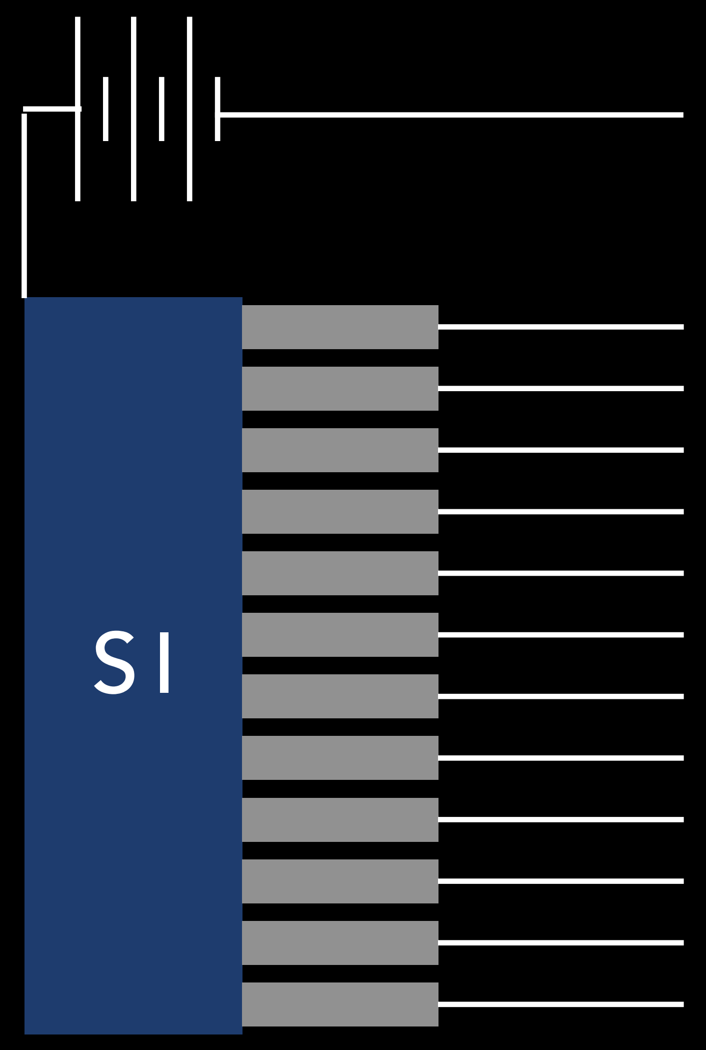
PHOTON COUNTING DETECTORS



PHOTON COUNTING DETECTORS

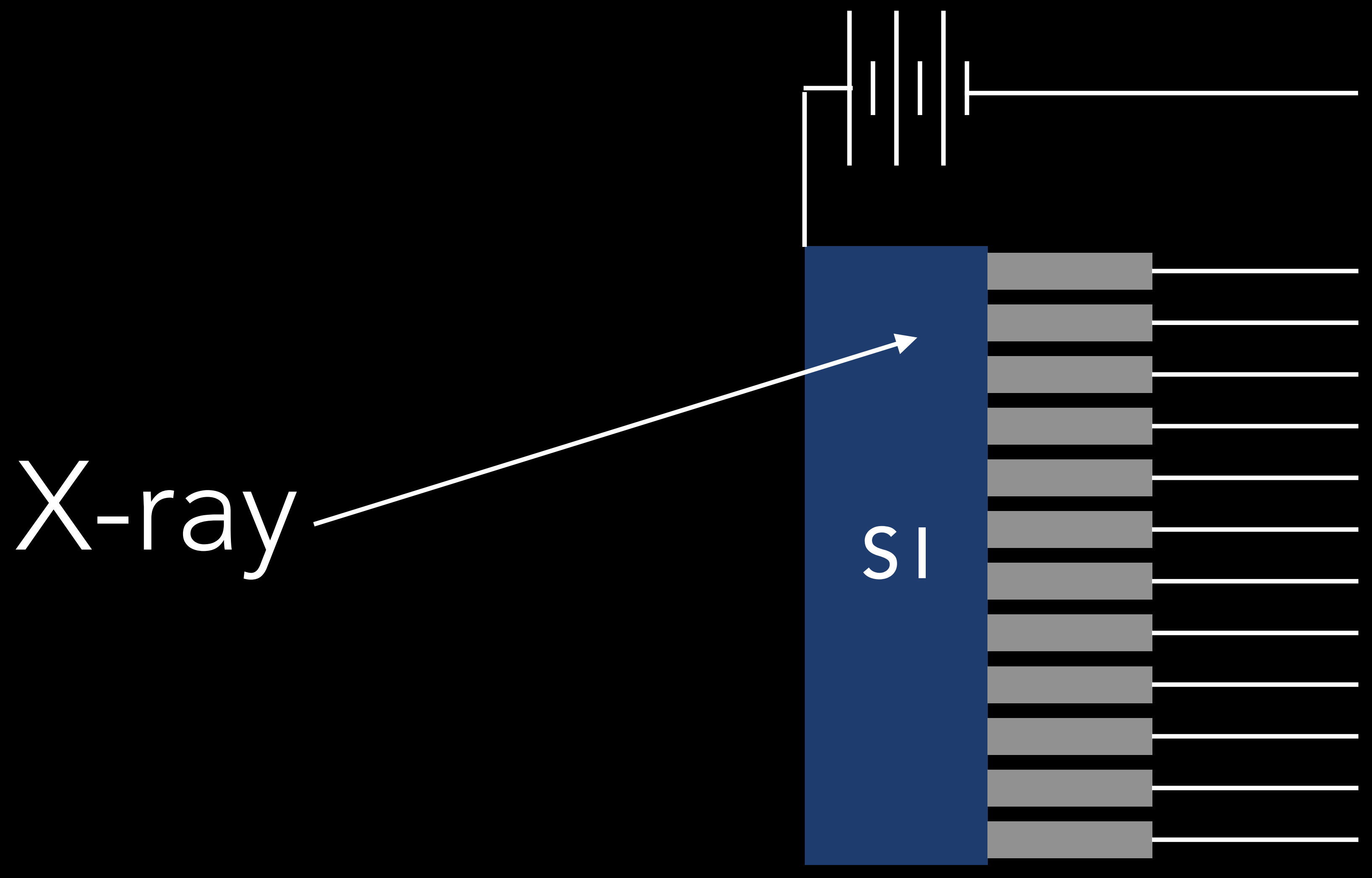
- Direct conversion of photon energy to charge in sensor - derived from sensors from particle physics
- Readout electronics for every pixel (e.g. unlike CCD)
- Enables shutterless data collection as readout is (near-)instant (ms to μs)
- Dectris PILATUS and EIGER most common

PHOTON COUNTING DETECTORS



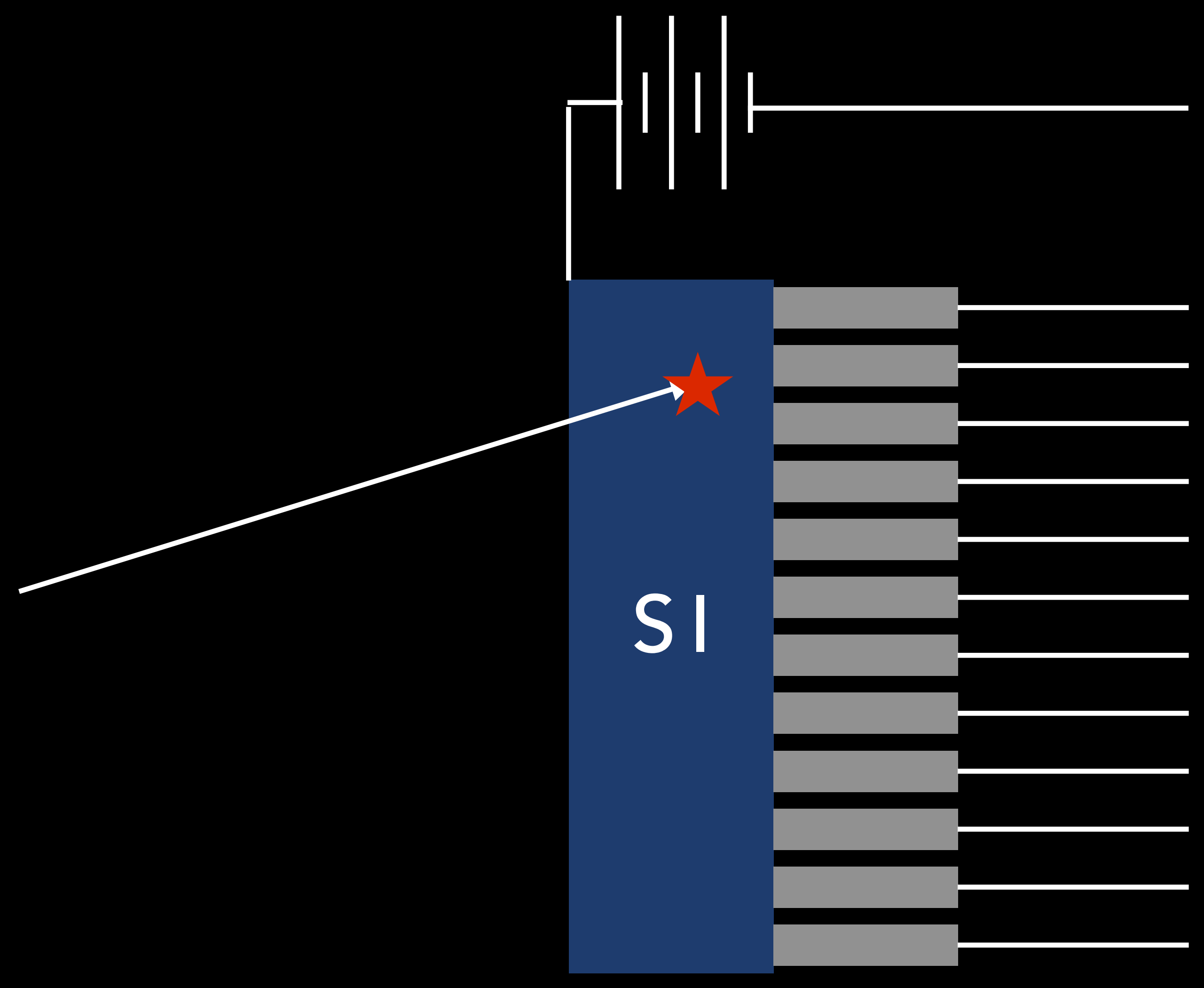
Detector: Si sensor material connected to readout electronics - large potential difference across sensor

PHOTON COUNTING DETECTORS



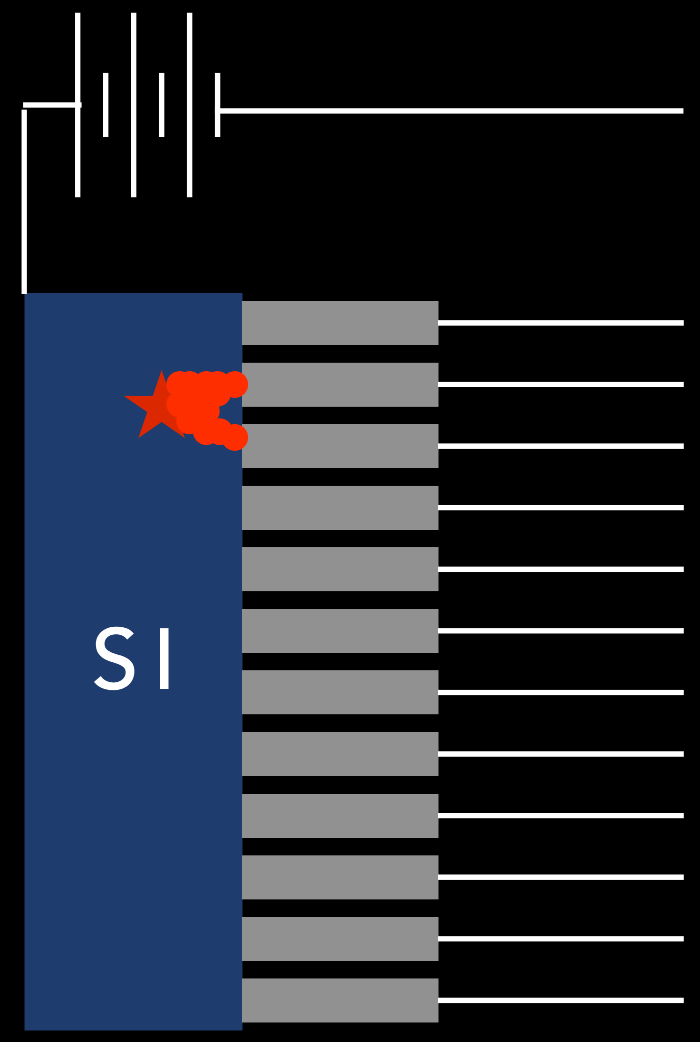
Incoming X-ray

PHOTON COUNTING DETECTORS



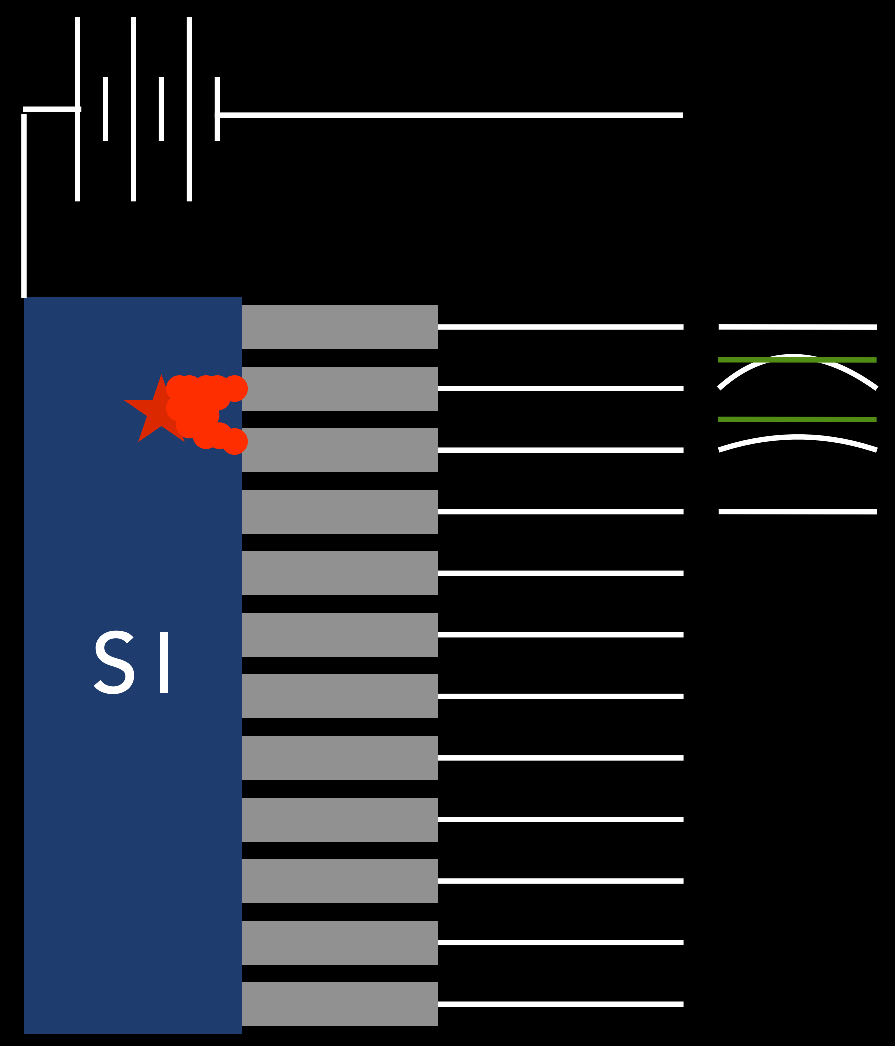
X-ray ionizes Si atom

PHOTON COUNTING DETECTORS



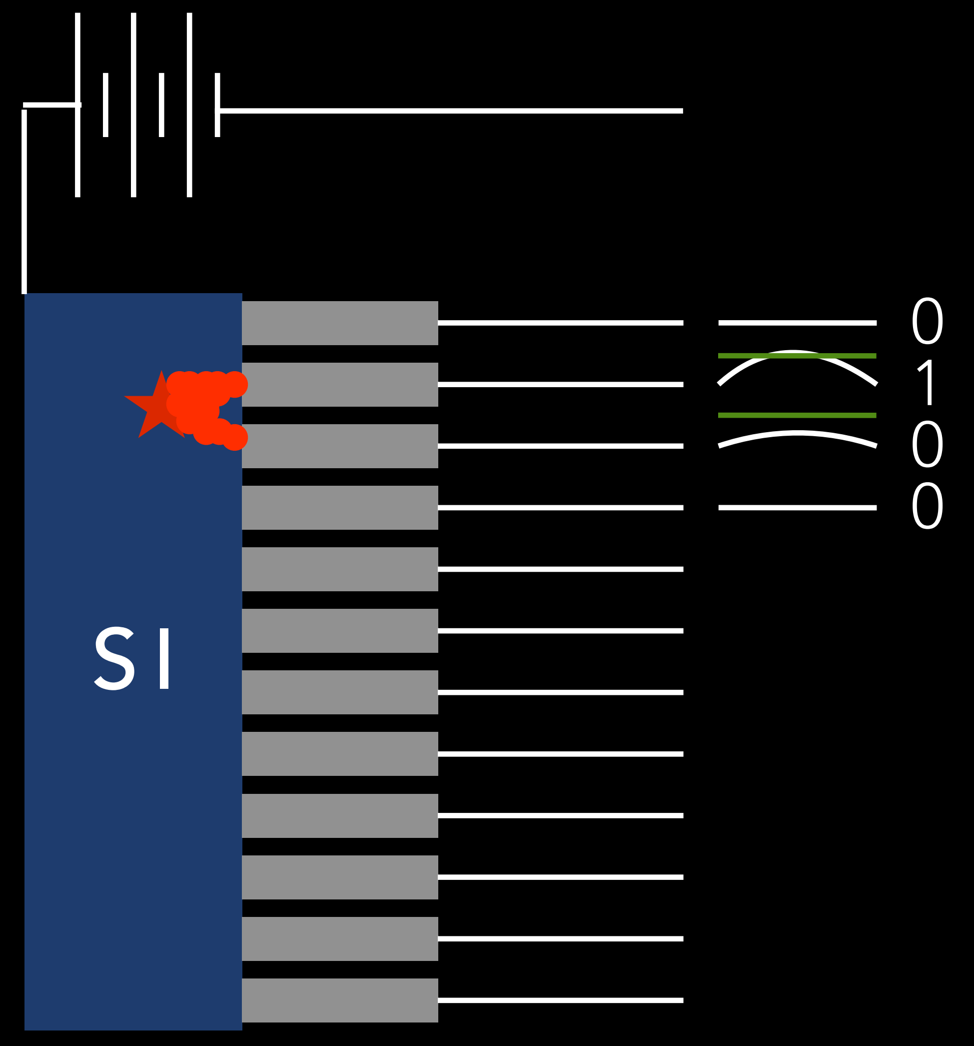
e- cascade
(3500 e-/hole pairs for
12.7 keV photon)

PHOTON COUNTING DETECTORS



Integrated current at anode ~ charge

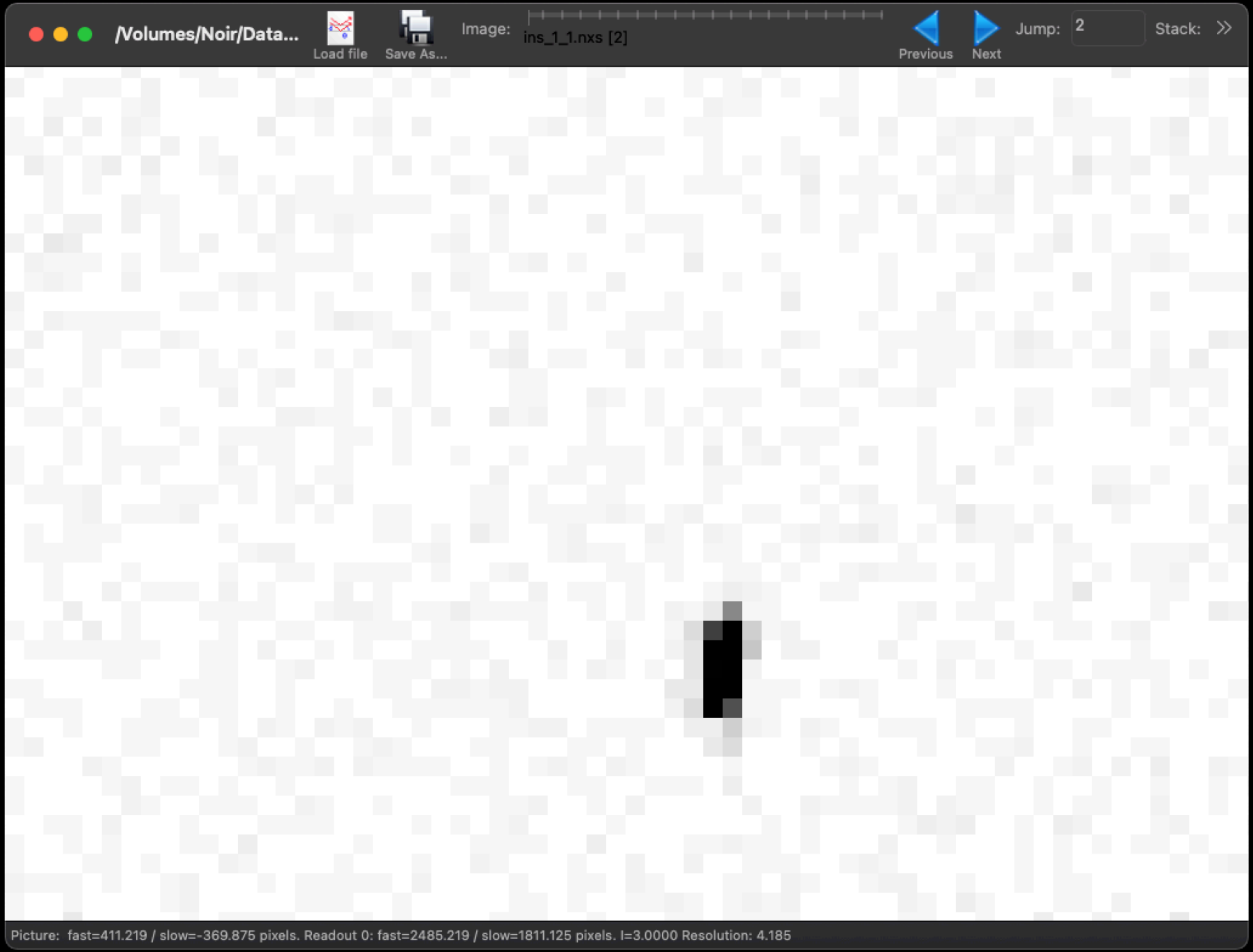
PHOTON COUNTING DETECTORS



Total charge > threshold
(typically 0.5 photon)

PHOTON COUNTING BACKGROUND

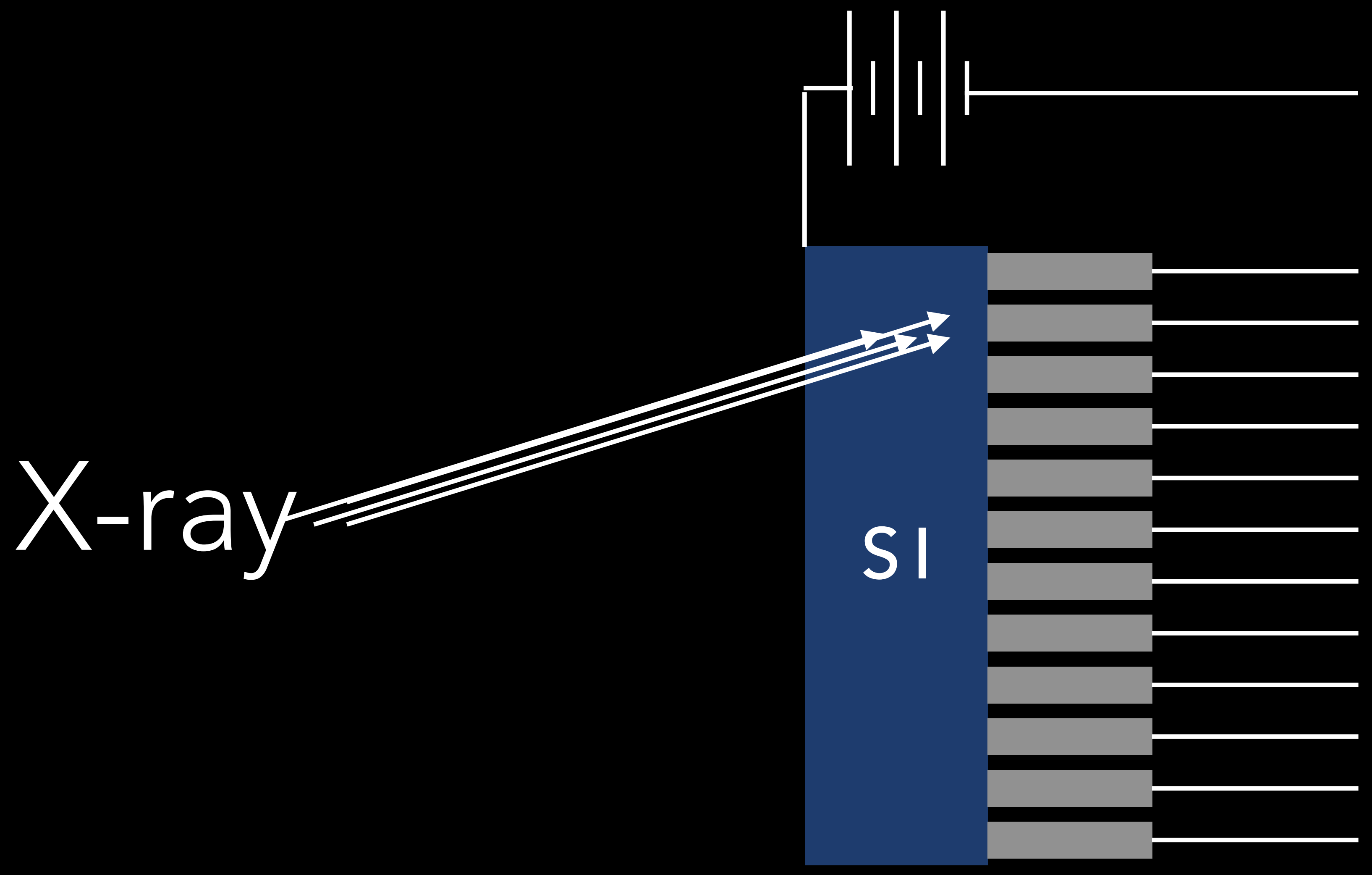
- Thresholding eliminates read-out noise
- Background caused by atoms in beam which are not crystal
- Well mounted samples will have lower background
- Poorly mounted, or in situ, may have high background



DETECTOR PARAMETERS

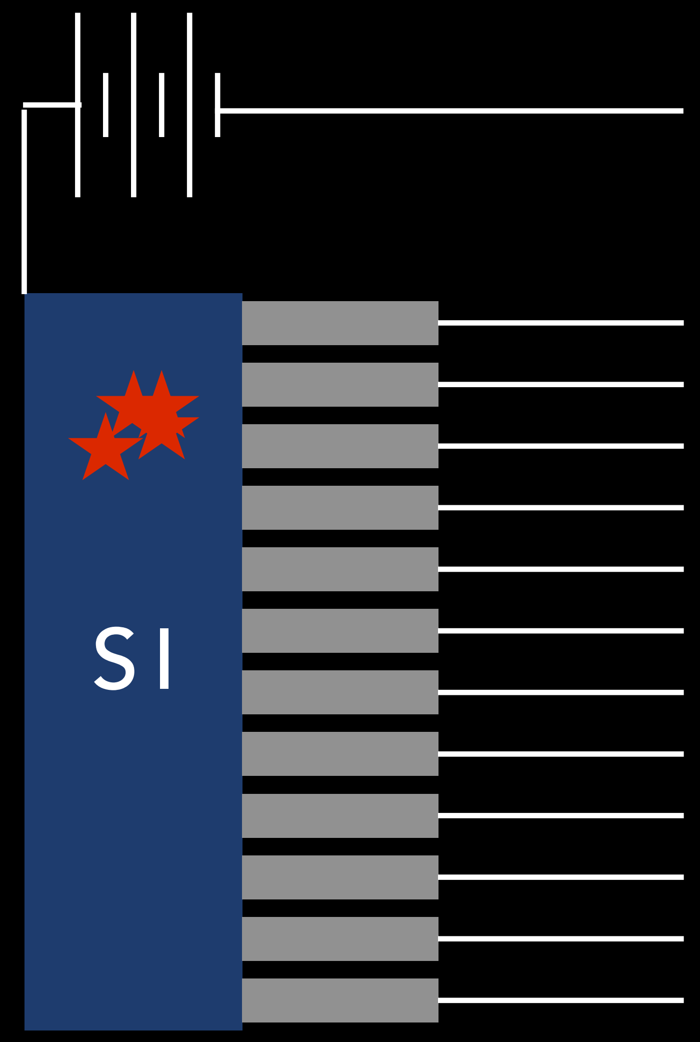
- Pilatus: 172μm square pixels, 320μm to 1mm thick
- Eiger: 75μm square pixels, typically 450μm thick
- Material: Si common, good at lower energies, transparent at high energy, CdTe becoming more common, excellent at high energy (> 20 keV)
- Pixels are very "small and deep" so spots can smear particularly at higher energies as photons pass through the sensor material

PHOTON COUNTING DETECTORS - “PILEUP”



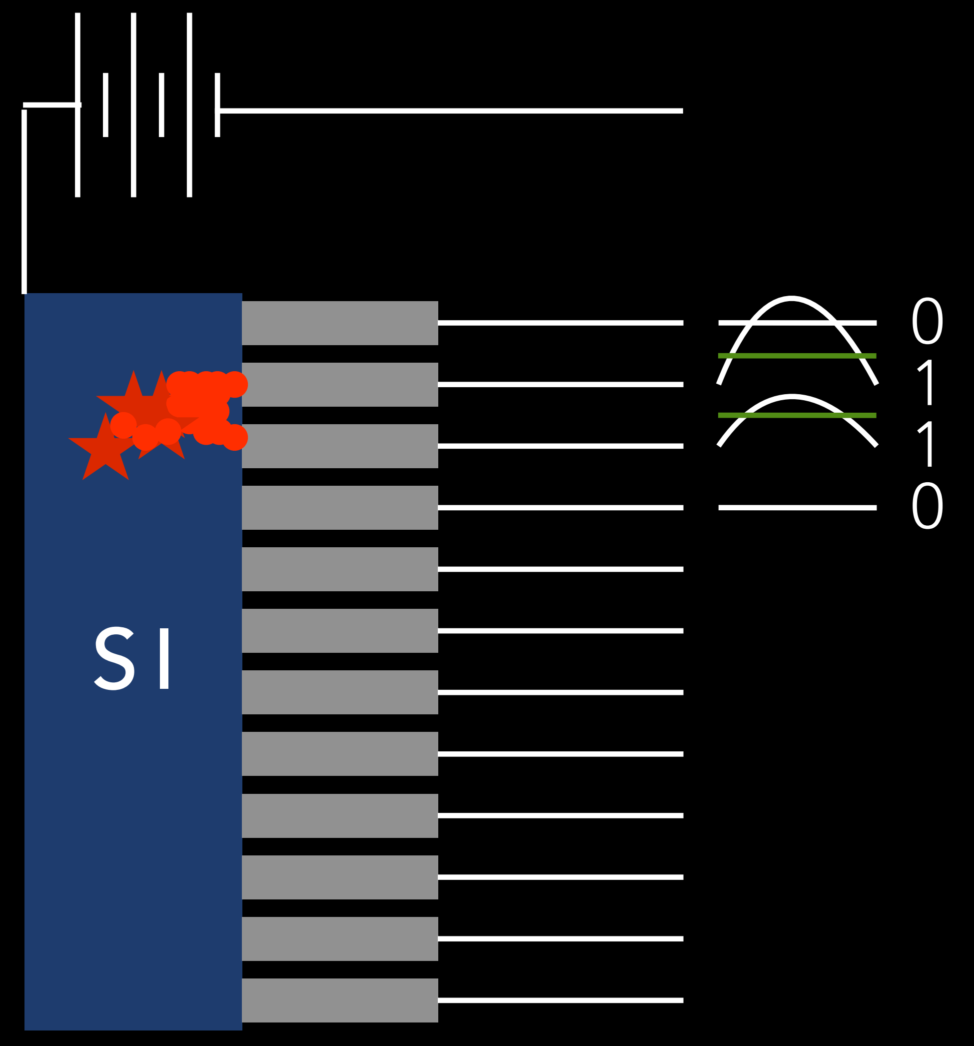
Many photons arrive
within detection interval

PHOTON COUNTING DETECTORS - "PILEUP"



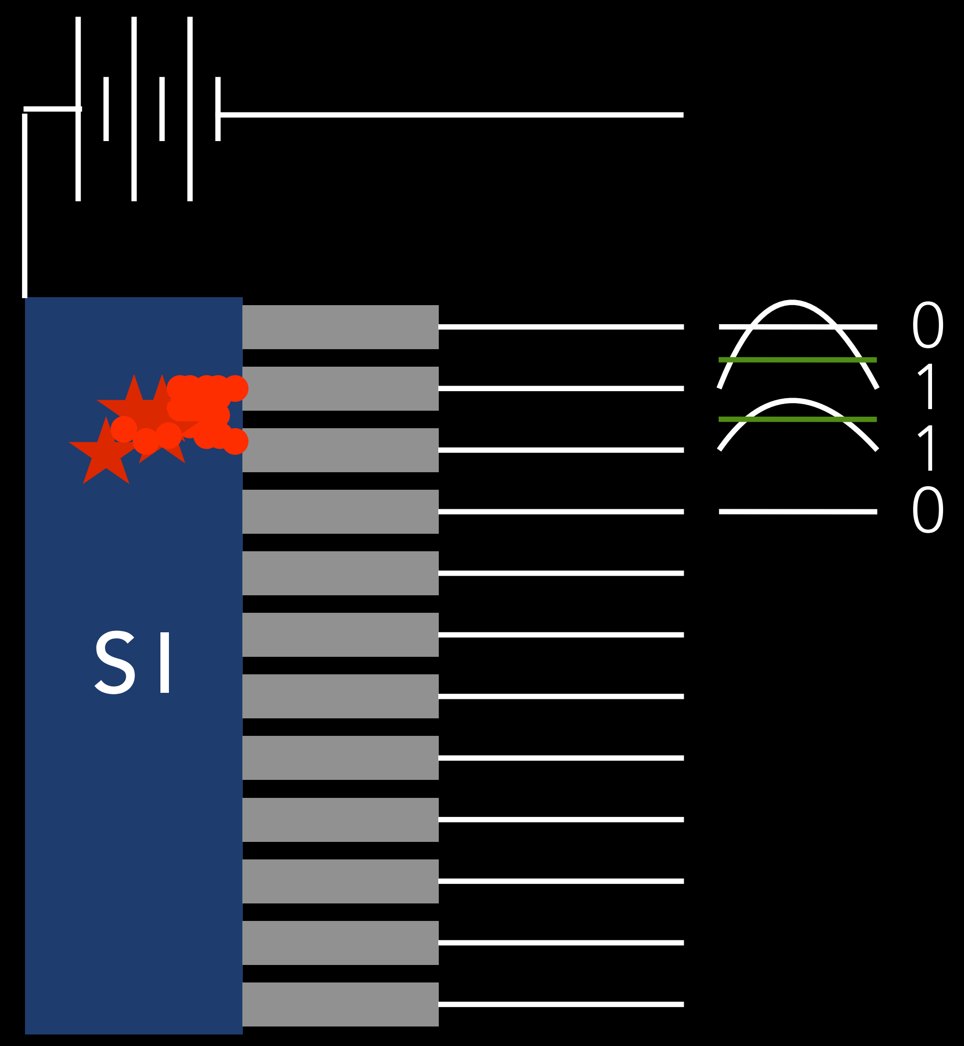
Much ionisation

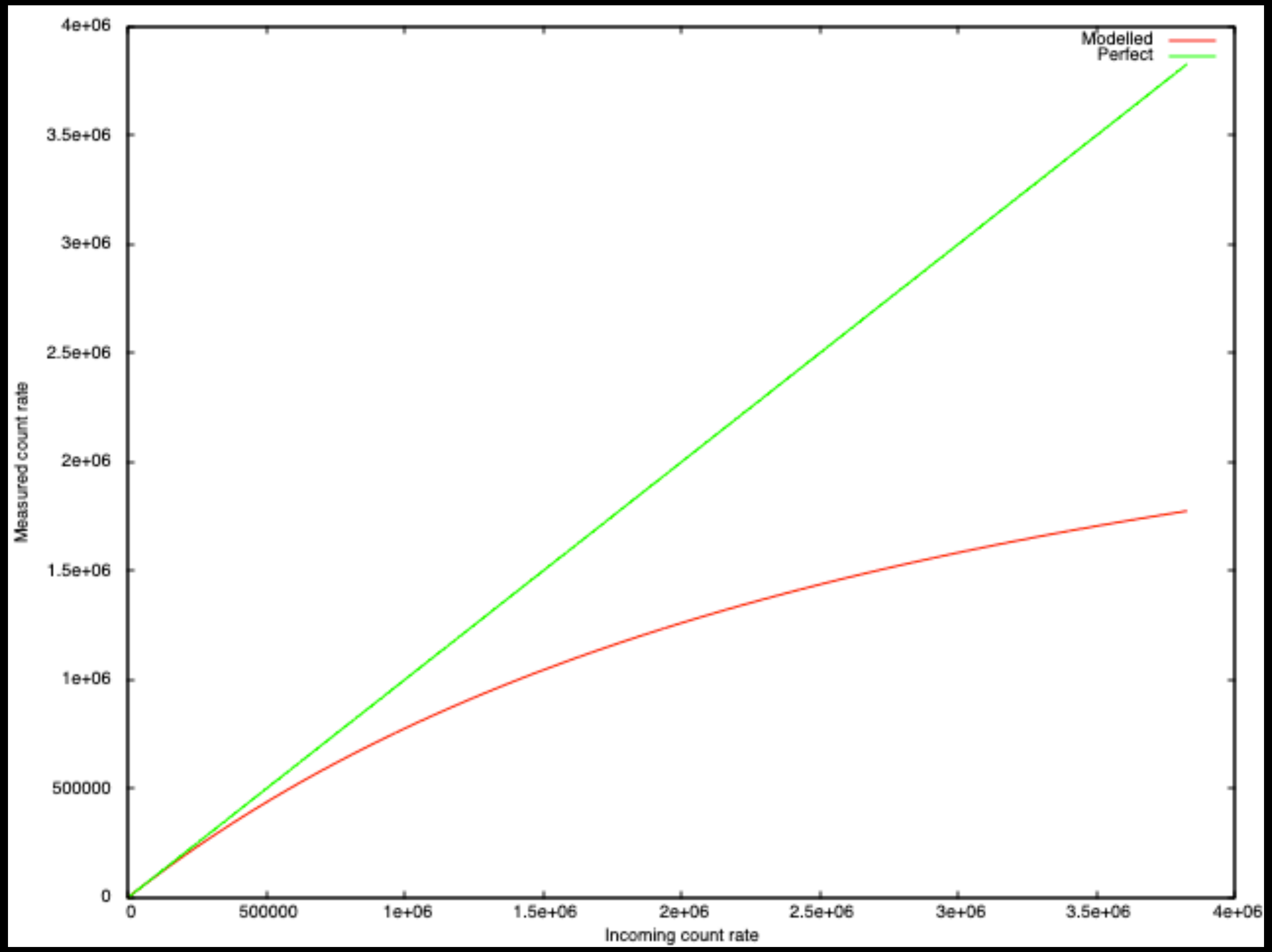
PHOTON COUNTING DETECTORS - "PILEUP"



Lots of charge - but still only passes above threshold so 1 count

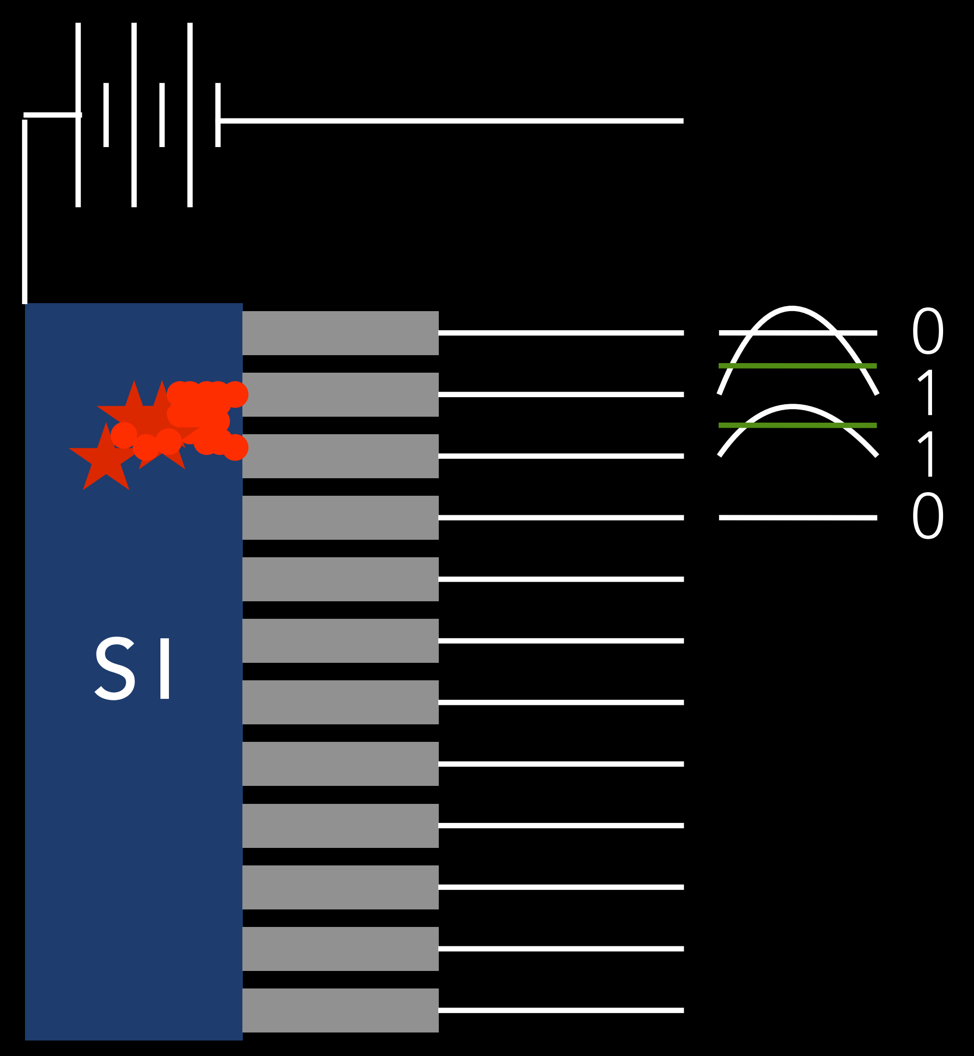
PHOTON COUNTING DETECTORS - "PILEUP"

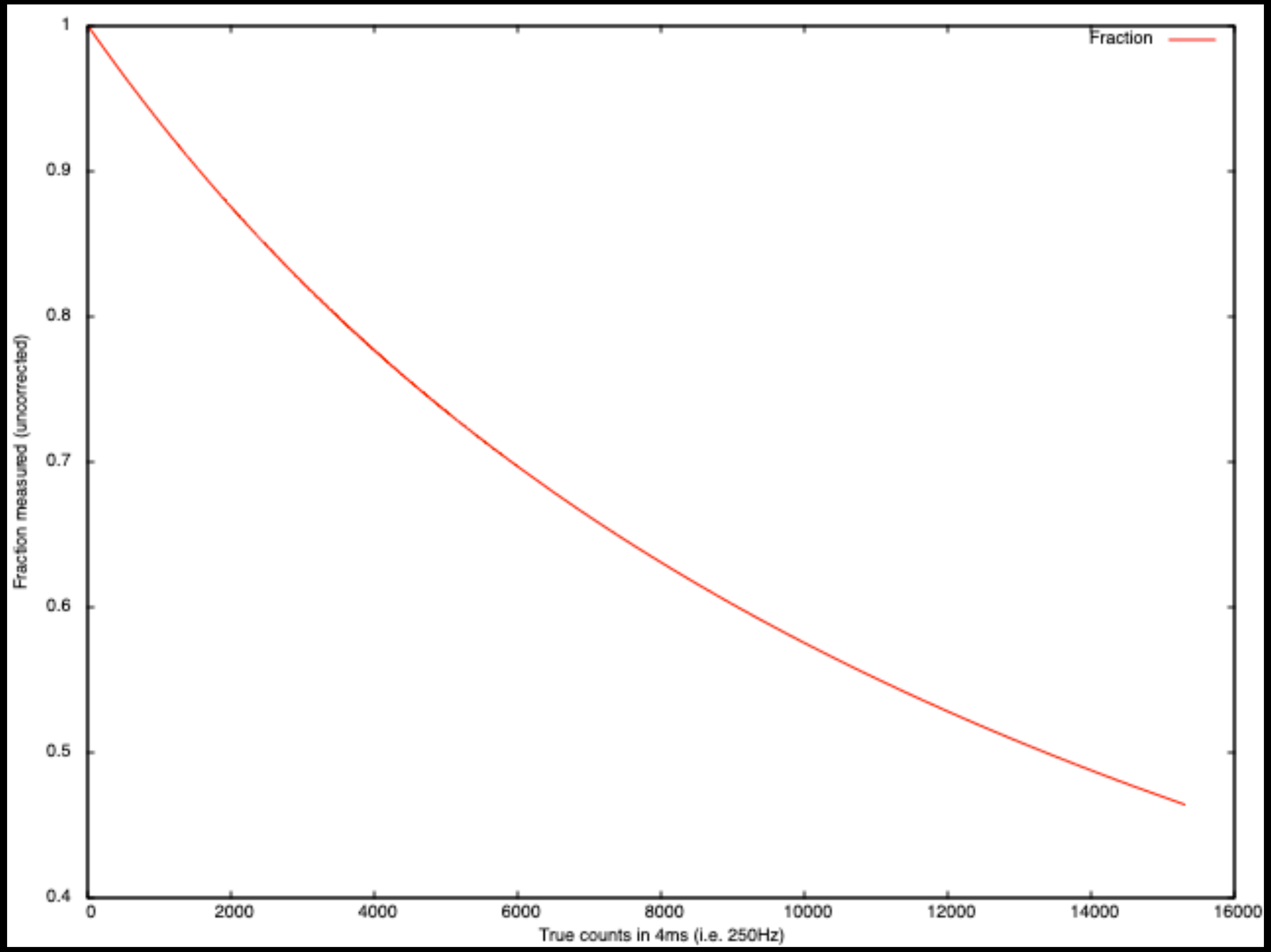




Incoming vs. uncorrected count rate
(photons / s / pixel)

PHOTON COUNTING DETECTORS - "PILEUP"



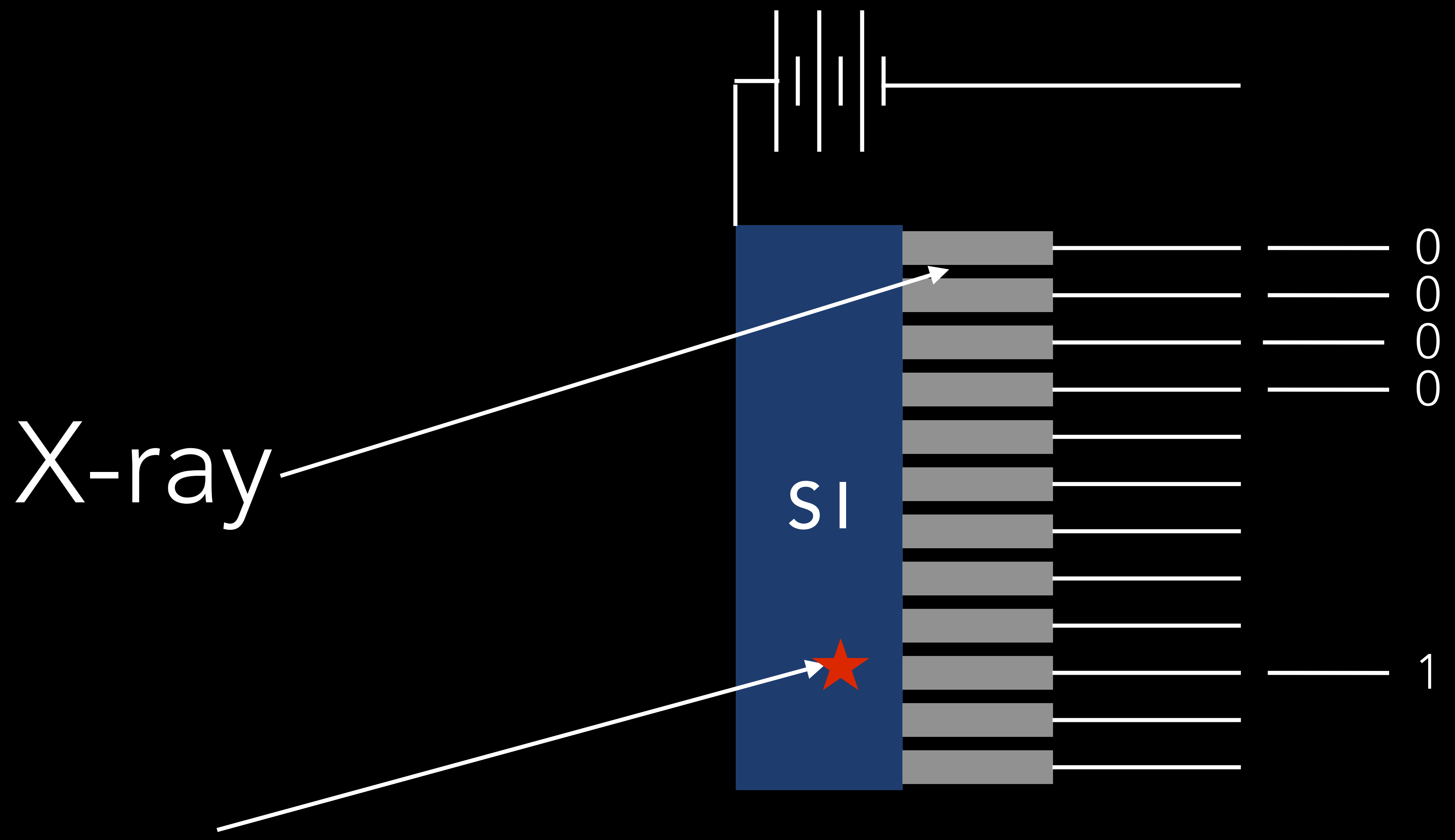


Fraction recorded (uncorrected) vs. pixel count in 4ms

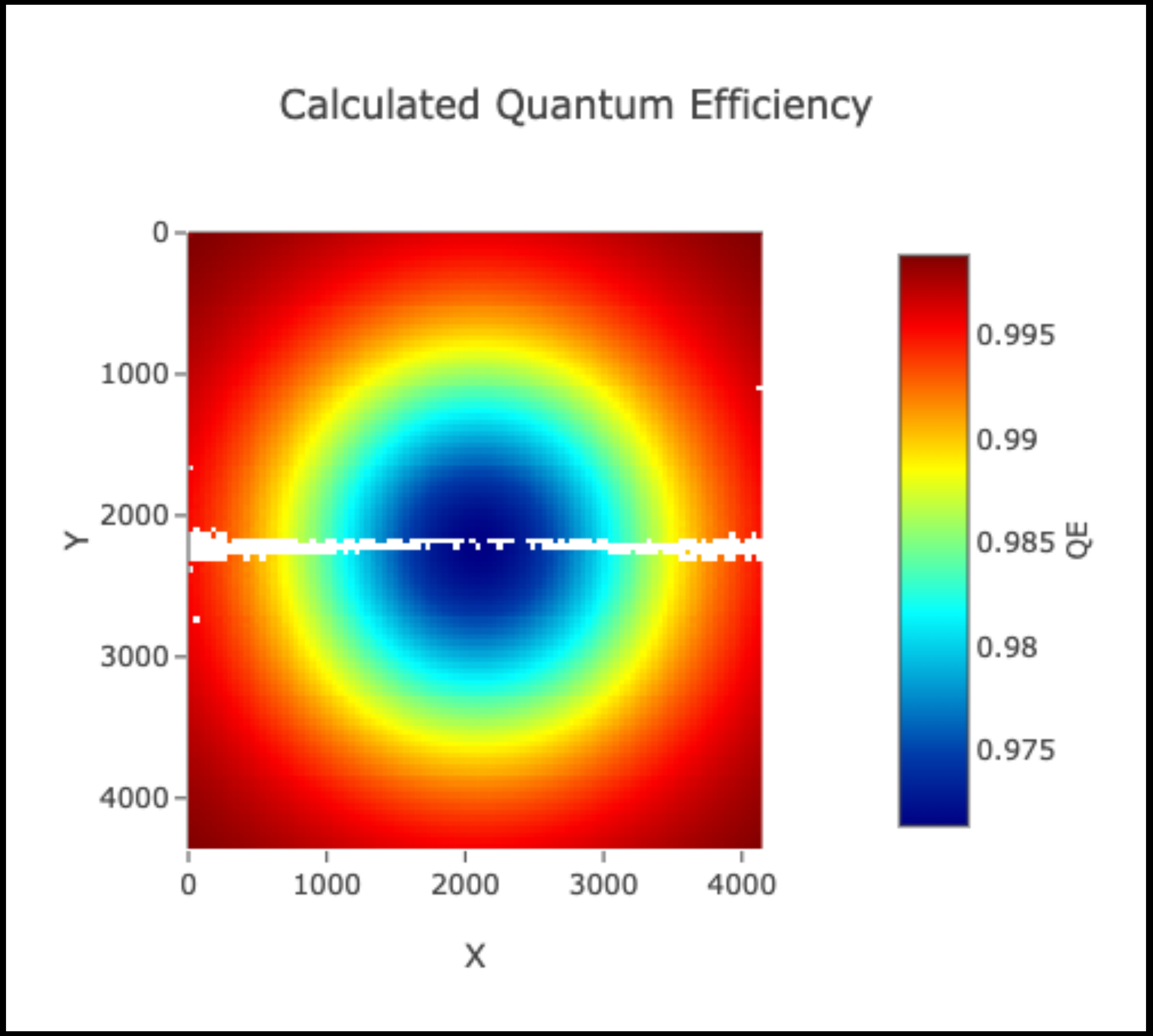
KEY POINTS

- This *is* corrected for but that correction assumes constant count rate in pixel
- If you have finely sliced data, this assumption is probably sound
- If you have recorded data with rotation ≫ mosaic spread the assumption is probably rather poor (more on this later)

PHOTON COUNTING DETECTORS - EFFICIENCY

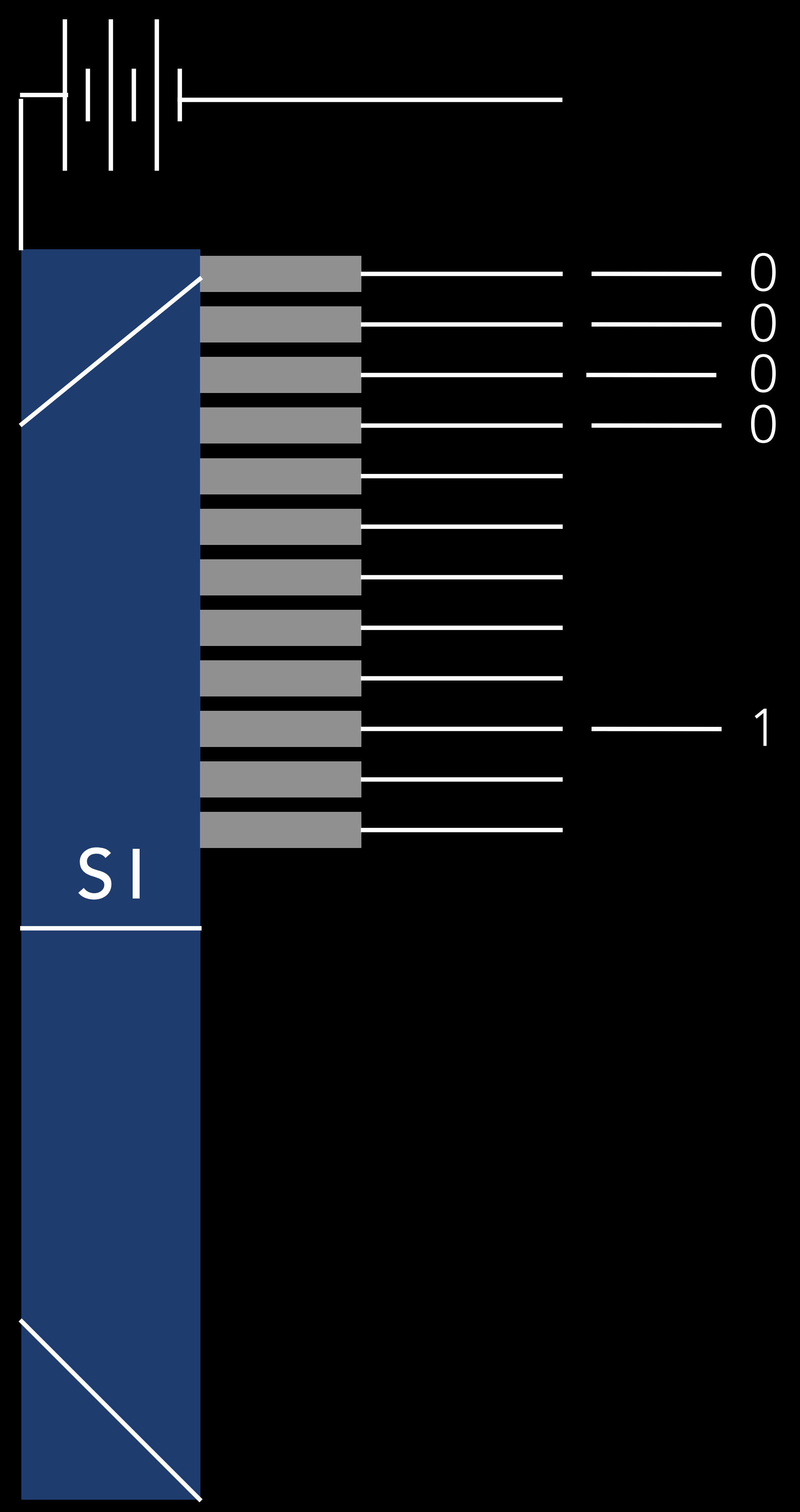


X-ray *can* pass through sensor unmeasured (energy dependent)

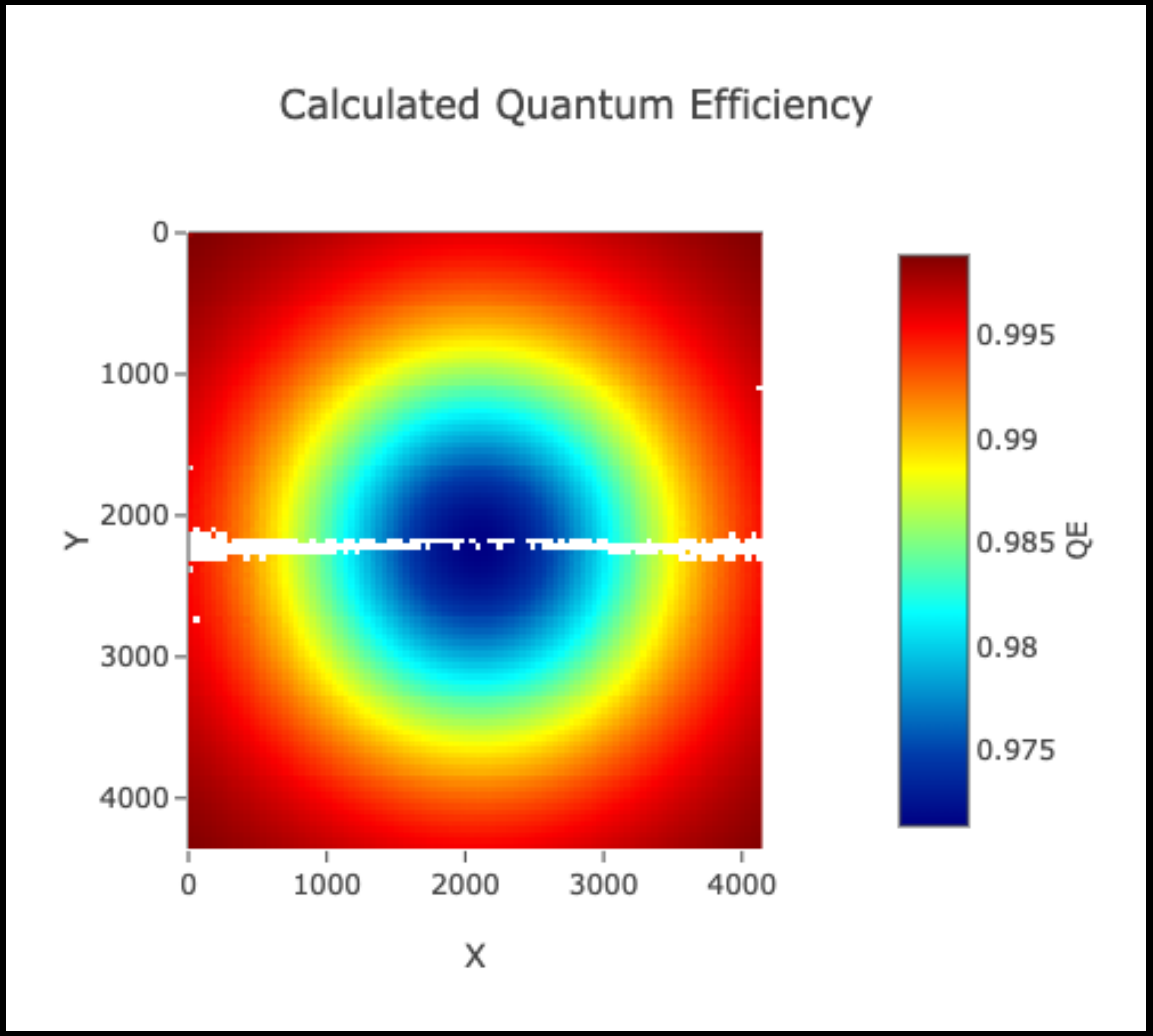


i03 Eiger at 10keV

PHOTON COUNTING DETECTORS - EFFICIENCY



Path length through sensor longer around edges of detector than middle



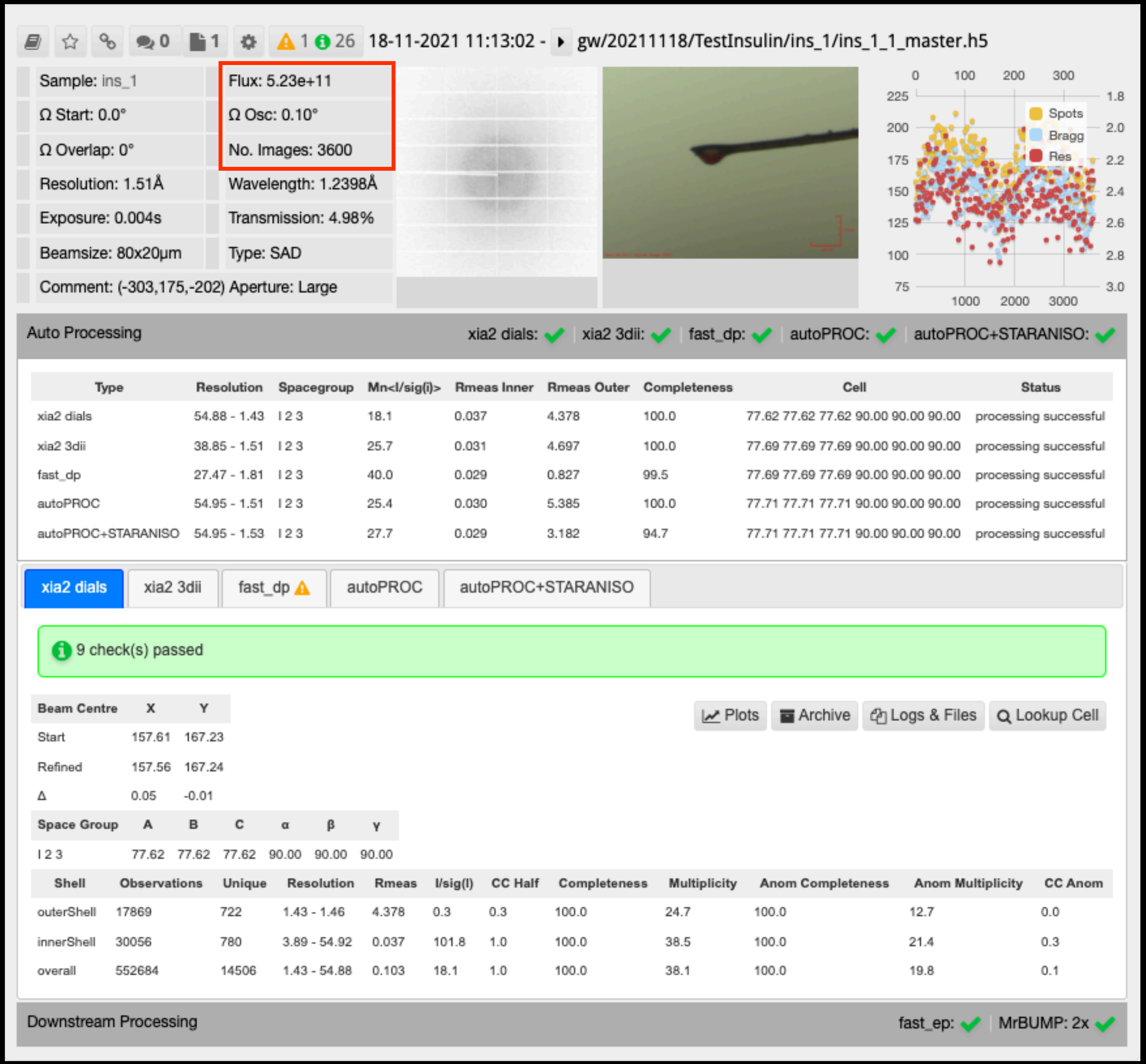
i03 Eiger at 10keV

IDEA #1

- Be aware of saturating the detector
- It is quite hard, but a good crystal on a bright beamline is able to do this
- It is hard to see if this has happened
- For a photon counting detector, too weak is rarely a problem
- Be aware of detector efficiency at higher energies
- Talking to beamline staff a good way to go

DATA INTENSITY / MULTIPLICITY

- The “quality” of data depends on many factors - including the individual spot intensity and number of observations
- Measuring twice as many copies of data half as strong “averages out” to give essentially the same data - with a photon counting detector
- Twice as many, half as strong, measurements also gives opportunity to truncate the data if damage is present



0 | 1 | 1 | 26 | 18-11-2021 11:13:02 - gw/20211118/TestInsulin/ins_1/ins_1_1_master.h5

Sample: ins_1	Flux: 5.23e+11
Ω Start: 0.0°	Ω Osc: 0.10°
Ω Overlap: 0°	No. Images: 3600
Resolution: 1.51Å	Wavelength: 1.2398Å
Exposure: 0.004s	Transmission: 4.98%
Beamsize: 80x20µm	Type: SAD
Comment: (-303,175,-202) Aperture: Large	

Auto Processing — xia2 dials: ✓ | xia2 3dii: ✓ | fast_dp: ✓ | autoPROC: ✓ | autoPROC+STARANISO: ✓

Type	Resolution	Spacegroup	Mn<I/sig(i)>	Rmeas Inner	Rmeas Outer	Completeness	Cell	Status
xia2 dials	54.88 - 1.43	I 2 3	18.1	0.037	4.378	100.0	77.62 77.62 77.62 90.00 90.00 90.00	processing successful
xia2 3dii	38.85 - 1.51	I 2 3	25.7	0.031	4.697	100.0	77.69 77.69 77.69 90.00 90.00 90.00	processing successful
fast_dp	27.47 - 1.81	I 2 3	40.0	0.029	0.827	99.5	77.69 77.69 77.69 90.00 90.00 90.00	processing successful
autoPROC	54.95 - 1.51	I 2 3	25.4	0.030	5.385	100.0	77.71 77.71 77.71 90.00 90.00 90.00	processing successful
autoPROC+STARANISO	54.95 - 1.53	I 2 3	27.7	0.029	3.182	94.7	77.71 77.71 77.71 90.00 90.00 90.00	processing successful

xia2 dials | xia2 3dii | fast_dp | autoPROC | autoPROC+STARANISO

9 check(s) passed

Plots | Archive | Logs & Files | Lookup Cell

Beam Centre	X	Y
Start	157.61	167.23
Refined	157.56	167.24
Δ	0.05	-0.01

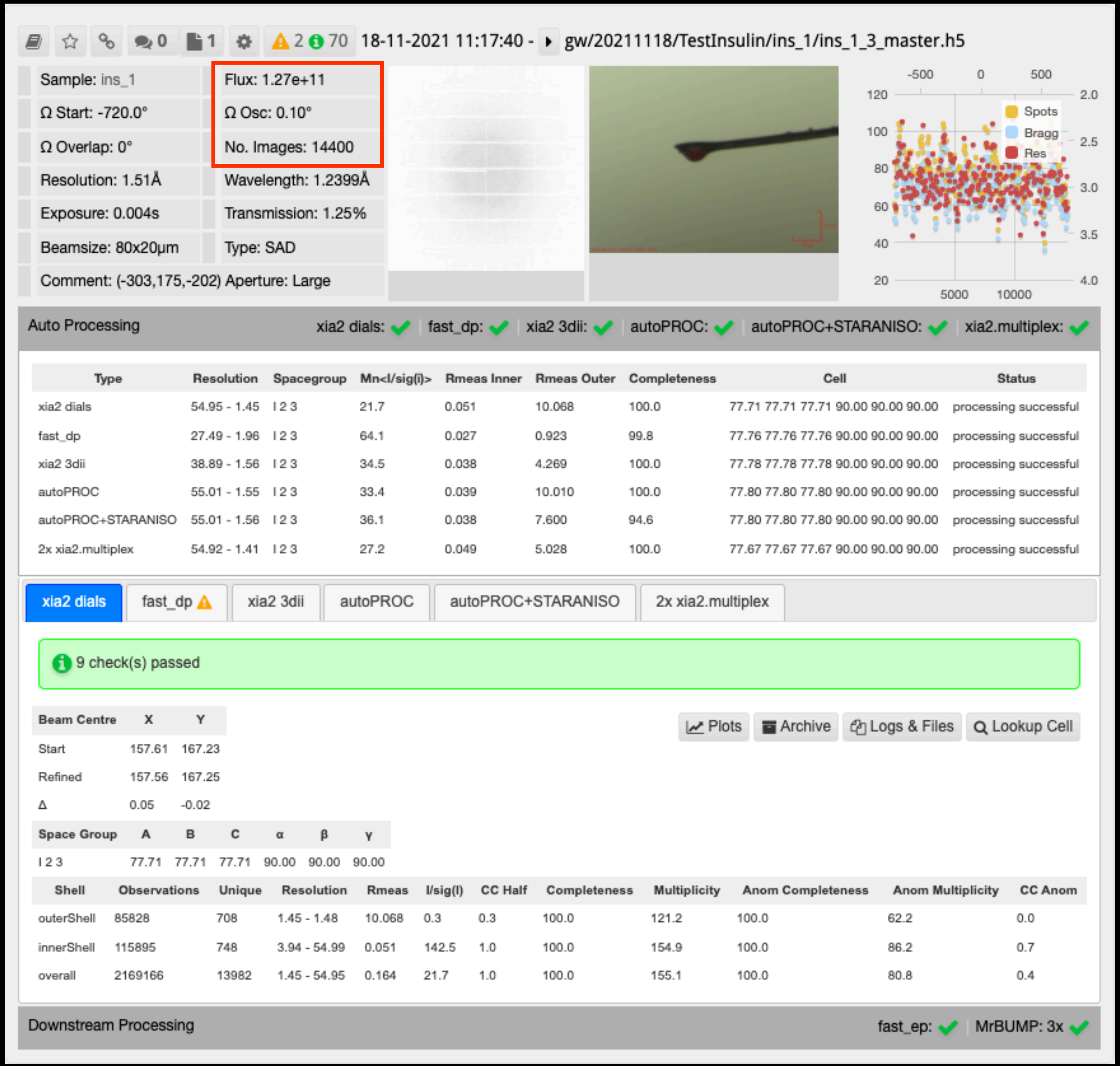
Space Group	A	B	C	α	β	γ
I 2 3	77.62	77.62	77.62	90.00	90.00	90.00

Shell	Observations	Unique	Resolution	Rmeas	I/sig(I)	CC Half	Completeness	Multiplicity	Anom Completeness	Anom Multiplicity	CC Anom
outerShell	17869	722	1.43 - 1.46	4.378	0.3	0.3	100.0	24.7	100.0	12.7	0.0
innerShell	30056	780	3.89 - 54.92	0.037	101.8	1.0	100.0	38.5	100.0	21.4	0.3
overall	552684	14506	1.43 - 54.88	0.103	18.1	1.0	100.0	38.1	100.0	19.8	0.1

Downstream Processing — fast_ep: ✓ | MrBUMP: 2x ✓

x4

/4



0 | 1 | 2 | 70 | 18-11-2021 11:17:40 - gw/20211118/TestInsulin/ins_1/ins_1_3_master.h5

Sample: ins_1	Flux: 1.27e+11
Ω Start: -720.0°	Ω Osc: 0.10°
Ω Overlap: 0°	No. Images: 14400
Resolution: 1.51Å	Wavelength: 1.2399Å
Exposure: 0.004s	Transmission: 1.25%
Beamsize: 80x20µm	Type: SAD
Comment: (-303,175,-202) Aperture: Large	

Auto Processing — xia2 dials: ✓ | fast_dp: ✓ | xia2 3dii: ✓ | autoPROC: ✓ | autoPROC+STARANISO: ✓ | xia2.multiplex: ✓

Type	Resolution	Spacegroup	Mn<I/sig(i)>	Rmeas Inner	Rmeas Outer	Completeness	Cell	Status
xia2 dials	54.95 - 1.45	I 2 3	21.7	0.051	10.068	100.0	77.71 77.71 77.71 90.00 90.00 90.00	processing successful
fast_dp	27.49 - 1.96	I 2 3	64.1	0.027	0.923	99.8	77.76 77.76 77.76 90.00 90.00 90.00	processing successful
xia2 3dii	38.89 - 1.56	I 2 3	34.5	0.038	4.269	100.0	77.78 77.78 77.78 90.00 90.00 90.00	processing successful
autoPROC	55.01 - 1.55	I 2 3	33.4	0.039	10.010	100.0	77.80 77.80 77.80 90.00 90.00 90.00	processing successful
autoPROC+STARANISO	55.01 - 1.56	I 2 3	36.1	0.038	7.600	94.6	77.80 77.80 77.80 90.00 90.00 90.00	processing successful
2x xia2.multiplex	54.92 - 1.41	I 2 3	27.2	0.049	5.028	100.0	77.67 77.67 77.67 90.00 90.00 90.00	processing successful

xia2 dials | fast_dp | xia2 3dii | autoPROC | autoPROC+STARANISO | 2x xia2.multiplex

9 check(s) passed

Plots | Archive | Logs & Files | Lookup Cell

Beam Centre	X	Y
Start	157.61	167.23
Refined	157.56	167.25
Δ	0.05	-0.02

Space Group	A	B	C	α	β	γ
I 2 3	77.71	77.71	77.71	90.00	90.00	90.00

Shell	Observations	Unique	Resolution	Rmeas	I/sig(I)	CC Half	Completeness	Multiplicity	Anom Completeness	Anom Multiplicity	CC Anom
outerShell	85828	708	1.45 - 1.48	10.068	0.3	0.3	100.0	121.2	100.0	62.2	0.0
innerShell	115895	748	3.94 - 54.99	0.051	142.5	1.0	100.0	154.9	100.0	86.2	0.7
overall	2169166	13982	1.45 - 54.95	0.164	21.7	1.0	100.0	155.1	100.0	80.8	0.4

Downstream Processing — fast_ep: ✓ | MrBUMP: 3x ✓

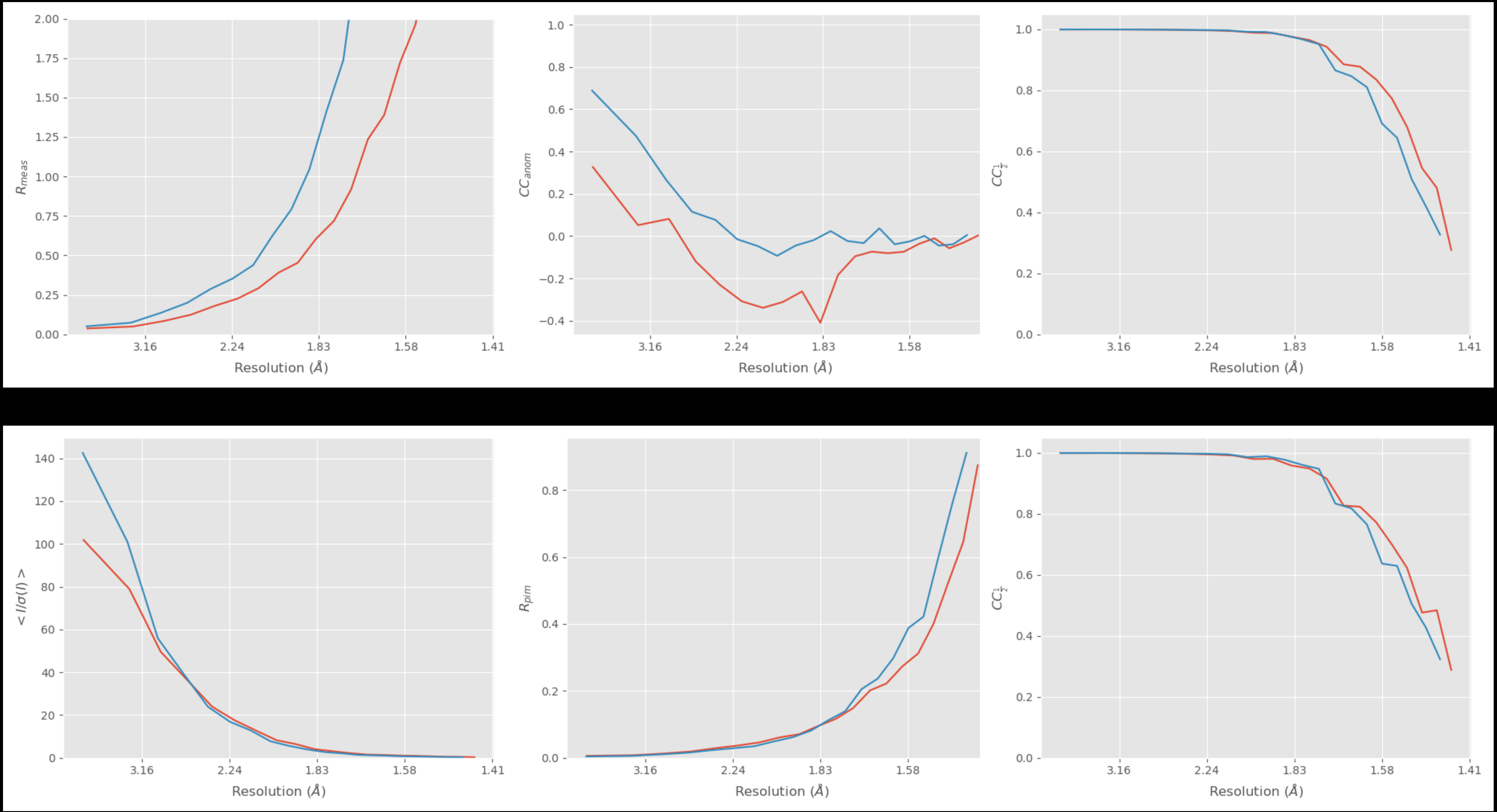
TABLE 1 NOT SO HELPFUL

Overall

	Overall	Low resolution	High resolution
Resolution (Å)	54.88 - 1.43	54.92 - 3.89	1.46 - 1.43
Observations	552684	30056	17869
Unique reflections	14506	780	722
Multiplicity	38.1	38.5	24.7
Completeness	100.00%	100.00%	100.00%
Mean I/σ(I)	18.1	101.8	0.3
R_{merge}	0.102	0.037	4.289
R_{meas}	0.103	0.037	4.378
R_{pim}	0.017	0.006	0.874
$CC_{½}$	1.000	0.999	0.288

Overall

	Overall	Low resolution	High resolution
Resolution (Å)	54.95 - 1.45	54.99 - 3.94	1.48 - 1.45
Observations	2169166	115895	85828
Unique reflections	13982	748	708
Multiplicity	155.1	154.9	121.2
Completeness	100.00%	100.00%	100.00%
Mean I/σ(I)	21.7	142.5	0.3
R_{merge}	0.164	0.051	10.026
R_{meas}	0.164	0.051	10.068
R_{pim}	0.013	0.004	0.911
$CC_{½}$	1.000	1.000	0.324



R_meas
2.00
1.75
1.50
1.25
1.00
0.75
0.50
0.25
0.00
CC_anom
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
CC_1/2
0.0
0.2
0.4
0.6
0.8
1.0
< I/σ(I) >
140
120
100
80
60
40
20
0
R_pim
3.16
2.24
1.83
1.58
1.41
Resolution (Å)

IDEA #2

- Measure more data with less transmission (at least until you know your samples well)
- Easy way to achieve this is stepped transmission experiment - start off weak, then 4x dose, 4x dose etc. (4 x should increase $I/\sigma(I)$ by ~ 2)
- Once you know your sample lifetimes, collect more data that are weak to give room for truncation and reduce impact of count rate on data

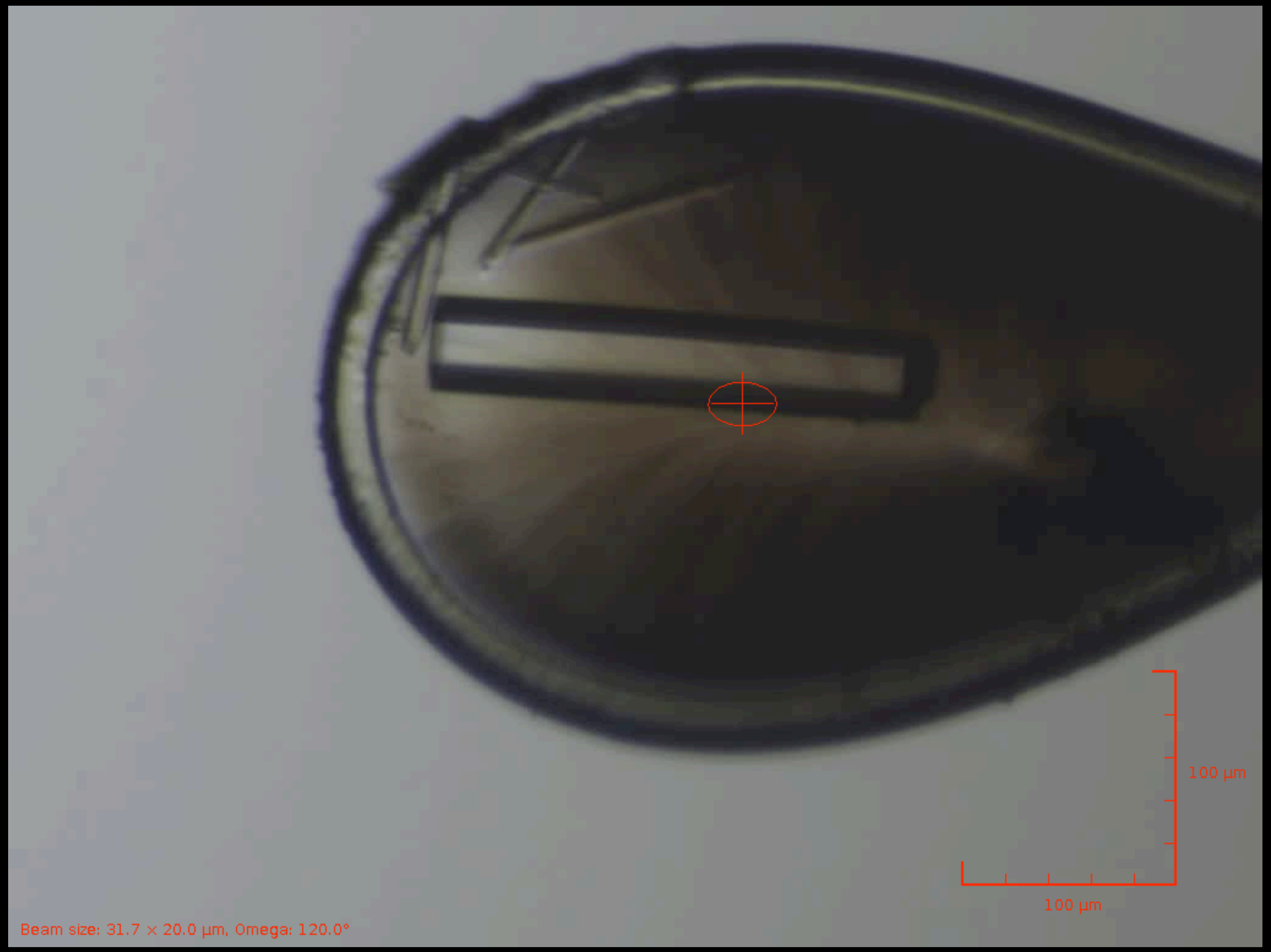
COLLECTION ADVICE

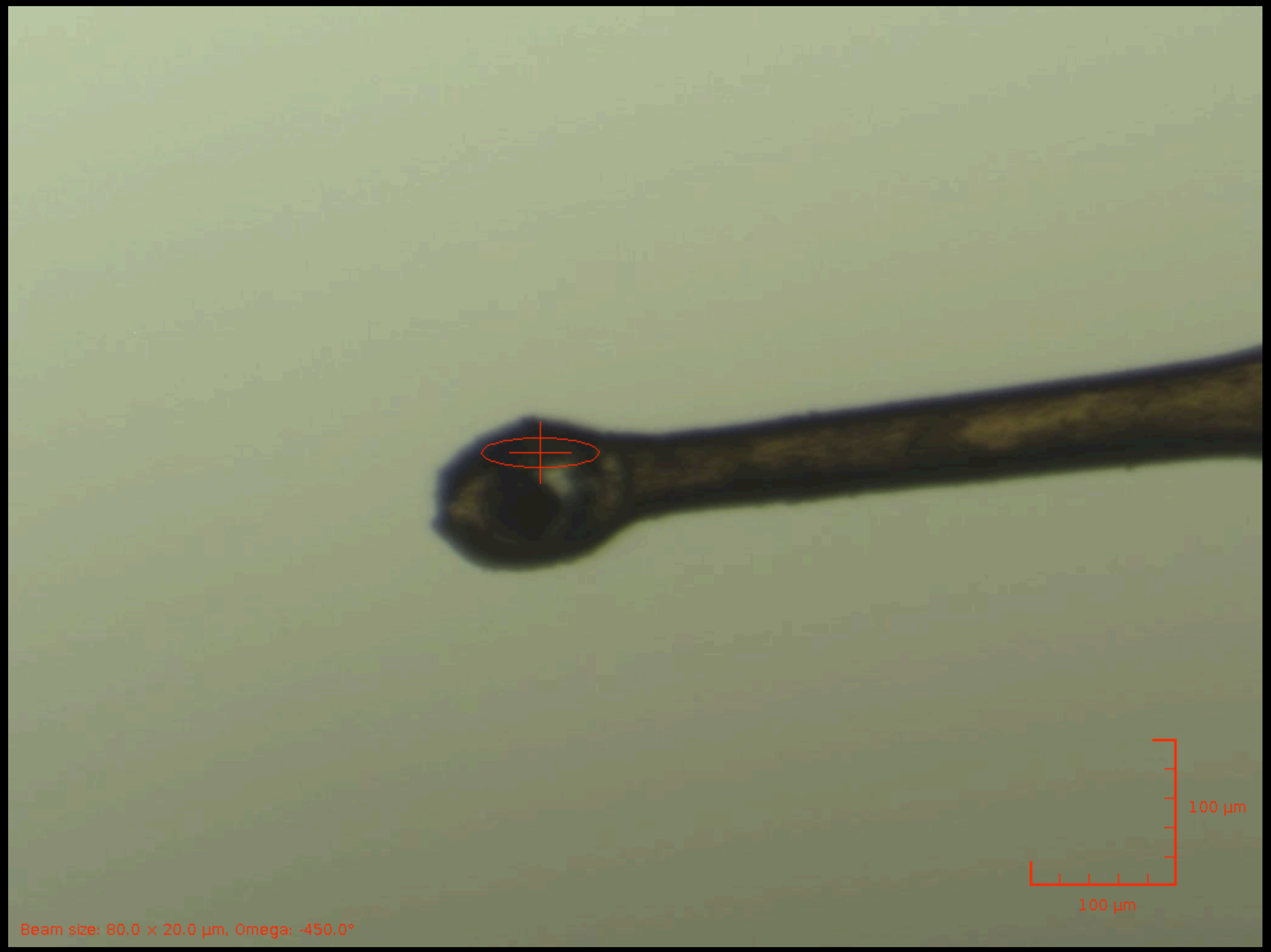
So I am collecting gently, collecting complete and high multiplicity data, what now?

- Background: ideally minimise this - mount sample carefully
- Also collect data carefully - for example sample alignment

SAMPLE MOUNTING

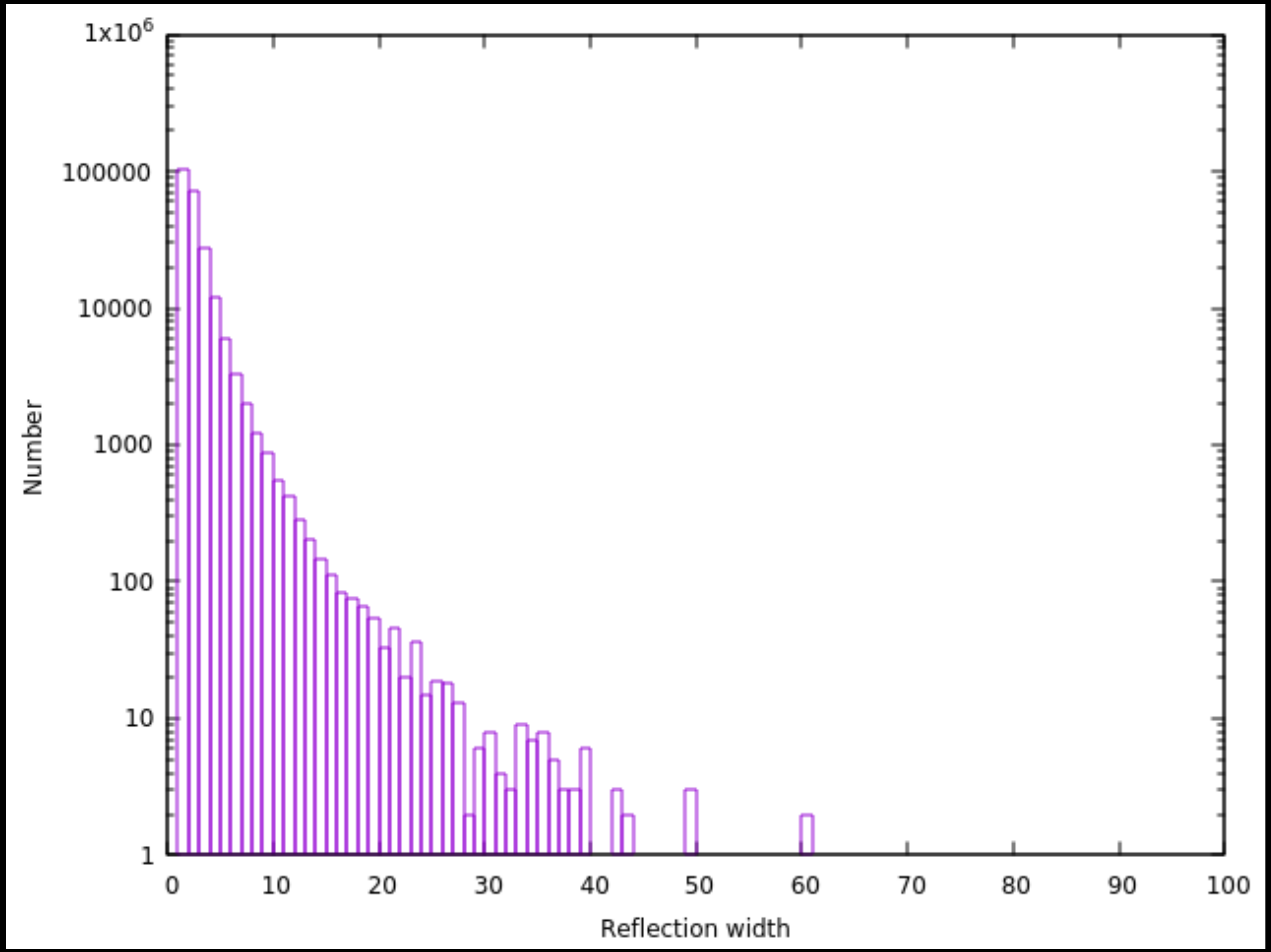
- Use suitable mount - similar size and ideally shape as sample
- Ideally only one sample
- Less liquid means better cooling (less ice) and lower background





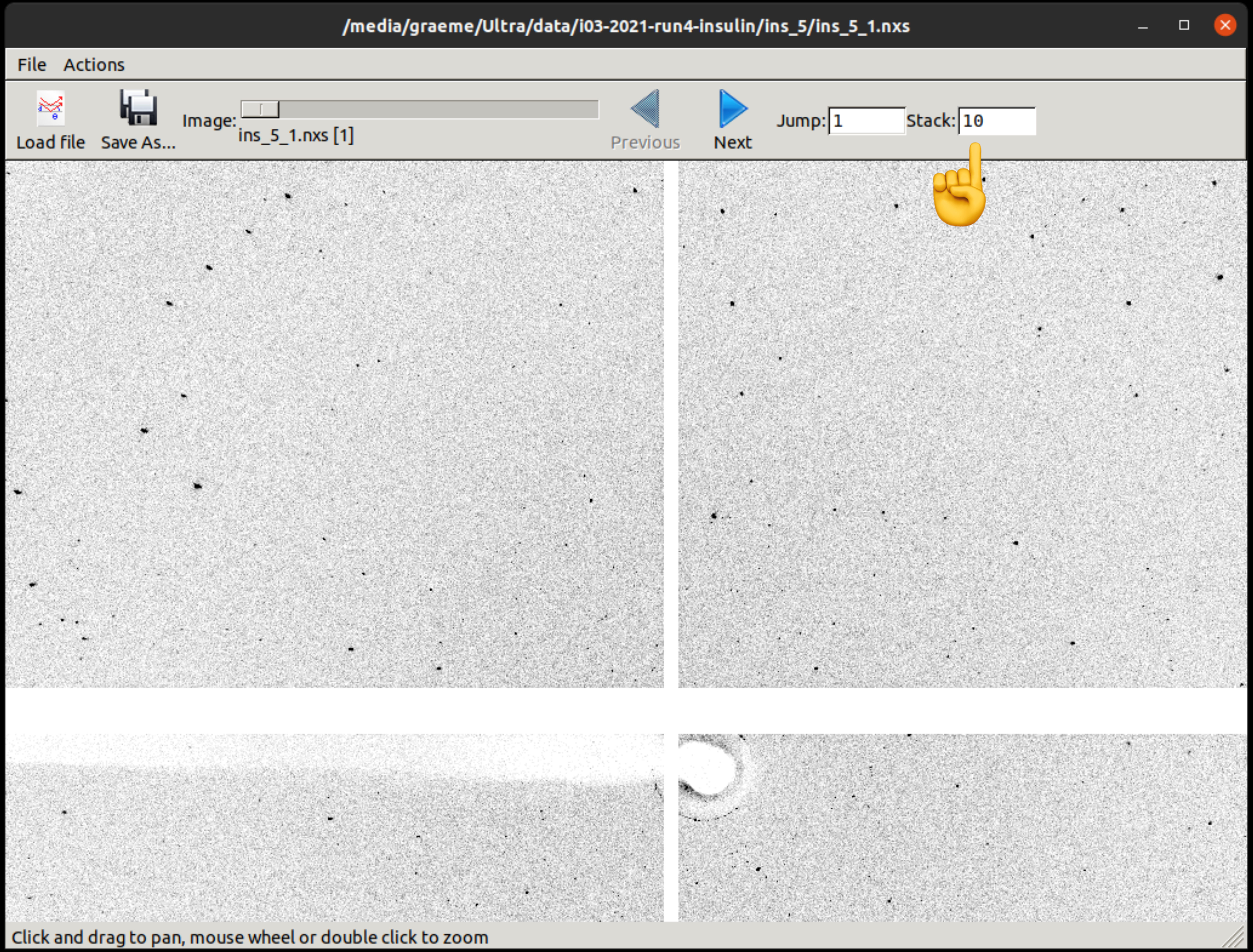
MINIMISE BACKGROUND

- The background is under the spot
- If we keep the rotation narrow, we have less background “before and after” the spot
- Strong spots taken from spot finding - most are rather narrow even with 0.1° images (median / mean ~ 2 images)
- 1° vs. 0.1° would have 10x background



Histogram of strong spot widths for insulin crystal with 0.1° / frame - most spots 2 pixels wide

LOOKING AT IMAGES



- Some people collect 1° images because "always have done"
- Some because they understand the data that way
- Better to use tools properly -> learn to use image viewers

KNOW YOUR EXPERIMENT

- Collecting data is there to meet experiment goals
- If you know what you are trying to do, you can collect “sub optimal” data which still meets your needs but could be more efficient e.g. wider oscillations
- Keep in mind the count rate correction
- All other things being equal however want rotation width < half mosaic spread

IDEA #3

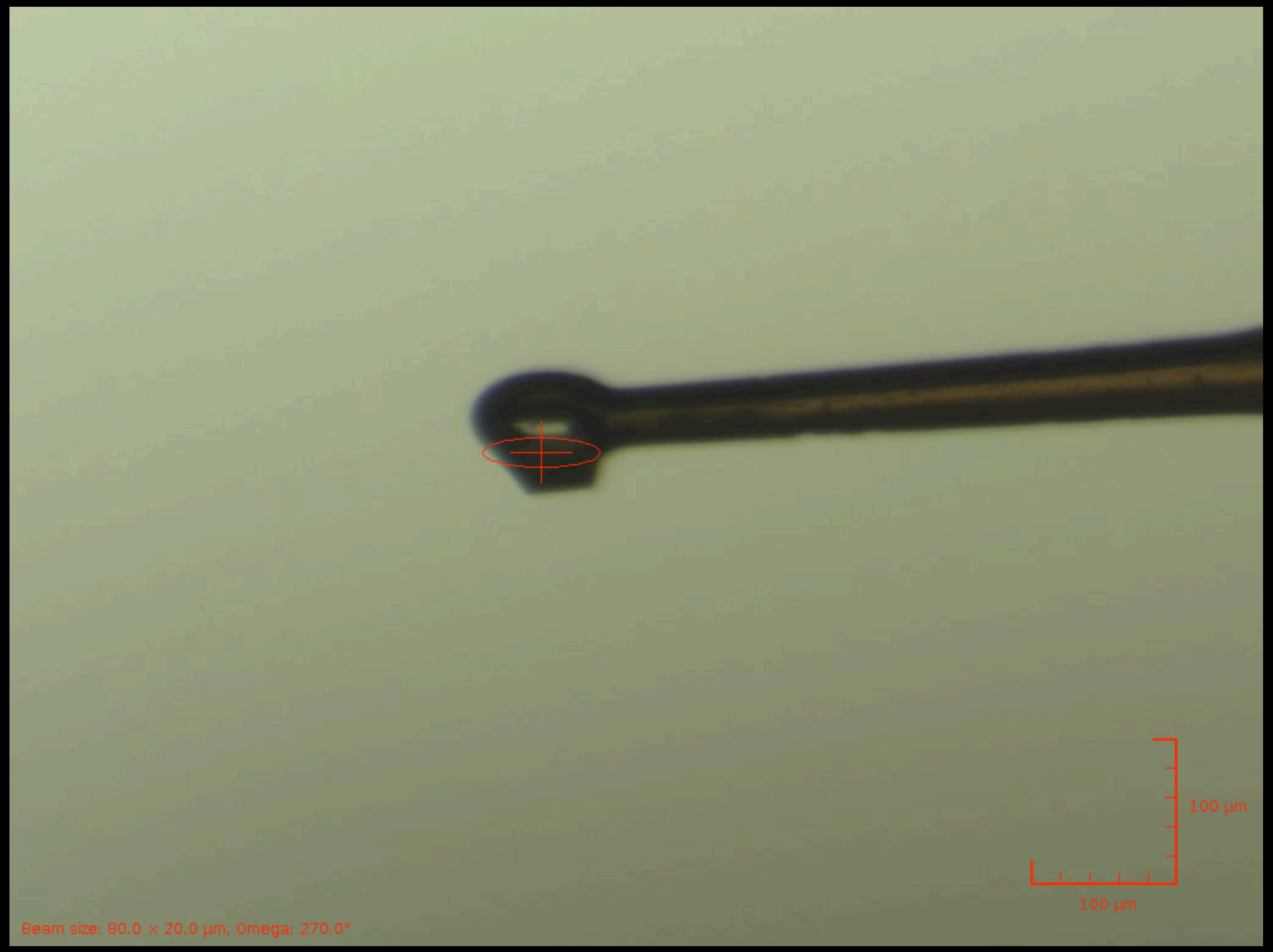
- Know your experiment goals and collect data to achieve them
- If you don't know anything specific, try to collect 360°+ of low dose finely sliced data
- Learn to look at the data properly

RADIATION DAMAGE

- Typically single biggest problem
- Cannot really be corrected for, or undone - perform experiment carefully
- Usually at workshop have “tricky data sets with added radiation damage”
- The data collection is the last experimental step -> everything later can be repeated etc.
- Where possible avoid radiation damage rather than hoping to correct for it in software

MULTIPLE SAMPLES

- Assumption is your structure / result is indicative of the “truth”
- Combining data from many samples should get you closer to the truth
- Can also help to reduce impact of radiation damage / other experimental effects



IDEA #4

- Radiation damage big problem - may need to combine data from multiple samples to achieve goals
- Carefully collecting more than you need allows options later
- Avoiding radiation damage much easier than correcting for it

CONCLUSIONS

- Ask the beamline staff for advice
- Do not destroy your samples
- Detectors can be saturated - lots of weak data better than a little very strong data
- Know your experimental goals - learn how your samples behave then apply this knowledge
- Where possible, avoid radiation damage and don’t be afraid of combining data from many samples

ACKNOWLEDGEMENTS

- Folks from DECTRIS and elsewhere on writing up how the detectors actually work
- Time on Diamond beamlines to collect example data, characterise detectors
- Samples! Thank you to team at Diamond for the steady supply of good quality test crystals

CONCLUSIONS

- Ask the beamline staff for advice
- Do not destroy your samples
- Detectors can be saturated - lots of weak data better than a little very strong data
- Know your experimental goals - learn how your samples behave then apply this knowledge
- Where possible, avoid radiation damage and don’t be afraid of combining data from many samples