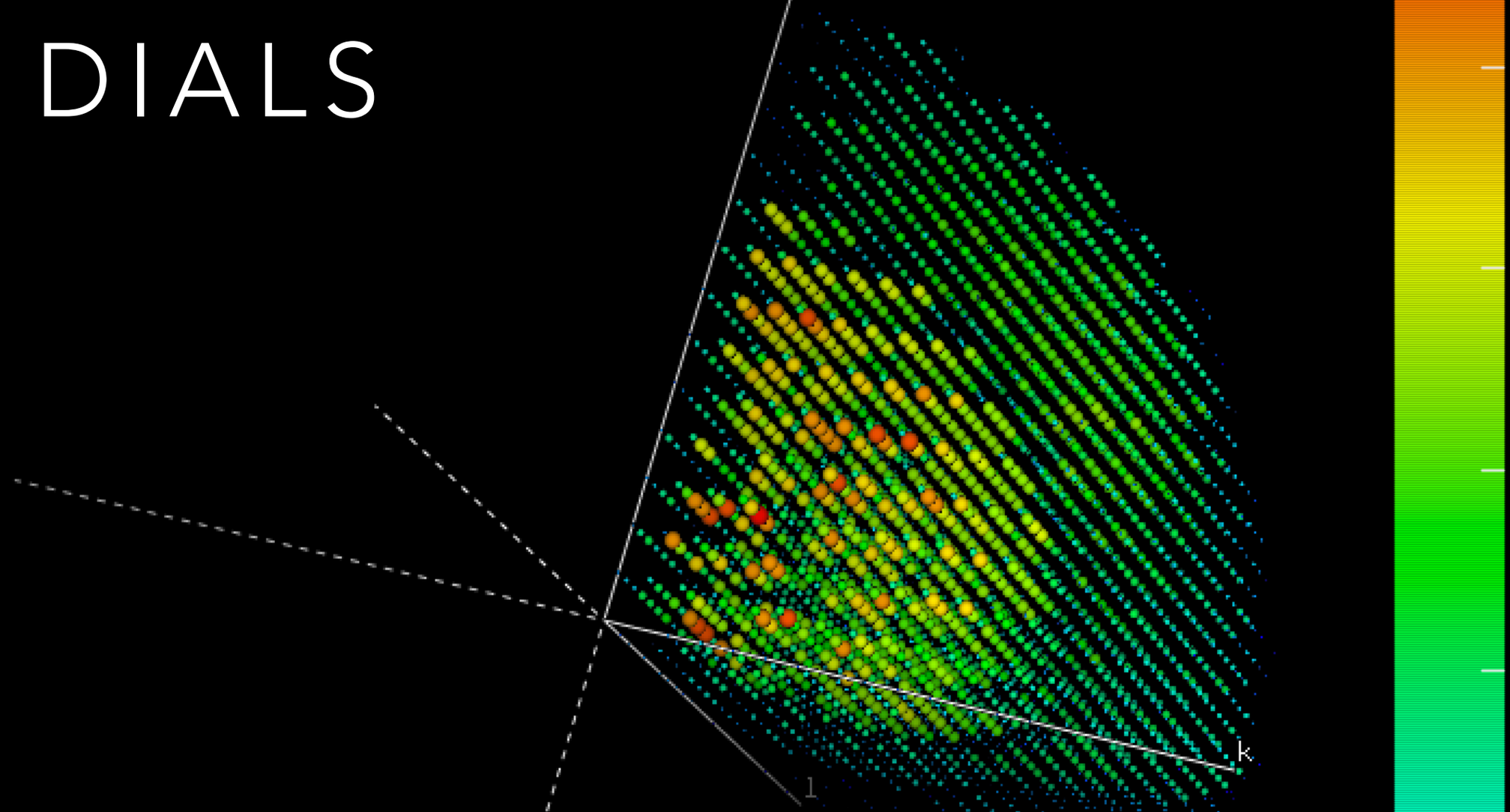
DATA PROCESSING WITH DIALS



NICHOLAS DEVENISH

diamond

DIALS
Diffraction Integration for Advanced Light Sources

CCP4 WORKSHOP - 1ST DECEMBER 2019

DATA PROCESSING

DATA PROCESSING PROBLEMS

Most can be solved before you run a single program -

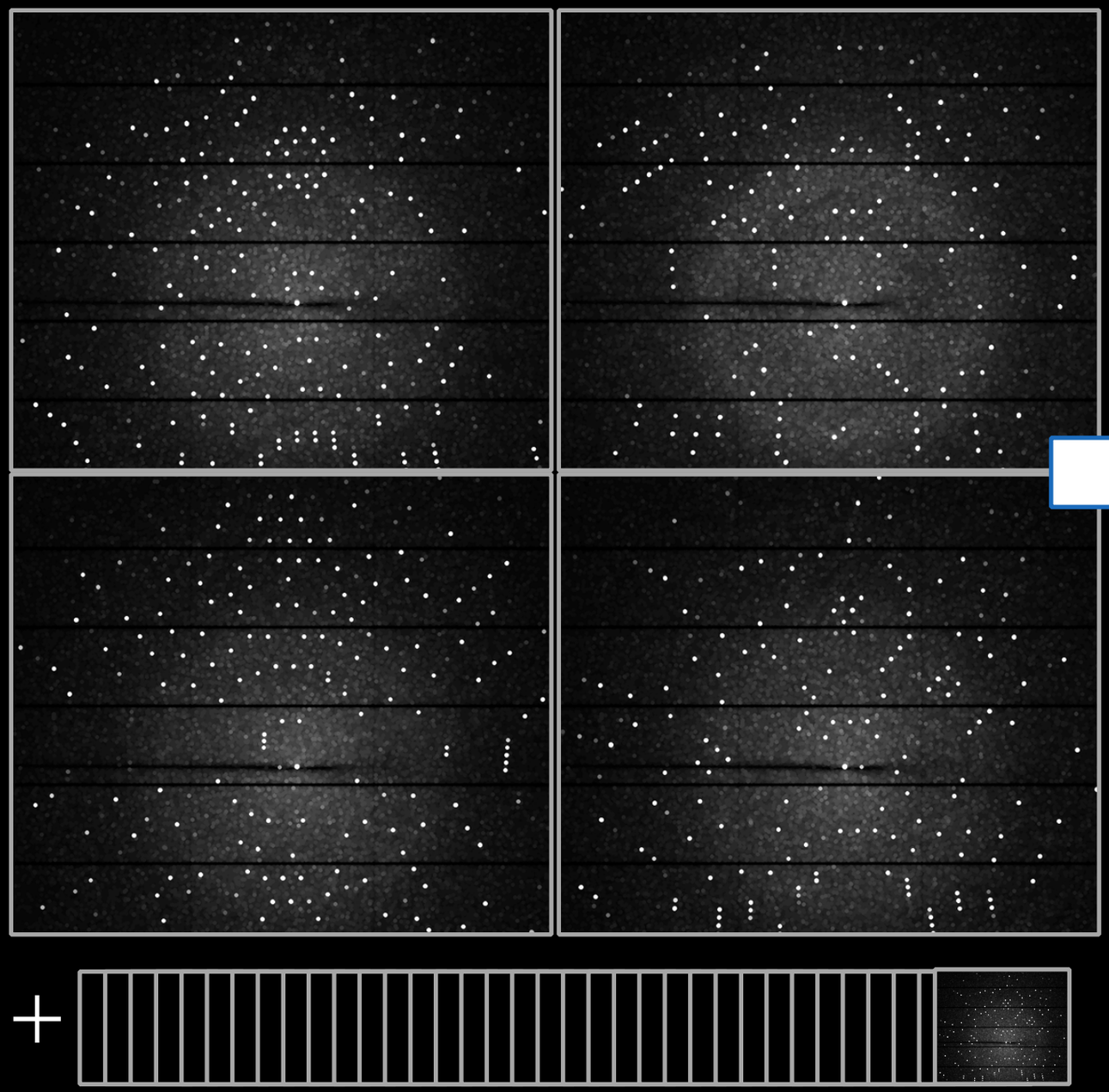
- mount your sample carefully
- collect your data carefully
- avoid / minimise radiation damage

DATA PROCESSING PROBLEMS

Most can be solved before you run a single program -

- mount your sample carefully
- collect your data carefully
- avoid / minimise radiation damage
    - Pixel array detectors (pilatus, eiger) have no readout noise - collecting weaker but more data can help

WHAT DO WE WANT TO KNOW?





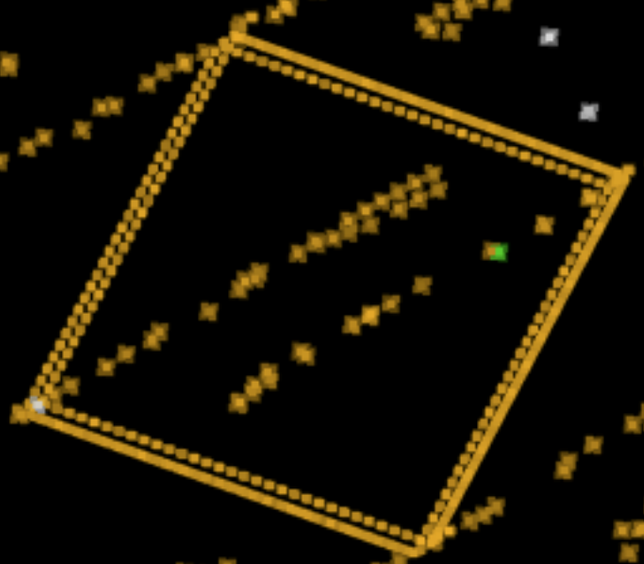
(EVENTUALLY)

NAME THAT SPOT (MILLER INDICES)

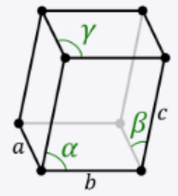
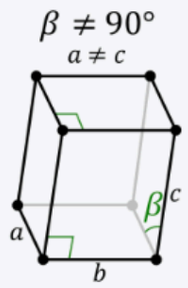
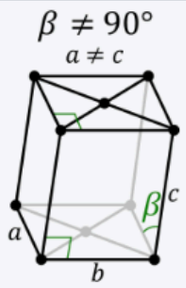
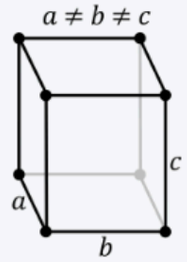
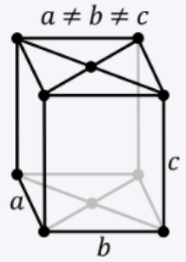
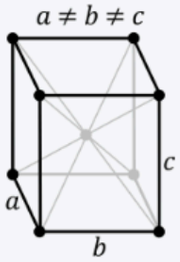
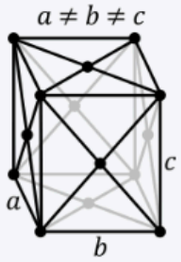
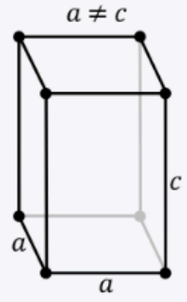
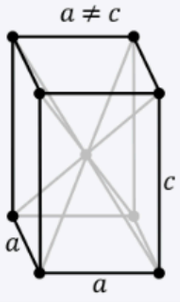
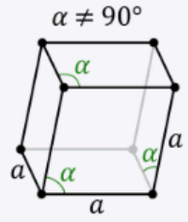
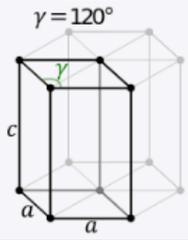
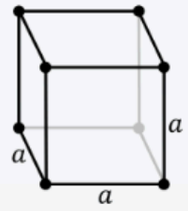
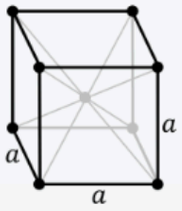
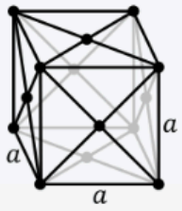
WHICH SPOT IS WHICH?



WHAT SHAPE IS THE UNIT CELL?

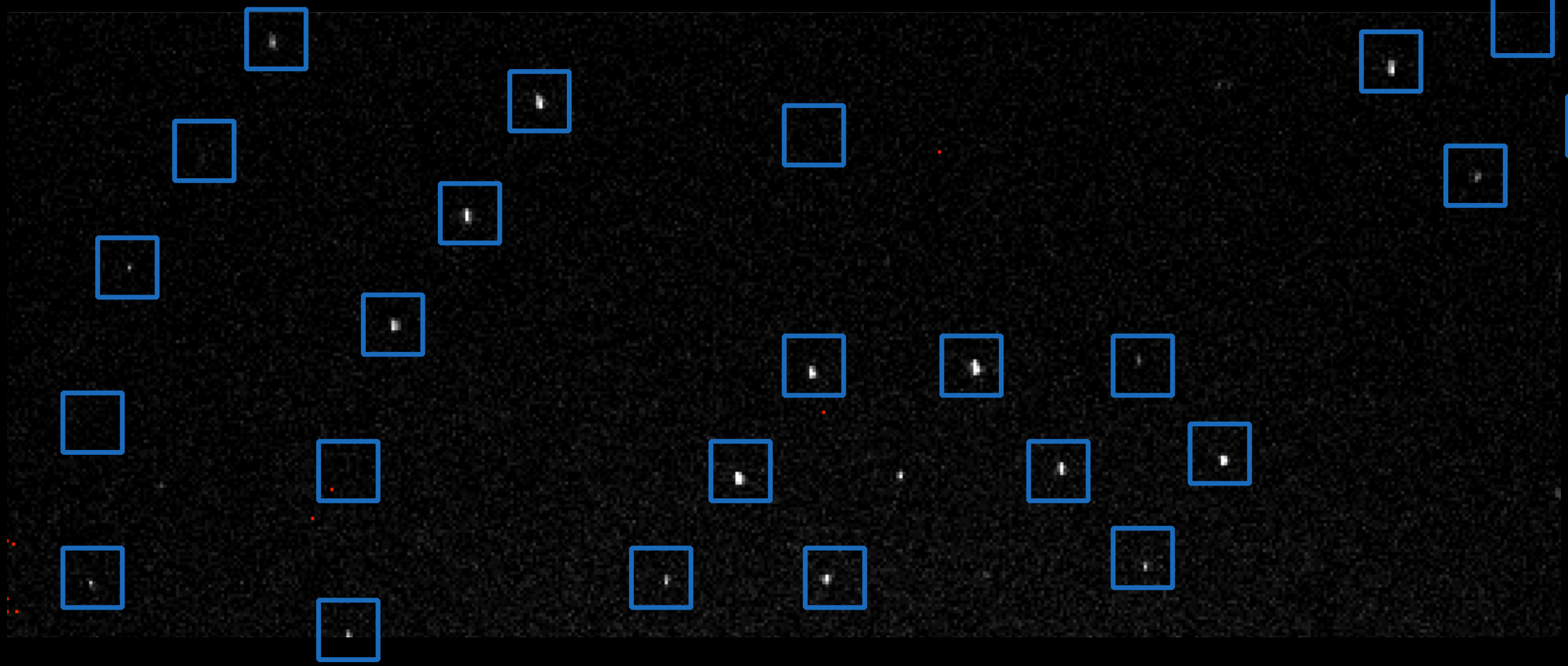


WHAT IS THE CORRECT BRAVAIS LATTICE?

Crystal Family	Lattice System	Schönflies	14 Bravais Lattices			
			Primitive	Base-centered	Body-centered	Face-centered
Triclinic		C_i				
Monoclinic		C_{2h}				
Orthorhombic		D_{2h}				
Tetragonal		D_{4h}				
Hexagonal	Rhombohedral	D_{3d}				
	Hexagonal	D_{6h}				
Cubic		O_h				

DATA INTEGRATION

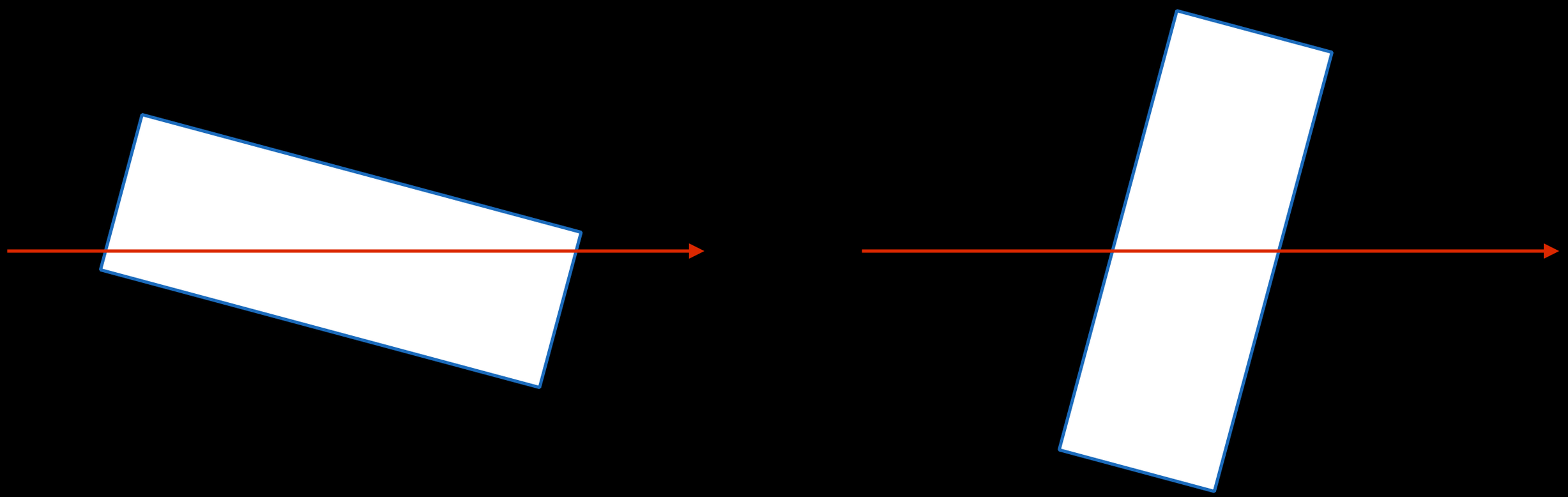
HOW MANY PHOTONS?



...EVEN FOR SPOTS WE COULDN'T SEE

CORRECTIONS FOR EXPERIMENTAL EFFECTS

PLACE ON COMMON SCALE



DIALS

~~DIALS~~

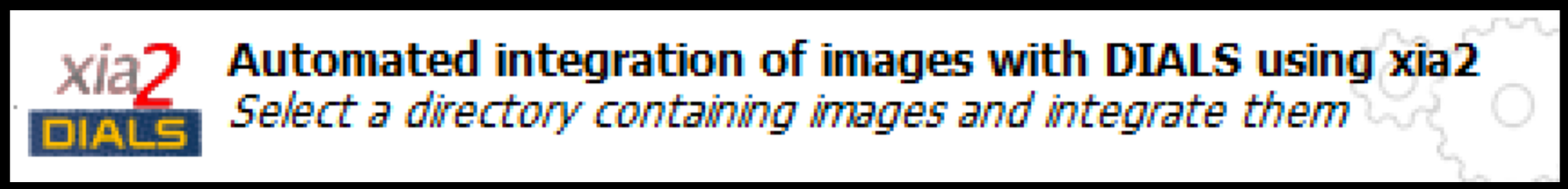
Fastest route to Intensities:
Automation!

generate_XDS.INP / autoPROC /
xia2 / fast_dp / …. most facilities have
something

For dials: xia2 =>

```
xia2 atom=Zn image=../data/C2_sum_0001.cbf.gz
```

or CCP4i2



PROJECT SUPPORT

DIALS
Diffraction Integration for Advanced Light Sources

diamond

wellcome

BERKELEY LAB
Lawrence Berkeley National Laboratory

BioStructX

Science & Technology Facilities Council

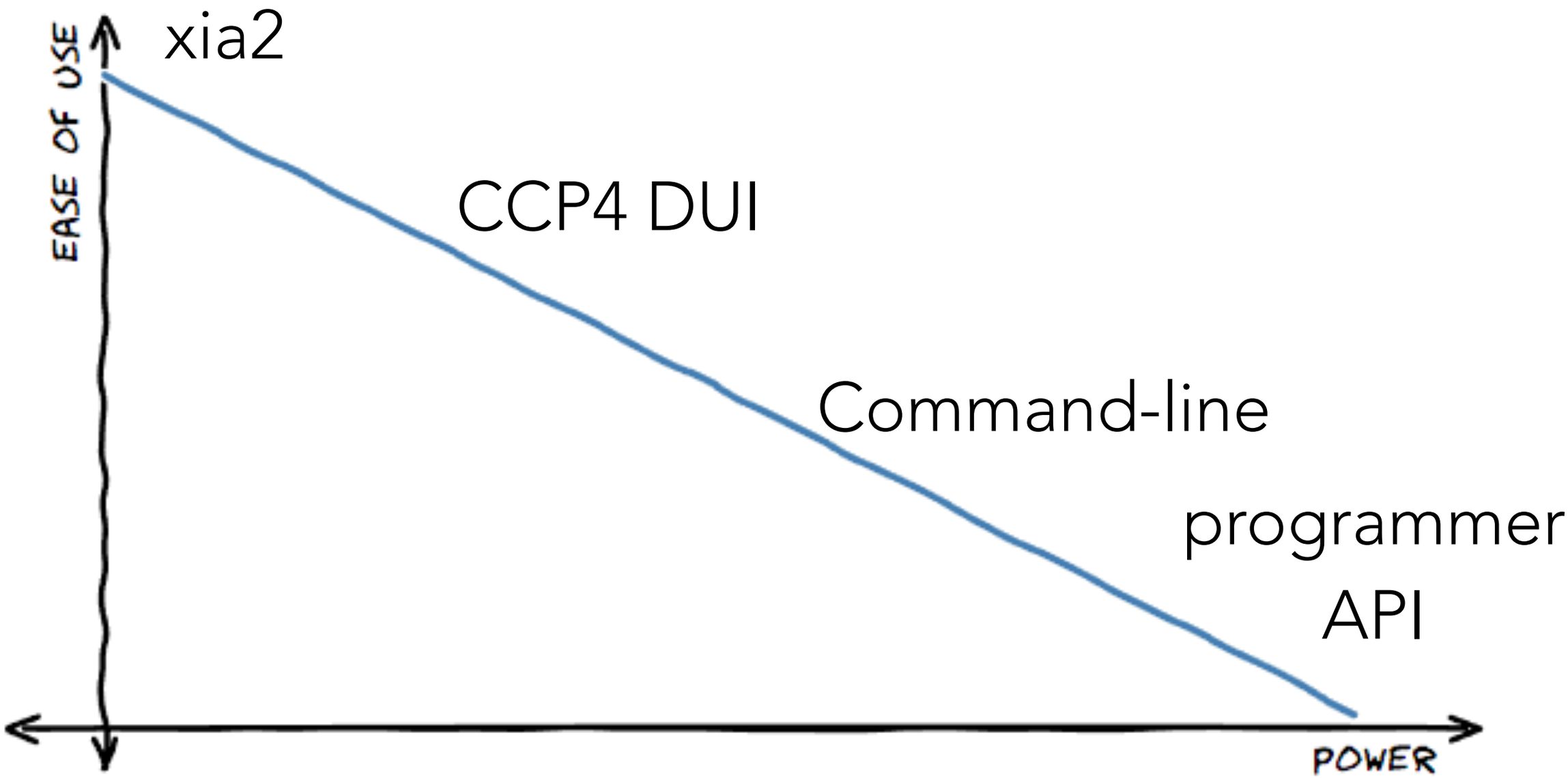
CCP4

EU FP7: #283570 NIH: GM095887 & GM102520

PHILOSOPHY: TOOLBOX



PHILOSOPHY: LEVELS OF INTERACTION



Free, open source (BSD Licence)

Download (Mac, Linux), find tutorials on:

https://dials.github.io

For Windows, in CCP4



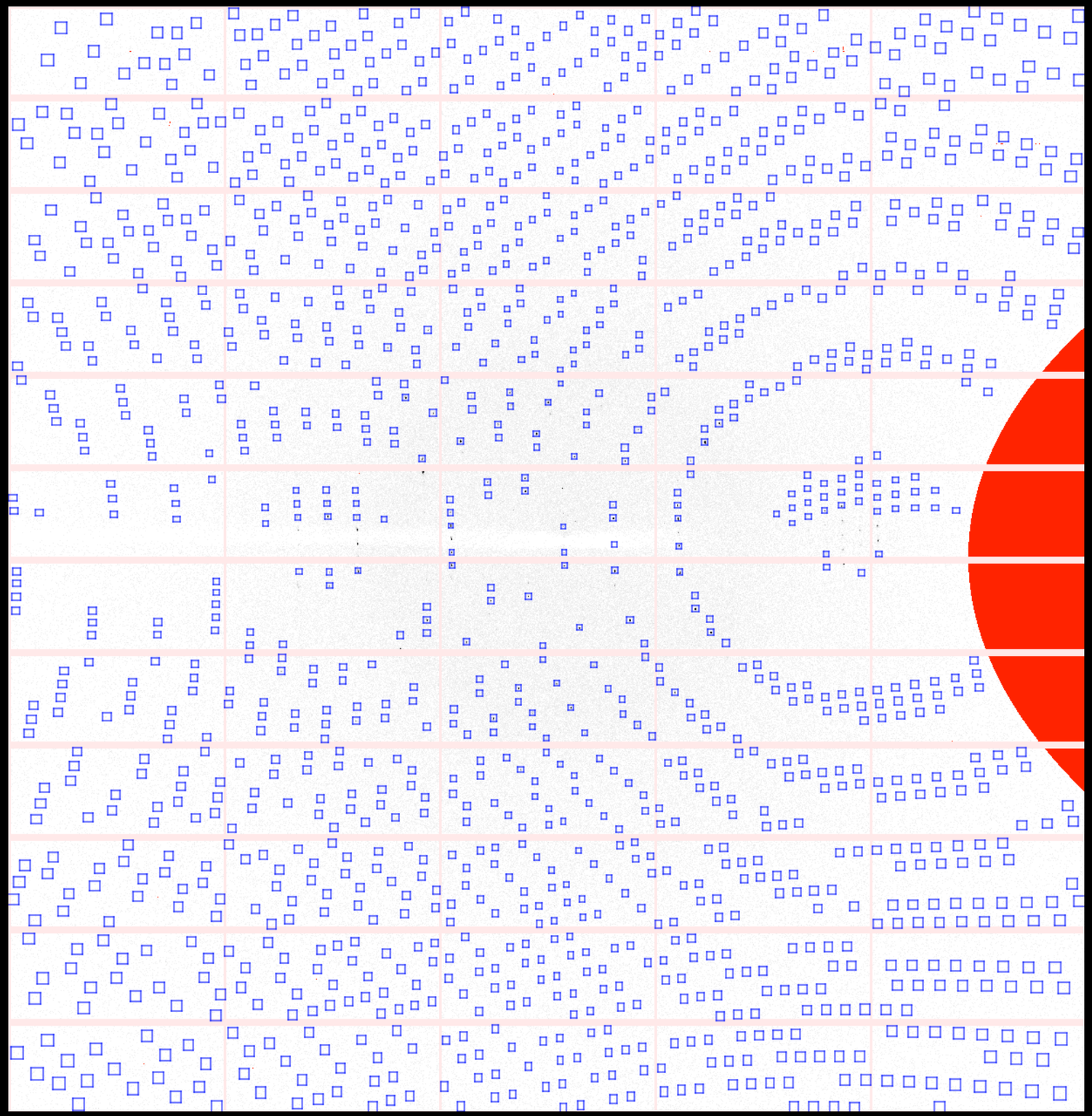
Home
About
Installation
Documentation
Tutorials
How-to
Workshops
Publications
Links
License

DIALS: Diffraction Integration for Advanced Light Sources

X-ray crystallography for structural biology has benefited greatly from a number of advances in recent years including high performance pixel array detectors, new beamlines capable of delivering micron and sub-micron focus and new light sources such as X-FELs. The DIALS project is a collaborative endeavour to develop new diffraction integration software to meet the data analysis requirements presented by these recent advances. There are three end goals: to develop an extensible framework for the development of algorithms to analyse X-ray diffraction data; the implementation of algorithms within this framework and finally a set of user facing tools using these algorithms to allow integration of data from diffraction experiments on synchrotron and free electron sources.

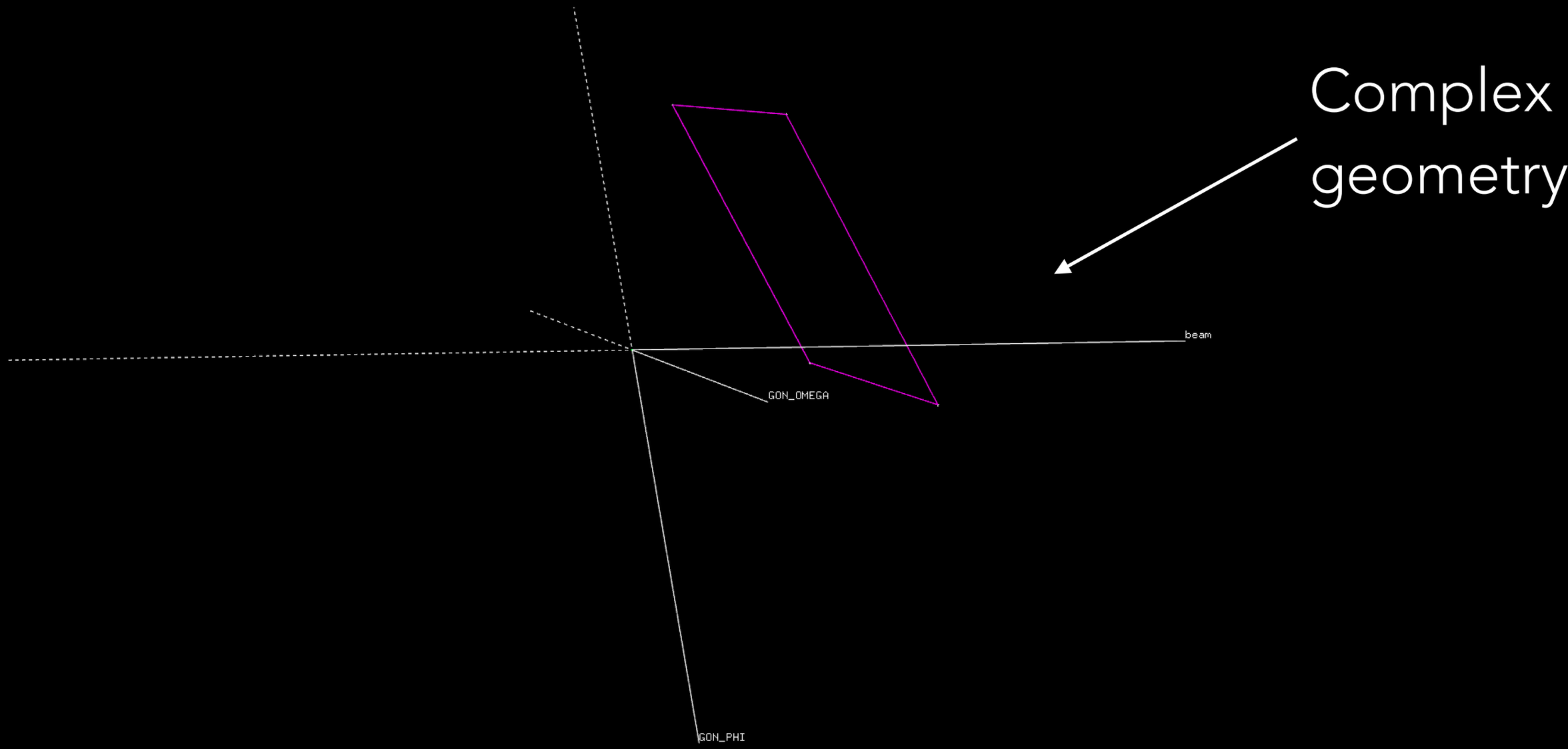
Contact

WHAT CAN DIALS COPE WITH?

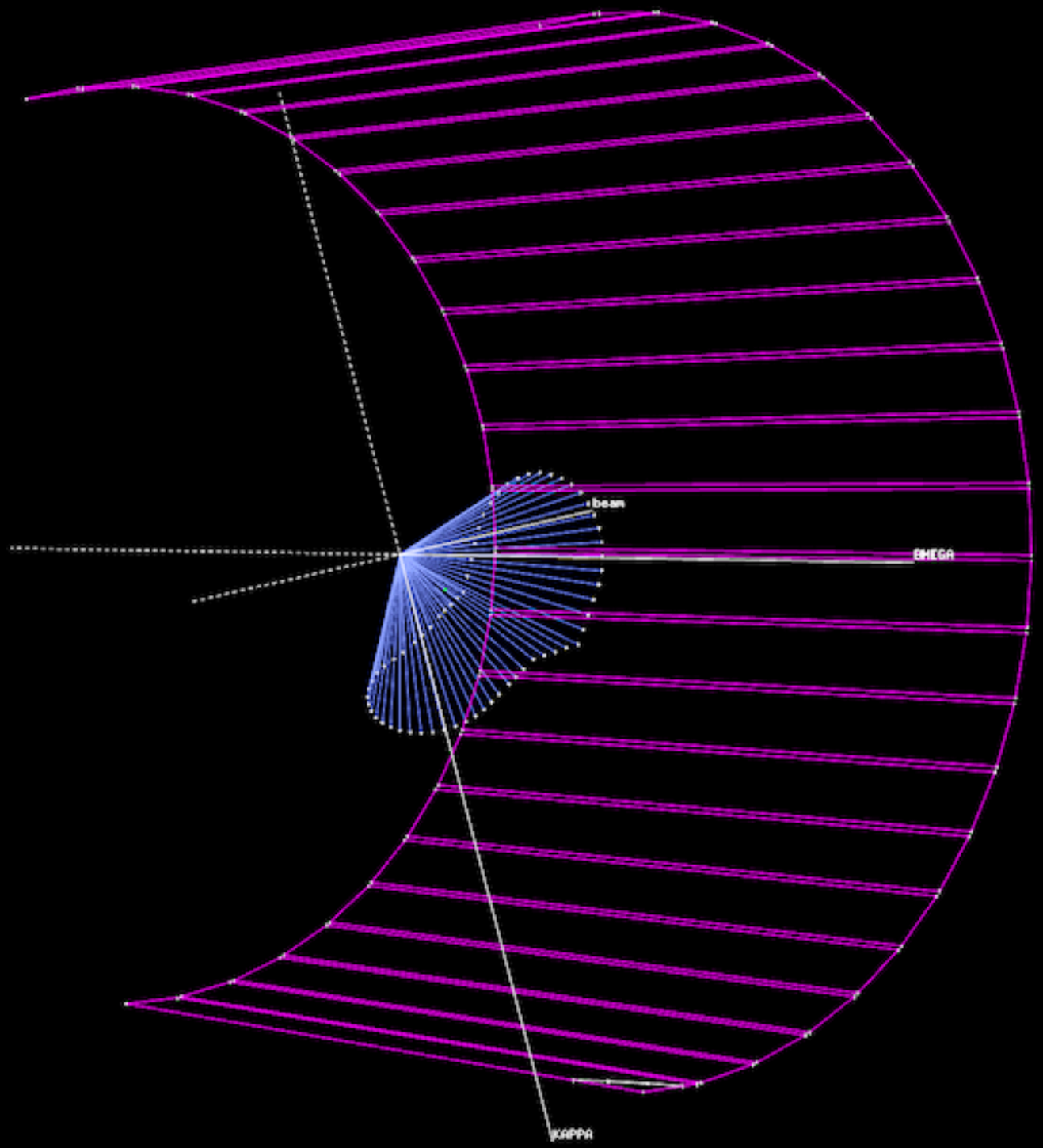


Moving shadows!

WHAT CAN DIALS COPE WITH?

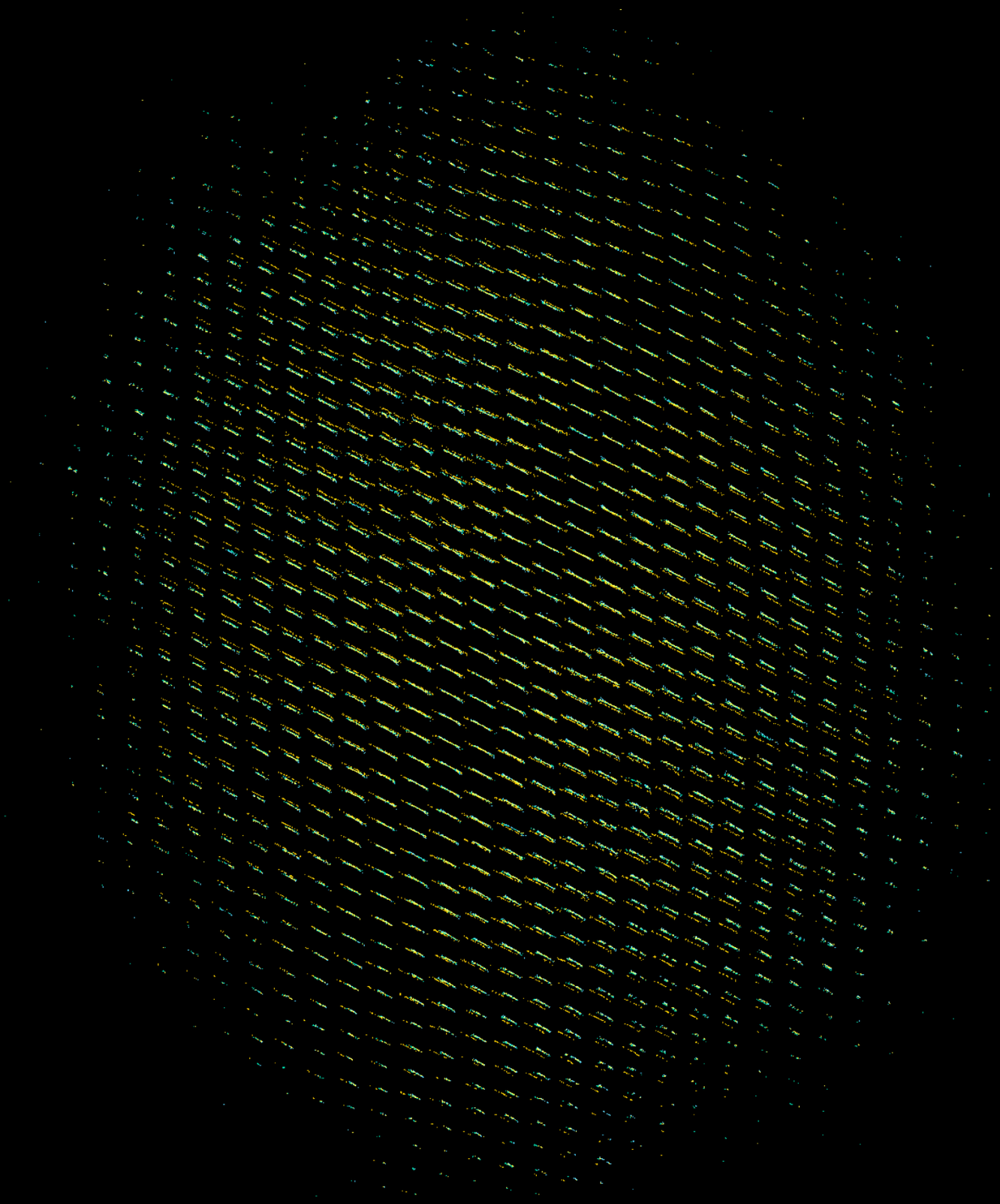


WHAT CAN DIALS COPE WITH?



Complex detectors

WHAT CAN DIALS COPE WITH?



Multiple orientations

WEAK DIFFRACTION

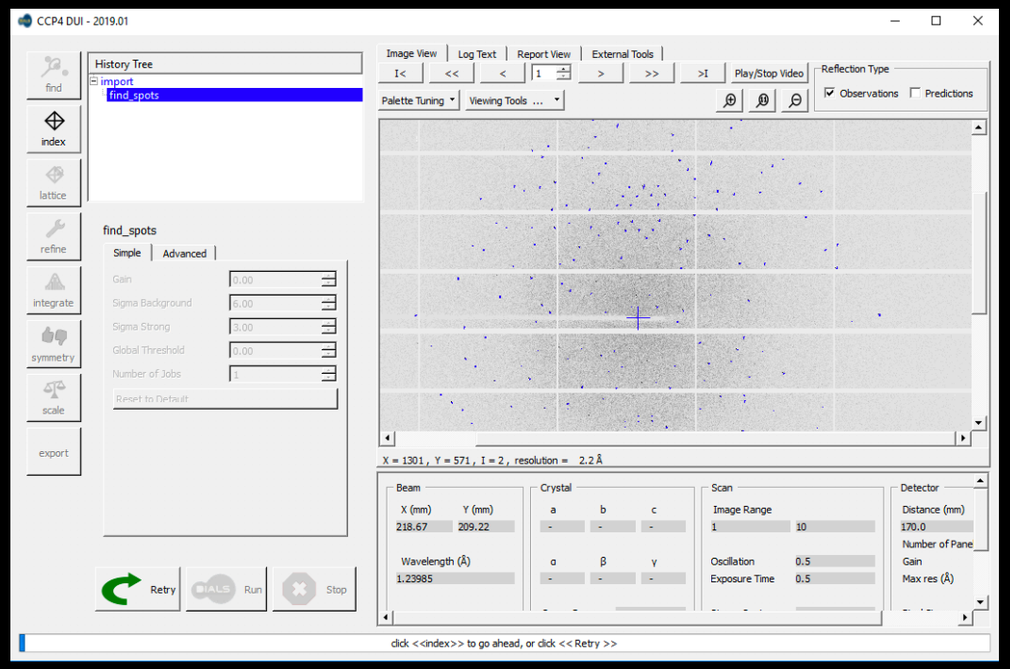


RUNNING DIALS

DIALS



CCP4 DUI



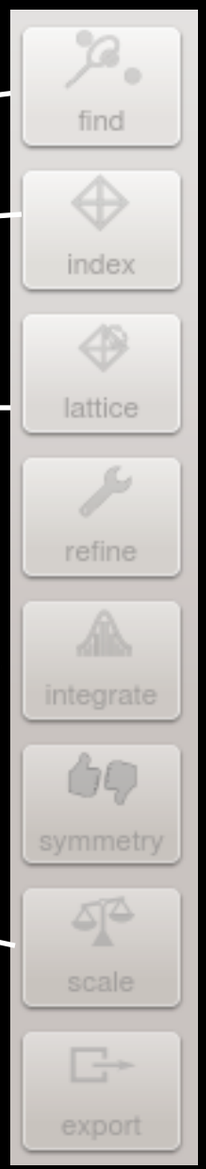
KEY STEPS IN DATA PROCESSING

- Look at your images <—— people often miss this one
- Spot finding
- Indexing (optionally selecting lattice)
- Refinement (may be part of indexing)
- Integration
- Post integration refinement
- Point group identification
- Scaling

MAIN DIALS PROGRAMS

DUI

dials.import
dials.find_spots
dials.index
dials.refine_bravais_settings
dials.refine
dials.integrate
dials.scale
dials.symmetry
dials.export



More than 50 other dials.* commands

IMAGE VIEWER

```
$ dials.image_viewer imported.expt
```

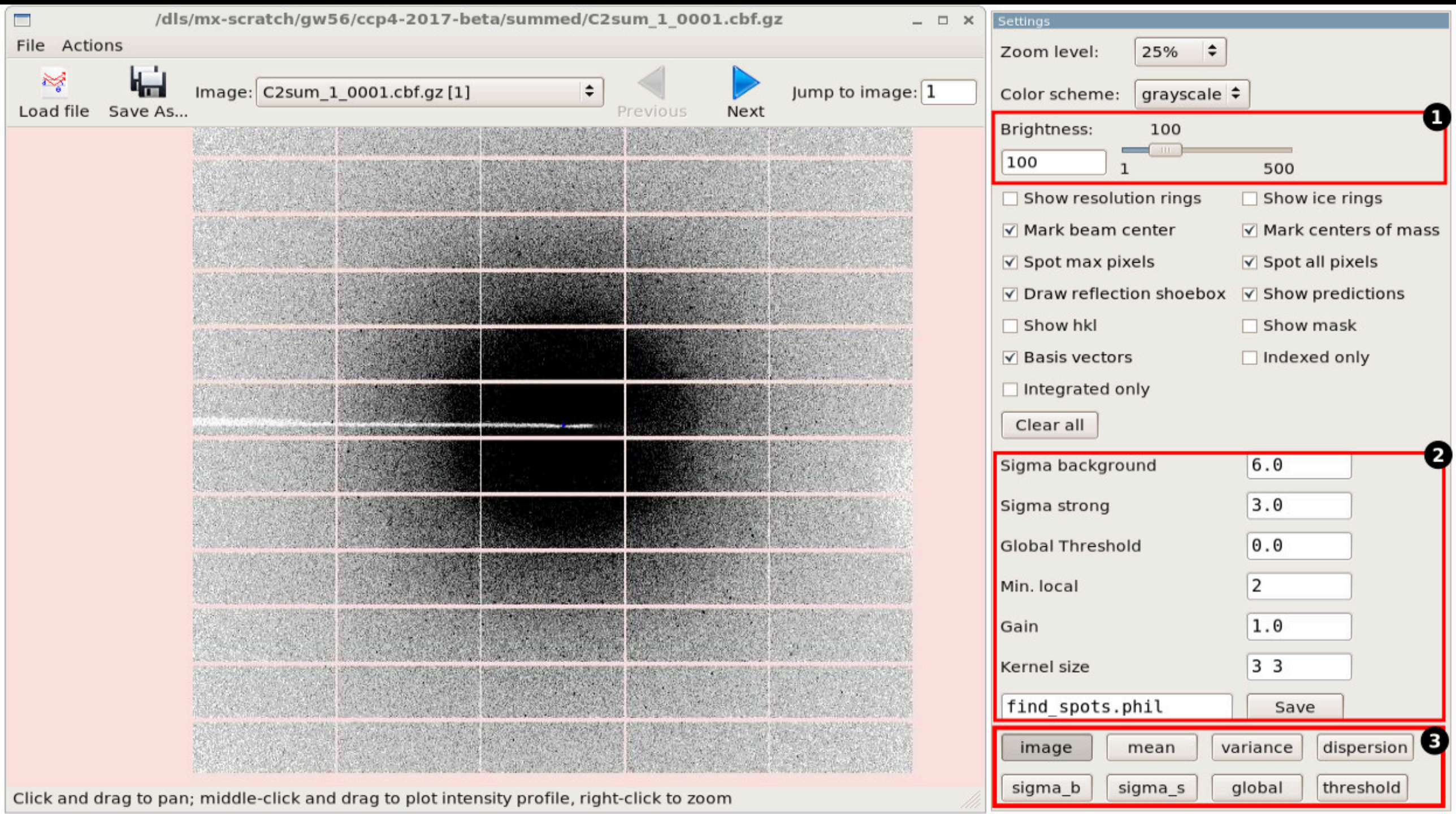
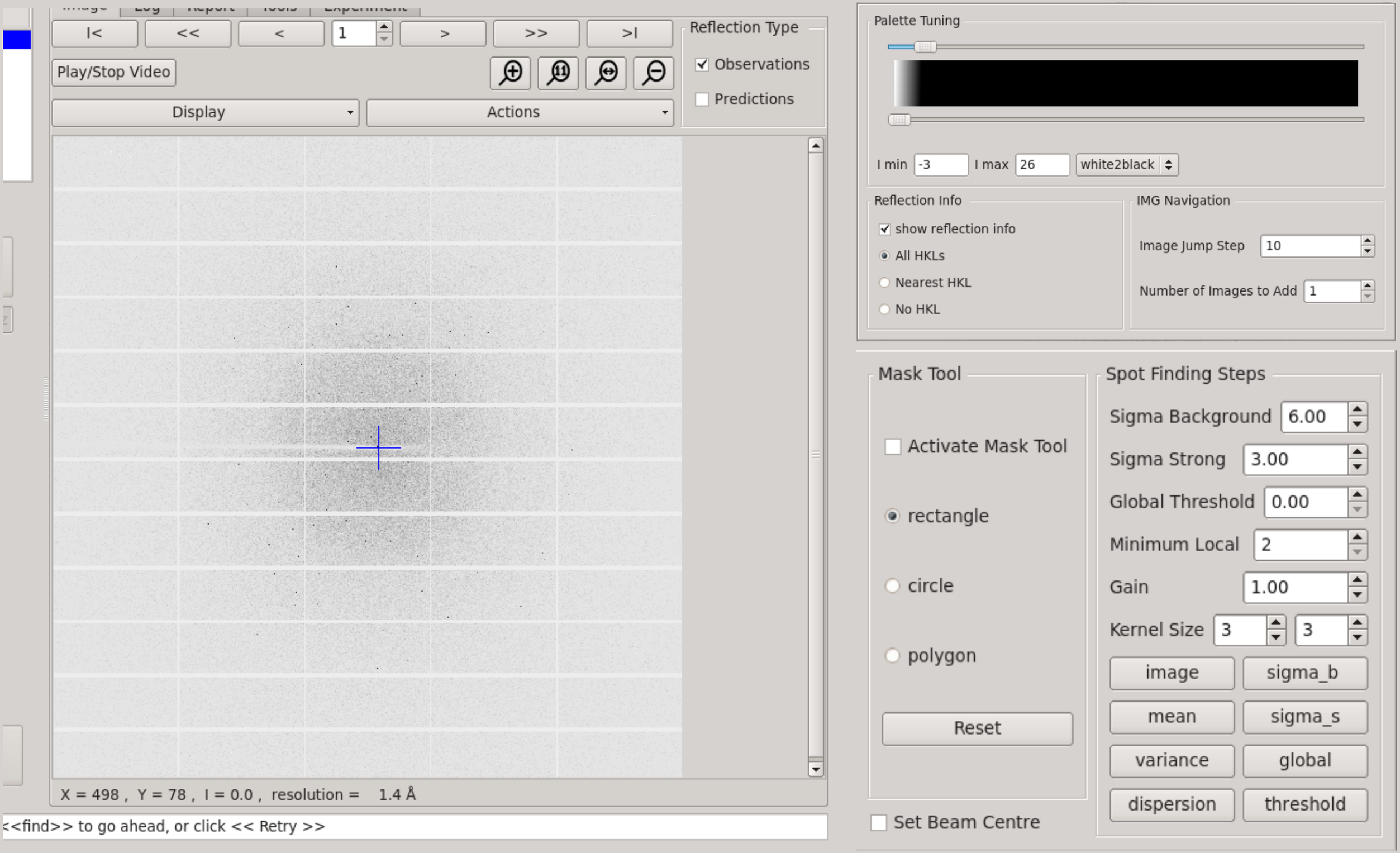


IMAGE VIEWER IN DUI



SPOT FINDING

DUI




```
$ dials.find_spots
```

Find spots by looking for pixels whose intensity are larger than the local area

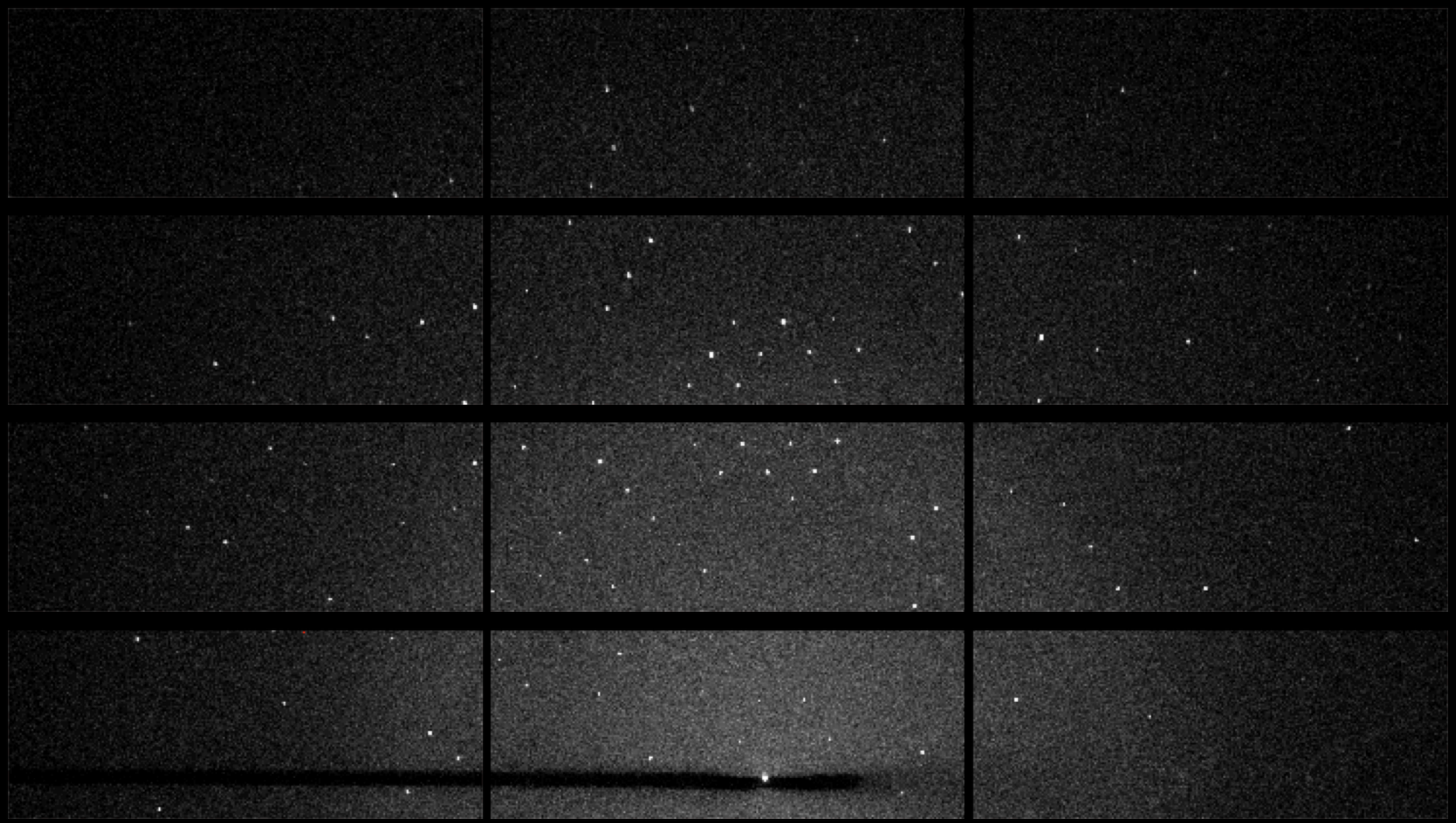
Images are analysed in 3D

Candidates can be filtered on:

- number of pixels
- peak-centroid distance
- resolution
- ice rings
- untrusted regions

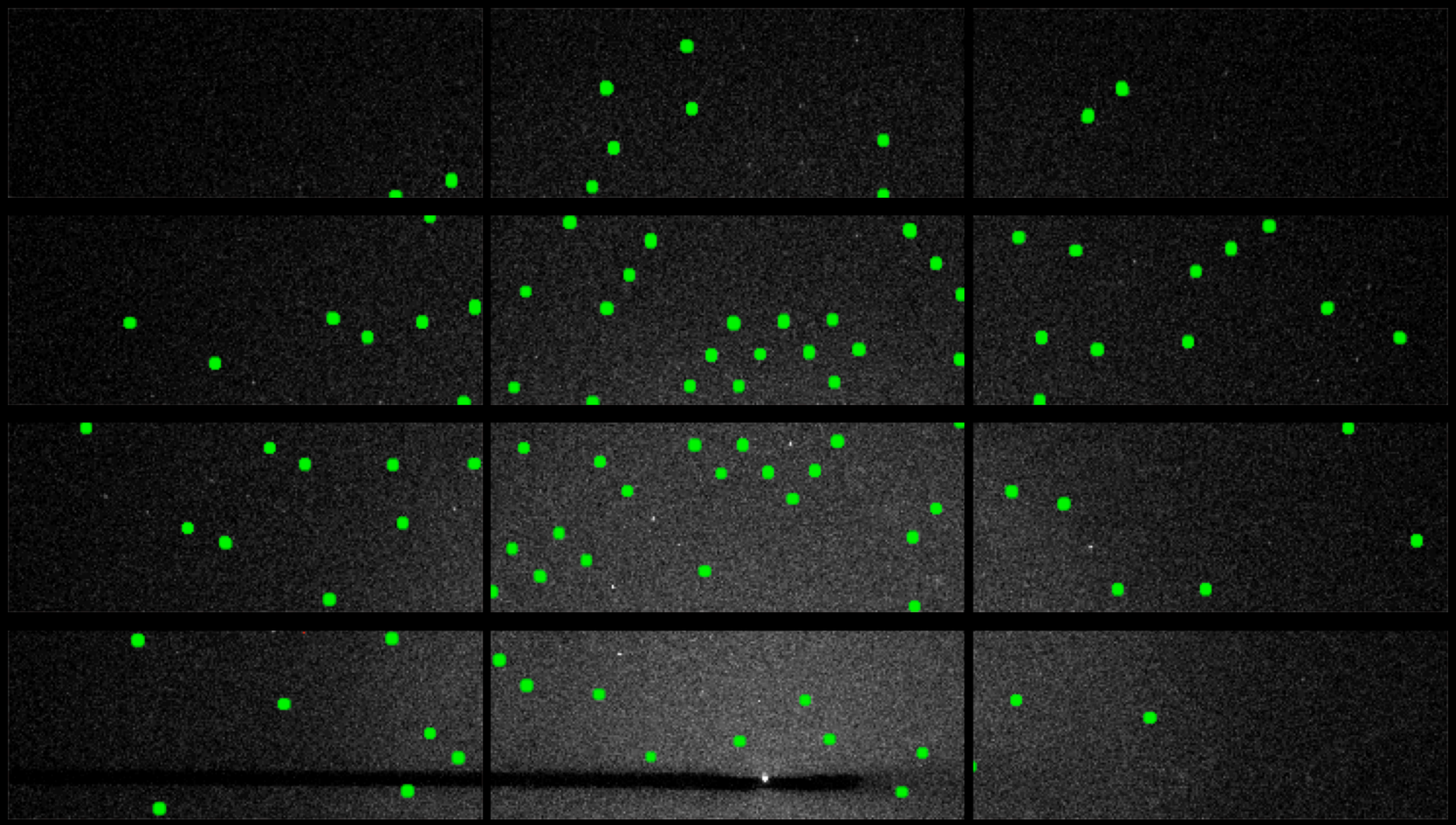
SPOT FINDING

dials.find_spots



Start with the image

SPOT FINDING



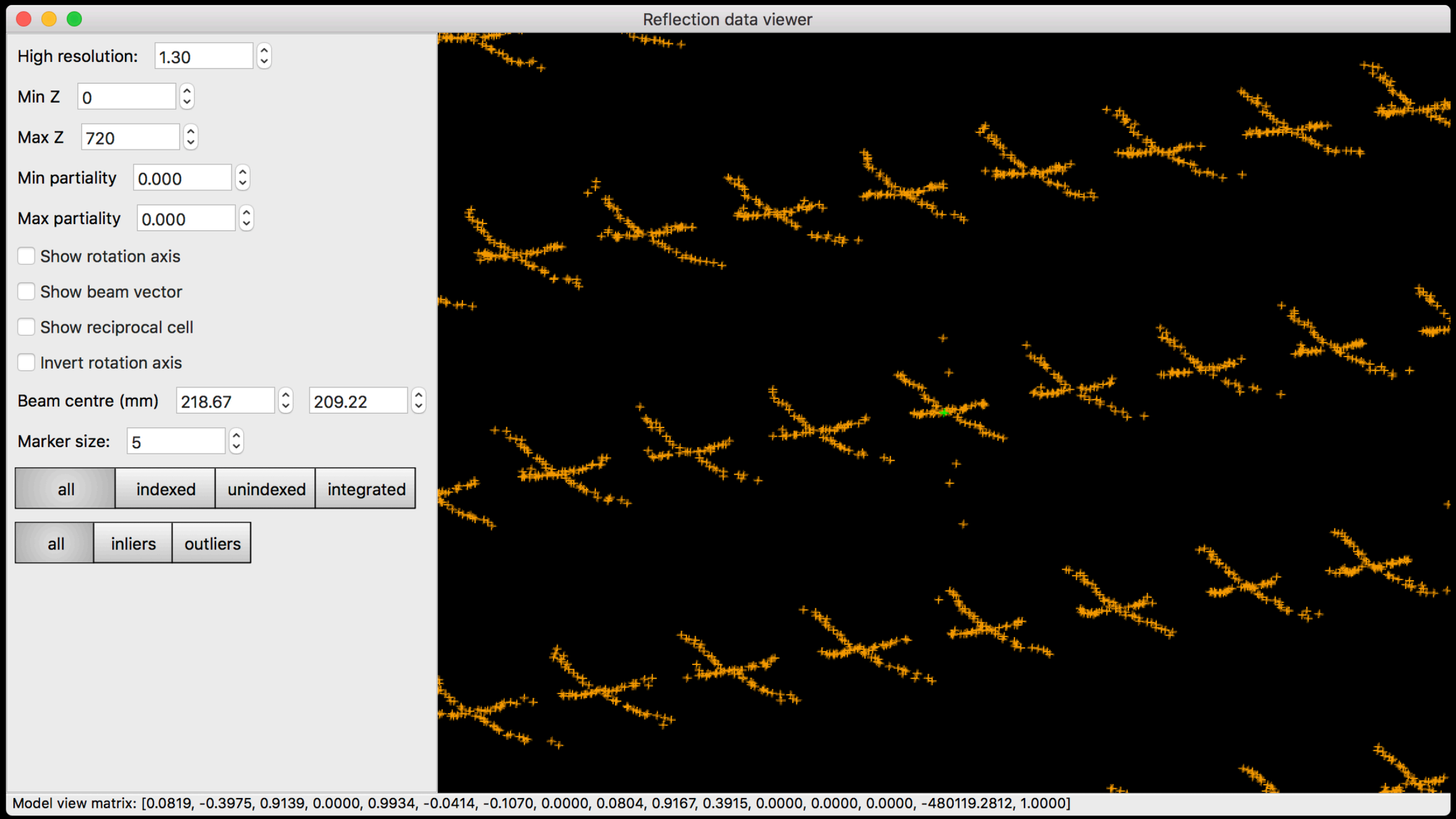
Identify strong reflections

MAP TO RECIPROCAL SPACE



(ASIDE)

RECIPROCAL LATTICE VIEWER

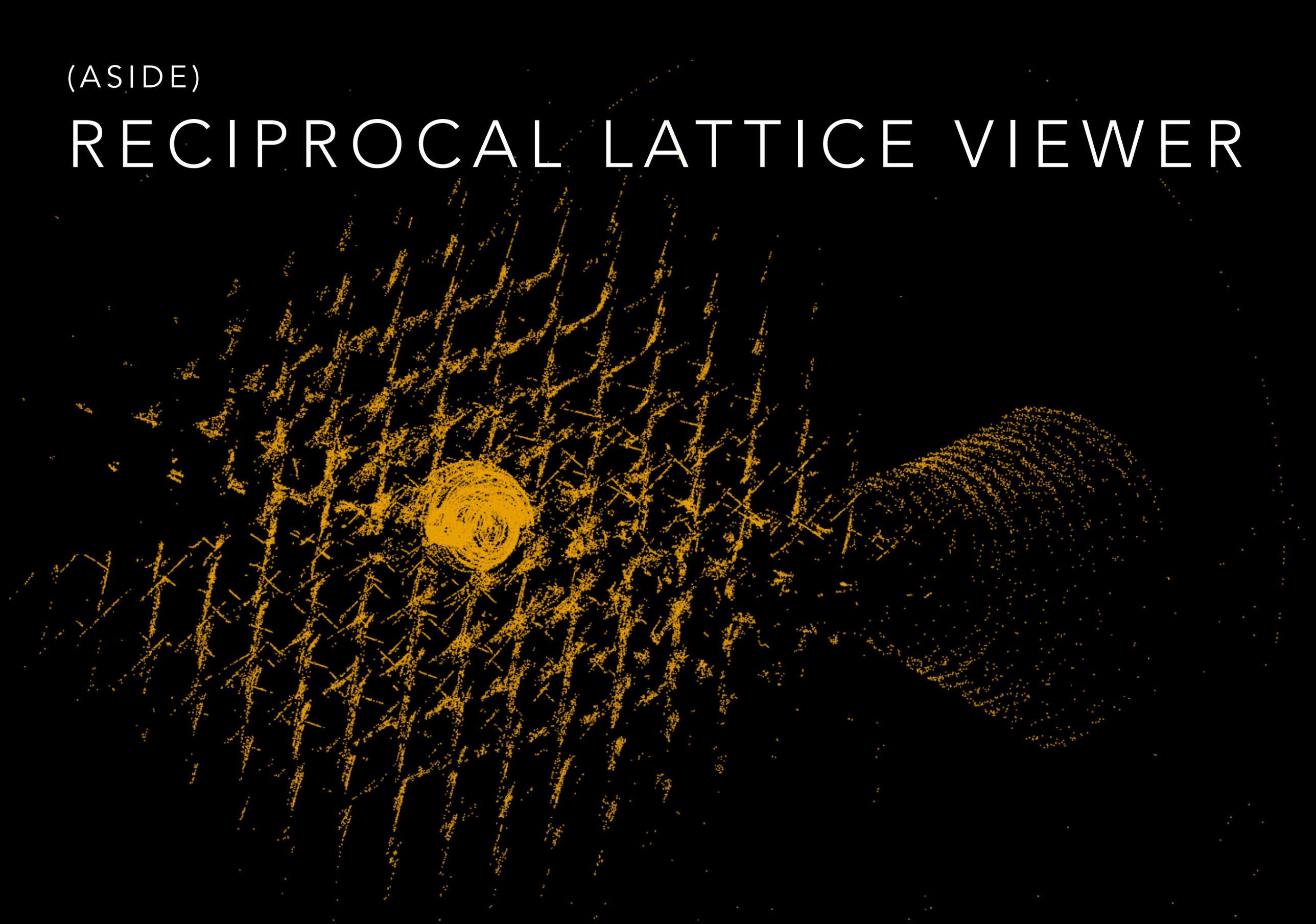


(ASIDE)

RECIPROCAL LATTICE VIEWER

(ASIDE)

RECIPROCAL LATTICE VIEWER



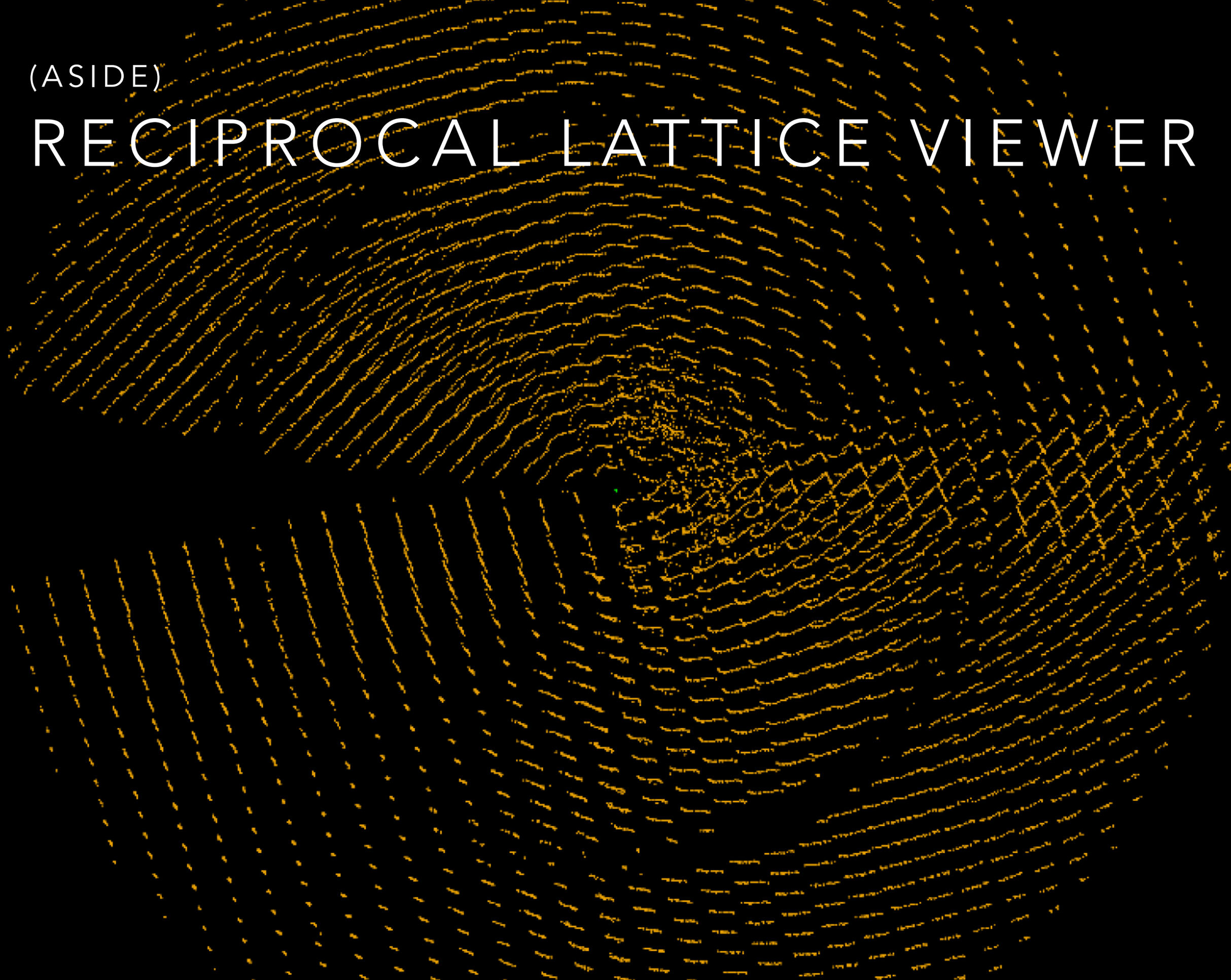
(ASIDE)

RECIPROCAL LATTICE VIEWER



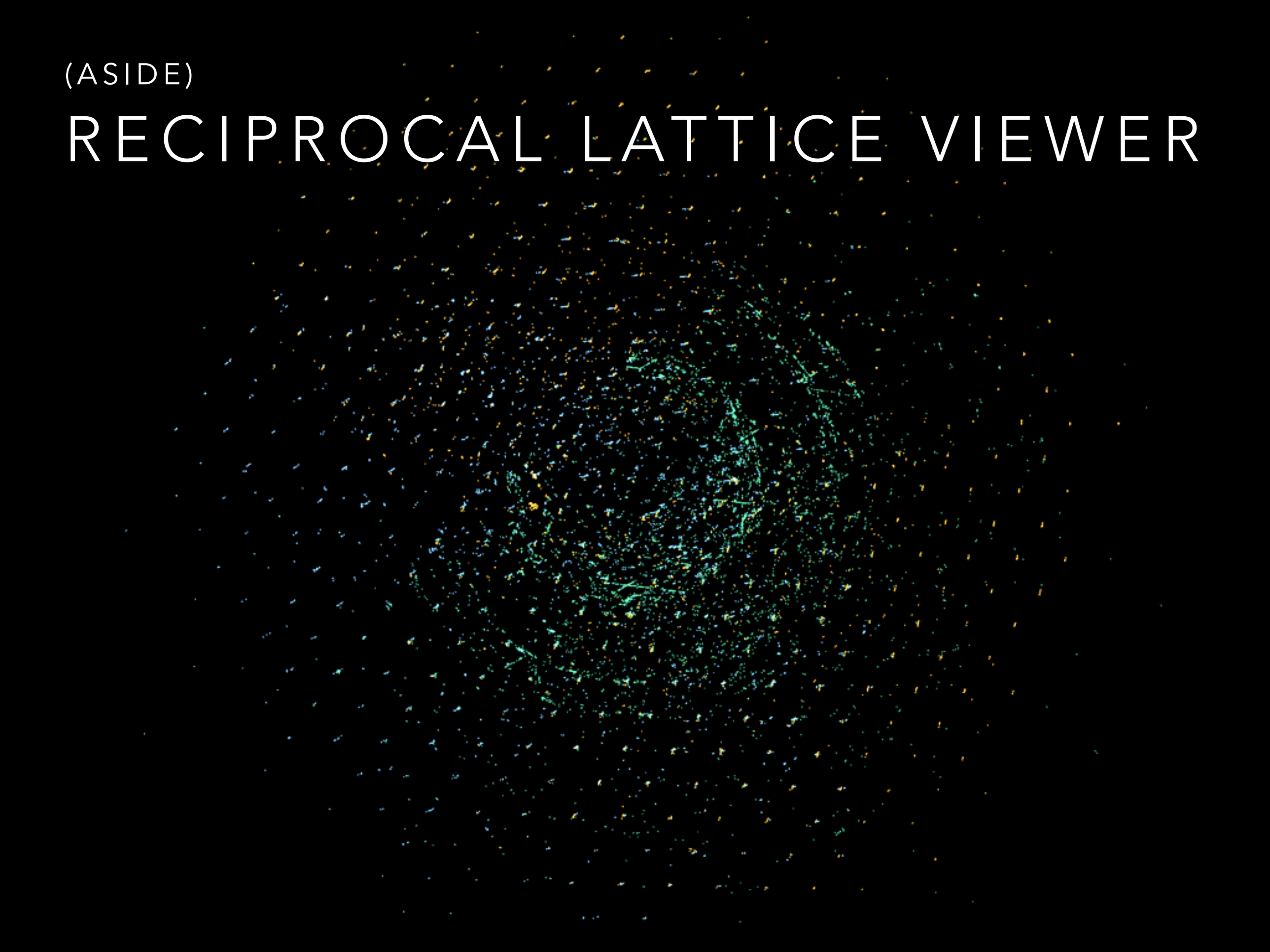
(ASIDE)

RECIPROCAL LATTICE VIEWER



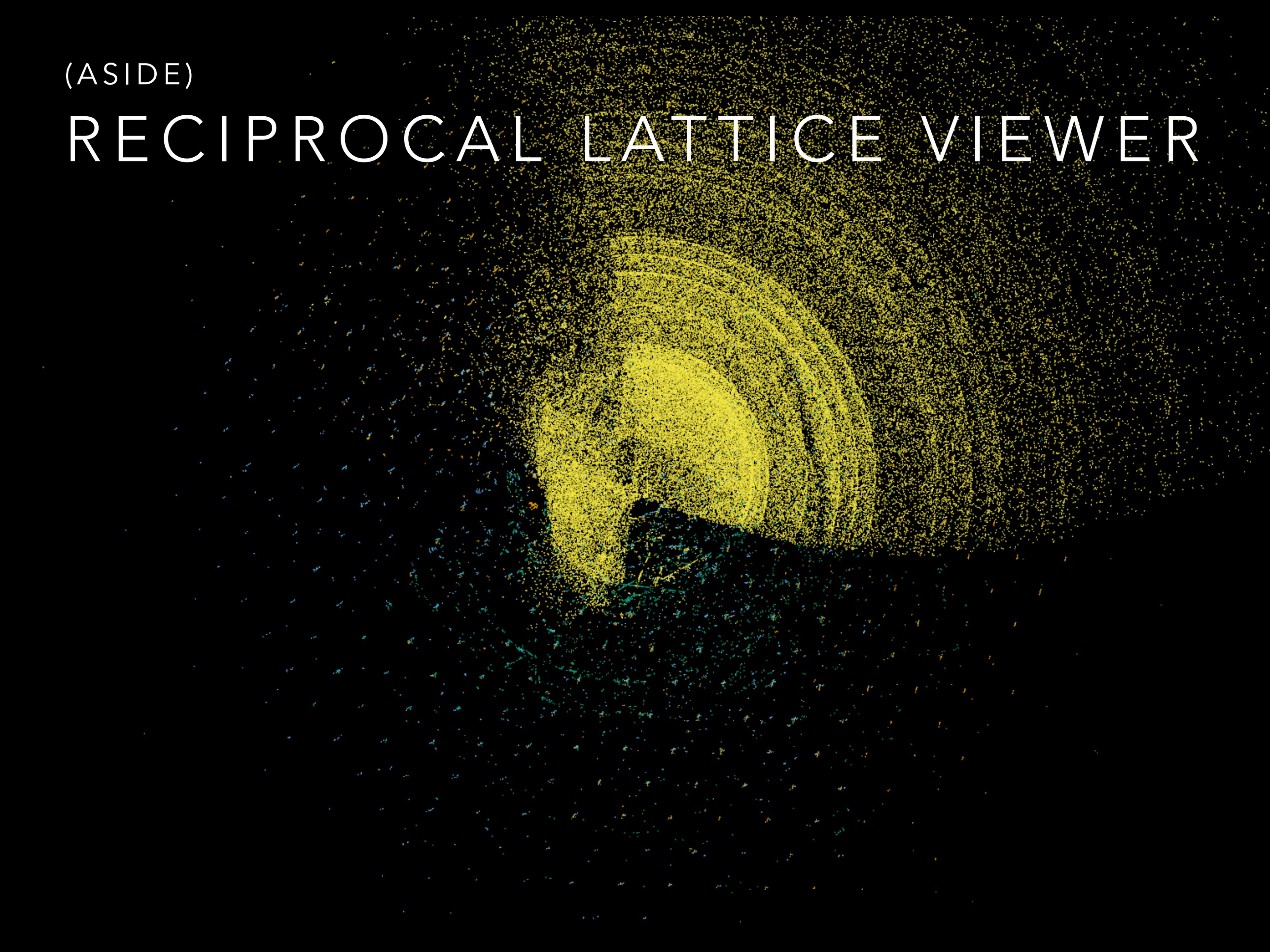
(ASIDE)

RECIPROCAL LATTICE VIEWER



(ASIDE)

RECIPROCAL LATTICE VIEWER



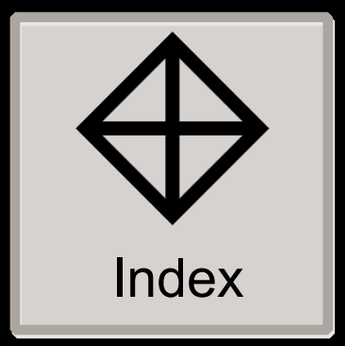
INDEXING

Map centroids to reciprocal space, analyse for periodicity and determine basis vectors for the reduced cell

- Choice of method:
  - 1D FFT
  - 3D FFT - default
  - real space grid search algorithm
- Optionally provide known unit cell and space group
- Begins refinement process

```
$ dials.index method=fft3d …
```

DUI



Reflection data viewer

High resolution: 1.31

Min Z 0

Max Z 720

Min Pixels 3

Max Pixels 91

Min partiality 0.000

Max partiality 0.000

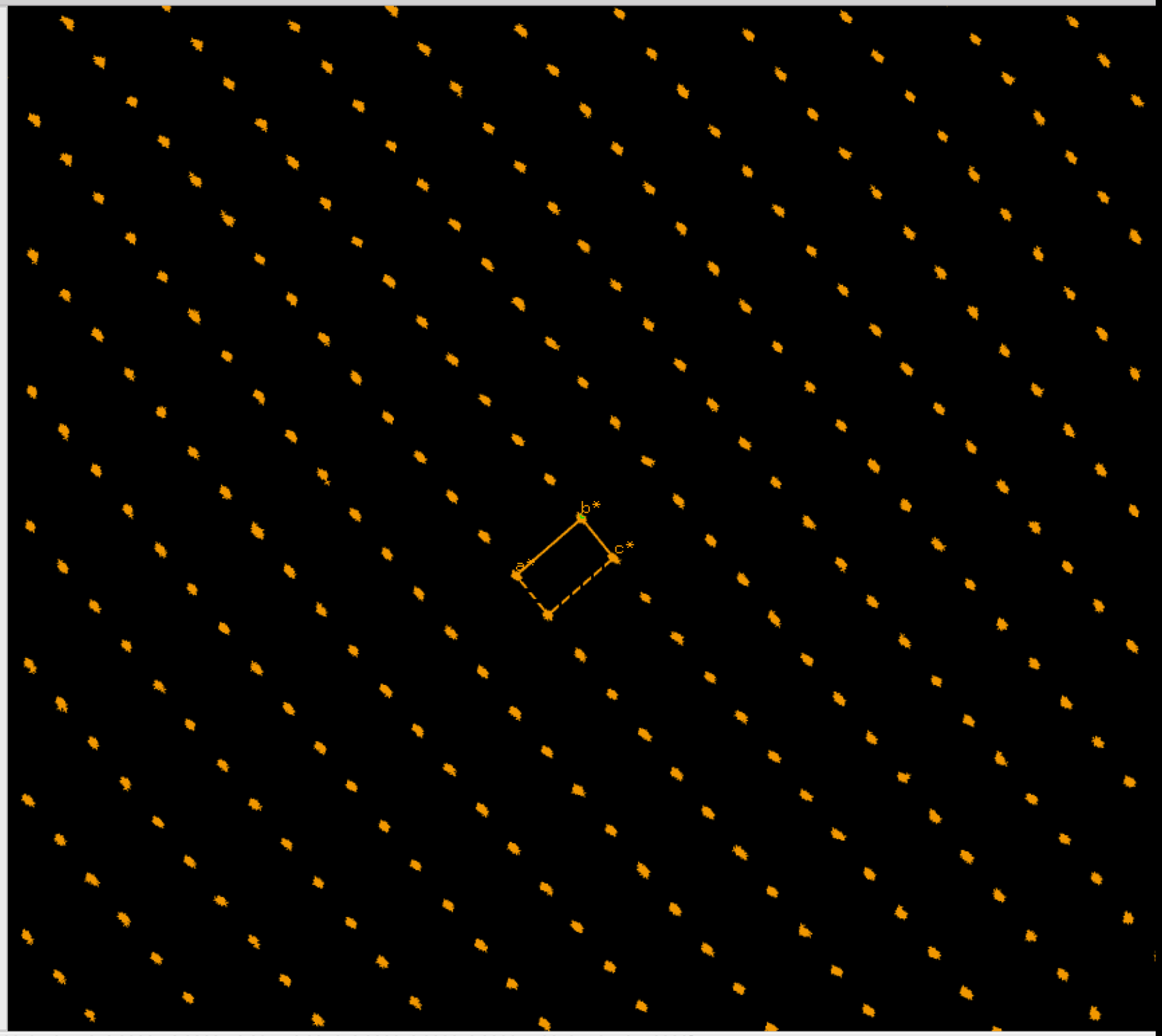
- [ ] Show rotation axis
- [ ] Show beam vector
- [x] Show reciprocal cell
- [ ] Invert rotation axis

Beam centre (mm) 218.57 209.18

Marker size: 10

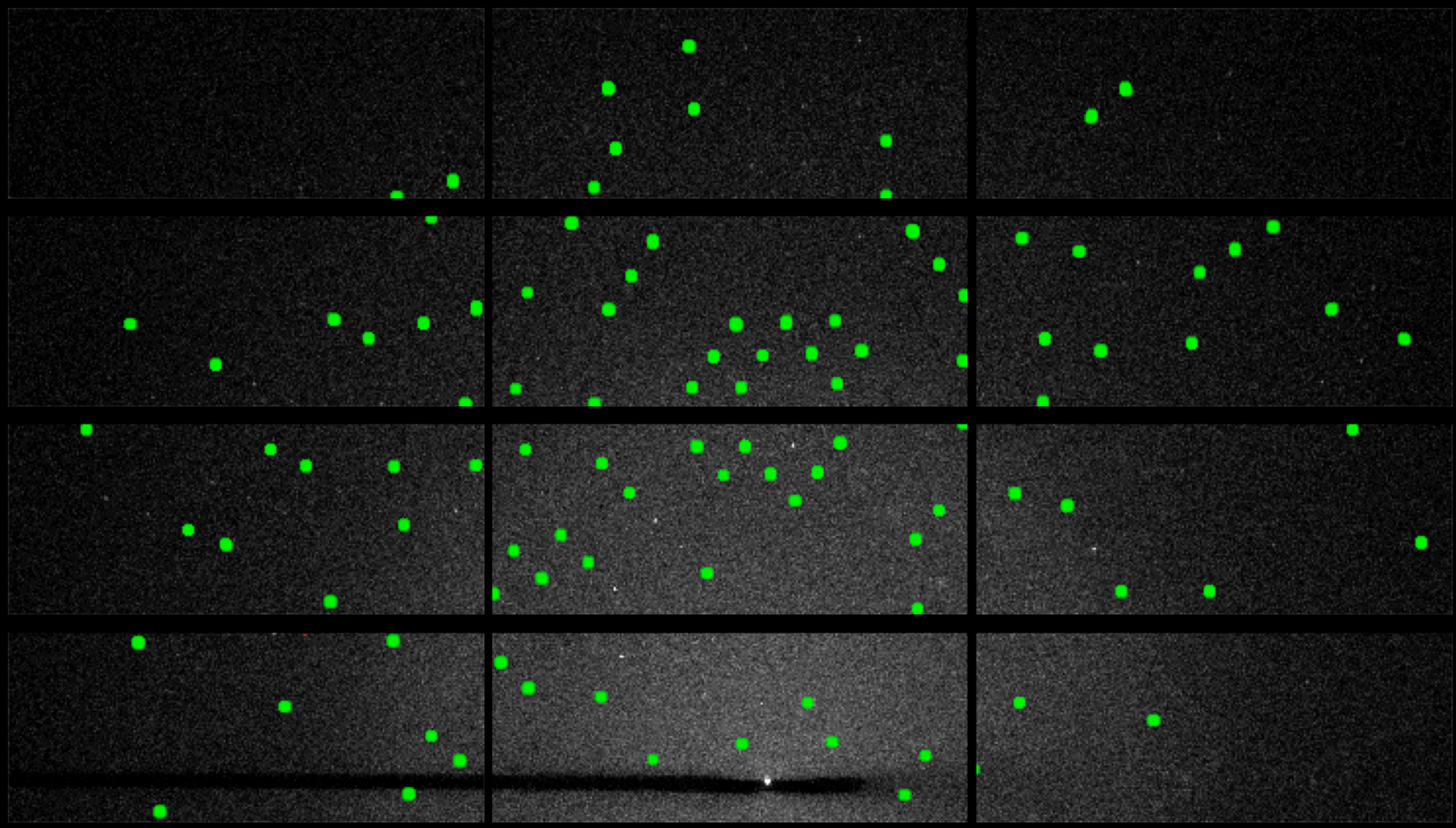
all | indexed | unindexed | integrated

all | inliers | outliers



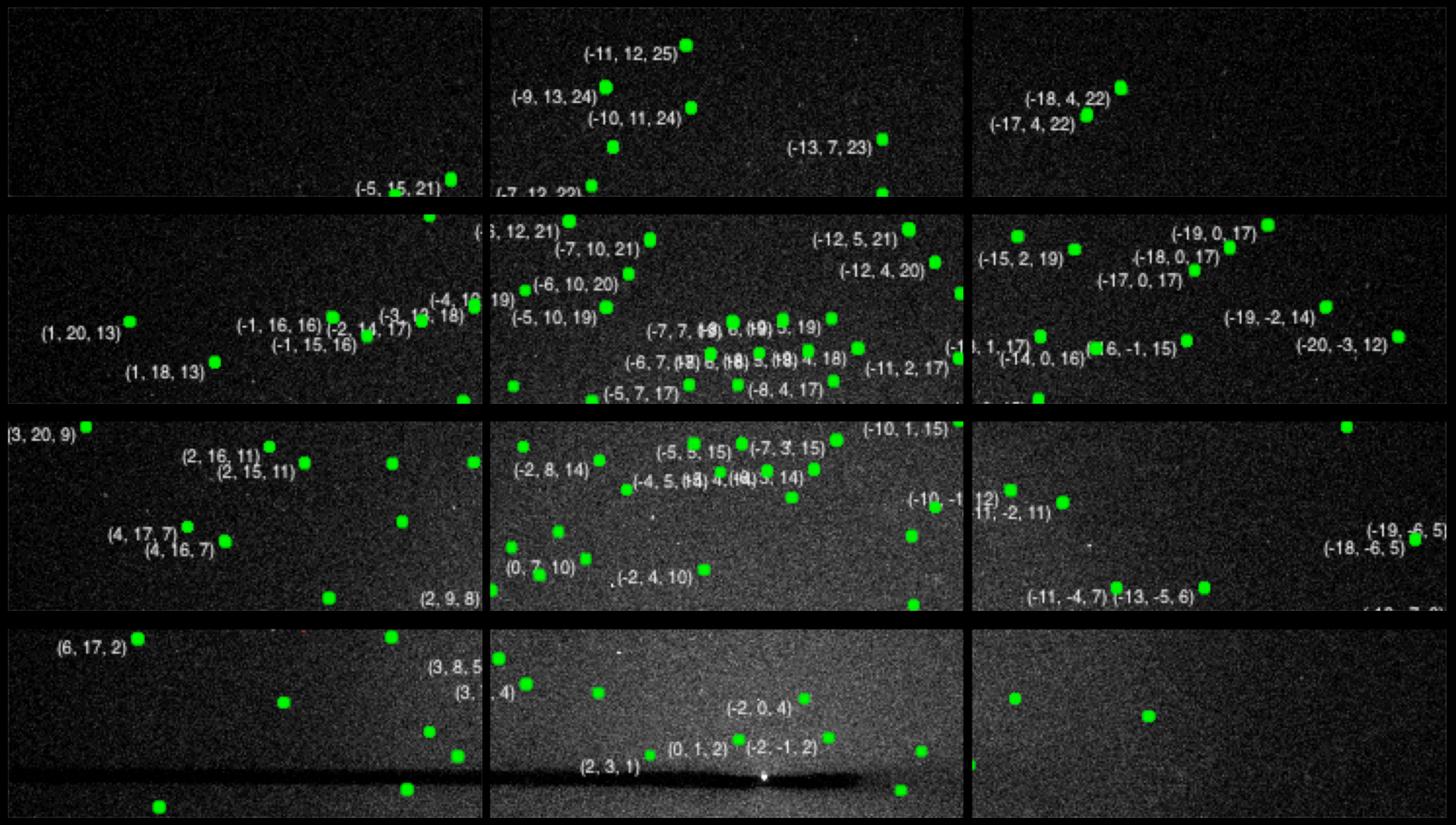
Model view matrix: [0.3025, 0.2791, 0.9114, 0.0000, -0.4578, 0.8812, -0.1179, 0.0000, -0.8360, -0.3815, 0.3943, 0.0000, 0.0000, 0.0000, -1667914.8750, 1.0000]

INDEXING



From found spots

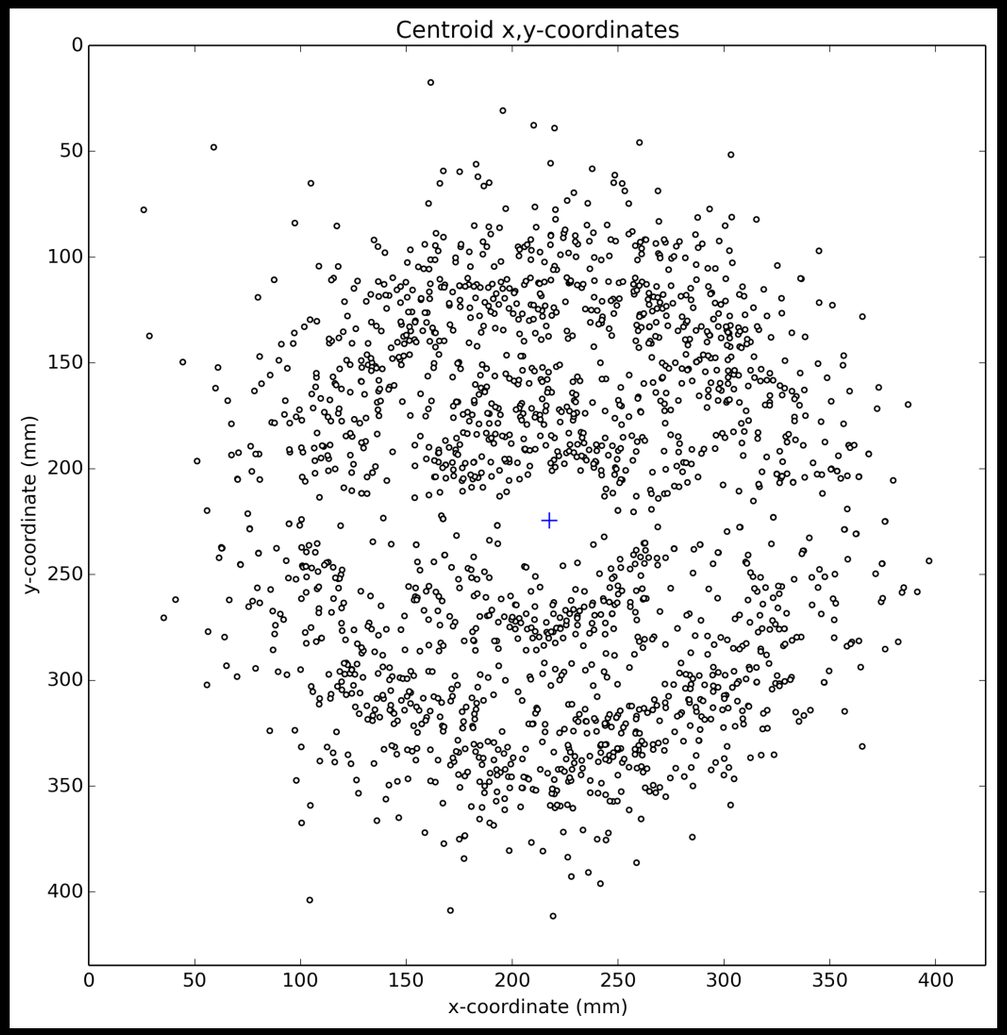
INDEXING

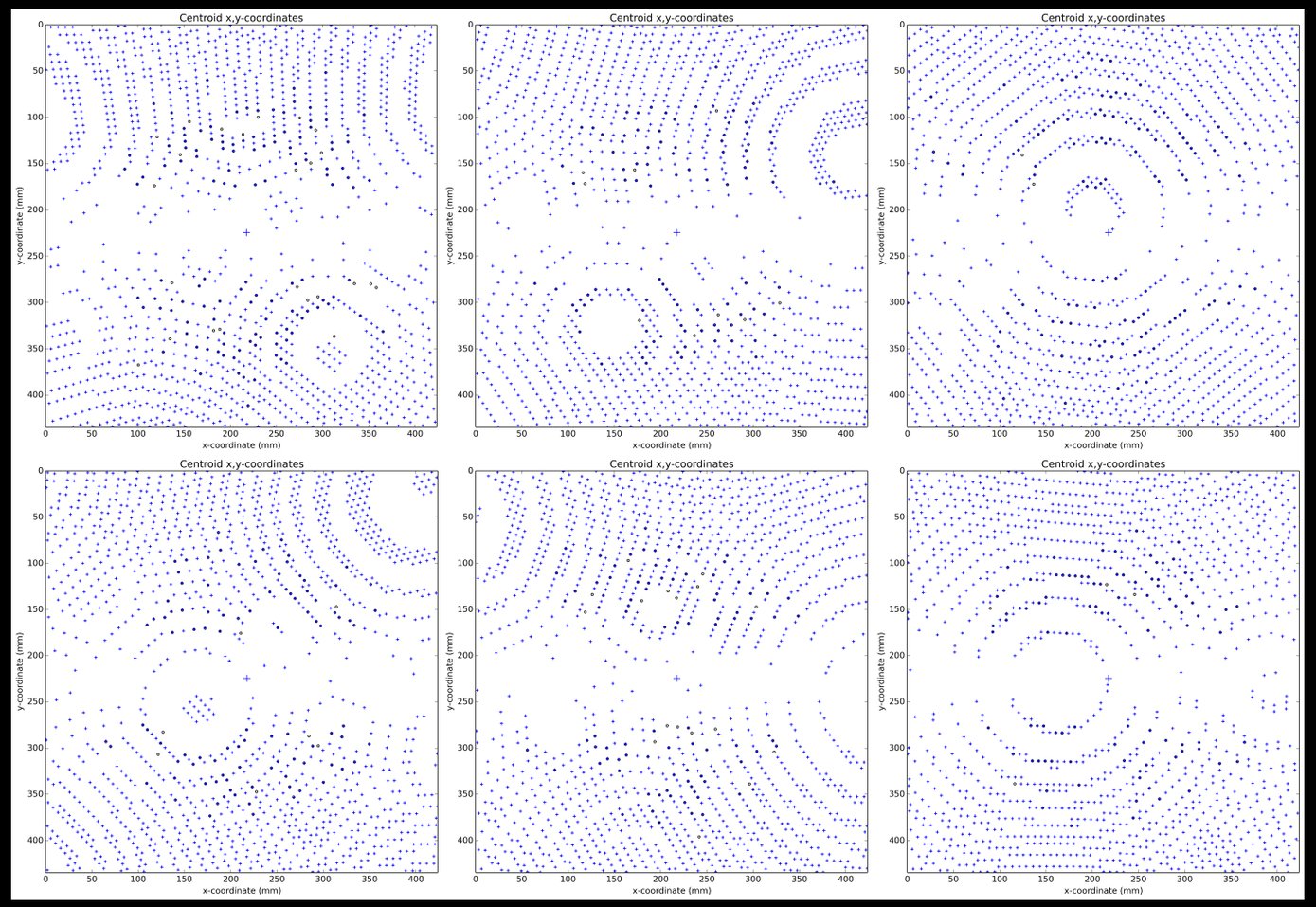


Find Miller indices

MULTIPLE LATTICES

Also works well for multiple lattices





1° wedge of data
1858 spots

6 Lattices identified

BASIC LATTICE DETERMINATION

```
$ dials.refine_bravais_lattice
```

DUI





Reindex

Chiral space groups corresponding to each Bravais lattice:
aP: P1
mP: P2 P21
mC: C2
oC: C2221 C222

	max δ	rmsd	min cc	max cc	latt	a	b	c	α	β	γ	Ok
1	0.0	0.07	No	No	aP	40.6	40.6	69.3	92.02	91.97	98.07	Y
2	0.031	0.072	0.97	0.97	mC	53.2	61.3	69.3	90	93.05	90	Y
3	3.0	2.3	0.62	0.62	mP	40.2	68.6	40.2	90	98.15	90	N
4	3.0	2.3	0.62	0.62	mC	60.8	52.7	68.6	90	89.96	90	N
5	3.0	2.3	0.62	0.97	oC	52.6	60.7	68.5	90	90	90	N

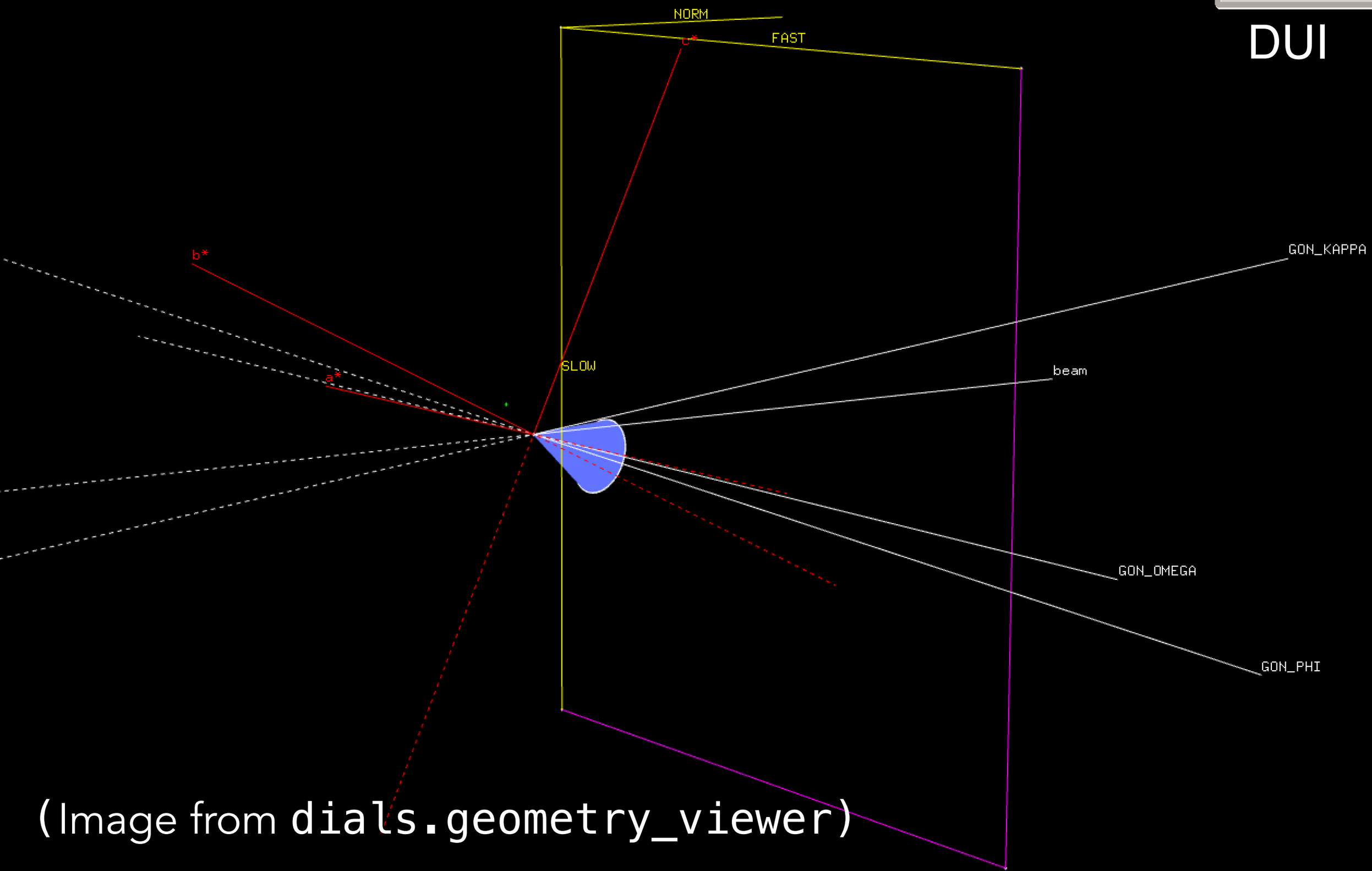
Select a bravais lattice to enforce:
(best guess solution = row 2)

OK

REFINEMENT dials.refine

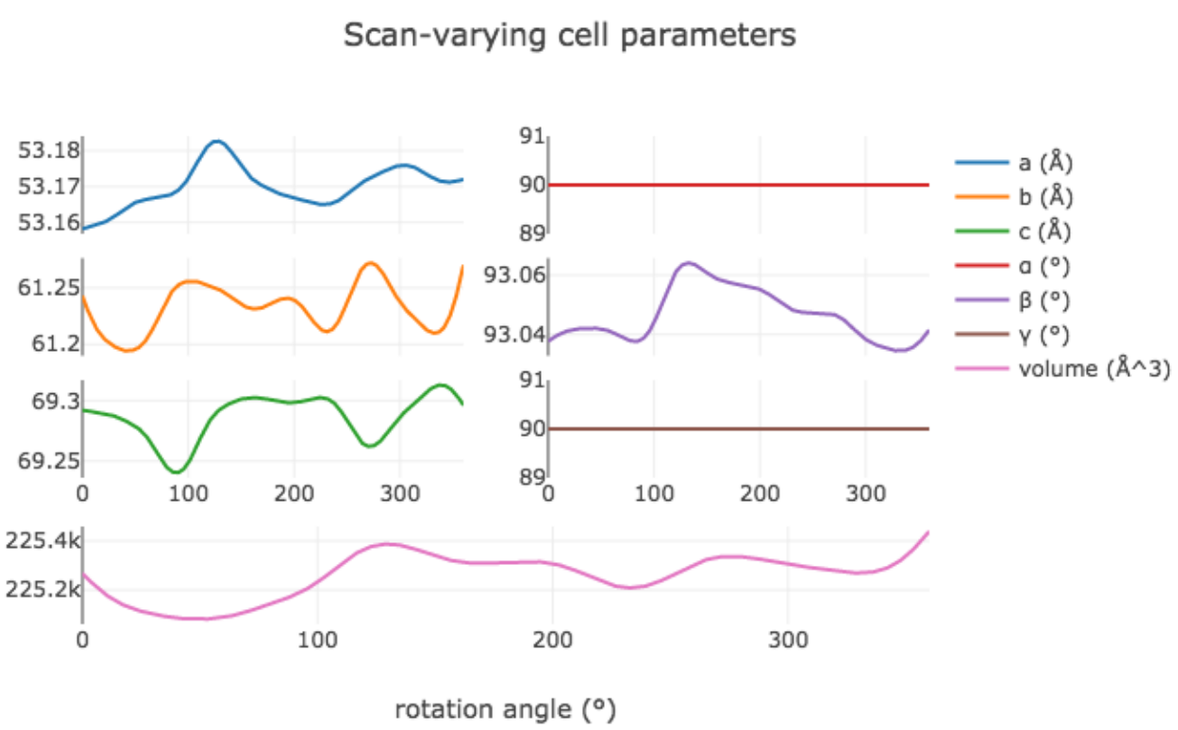


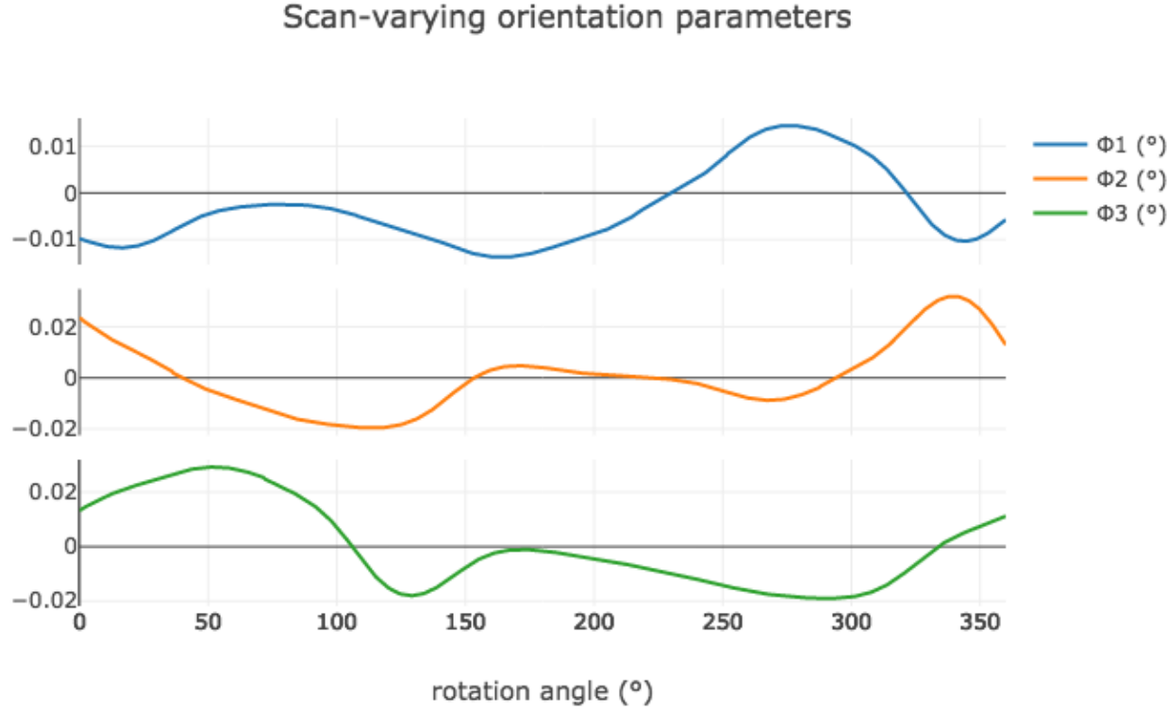
DUI



(Image from dials.geometry_viewer)

SCAN VARYING REFINEMENT



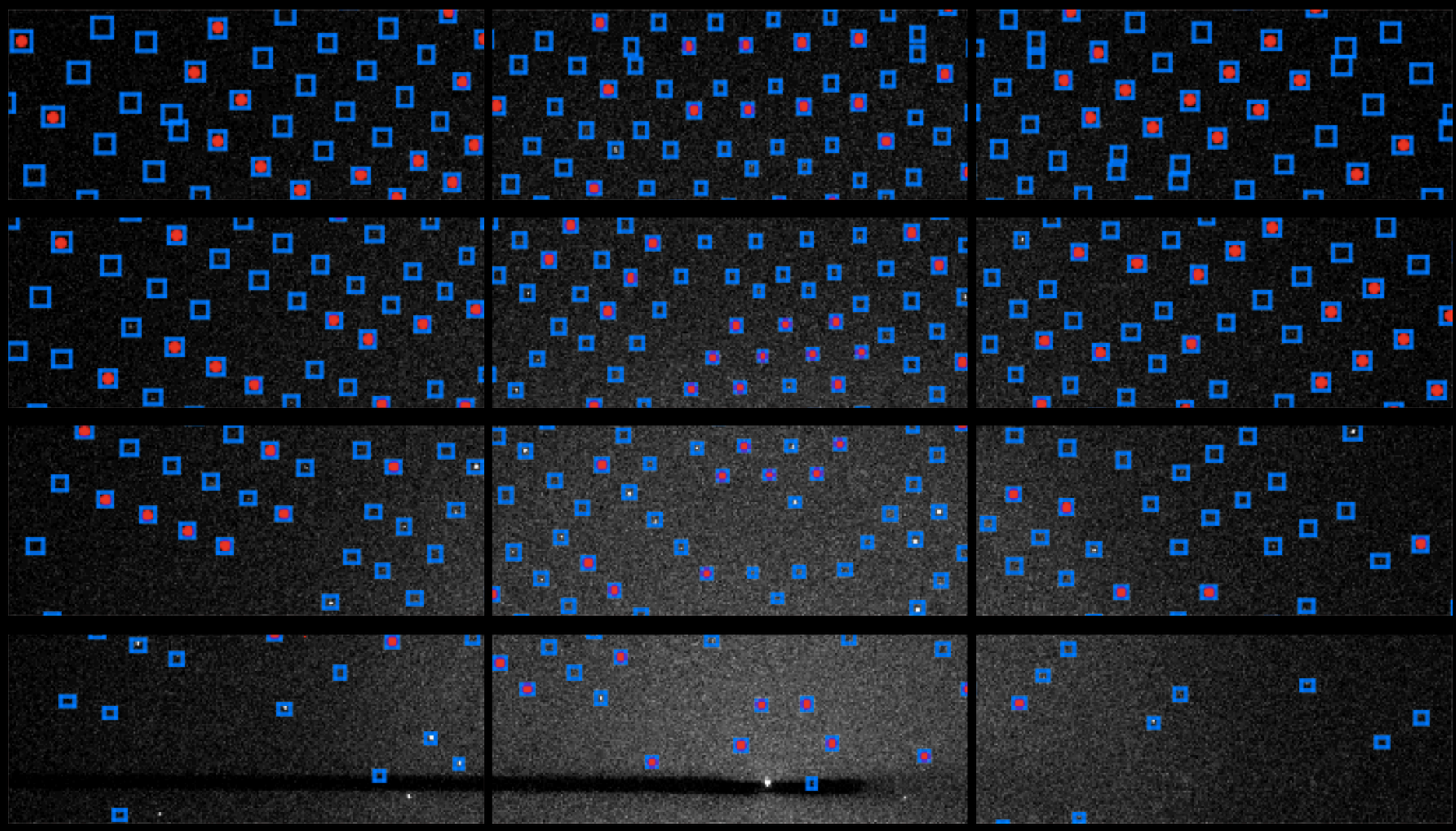


INTEGRATION

dials.integrate

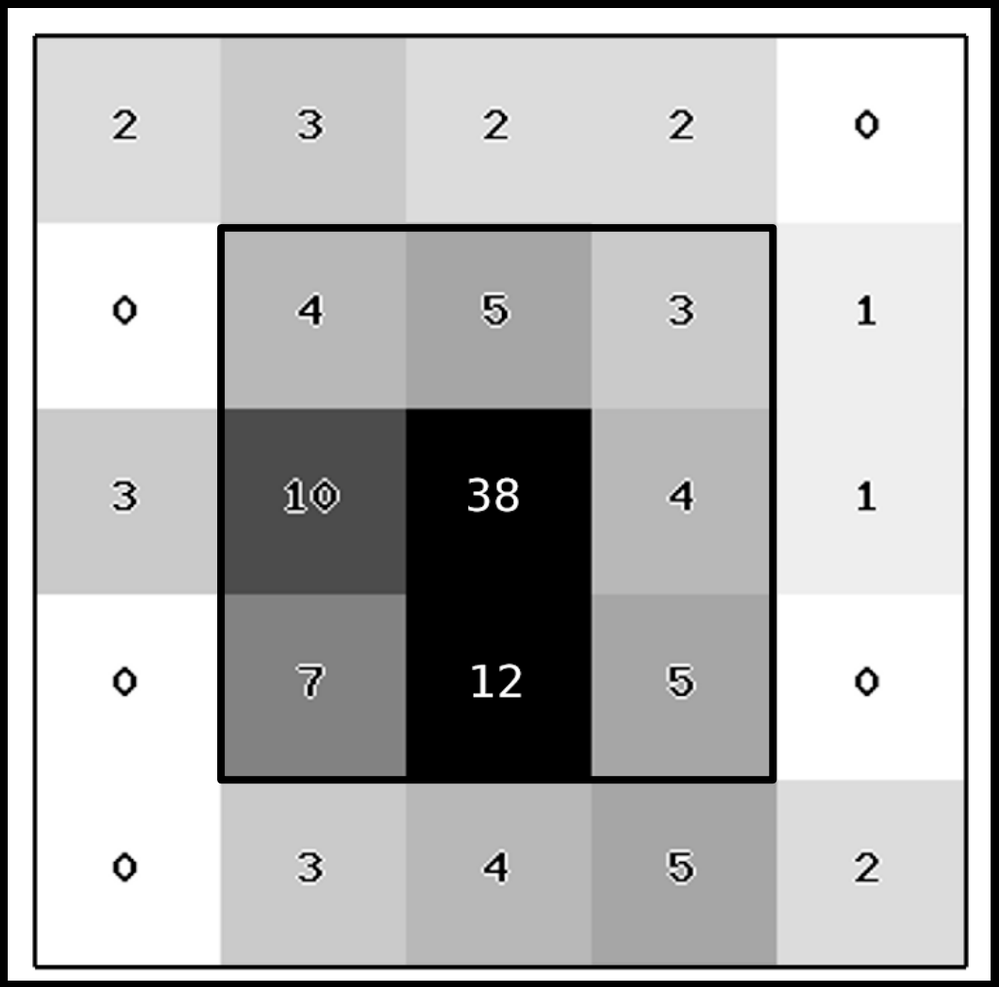
DUI





Calculate the intensity of every possible reflection

INTENSITY DETERMINATION

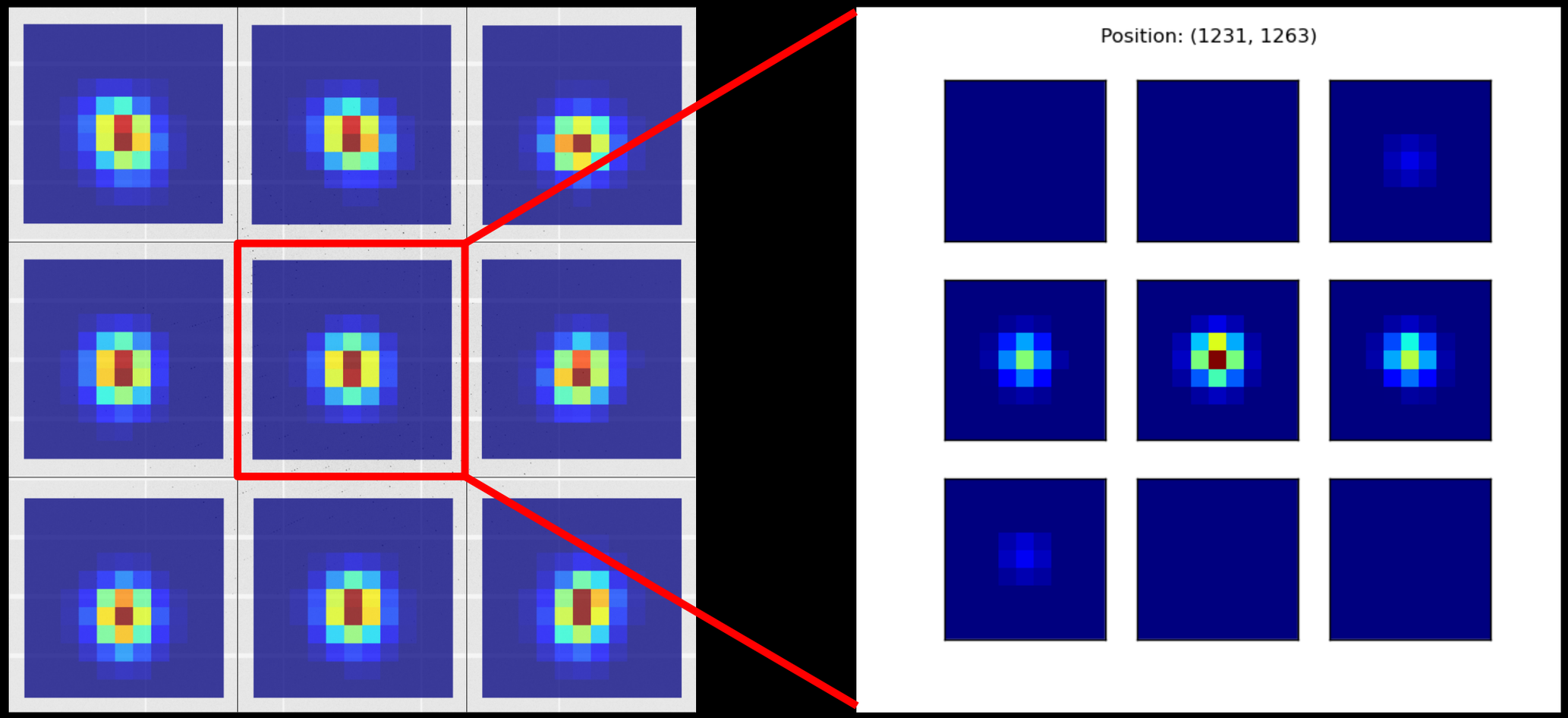


Need to separate signal and background under reflection peak

Don’t know background in signal region so estimate from the surrounding pixels

- Summation integration: estimate the reflection intensity by summing the counts contributing to the reflection and subtracting the background
- Profile fitting: fit a known profile shape to the reflection to estimate the intensity

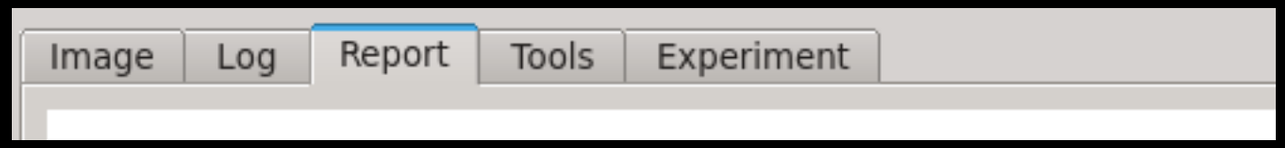
BUILDING REFERENCE PROFILES



- **Building reference profiles - detector space**
- **DIALS** constructs a number of reference profiles using a rectangular grid on the detector

REPORTING dials.report

DUI



Experiments

Crystal: Space group: P 4 (No. 75)

Unit cell: $(57.785 \quad 57.785 \quad 150.006 \quad 90.000 \quad 90.000 \quad 90.000)$

U matrix:
$$\begin{pmatrix} 0.3455 & -0.2590 & -0.9020 \\ 0.8913 & 0.3911 & 0.2292 \\ 0.2935 & -0.8831 & 0.3660 \end{pmatrix}$$

B matrix:
$$\begin{pmatrix} 0.0173 & 0.0000 & 0.0000 \\ -0.0000 & 0.0173 & 0.0000 \\ -0.0000 & 0.0000 & 0.0067 \end{pmatrix}$$

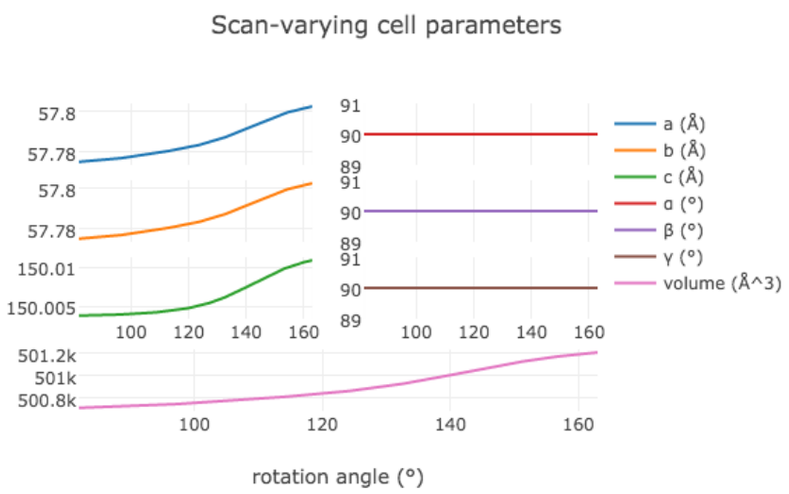
A = UB:
$$\begin{pmatrix} 0.0060 & -0.0045 & -0.0060 \\ 0.0154 & 0.0068 & 0.0015 \\ 0.0051 & -0.0153 & 0.0024 \end{pmatrix}$$

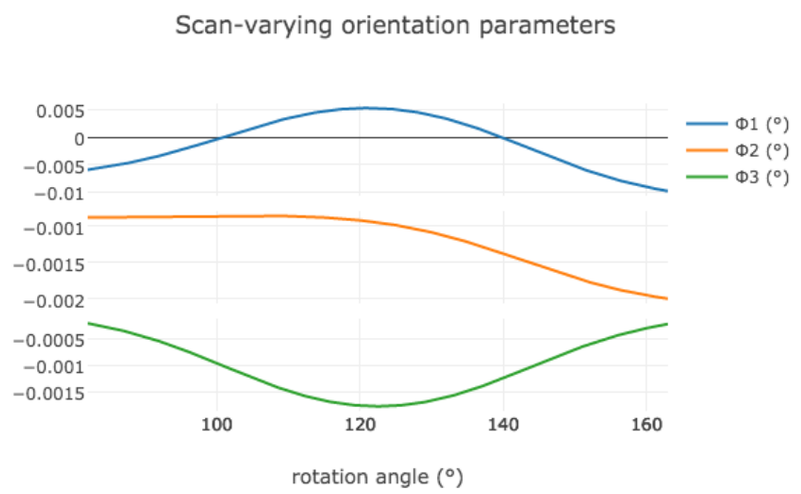
A sampled at 541 scan points

Average unit cell: $(57.786 \quad 57.786 \quad 150.006 \quad 90.000 \quad 90.000 \quad 90.000)$

Experimental geometry

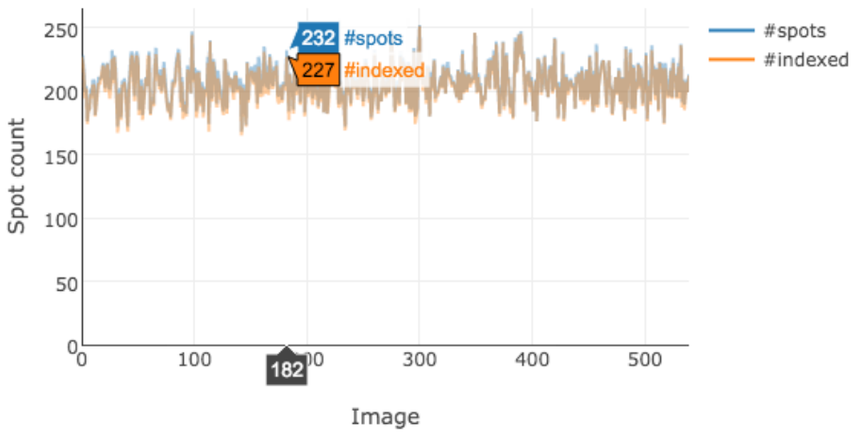
Analysis of scan-varying crystal model



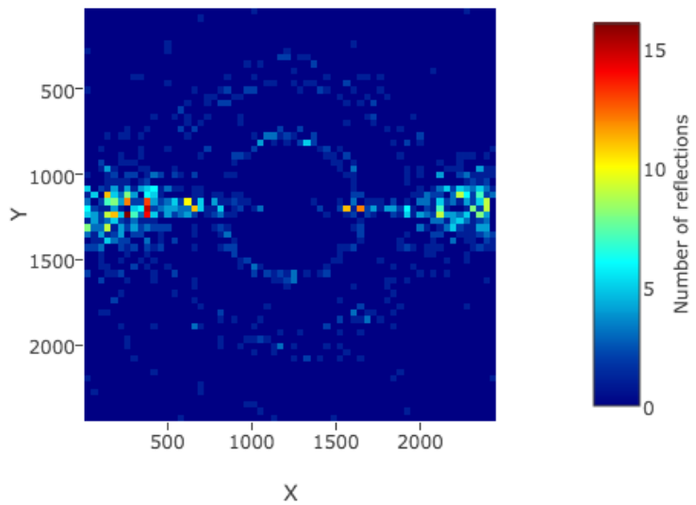


Analysis of strong reflections

Spot count per image

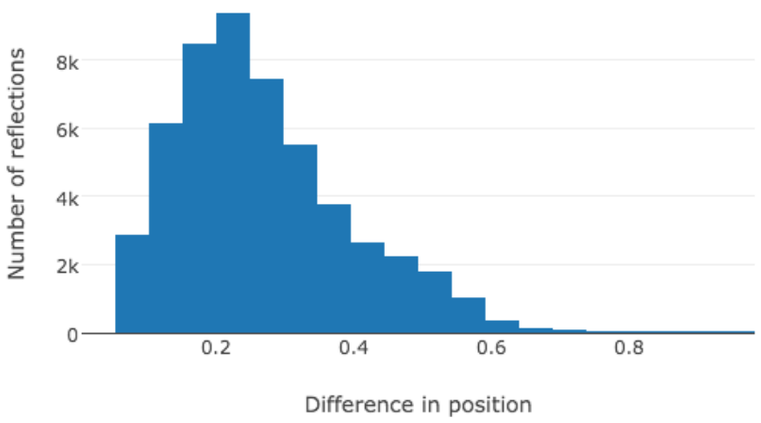


Number of unindexed reflections binned in X/Y

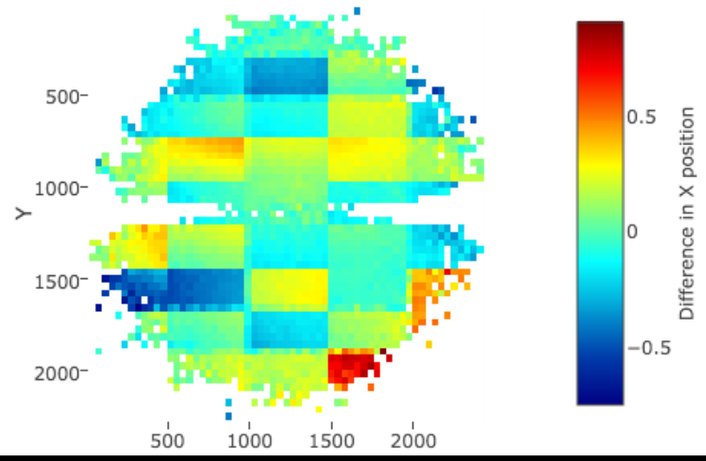


Analysis of reflection centroids

Difference between observed and calculated centroids

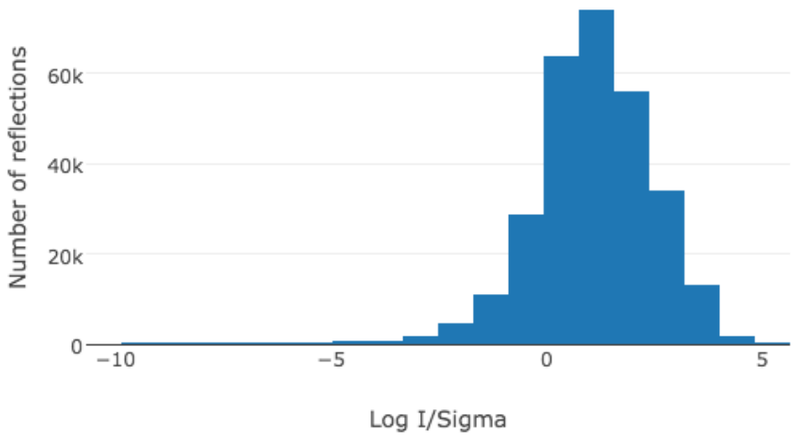


Difference between observed and calculated centroids in X

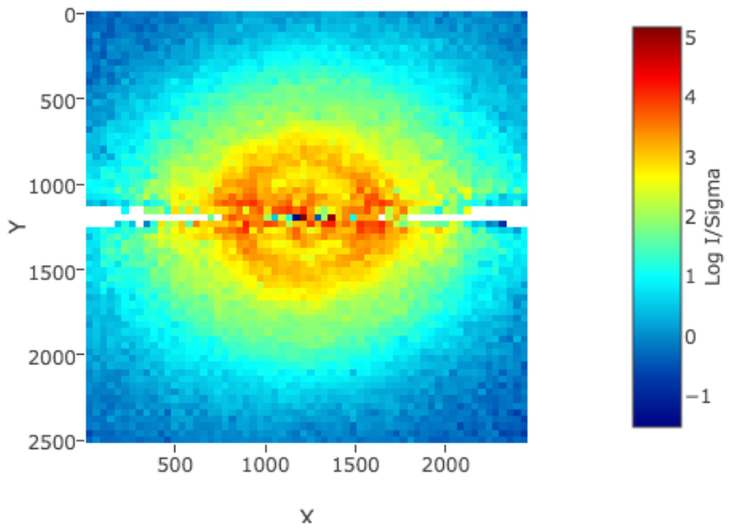


Analysis of reflection intensities

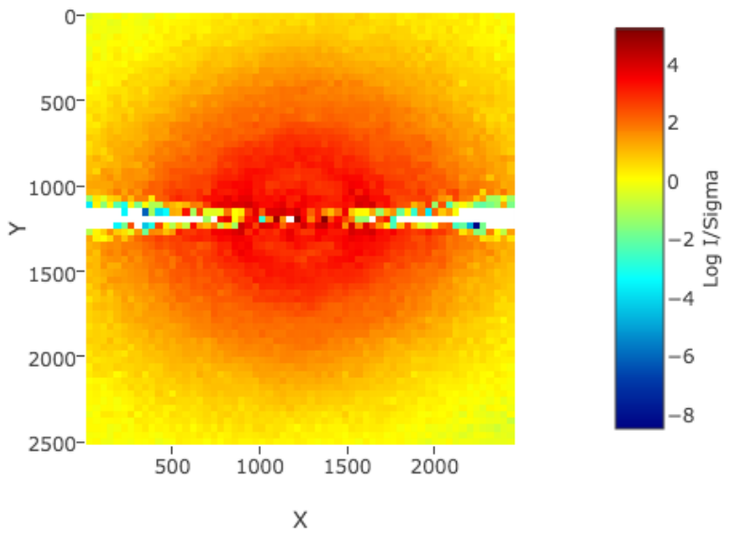
Log I/Sigma histogram



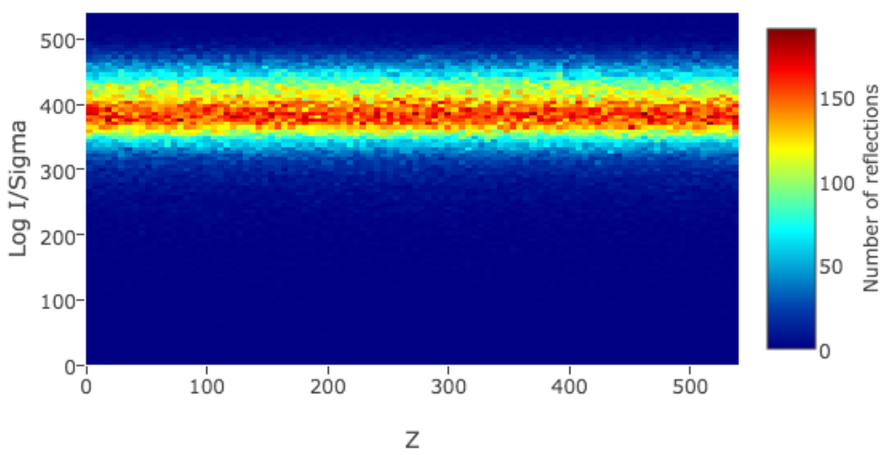
Distribution of I(sum)/Sigma vs X/Y



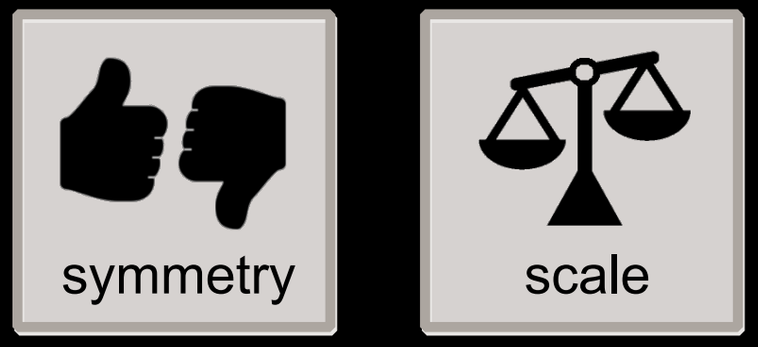
Distribution of I(prf)/Sigma vs X/Y



Distribution of I/Sigma vs Z



SCALING



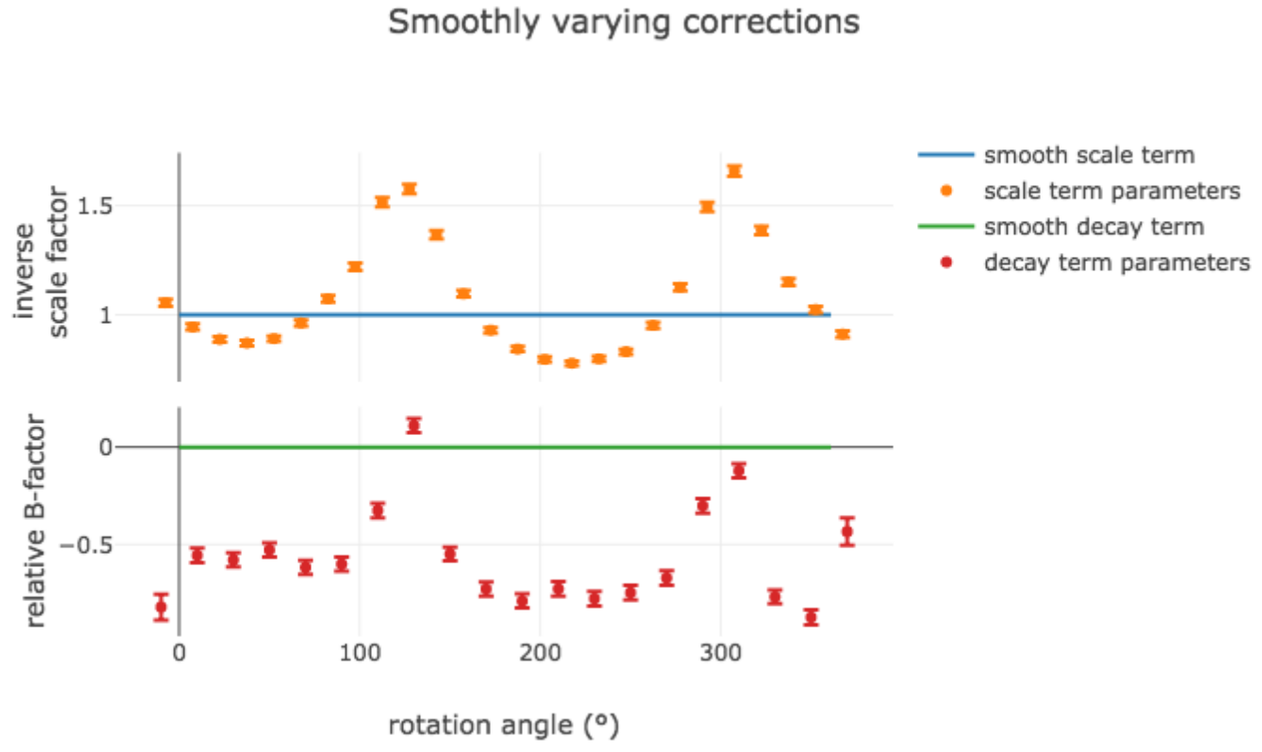
- With dials: dials.symmetry and dials.scale
- With CCP4: pointless hklin dials_integrated.mtz hklout sorted.mtz; aimless hklin sorted.mtz hklout scaled.mtz
- CCP4i; i2 helpful for this (particularly for log files)

SYMMETRY

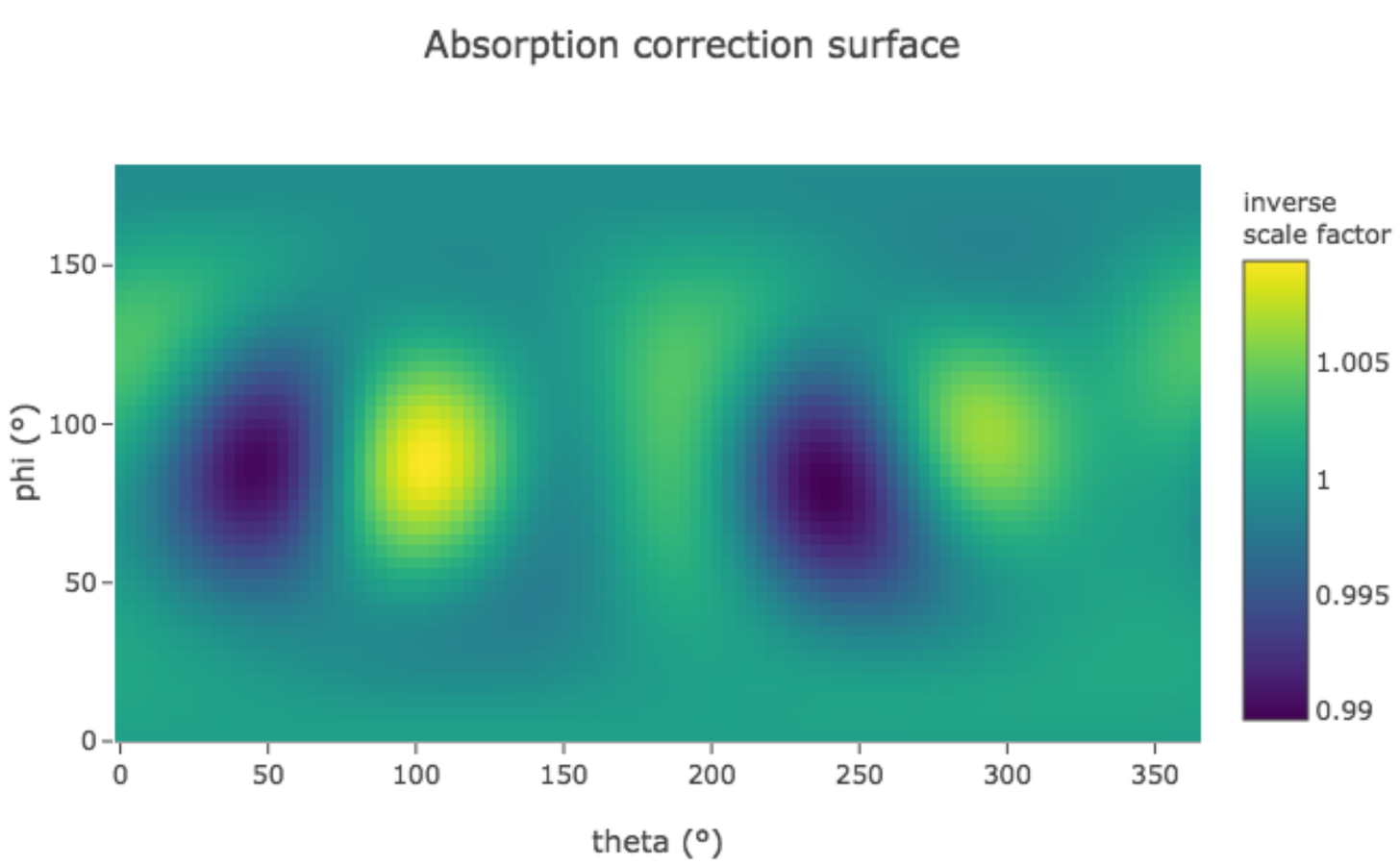
```
Scoring individual symmetry elements
---------------------------------------------
likelihood  Z-CC  CC    N            Operator
---------------------------------------------
0.916       9.81  0.98  297147  ***   1 |(0, 0, 0)
0.913       9.72  0.97  294871  ***   2 |(-1, 1, 0)
---------------------------------------------
Scoring all possible sub-groups
-------------------------------------------------------------------------------------------
Patterson group       Likelihood  NetZcc  Zcc+   Zcc-   CC     CC-    delta  Reindex
operator
-------------------------------------------------------------------------------------------
C 1 2/m 1         ***  0.913        9.76    9.76   0.00   0.98   0.00  0.0    -a,b,-c
P -1                   0.087        0.09    9.81   9.72   0.98   0.97  0.0    -x-y,-x+y,-z
-------------------------------------------------------------------------------------------
```

SCALING

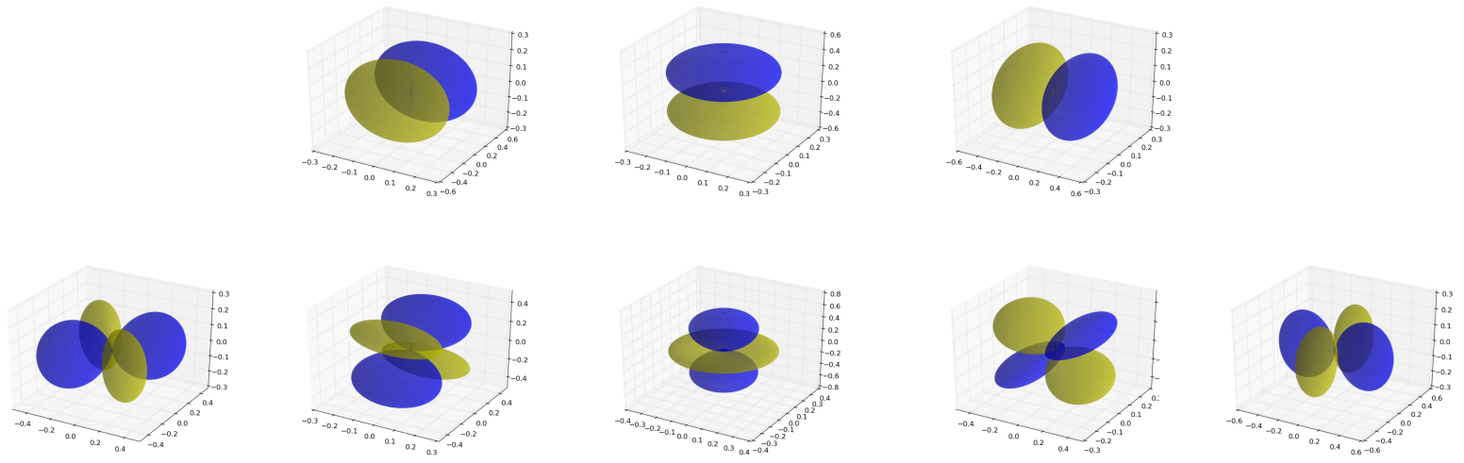
kB (smoothed)



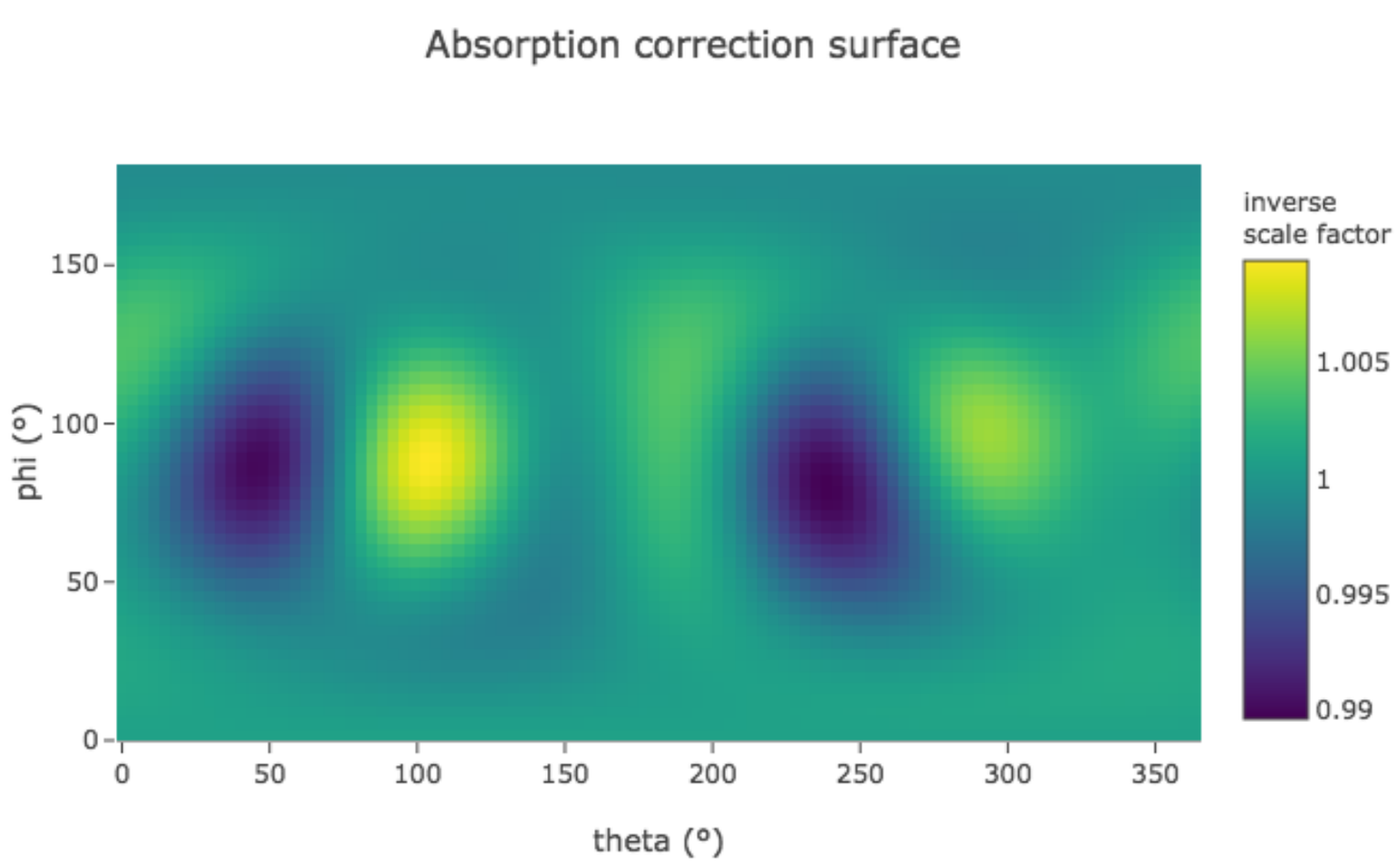
Absorption surface (as spherical harmonics)



SCALING



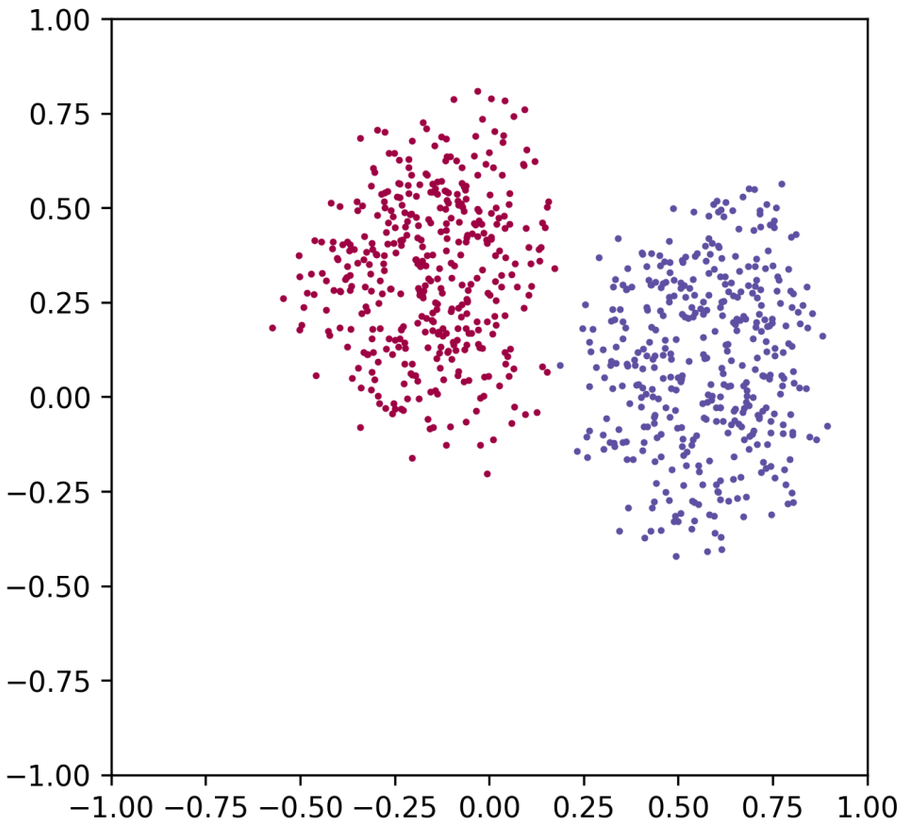
Absorption surface (as spherical harmonics)

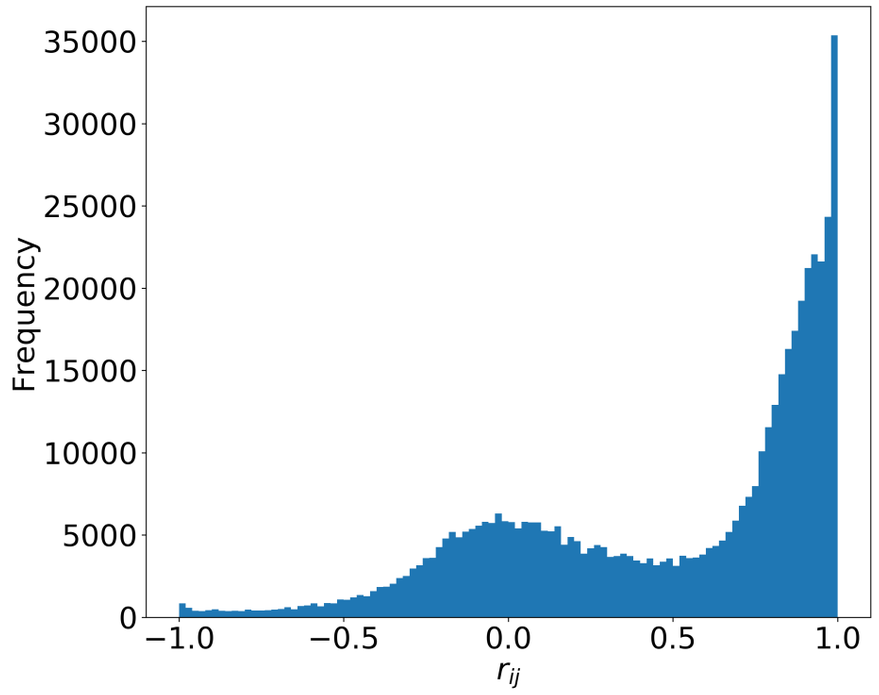


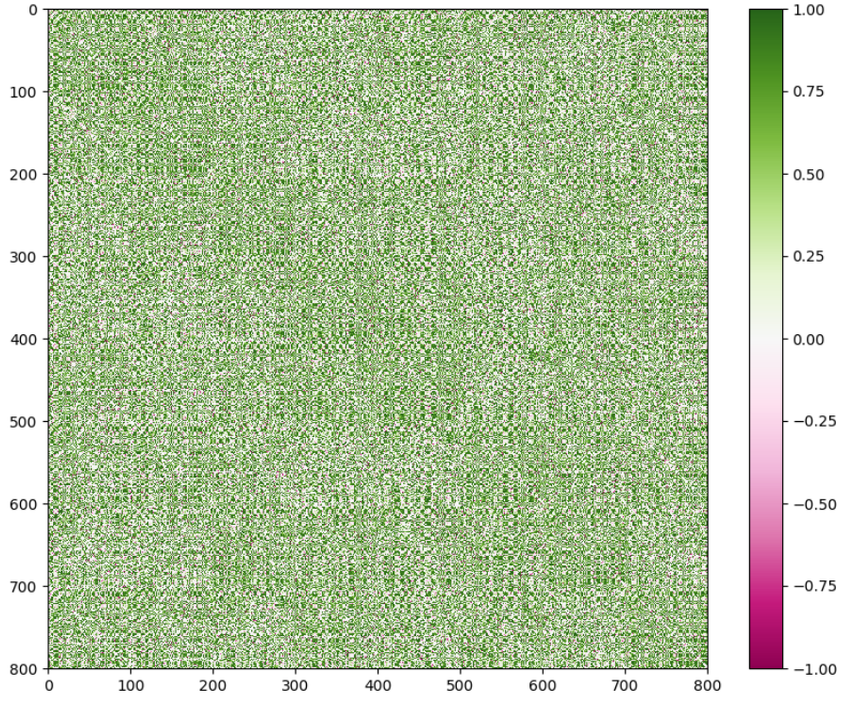
Summary of merging statistics

	Overall	Low resolution	High resolution
Resolution (Å)	69.19 - 1.22	69.31 - 3.31	1.24 - 1.22
Observations	308412	22287	296
Unique reflections	48399	3348	185
Multiplicity	6.4	6.7	1.6
Completeness	73.11%	98.73%	5.64%
Mean I/sigma(I)	12.0	50.8	0.3
Rmerge	0.066	0.038	0.917
Rmeas	0.072	0.041	1.205
Rpim	0.028	0.016	0.774
CC1/2	0.999	0.998	0.604

SYMMETRY & ISOMORPHISM - COSYM







DIALS.GITHUB.IO

XIA2.GITHUB.IO

research papers

Acta Cryst D STRUCTURAL BIOLOGY

ISSN 2059-7983

DIALS: implementation and evaluation of a new integration package

Graeme Winter,[a]* David G. Waterman,[c,d] James M. Parkhurst,[a,e] Aaron S. Brewster,[b] Richard J. Gildea,[a] Markus Gerstel,[a] Luis Fuentes-Montero,[a] Melanie Vollmar,[a] Tara Michels-Clark,[b] Iris D. Young,[b] Nicholas K. Sauter[b] and Gwyndaf Evans[a]*

[a]Diamond Light Source Ltd, Harwell Science and Innovation Campus, Didcot OX11 0DE, England, [b]Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA, [c]STFC Rutherford Appleton Laboratory, Didcot OX11 0FA, England, [d]CCP4, Research Complex at Harwell, Rutherford Appleton Laboratory, Didcot OX11 0FA, England, and [e]Laboratory of Molecular Biology, Francis Crick Avenue, Cambridge CB2 0QH, England. *Correspondence e-mail: graeme.winter@diamond.ac.uk, gwyndaf.evans@diamond.ac.uk

Received 14 June 2017
Accepted 30 November 2017

Keywords: X-ray diffraction; data processing; methods development; *DIALS*.

Supporting information: this article has supporting information at journals.iucr.org/d

The *DIALS* project is a collaboration between Diamond Light Source, Lawrence Berkeley National Laboratory and CCP4 to develop a new software suite for the analysis of crystallographic X-ray diffraction data, initially encompassing spot finding, indexing, refinement and integration. The design,