DIALS

Diffraction Integration for Advanced Light Sources

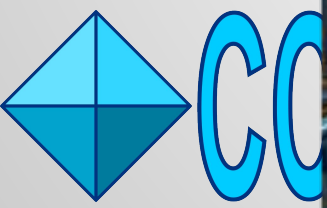
David Waterman
Instituto de Física de São Carlos/USP, Brazil, 2016

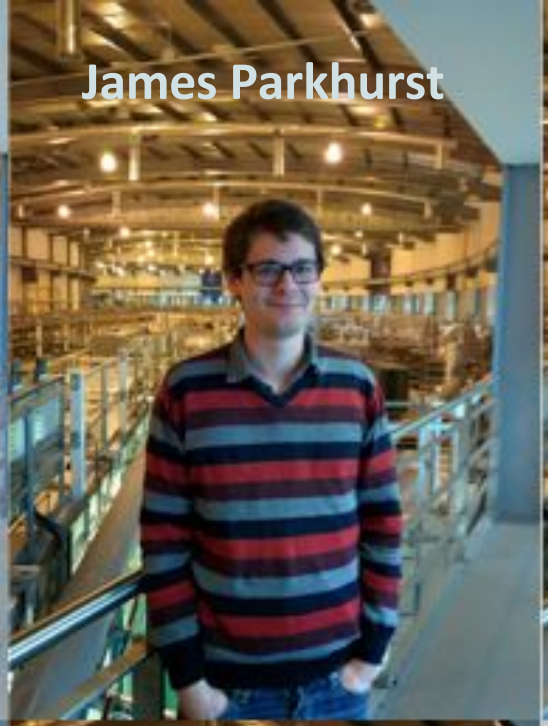
CCP4

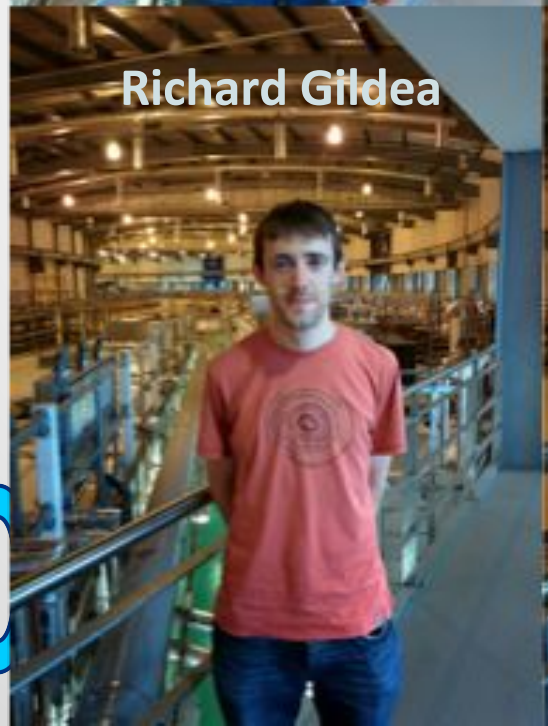
DIALS
Diffraction Integration for Advanced Light Sources

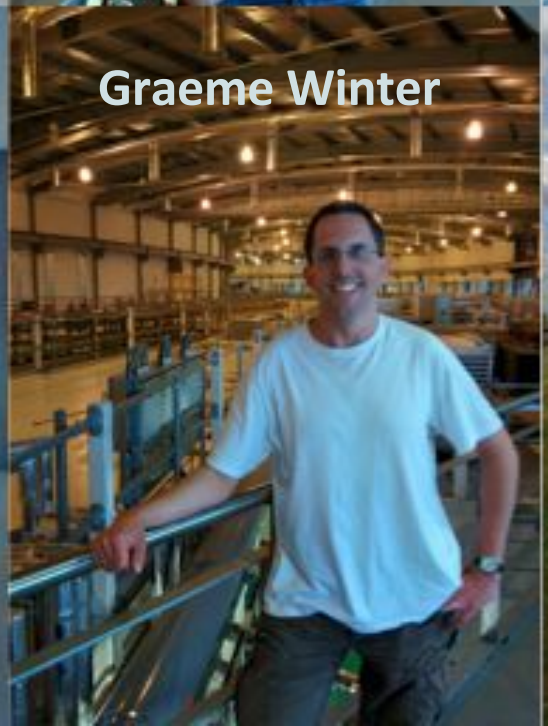
DIALS East

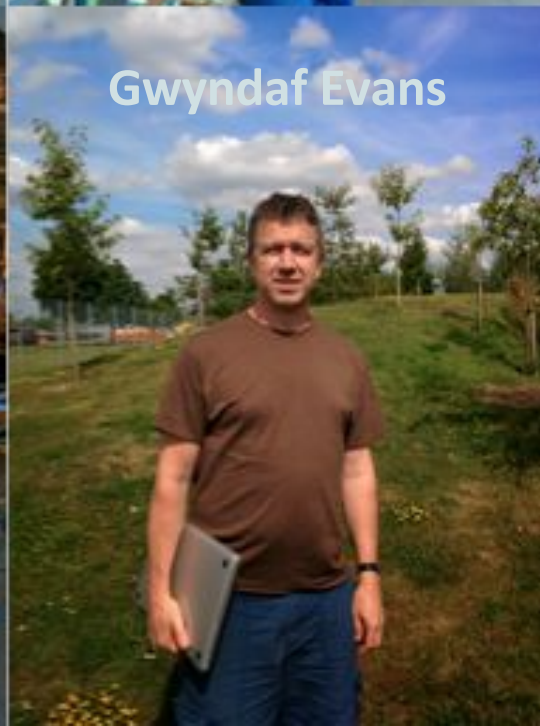
Diamond / CCP4







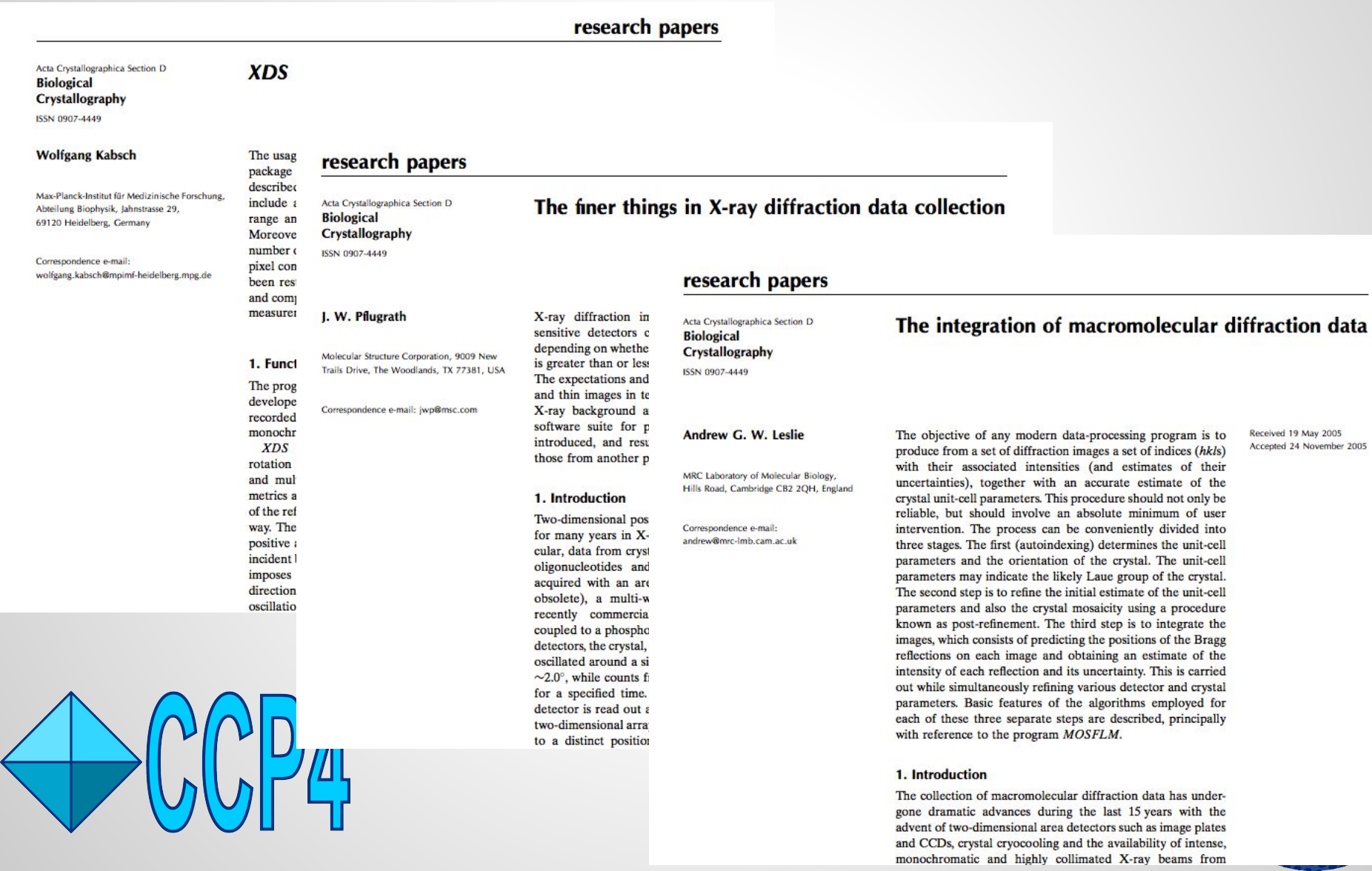




DIALS West – LBL / SLAC

Iris Young Tara Michels-Clark Nicholas Sauter Aaron Brewster Muhamed Amin

Acknowledgements



research papers

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

XDS

Wolfgang Kabsch

Max-Planck-Institut für Medizinische Forschung, Abteilung Biophysik, Jahnstrasse 29, 69120 Heidelberg, Germany

Correspondence e-mail: wolfgang.kabsch@mpimf-heidelberg.mpg.de

research papers

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The finer things in X-ray diffraction data collection

J. W. Pflugrath

Molecular Structure Corporation, 9009 New Trails Drive, The Woodlands, TX 77381, USA

Correspondence e-mail: jwp@msc.com

research papers

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The integration of macromolecular diffraction data

Andrew G. W. Leslie

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, England

Correspondence e-mail: andrew@mrc-lmb.cam.ac.uk

Received 19 May 2005
Accepted 24 November 2005

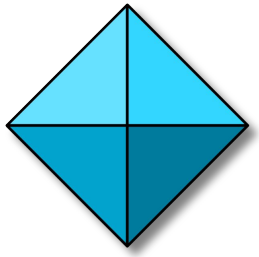
The objective of any modern data-processing program is to produce from a set of diffraction images a set of indices (*hkl*s) with their associated intensities (and estimates of their uncertainties), together with an accurate estimate of the crystal unit-cell parameters. This procedure should not only be reliable, but should involve an absolute minimum of user intervention. The process can be conveniently divided into three stages. The first (autoindexing) determines the unit-cell parameters and the orientation of the crystal. The unit-cell parameters may indicate the likely Laue group of the crystal. The second step is to refine the initial estimate of the unit-cell parameters and also the crystal mosaicity using a procedure known as post-refinement. The third step is to integrate the images, which consists of predicting the positions of the Bragg reflections on each image and obtaining an estimate of the intensity of each reflection and its uncertainty. This is carried out while simultaneously refining various detector and crystal parameters. Basic features of the algorithms employed for each of these three separate steps are described, principally with reference to the program *MOSFLM*.

1. Introduction

The collection of macromolecular diffraction data has undergone dramatic advances during the last 15 years with the advent of two-dimensional area detectors such as image plates and CCDs, crystal cryocooling and the availability of intense, monochromatic and highly collimated X-ray beams from

diamond

Research Complex
at Harwell

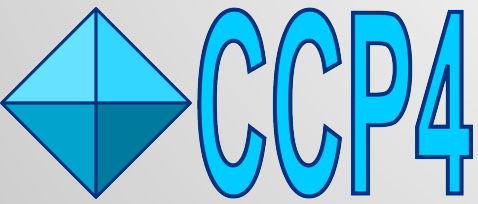


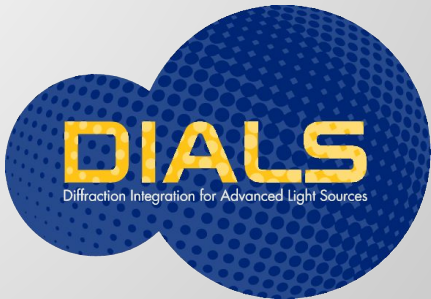
CCP4

Science & Technology
Facilities Council

BioStructX

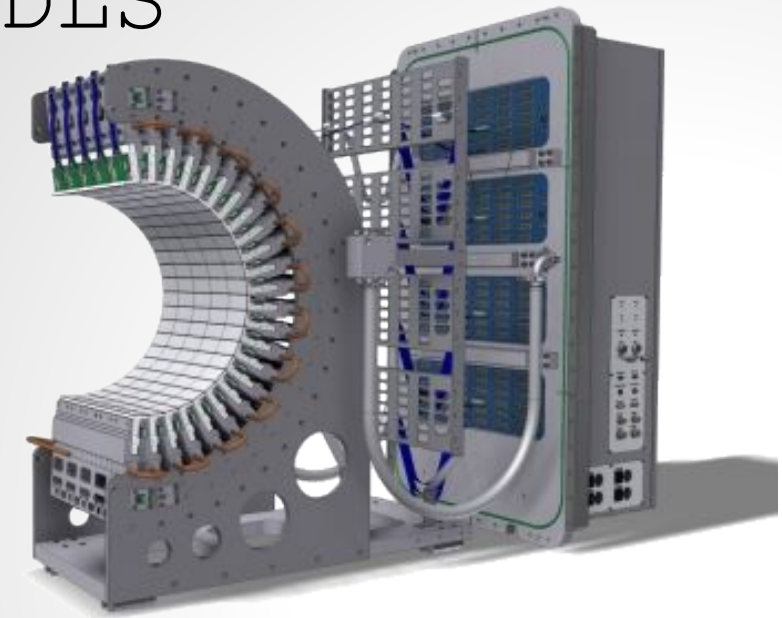
NEW SOFTWARE: DIALS



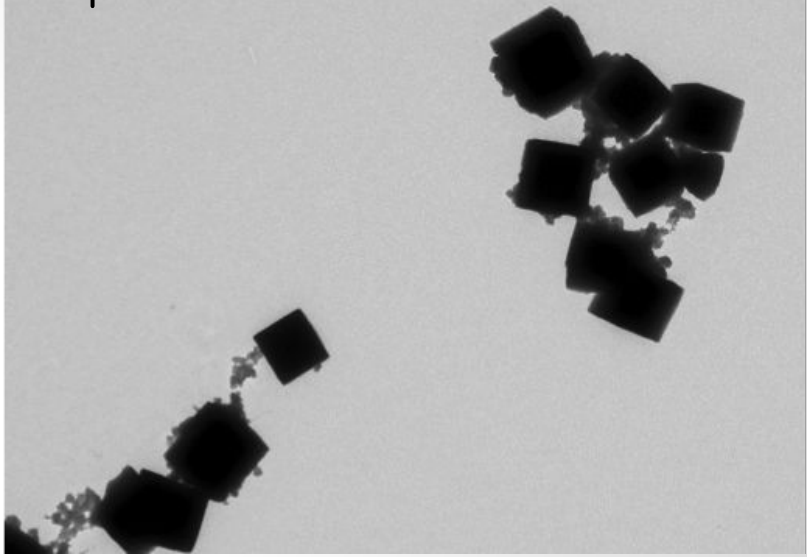


New challenges

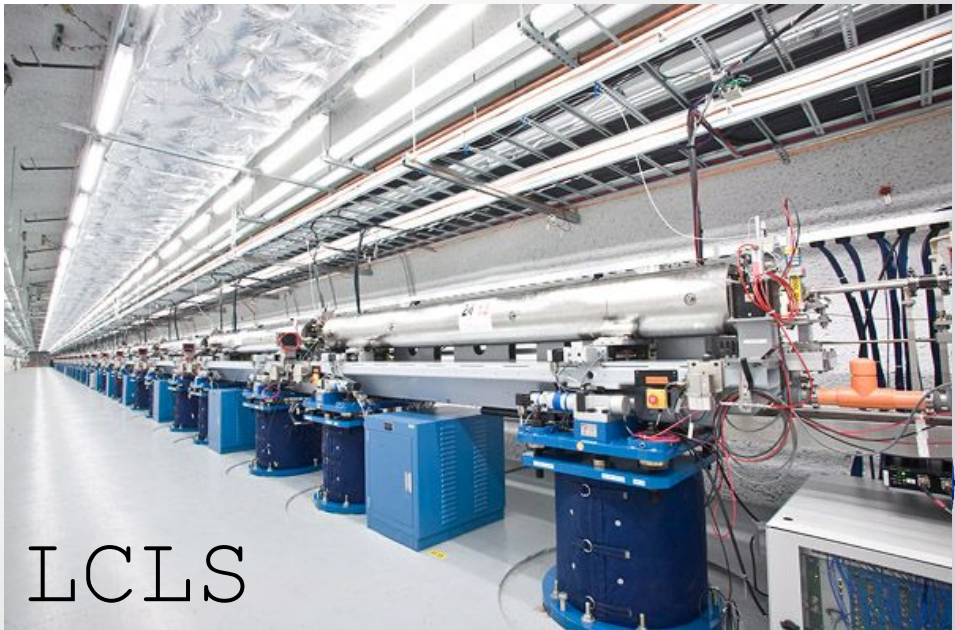
P12M-DLS



1µm CRYSTALS

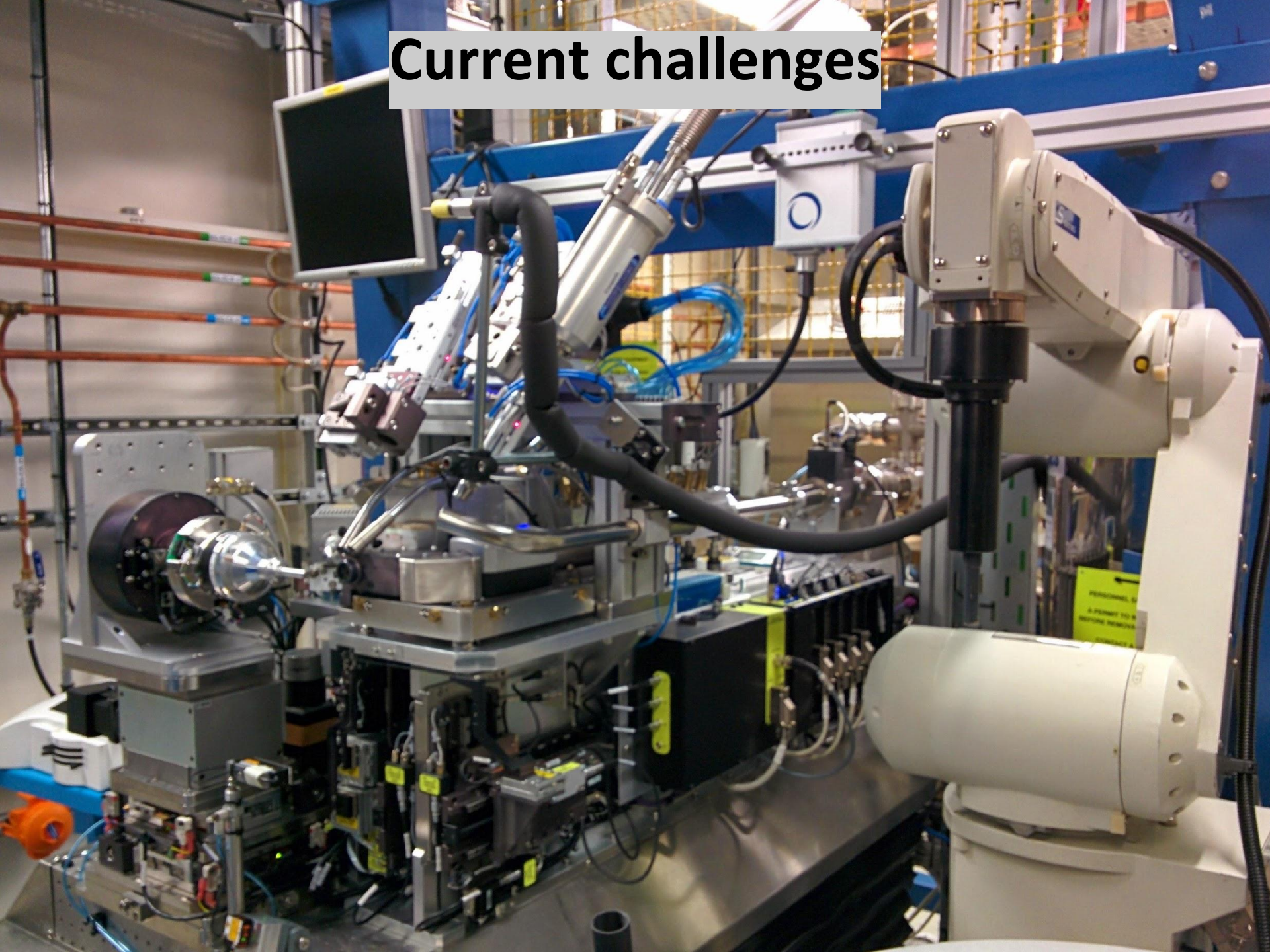


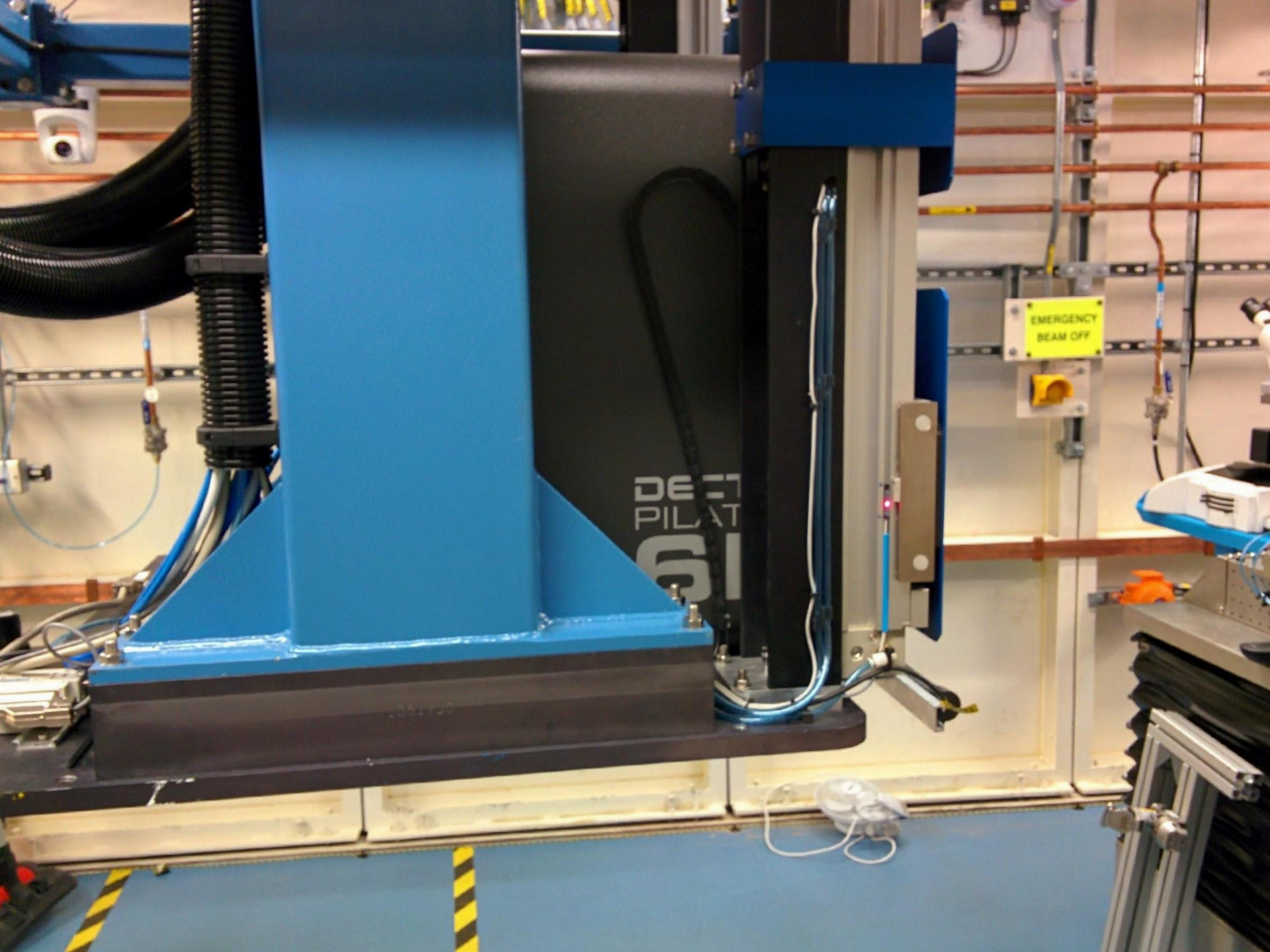
DIAMOND I24

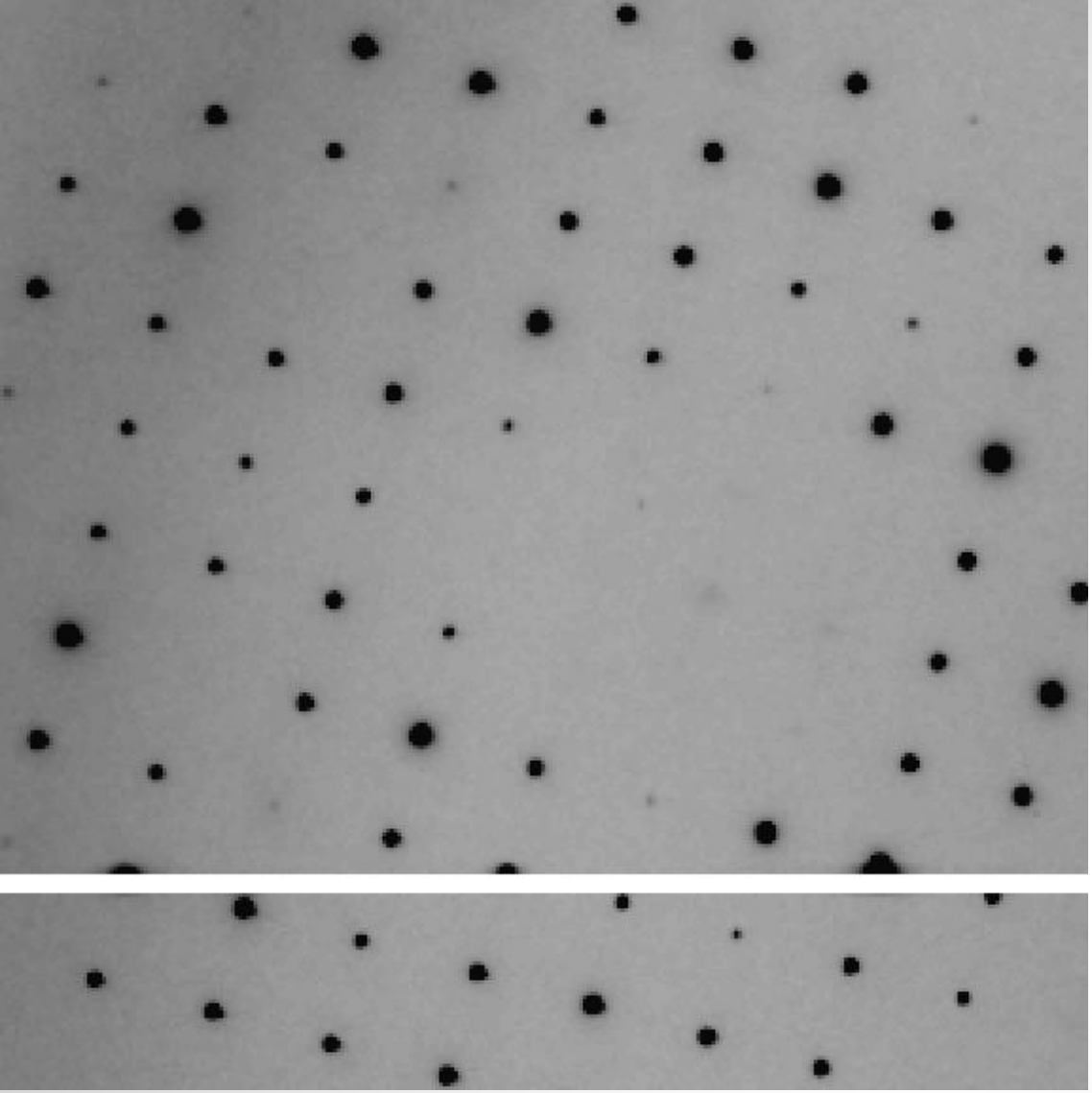


LCLS

Current challenges




EMERGENCY
BEAM OFF



Compute the intensity (and structure factor) of each Bragg spot in a set of diffraction images

$$|F_{hkl}| = \left(\frac{KI_{hkl}}{Lp}\right)^{\frac{1}{2}}$$

K = scale factor

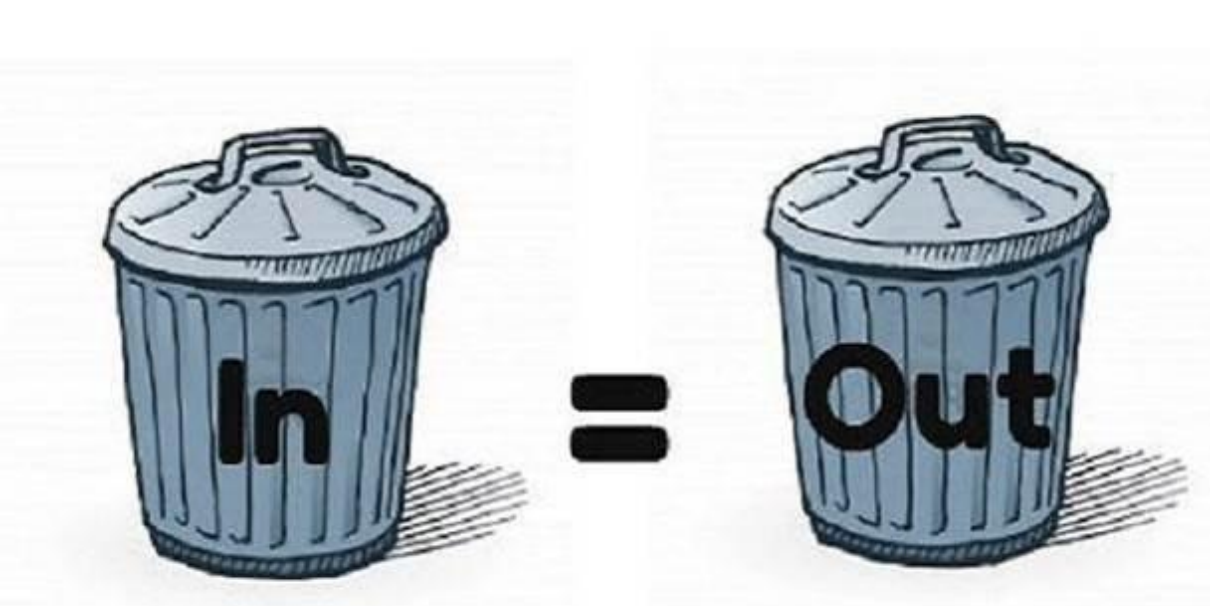
L = Lorentz factor

p = polarization factor

CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Warning: garbage in, garbage out



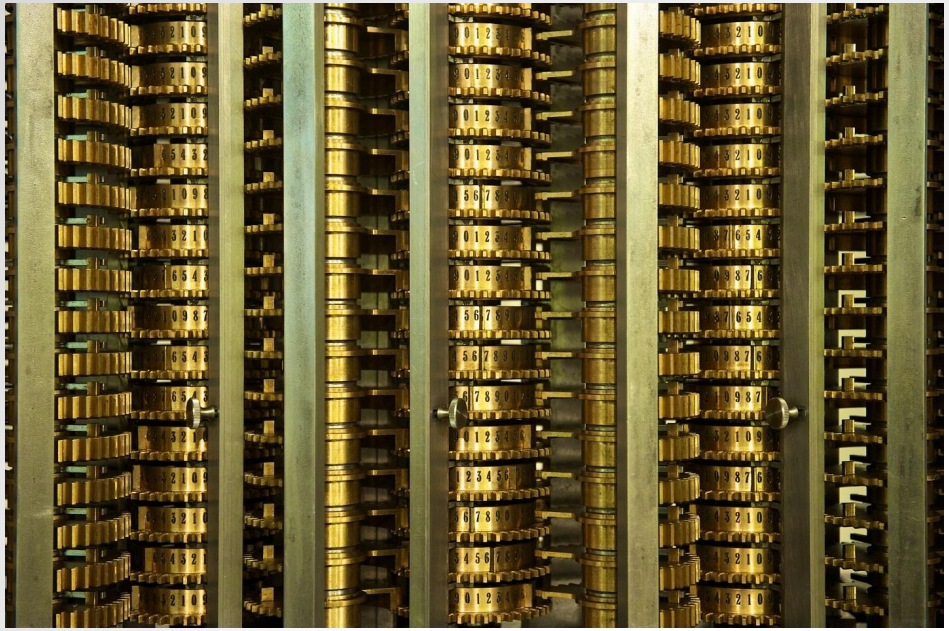
Data collection is the last experimental stage

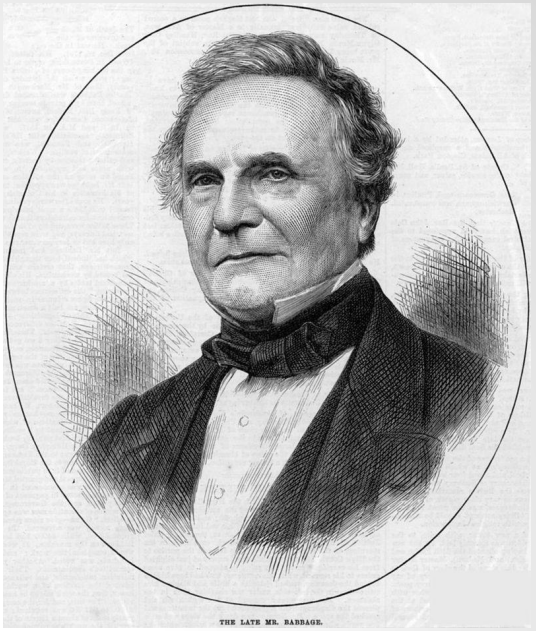
A fundamentally bad dataset cannot be rescued by any algorithm!

Don't spend months optimising crystals only to rush the data collection

CCP4

DIALS
Diffraction Integration for Advanced Light Sources





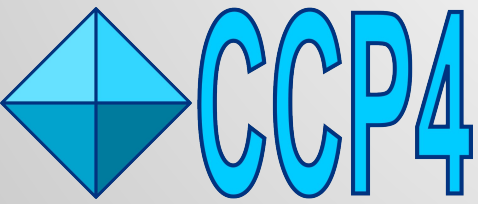
On two occasions I have been asked, "Pray, Mr. Babbage, if you put into the machine wrong figures, will the right answers come out?" ... I am not able rightly to apprehend the kind of confusion of ideas that could provoke such a question.

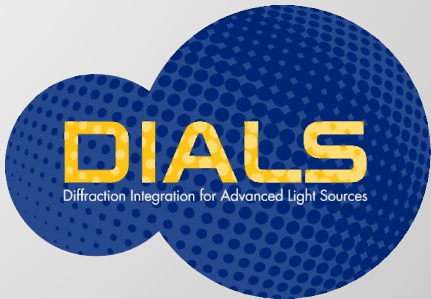
- Charles Babbage, Passages from the Life of a Philosopher

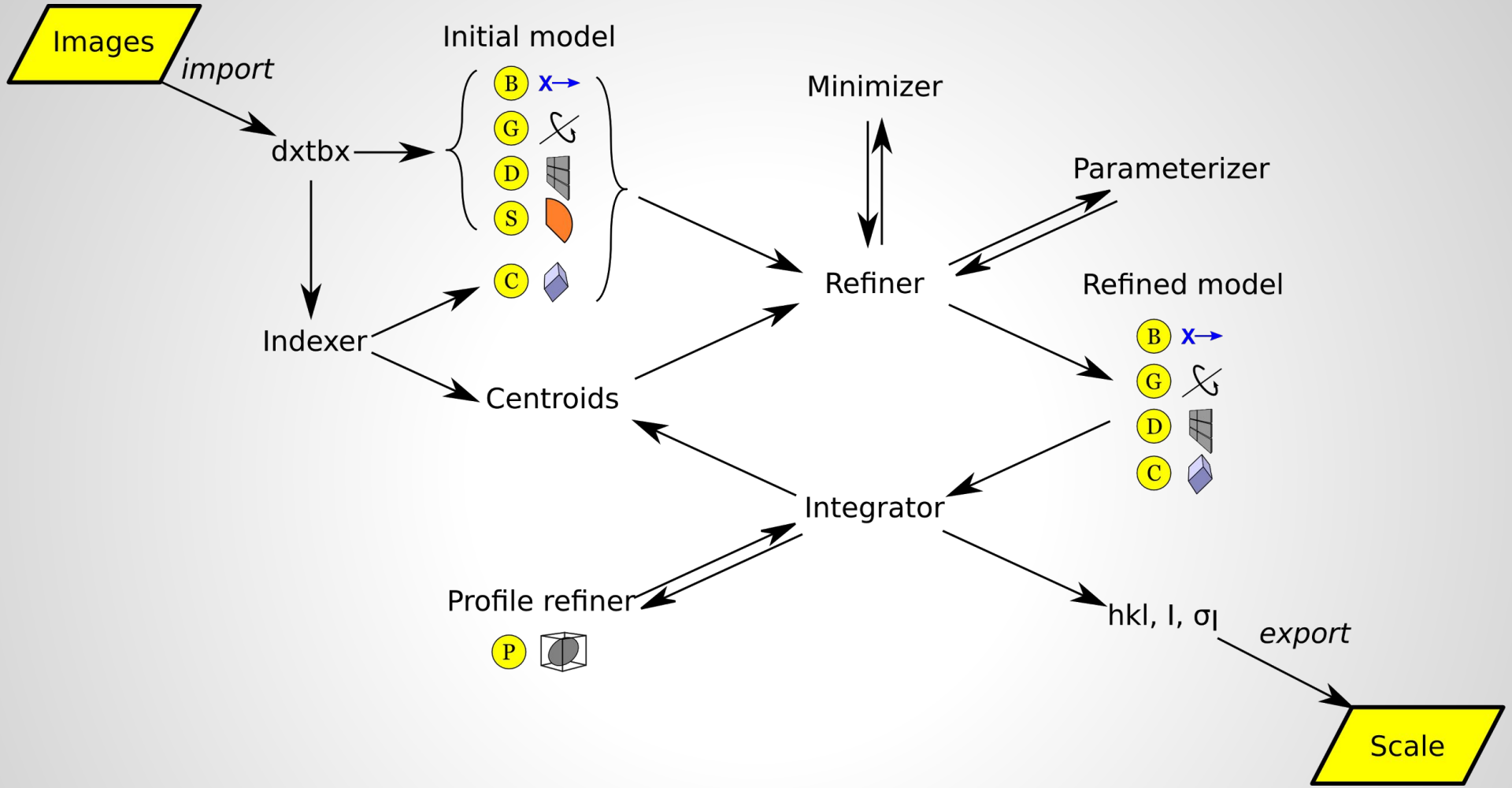
CCP4

DIALS
Diffraction Integration for Advanced Light Sources

DIALS OVERVIEW







CCP4

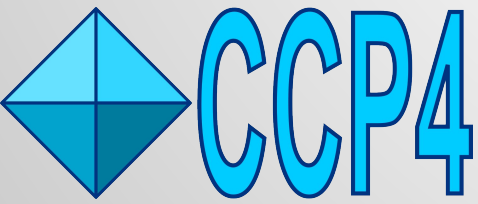
DIALS
Diffraction Integration for Advanced Light Sources

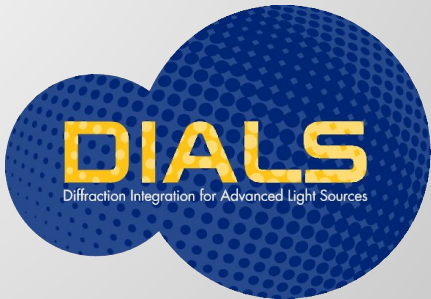
Main DIALS programs

- `dials.import`
- **`dials.find_spots`**
- **`dials.index`**
- `dials.refine_bravais_settings`
- **`dials.refine`**
- **`dials.integrate`**
- `dials.export_mtz`

Then onwards into the CCP4 data processing pipeline:

- POINTLESS → AIMLESS → ctruncate...





Principal user interface

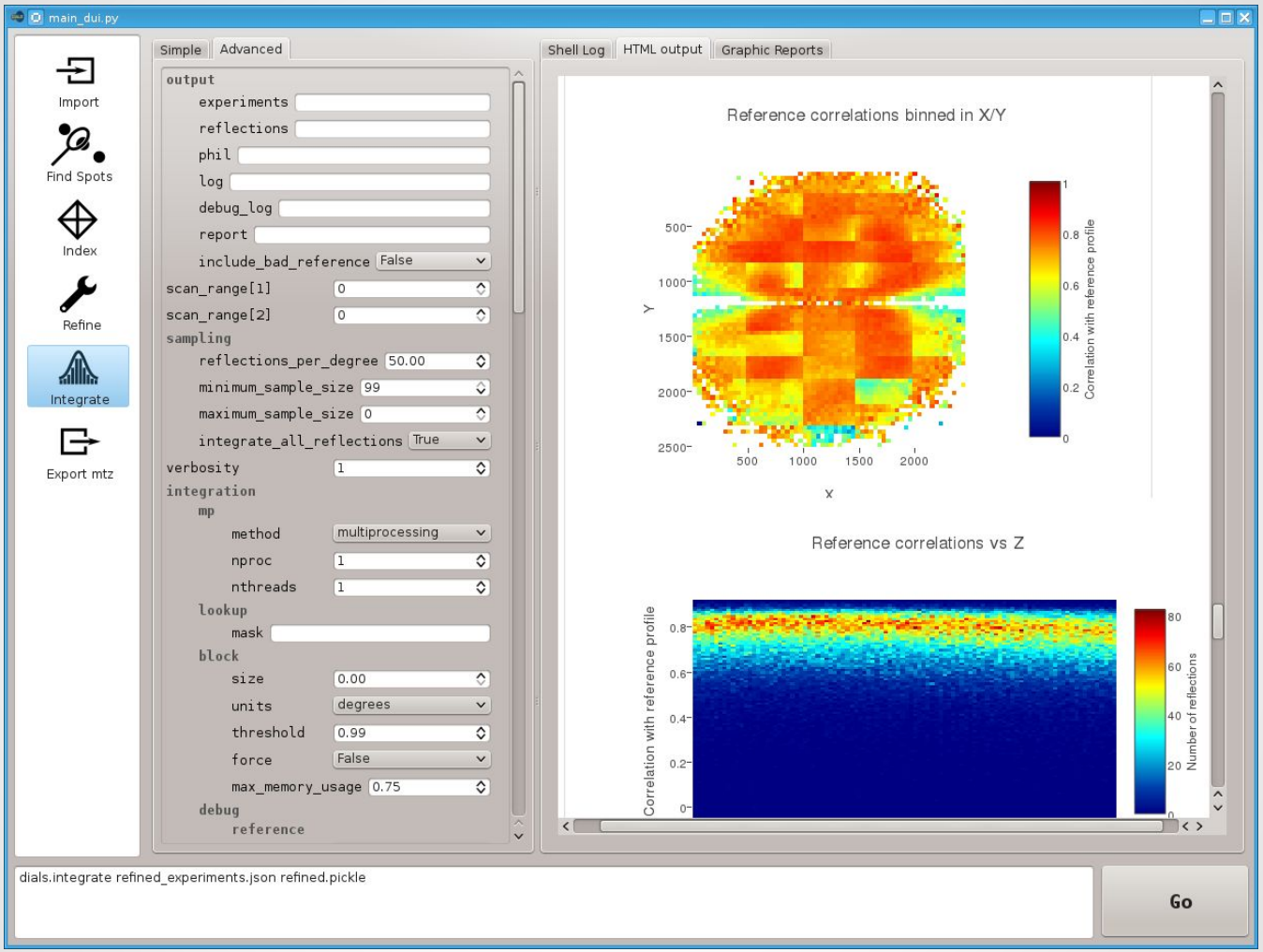
```
$ dials.import ${data_directory}/th_8_2_0*.cbf
$ dials.find_spots datablock.json nproc=8
$ dials.index datablock.json strong.pickle
$ dials.refine_bravais_settings experiments.json indexed.pickle
$ dials.reindex indexed.pickle change_of_basis_op=a,b,c
$ dials.refine bravais_setting_9.json reindexed_reflections.pickle
$ dials.refine refined_experiments.json refined.pickle scan_varying=true
$ dials.integrate refined_experiments.json refined.pickle nproc=8
$ dials.export integrated.pickle refined_experiments.json \
mtz.hklout=integrated.mtz
$ pointless hklin integrated.mtz hklout scaled.mtz > pointless.log
$ aimless hklin scaled.mtz hklout scaled.mtz > aimless.log <<eof
  resolution 1.3
  anomalous off
eof
$ ctruncate -hklin scaled.mtz -hklout truncated.mtz \
  -colin '/*/*/[IMEAN,SIGIMEAN]' > ctruncate.log
```

who needs a GUI?

CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Future: DIALS GUI (DUI)



CCP4

DIALS
Diffraction Integration for Advanced Light Sources

xia2 -dials

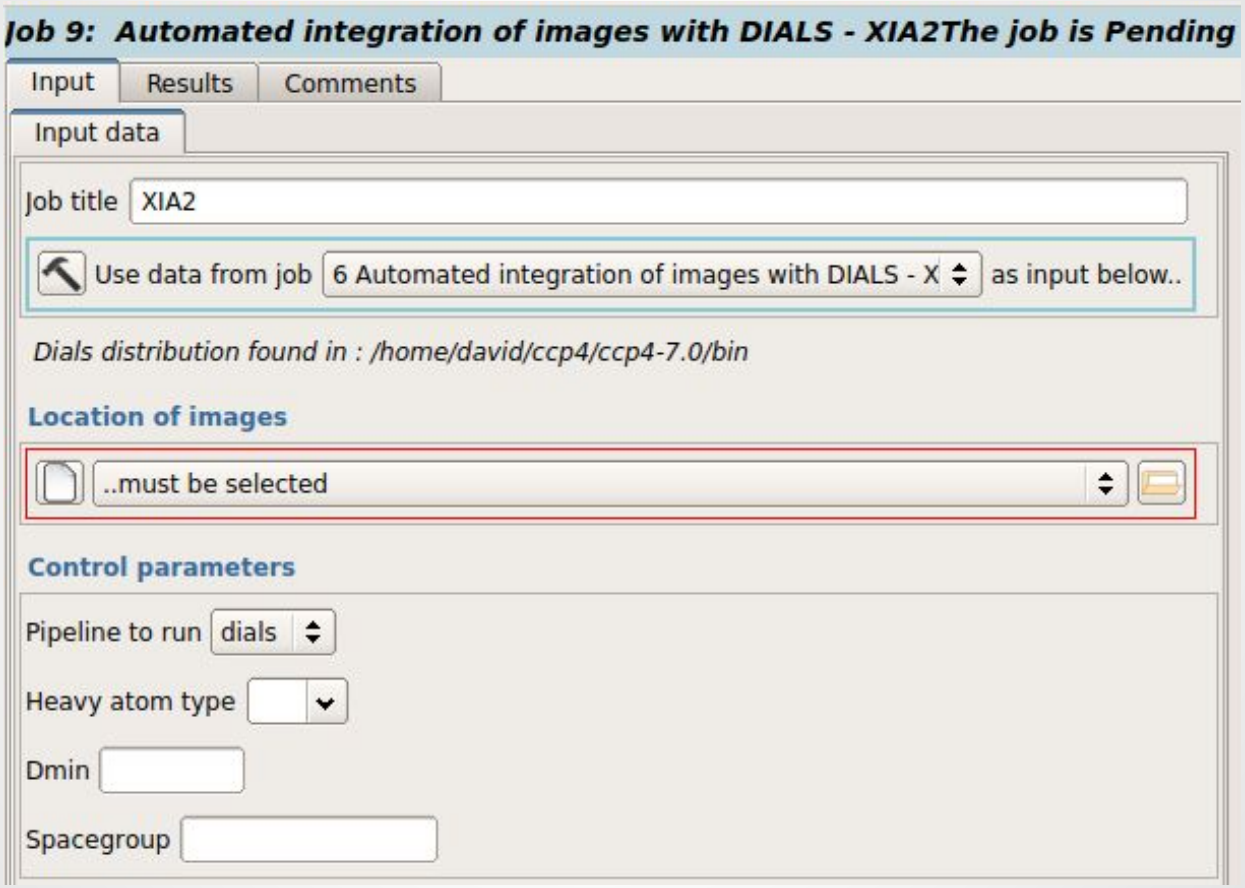
```
$ xia2 -dials ${data_directory}
```

```
For AUTOMATIC/DEFAULT/NATIVE
High resolution limit                          1.36   6.08   1.36
Low resolution limit                          53.92  53.92   1.40
Completeness                                  99.8   99.9   97.8
Multiplicity                                   5.3    4.9    3.1
I/sigma                                       11.7   26.8    2.2
Rmerge                                       0.061  0.026  0.370
Rmeas(I)                                     0.075  0.032  0.521
Rmeas(I+/-)                                  0.075  0.032  0.495
Rpim(I)                                      0.032  0.014  0.279
Rpim(I+/-)                                   0.042  0.018  0.326
CC half                                      0.999  0.998  0.818
Wilson B factor                              8.913
Anomalous completeness                        97.5  100.0   77.3
Anomalous multiplicity                         2.6    3.1    1.8
Anomalous correlation                        0.005  0.204 -0.019
Anomalous slope                              0.955  0.000  0.000
Total observations                          292123 3747   12262
Total unique                                55480  768    3919
Assuming spacegroup: P 41 21 2
Other likely alternatives are:
P 43 21 2
Unit cell:
57.781  57.781 149.995
90.000  90.000  90.000
```

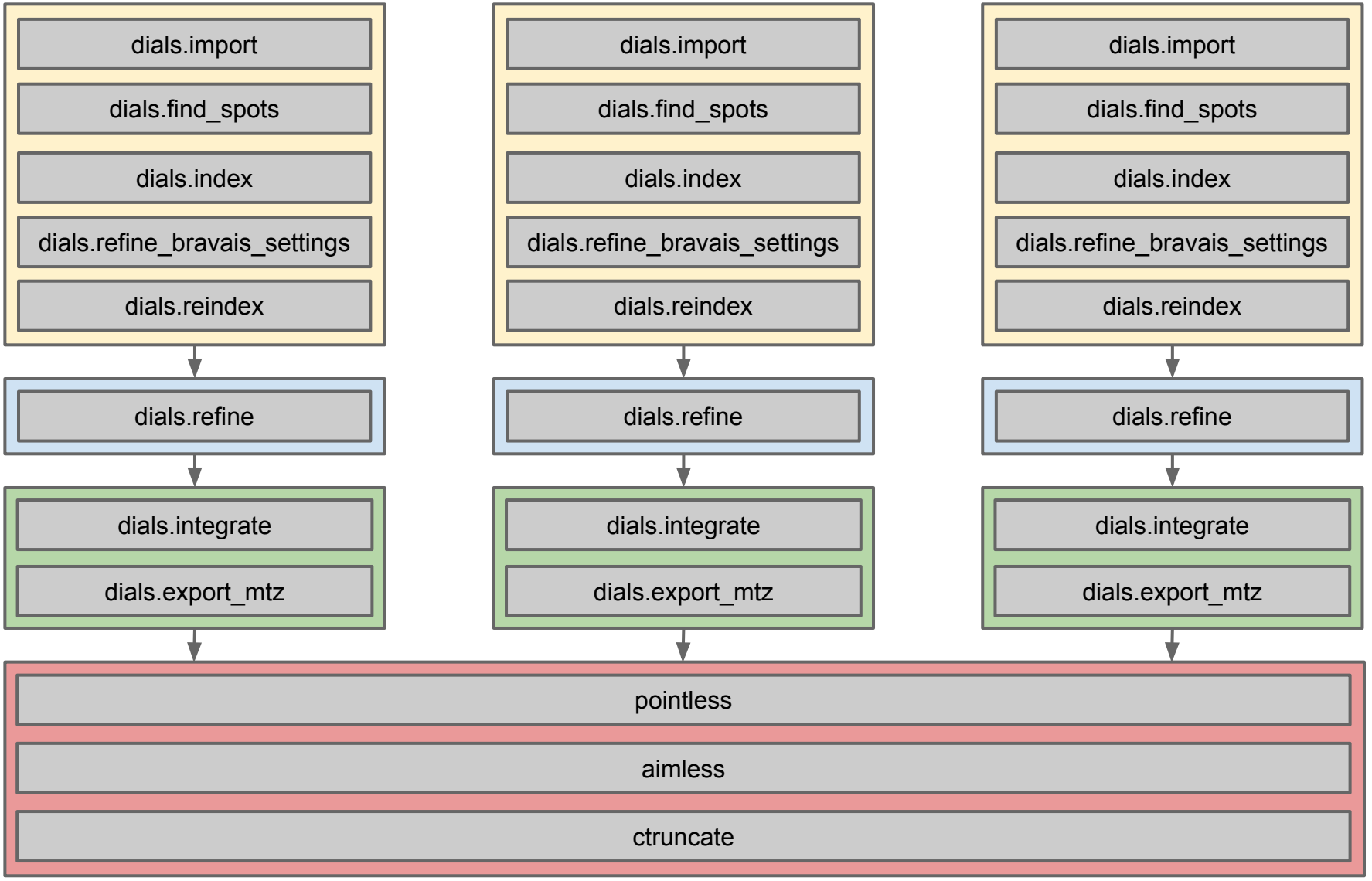
CCP4

DIALS
Diffraction Integration for Advanced Light Sources

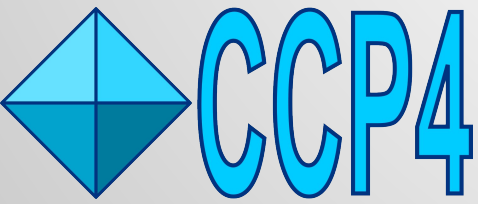
xia2 in CCP4 7.0

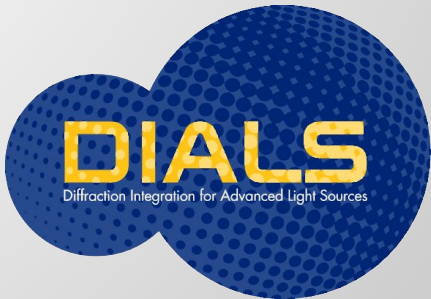


xia2 -dials


dials.import
dials.find_spots
dials.index
dials.refine_bravais_settings
dials.reindex
dials.refine
dials.integrate
dials.export_mtz
dials.import
dials.find_spots
dials.index
dials.refine_bravais_settings
dials.reindex
dials.refine
dials.integrate
dials.export_mtz
dials.import
dials.find_spots
dials.index
dials.refine_bravais_settings
dials.reindex
dials.refine
dials.integrate
dials.export_mtz
pointless
aimless
ctruncate

SPOTFINDING





dials.find_spots

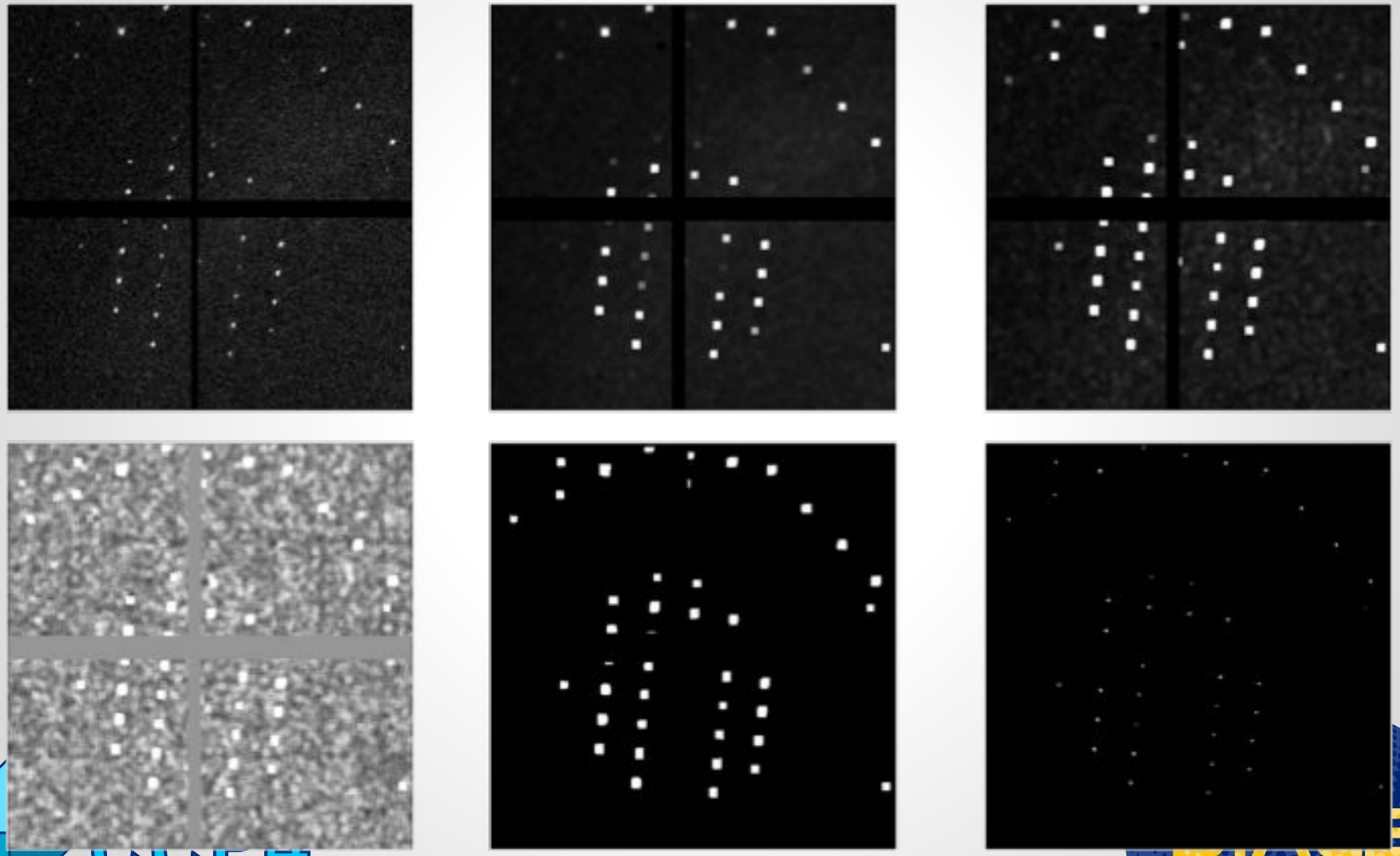
- Sequence of per-image filters to find strong pixels
- 3D analysis of strong pixels to identify strong spots
- Filter spots by
  - number of pixels
  - peak-centroid distance
  - resolution
  - ice rings
  - untrusted regions

```
$ dials.find_spots datablock.json nproc=8

Setting spotfinder.filter.min_spot_size=3
Configuring spot finder from input parameters
-----------------------------------------------
Finding strong spots in imageset 0
-----------------------------------------------
Finding spots in image 1 to 540...
Extracting strong pixels from images (may take a while)
Extracted strong pixels from images
Merging 8 pixel lists
Merged 8 pixel lists with 922120 pixels
Extracting spots
Extracted 219125 spots
Calculating 219125 spot centroids
Calculated 219125 spot centroids
Calculating 219125 spot intensities
Calculated 219125 spot intensities
Found 1 possible hot spots
Found 1 possible hot pixel(s)
Filtering 219125 spots by number of pixels
Filtered 116321 spots by number of pixels
Filtering 116321 spots by peak-centroid distance
Filtered 116082 spots by peak-centroid distance

-----------------------------------------------
Saving 116082 reflections to strong.pickle
Saved 116082 reflections to strong.pickle
Time Taken: 31.768495
```

CCP4

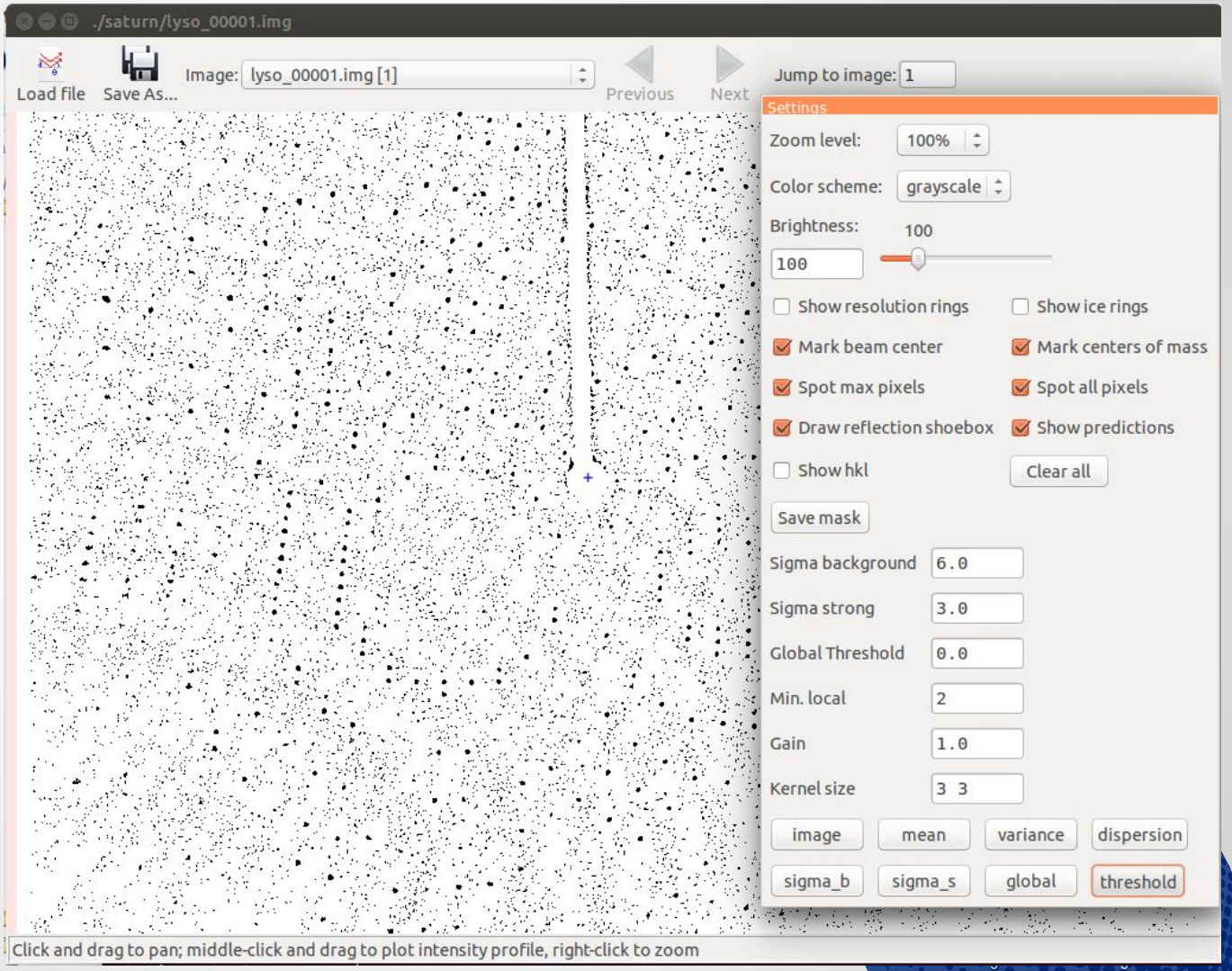
DIALS
Diffraction Integration for Advanced Light Sources

dials.find_spots



dials.find_spots

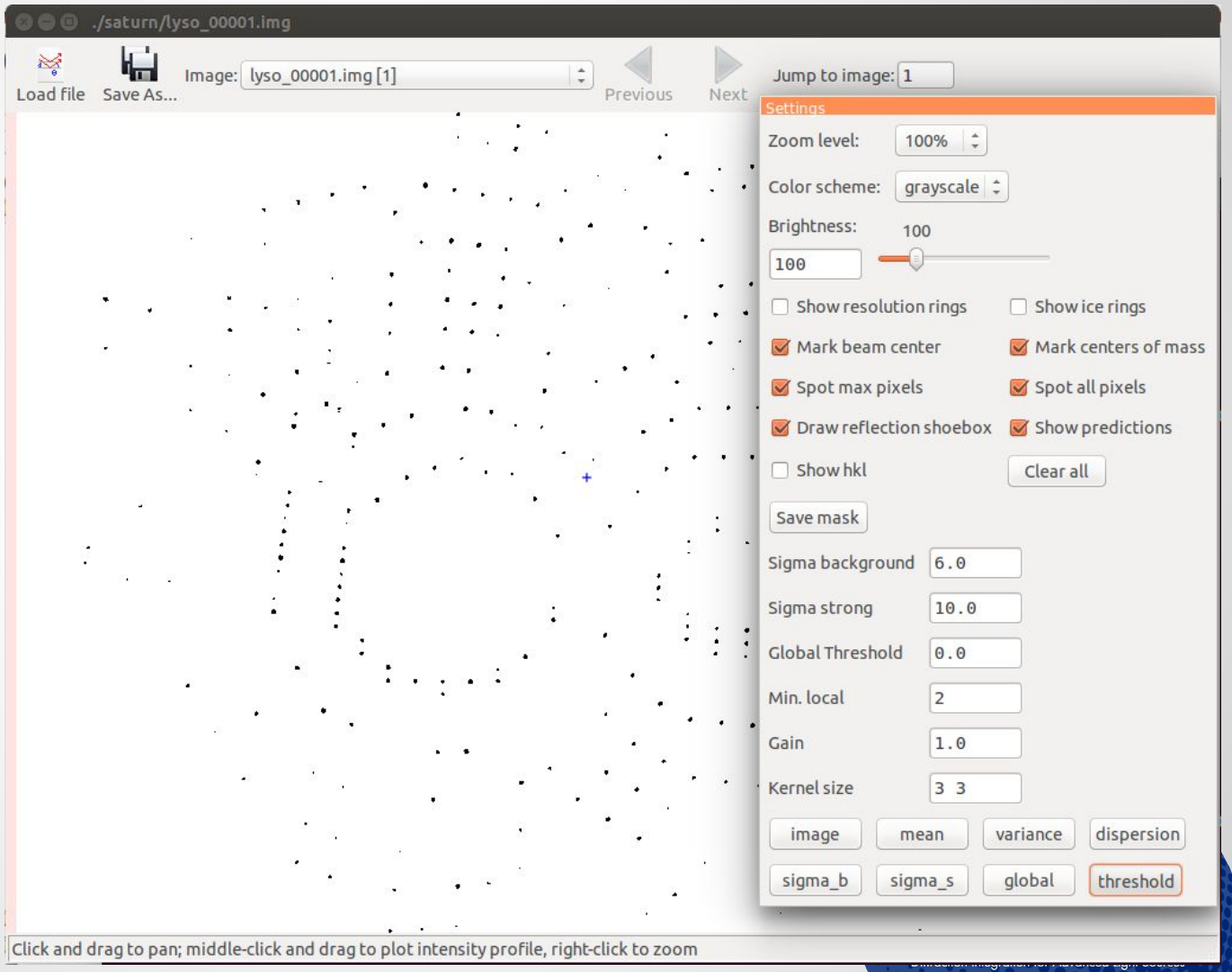
Default spot finding parameters are often not suitable for CCD images



CCP4

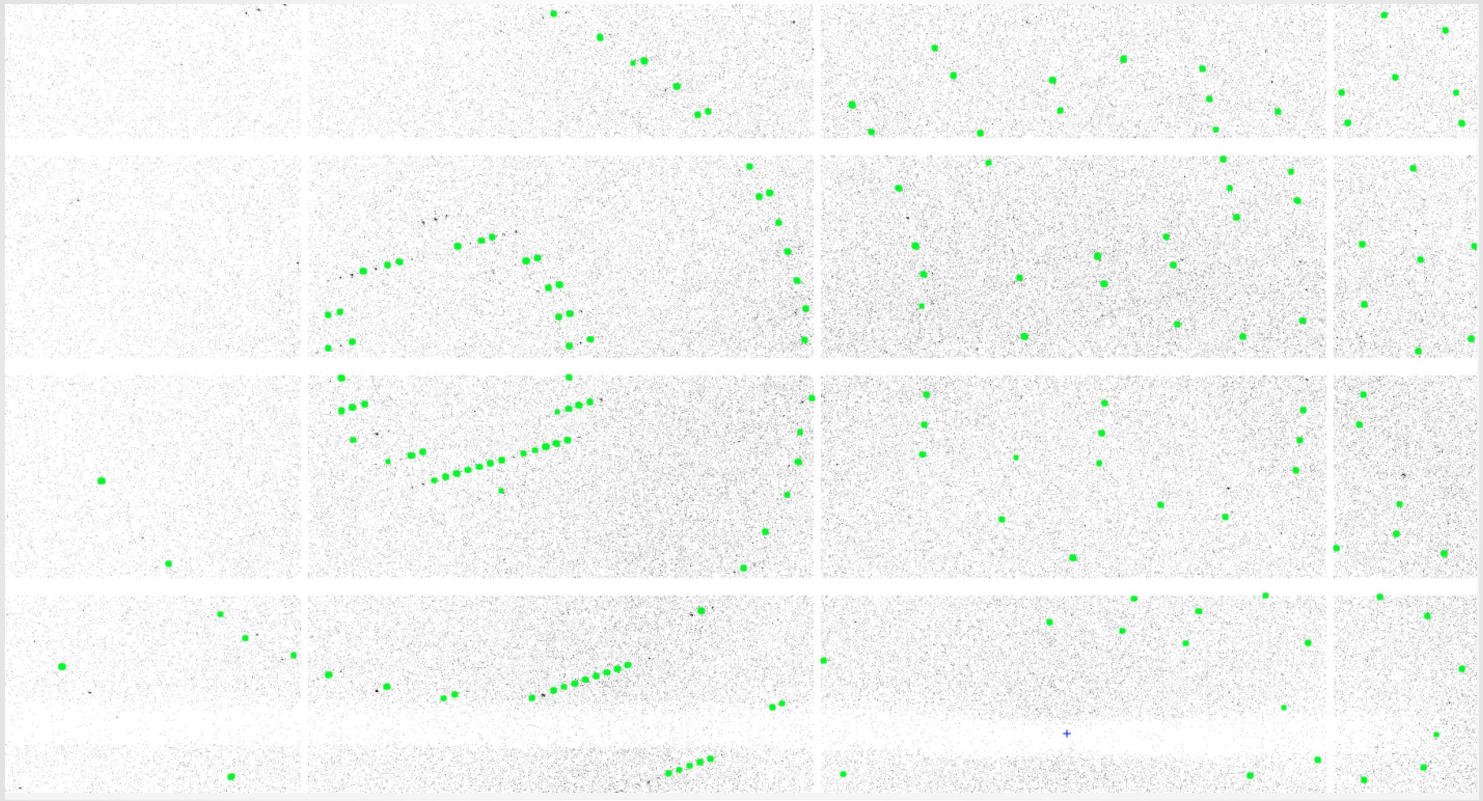
dials.find_spots

Default spot finding parameters are often not suitable for CCD images



dials.image_viewer

```
$ dials.image_viewer datablock.json strong.pickle
```

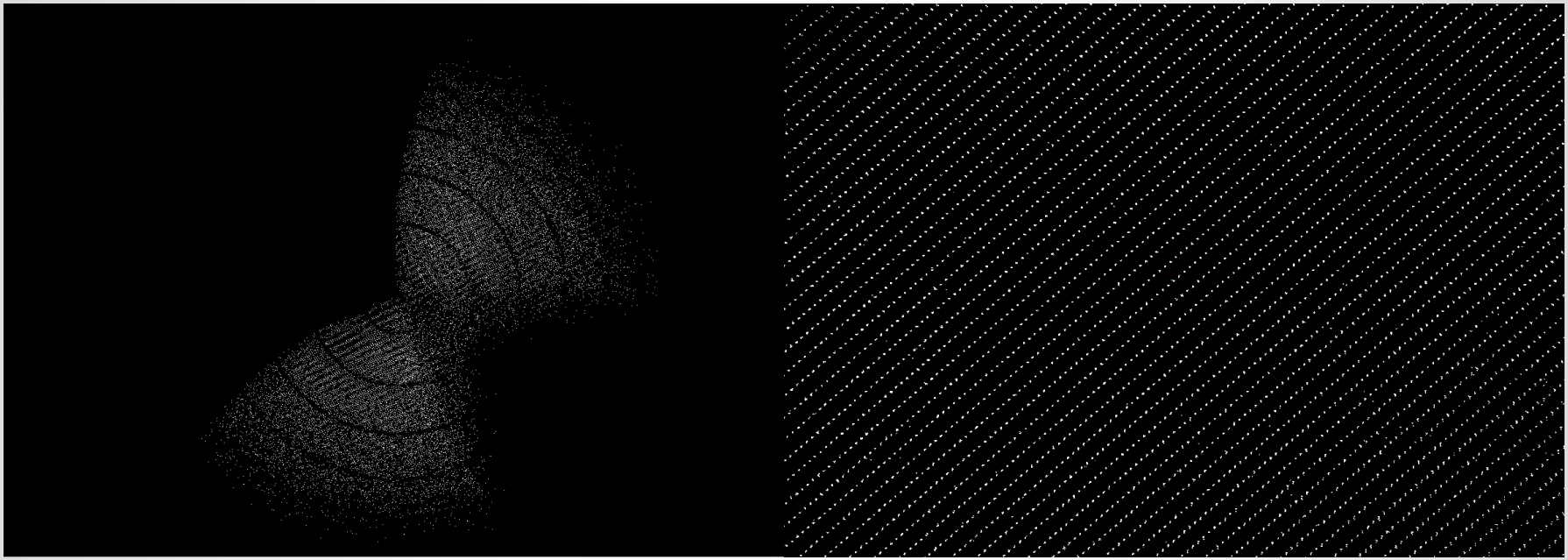


CCP4

DIALS
Diffraction Integration for Advanced Light Sources

dials.reciprocal_lattice_viewer

```
$ dials.reciprocal_lattice_viewer datablock.json strong.pickle
```

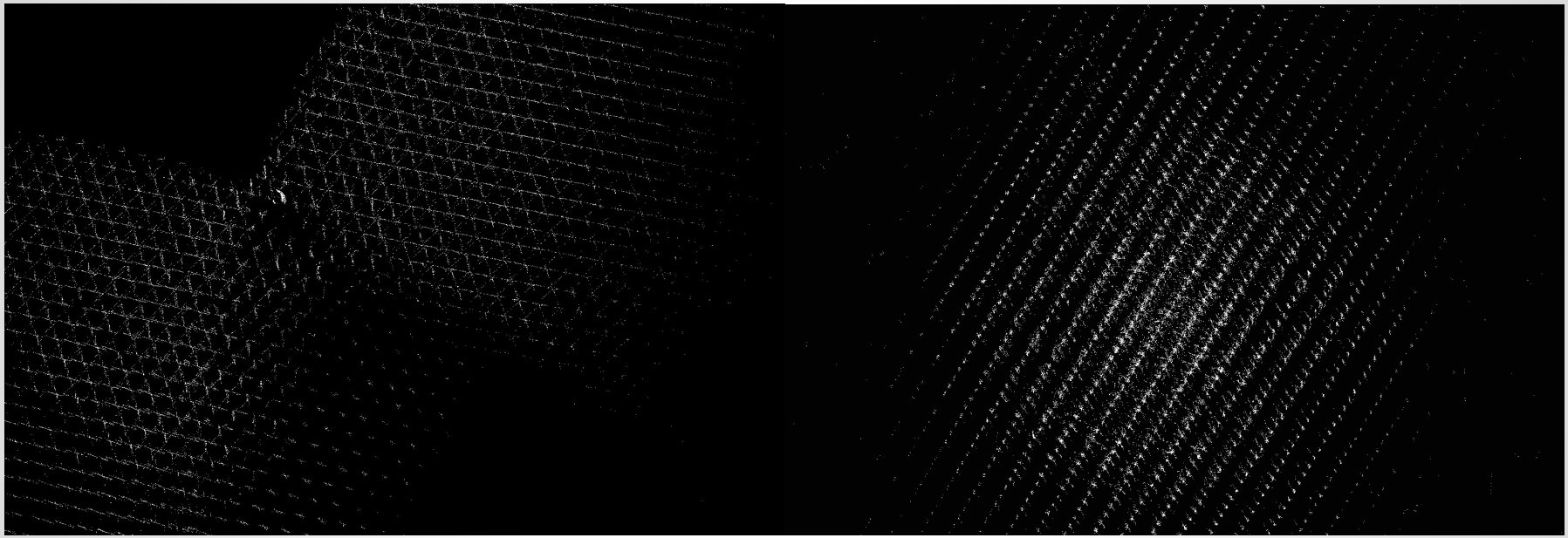


CCP4

DIALS
Diffraction Integration for Advanced Light Sources

dials.reciprocal_lattice_viewer

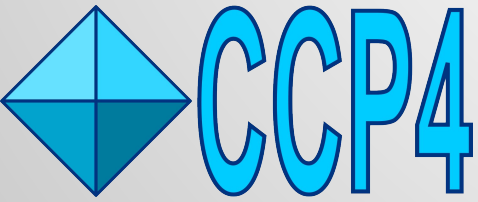
```
$ dials.reciprocal_lattice_viewer datablock.json strong.pickle
```

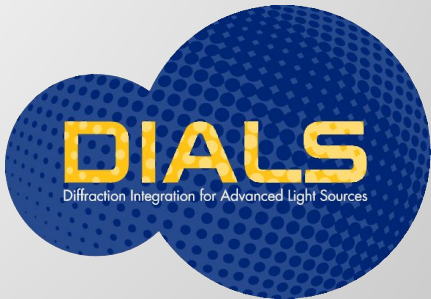


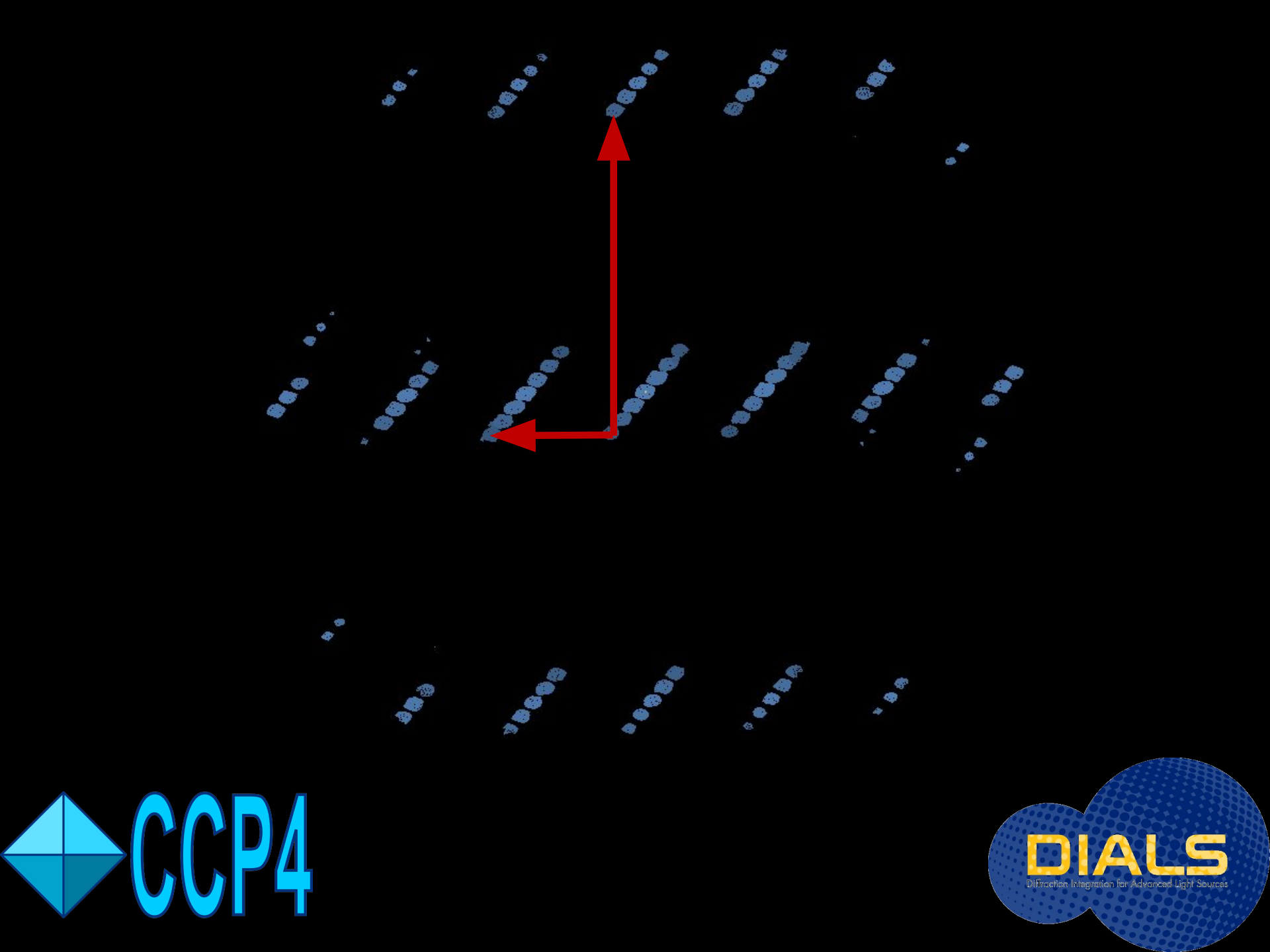
CCP4

DIALS
Diffraction Integration for Advanced Light Sources

INDEXING


CCP4

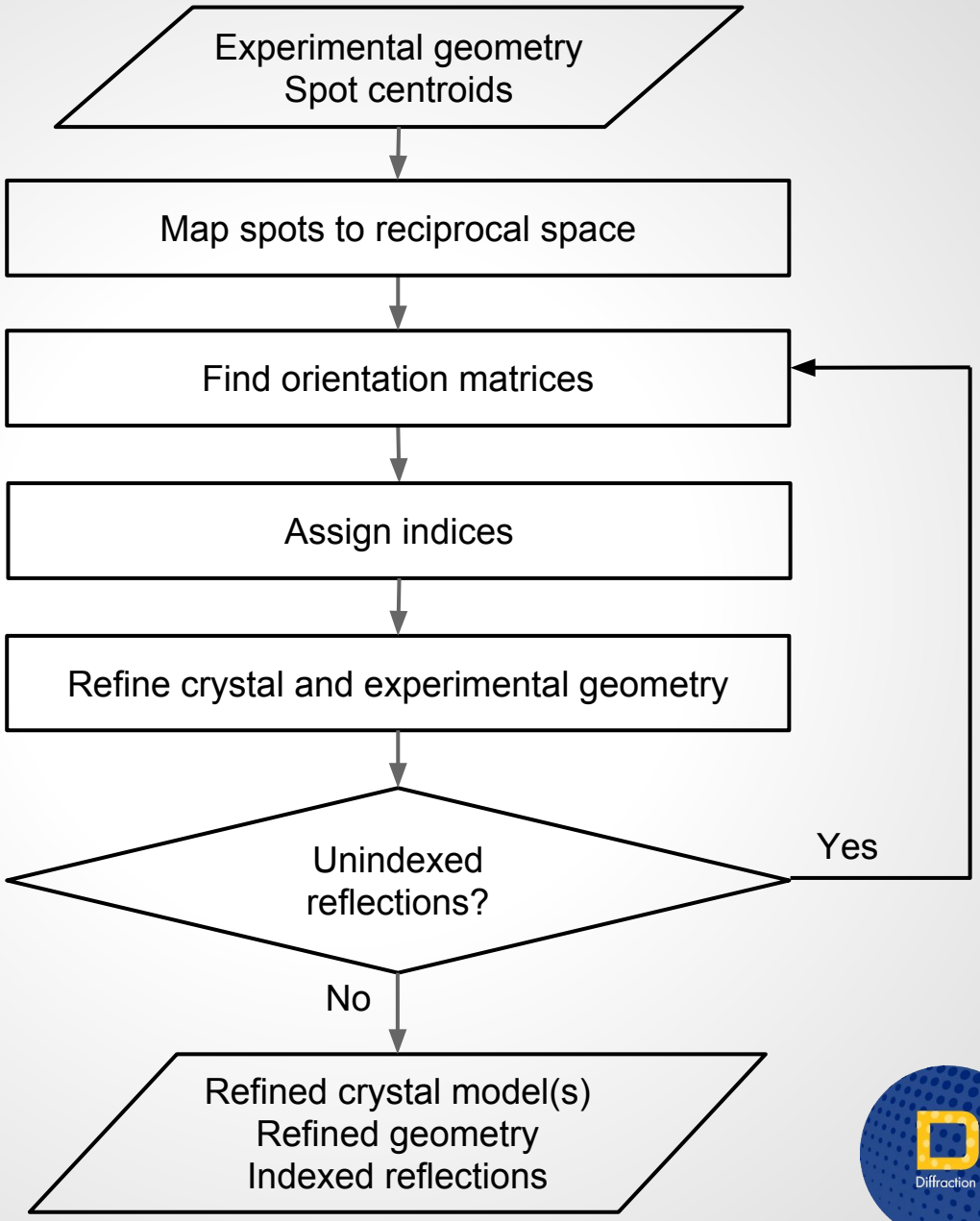

DIALS
Diffraction Integration for Advanced Light Sources



CCP4

DIALS
Diffraction Integration for Advanced Light Sources

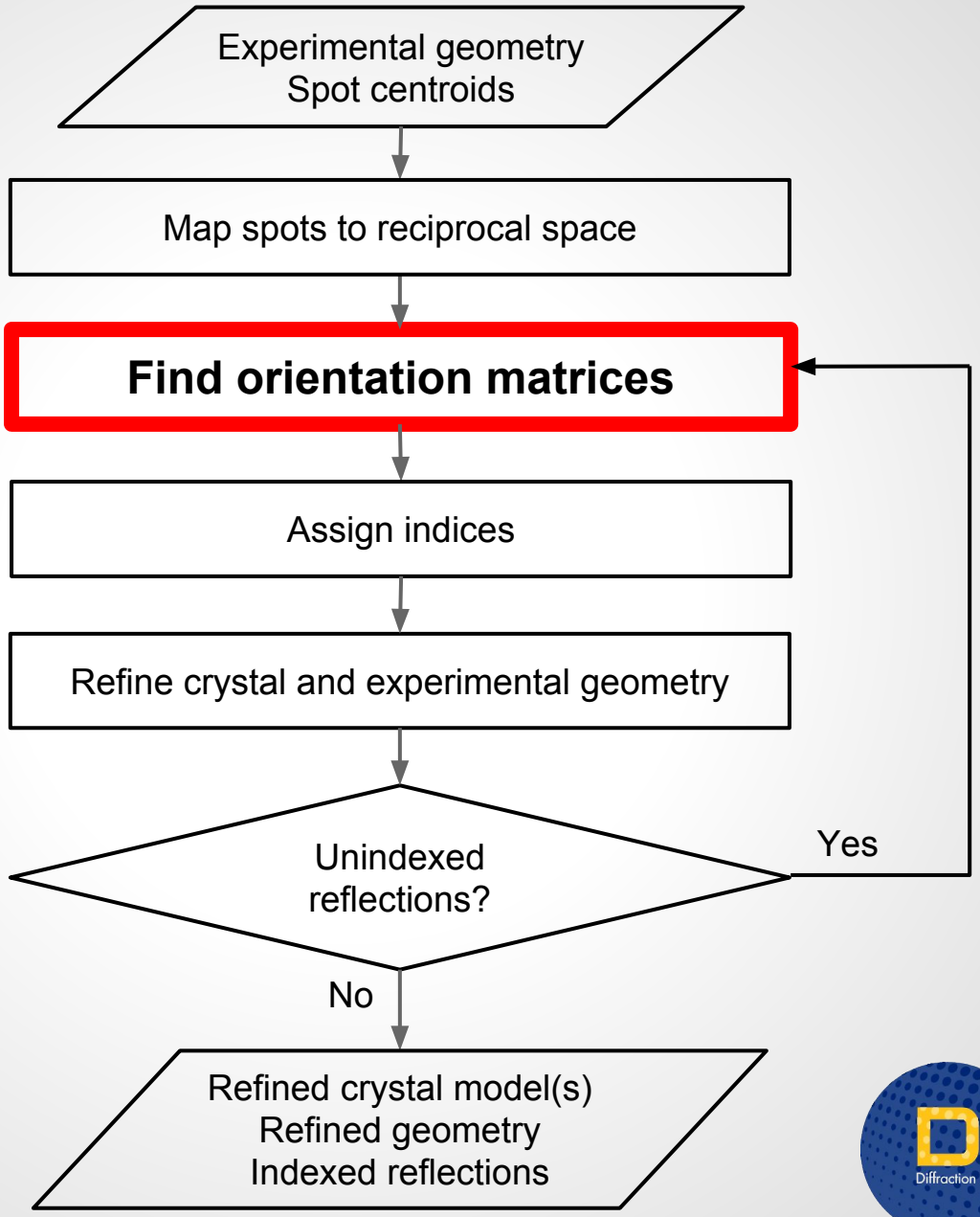
dials.index



CCP4

DIALS
Diffraction Integration for Advanced Light Sources

dials.index



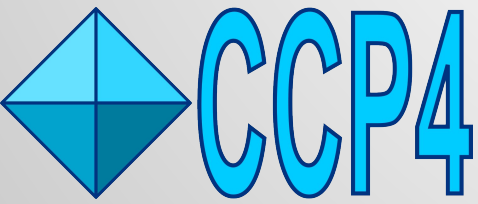
CCP4

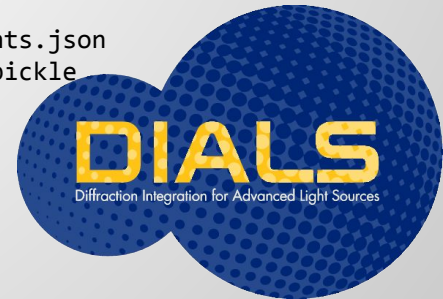
DIALS
Diffraction Integration for Advanced Light Sources

dials.index

- Choice of method:
  - 1D FFT (DPS)
  - 3D FFT
  - new real space grid search
- Optionally provide known unit cell and space group

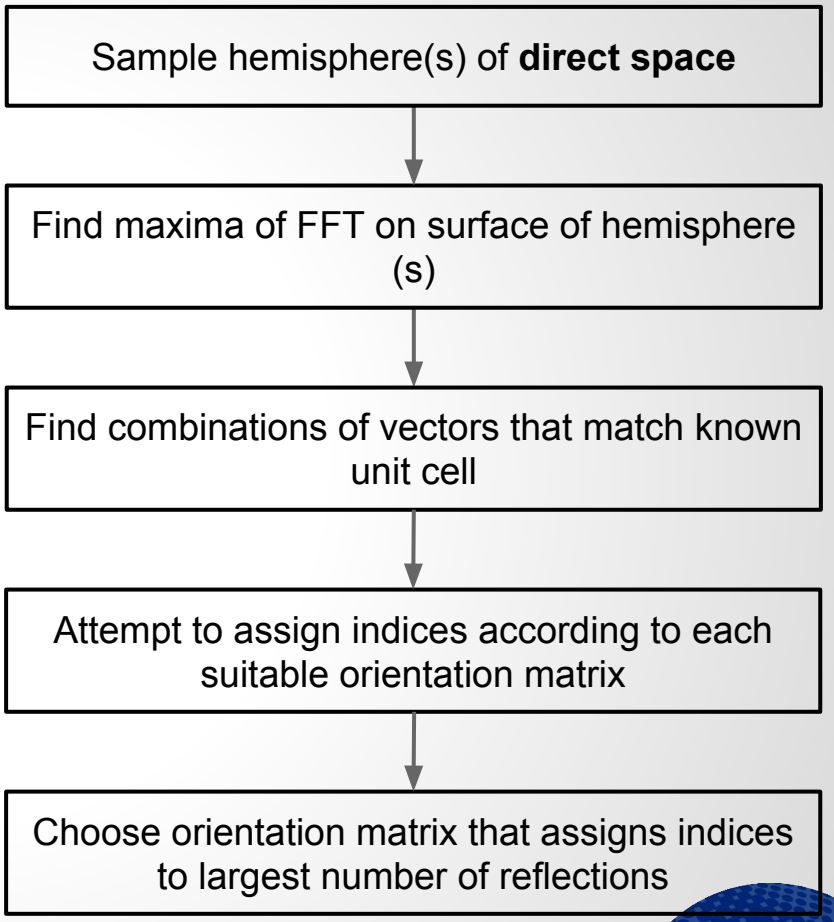
```
$ dials.index datablock.json strong.pickle
Found max_cell: 199.1 Angstrom
Setting d_min: 3.89
...
RMSDs by experiment:
-----------------------------------------
| Exp | Nref | RMSD_X | RMSD_Y  | RMSD_Z   |
|     |      | (px)   | (px)    | (images) |
-----------------------------------------
| 0   | 4049 | 0.2881 | 0.25838 | 0.17767  |
-----------------------------------------

Final refined crystal models:
model 1 (114690 reflections):
Crystal:
    Unit cell: (57.804, 57.782, 150.027, 90.009, 89.991, 89.990)
    Space group: P 1
    U matrix:  {{ 0.3455, -0.2589, -0.9020},
                { 0.8914,  0.3909,  0.2292},
                { 0.2933, -0.8833,  0.3659}}
    B matrix:  {{ 0.0173,  0.0000,  0.0000},
                {-0.0000,  0.0173,  0.0000},
                {-0.0000,  0.0000,  0.0067}}
    A = UB:    {{ 0.0060, -0.0045, -0.0060},
                { 0.0154,  0.0068,  0.0015},
                { 0.0051, -0.0153,  0.0024}}

Saving refined experiments to experiments.json
Saving refined reflections to indexed.pickle
```





Real space grid search

- The FFT algorithms try to determine the magnitude and direction of the basis vectors simultaneously
- If the unit cell is already known, then we only need to determine the direction of the basis vectors

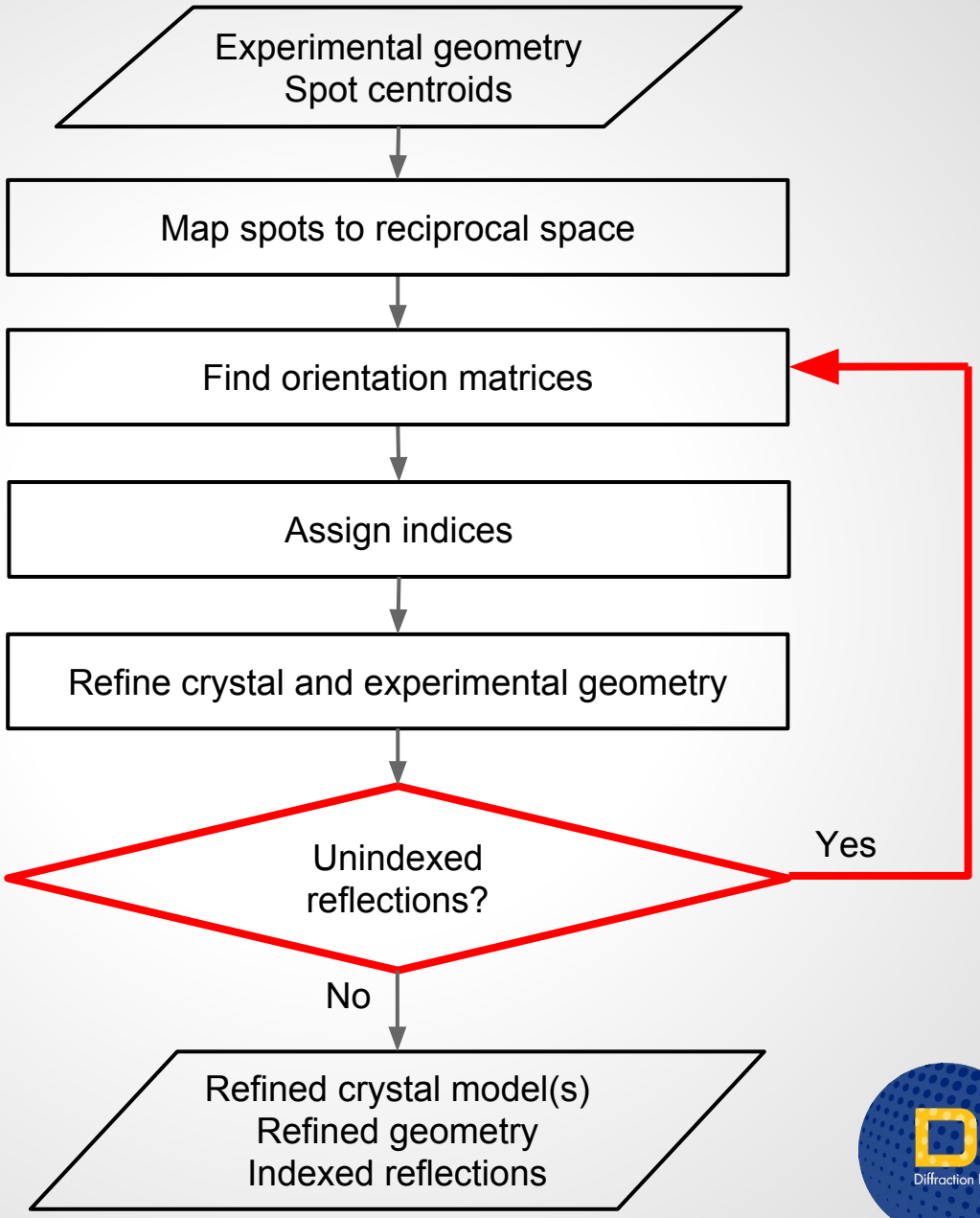


New methods for indexing multi-lattice diffraction data. Gildea *et al.*, *Acta Cryst.* (2014) **D**70, 2652-2666.

CCP4

DIALS
Diffraction Integration for Advanced Light Sources

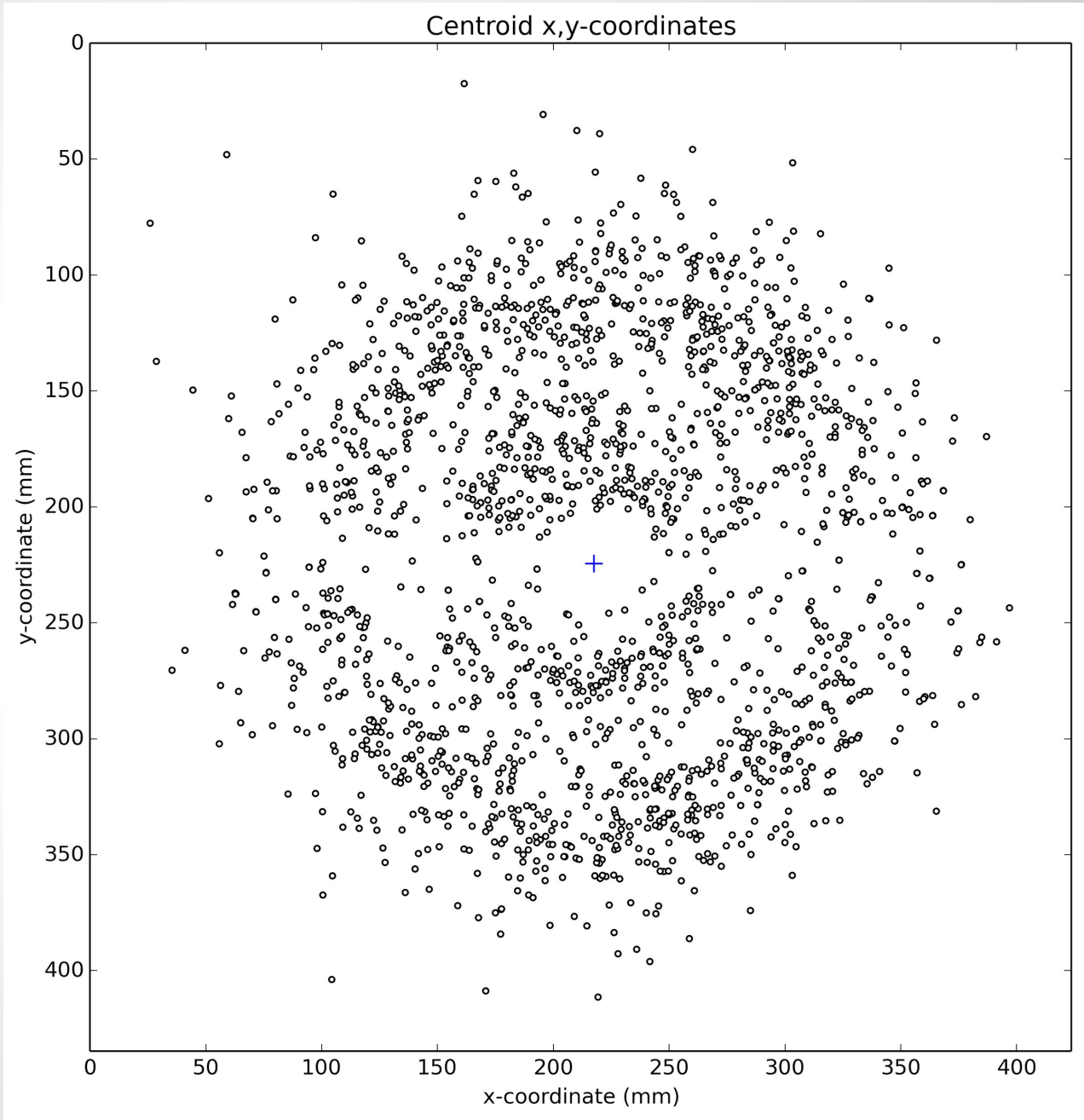
dials.index

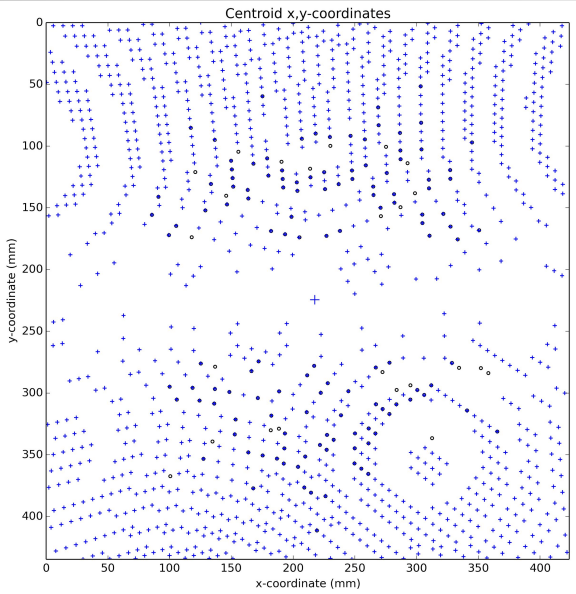


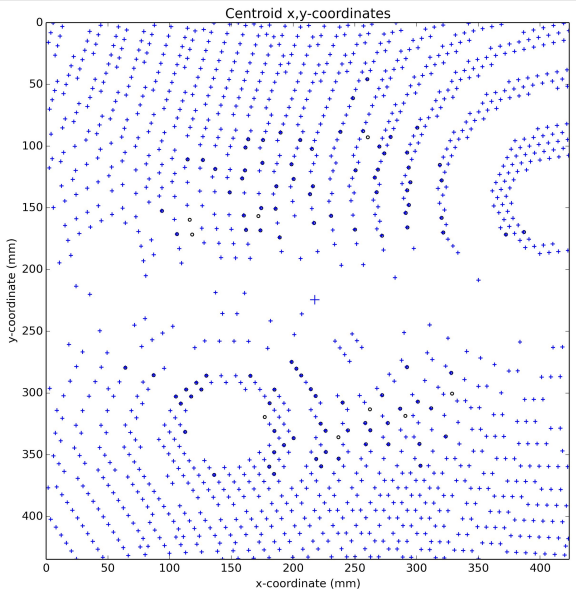
CCP4

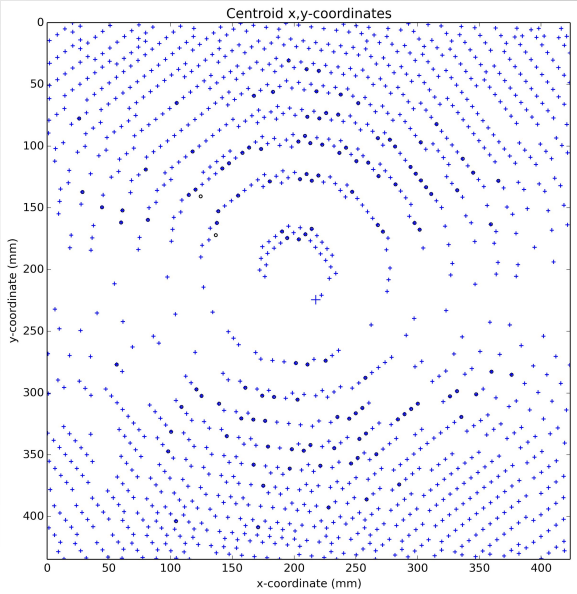
DIALS
Diffraction Integration for Advanced Light Sources

- 1° wedge of data
- 1858 spots
- How many lattices?

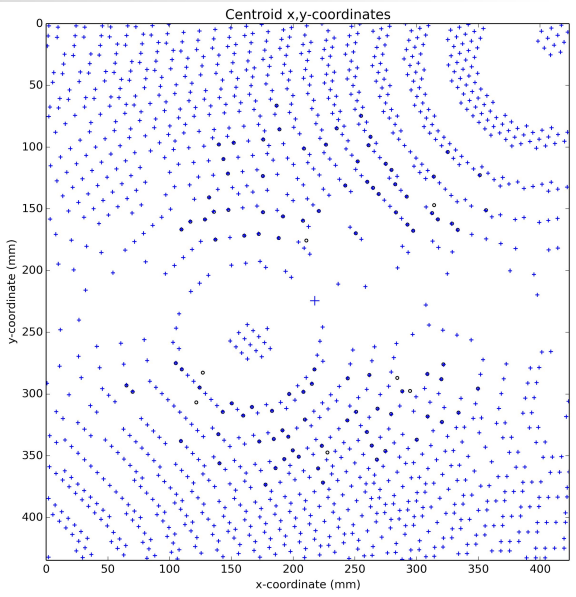


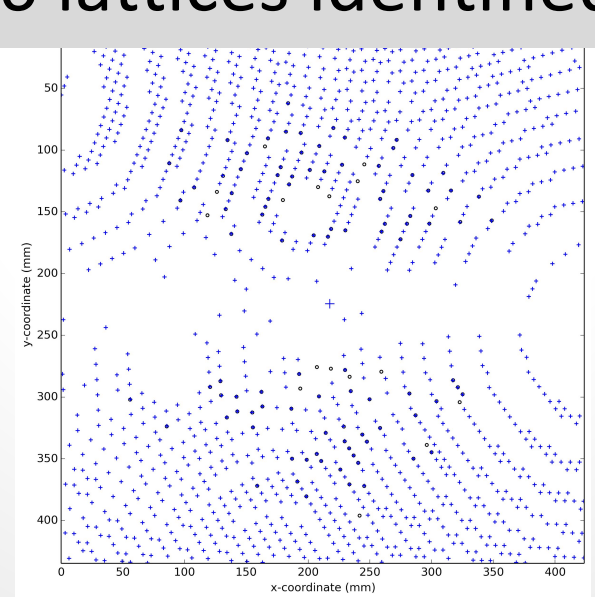


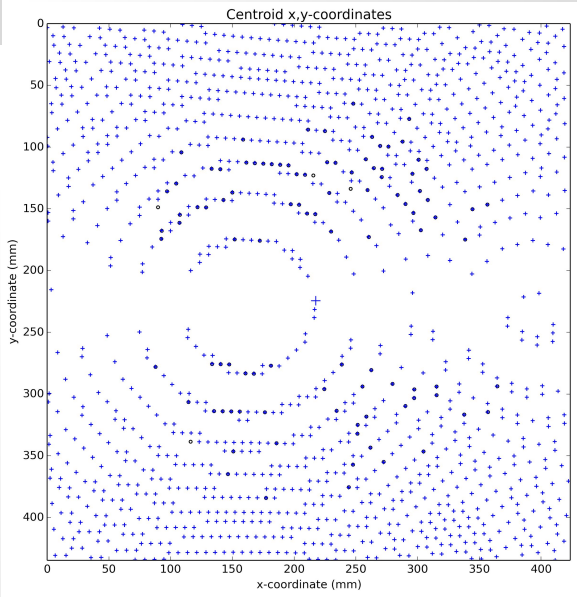




6 lattices identified!







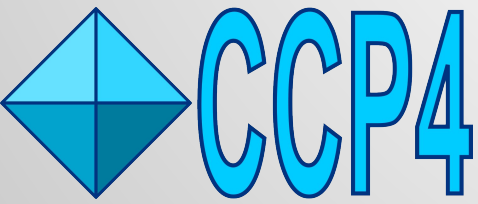
CCP4

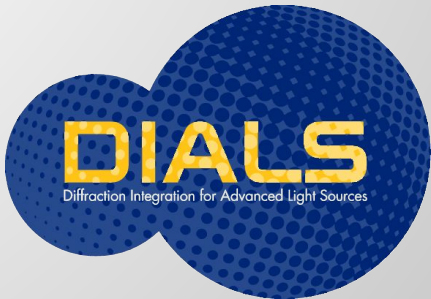
DIALS
Diffraction Integration for Advanced Light Sources

dials.refine_bravais_settings

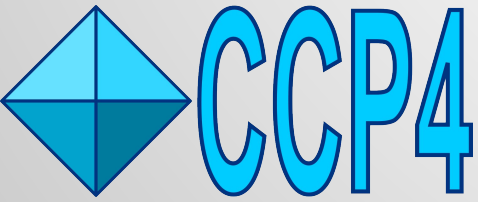
```
$ dials.refine_bravais_settings experiments.json indexed.pickle

---------------------------------------------------------------------------------------------------------------
Solution Metric fit  rmsd  min/max cc #spots lattice                                 unit_cell  volume      cb_op
---------------------------------------------------------------------------------------------------------------
       9     0.0250 0.073 0.787/0.848   4049      tP  57.78  57.78 149.99  90.00  90.00  90.00  500681       a,b,c
       8     0.0250 0.072 0.787/0.970   4049      oC  81.71  81.73 150.00  90.00  90.00  90.00 1001813  a-b,a+b,c
       7     0.0133 0.071 0.787/0.899   4049      oP  57.78  57.76 149.98  90.00  90.00  90.00  500537       a,b,c
       6     0.0217 0.071 0.970/0.970   4049      mC  81.72  81.74 150.02  90.00  89.99  90.00 1002178  a-b,a+b,c
       5     0.0250 0.072 0.795/0.795   4049      mC  81.73  81.71 150.01  90.00  89.99  90.00 1001809 a+b,-a+b,c
       4     0.0131 0.070 0.807/0.807   4049      mP  57.76  57.79 149.99  90.00  90.01  90.00  500676    -b,-a,-c
       3     0.0133 0.070 0.899/0.899   4049      mP  57.79  57.77 150.00  90.00  89.99  90.00  500732       a,b,c
       2     0.0125 0.071 0.787/0.787   4049      mP  57.77 149.99  57.79  90.00  89.99  90.00  500744       b,c,a
       1     0.0000 0.070         -/-   4049      aP  57.80  57.77 150.01  90.01  89.99  89.99  500927       a,b,c
---------------------------------------------------------------------------------------------------------------
```

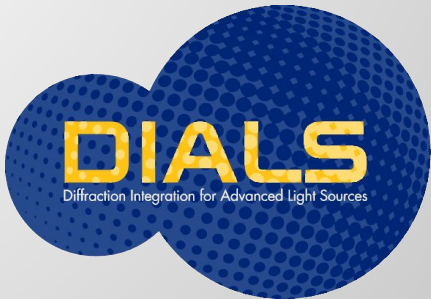
- After indexing, we can look for lattice symmetry
- All compatible Bravais lattices are tested
- Metric fit score, refined RMSD and symmetry element CCs are reported
- The user chooses which to take further





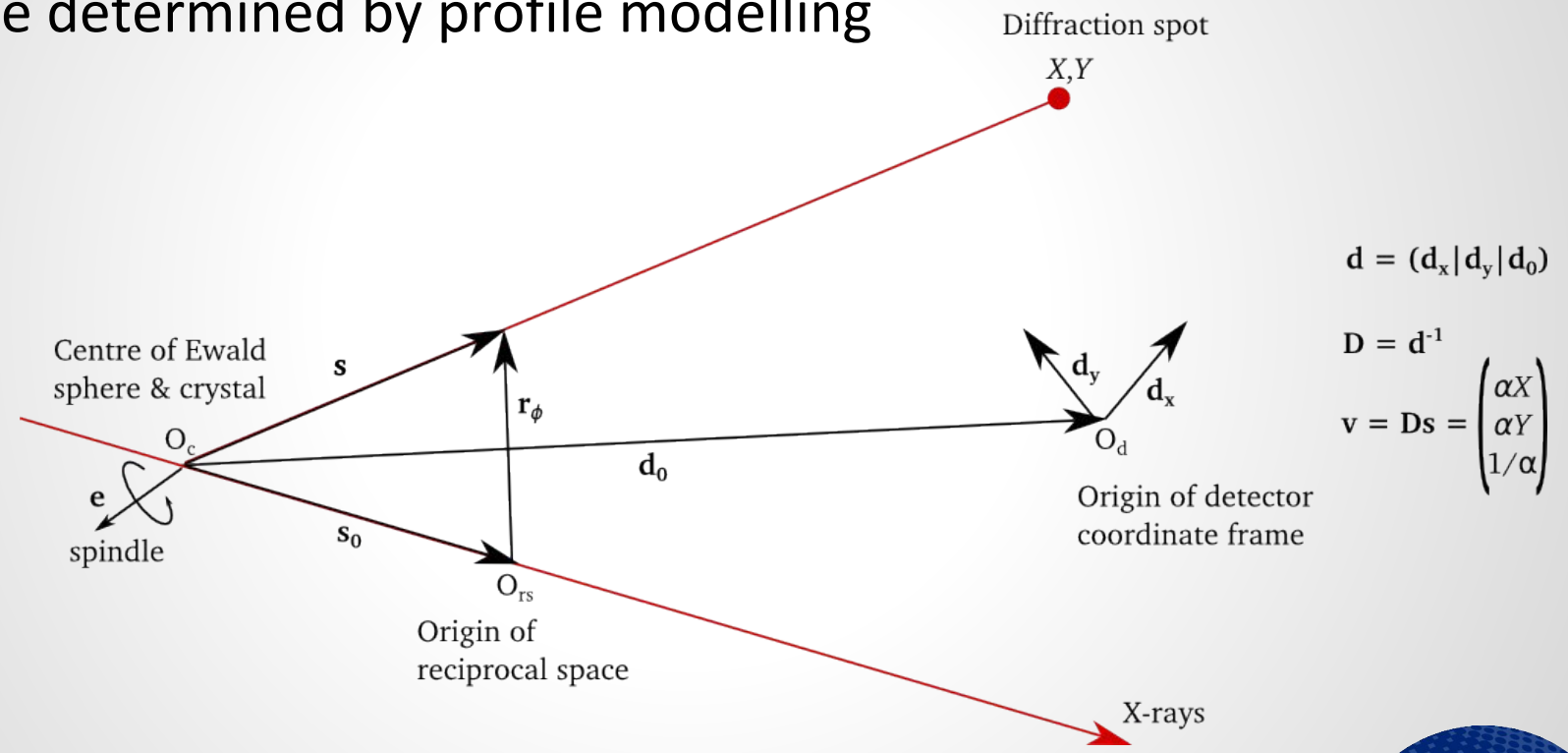
REFINEMENT


CCP4


DIALS
Diffraction Integration for Advanced Light Sources

Centroid refinement

- Refine parameters that affect *central impacts**
- Parameters that affect *general impacts* (mosaicity, Δλ, …) are determined by profile modelling



*cf. *EVAL* package
J. Appl. Cryst. **36**, (2003) 220-229

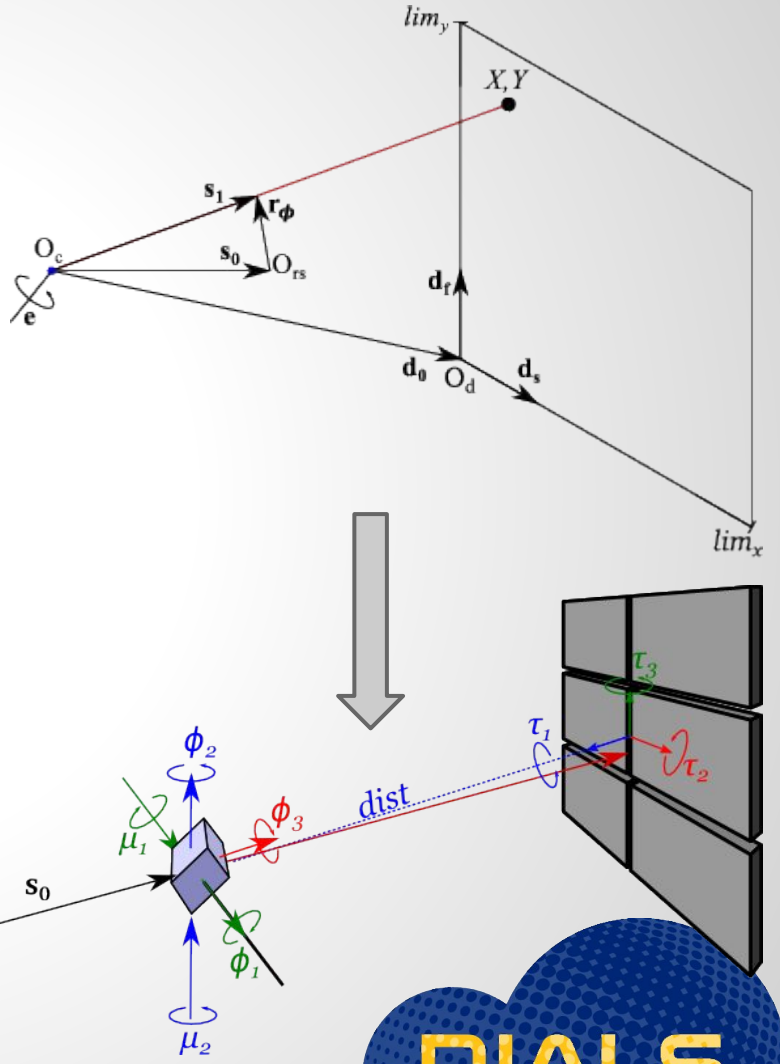
Parameterisation

There are 18 parameters in the *P* 1 case:

Table 1. *Default parameterisation in* dials.refine *for scan-static refinement using a single panel detector.*

Parameterisation	Model state	Parameters	Action
Beam	$\mathbf{s_0}$	μ_1	rotation about initial $\hat{\mu_2} \times \hat{\mathbf{s_0}}$
		μ_2	rotation about initial $\hat{\mathbf{s_0}} \times \hat{\mathbf{e}}$
		ν	set length of s_0 (wavenumber)
Crystal orientation	$\mathbf{U}$	ϕ_1	rotation about laboratory X
		ϕ_2	rotation about laboratory Y
		ϕ_3	rotation about laboratory Z
Crystal unit cell	$\mathbf{B}$	g^*_{11}	set metrical matrix elements
		g^*_{22}	
		g^*_{33}	
		g^*_{12}	
		g^*_{13}	
		g^*_{23}	
Detector	$\mathbf{d}$	p_0	set distance along initial $\hat{\mathbf{d_f}} \times \hat{\mathbf{d_s}}$
		t_1	translation along initial $\hat{\mathbf{d_f}}$
		t_2	translation along initial $\hat{\mathbf{d_s}}$
		τ_1	rotation about initial $\hat{\mathbf{d_f}} \times \hat{\mathbf{d_s}}$
		τ_2	rotation about initial $\hat{\mathbf{d_f}}$
		τ_3	rotation about initial $\hat{\mathbf{d_s}}$

Usually ν and μ_1 are fixed

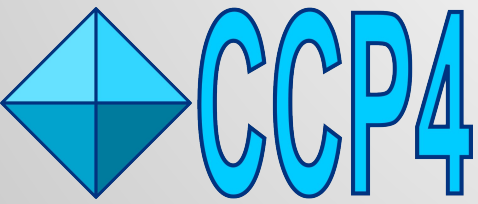


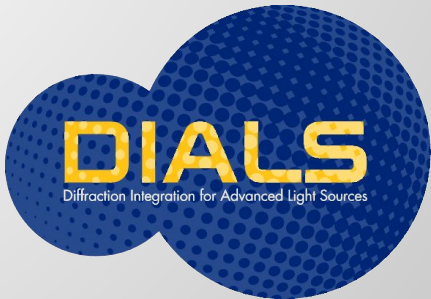
CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Scan-varying refinement

- We do global, not local, refinement
- How to model changes to the crystal model over time?
- Scan divided into equal-sized intervals
- Crystal parameterisation split over sample points
- Gaussian smoother, inspired by AIMLESS

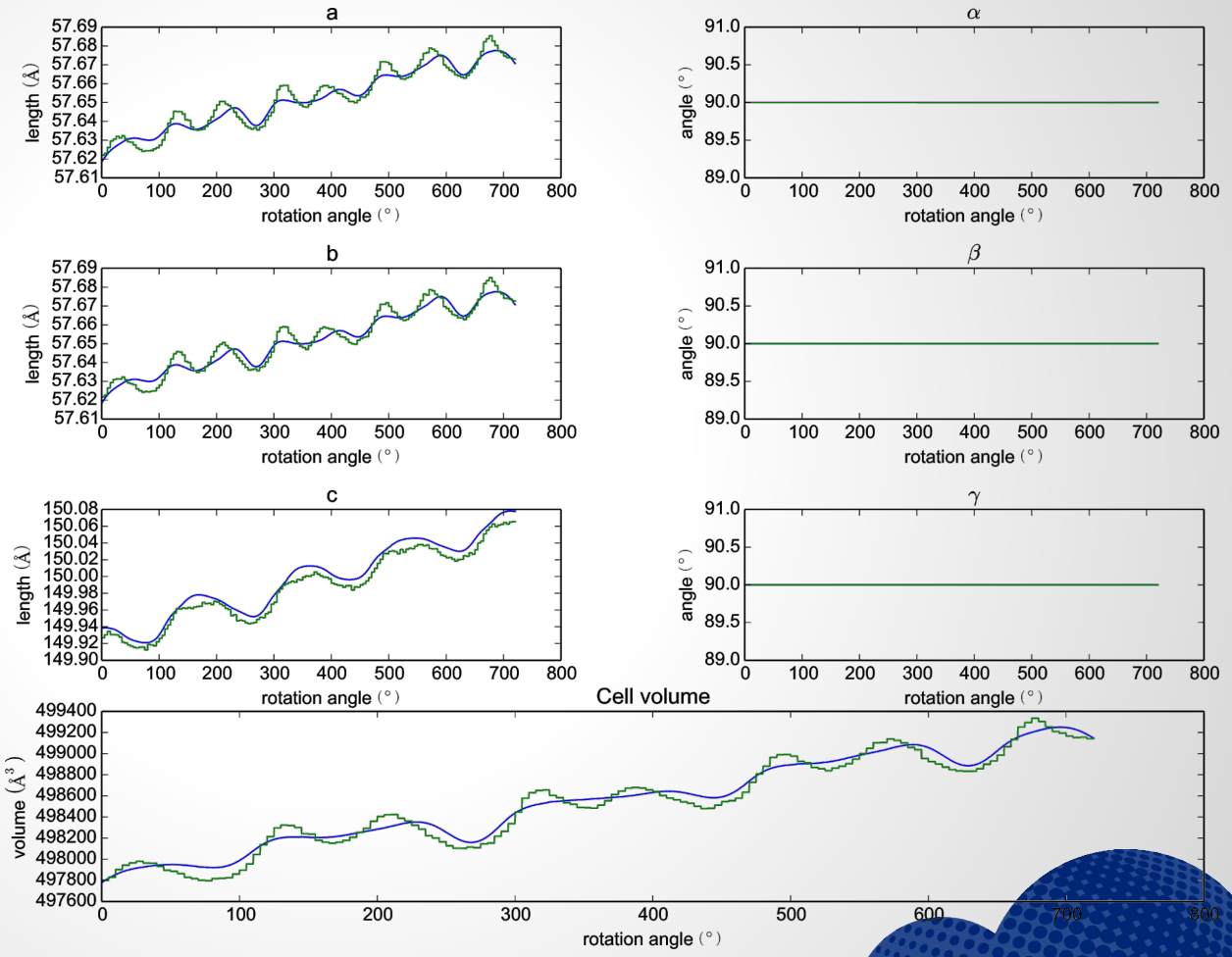




Scan-varying refinement

117 parameters:

- 6 detector
- 1 beam
- 3 crystal orientation × 22 "samples"
- 2 unit cell parameters × 22 "samples"



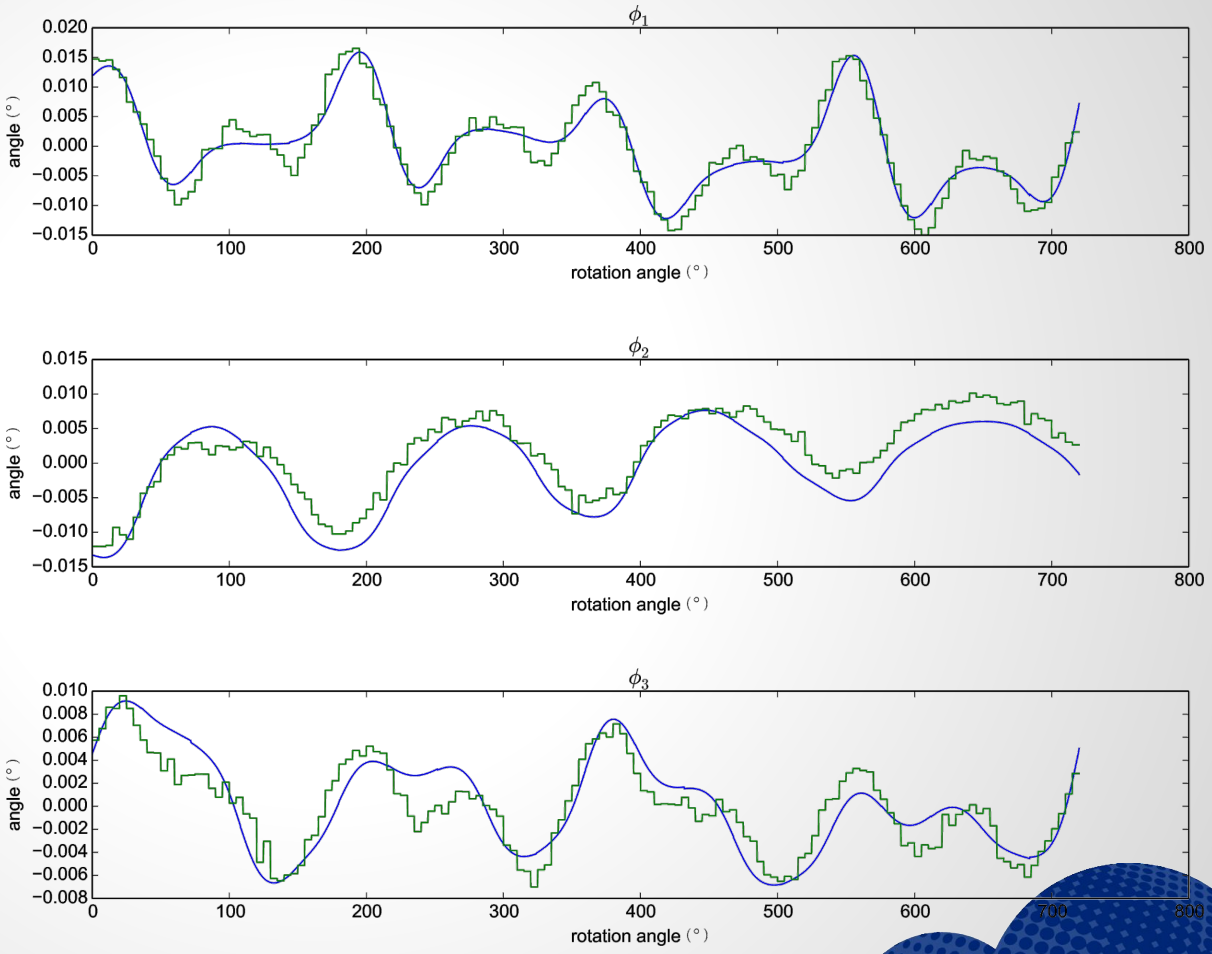
CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Scan-varying refinement

117 parameters:

- 6 detector
- 1 beam
- 3 crystal orientation × 22 "samples"
- 2 unit cell parameters × 22 "samples"

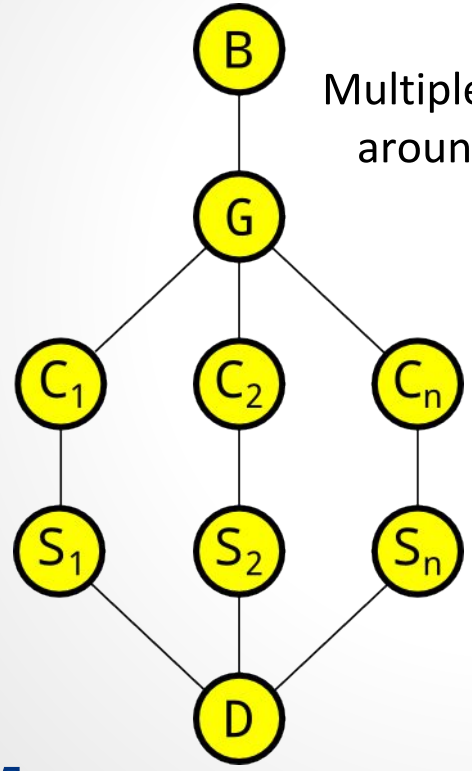


CCP4

DIALS
Diffraction Integration for Advanced Light Sources

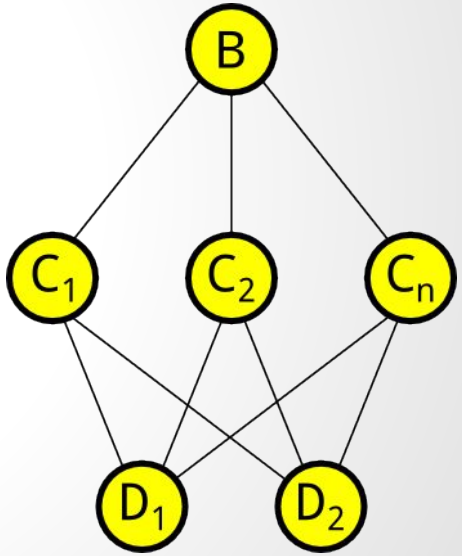
Multiple Experiments

- Global refinement across datasets that share some models
- Typical use cases involve multiple crystals



Multiple rotation scans around the same axis

XFEL shots with two detectors

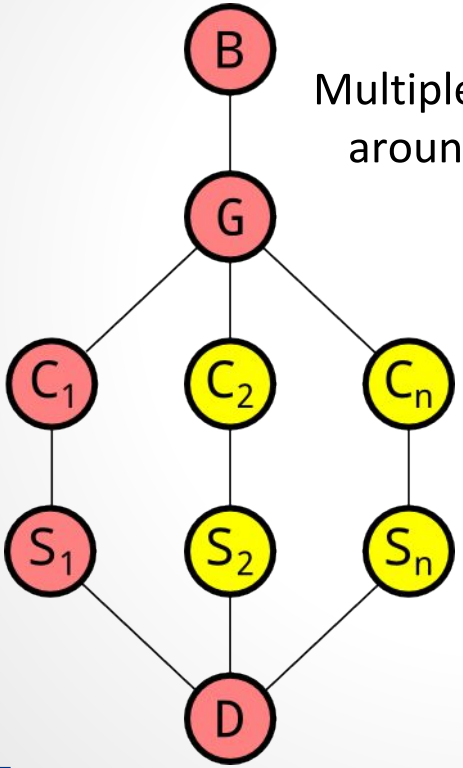


CCP4

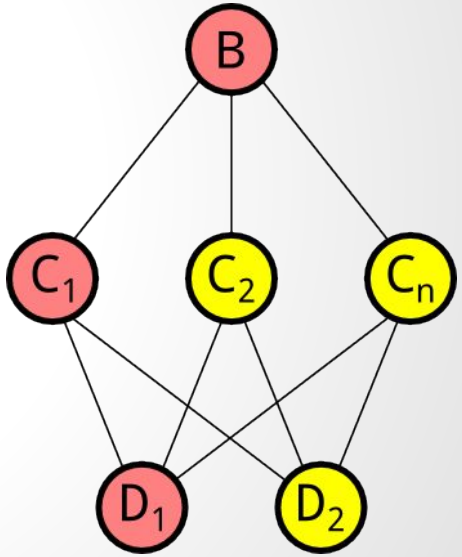
DIALS
Diffraction Integration for Advanced Light Sources

Multiple Experiments

- Global refinement across datasets that share some models
- Typical use cases involve multiple crystals



Multiple rotation scans around the same axis



XFEL shots with two detectors

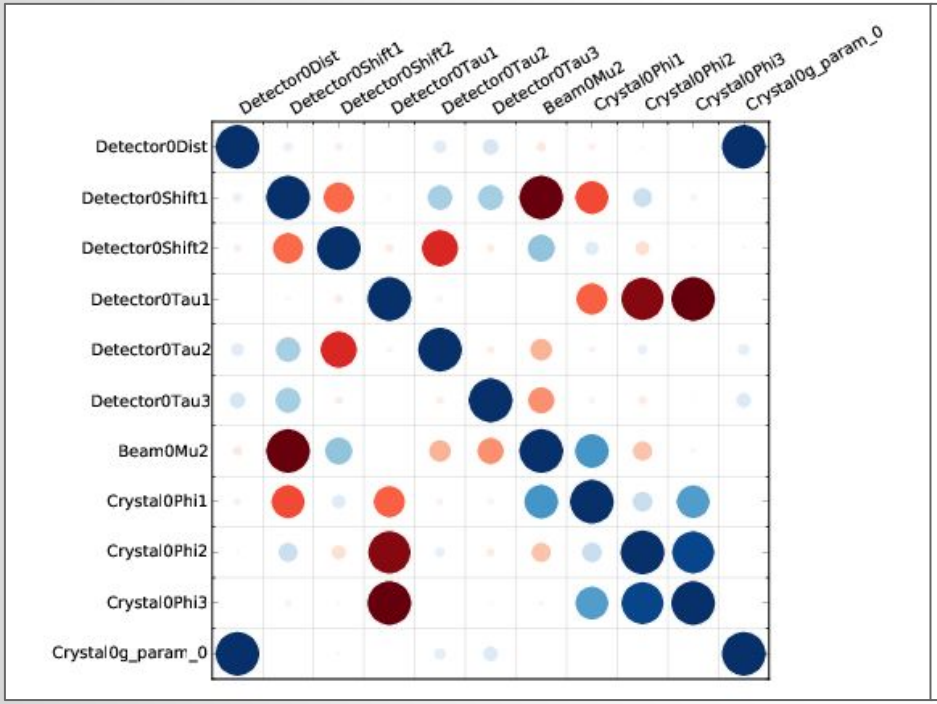
CCP4

DIALS
Diffraction Integration for Advanced Light Sources

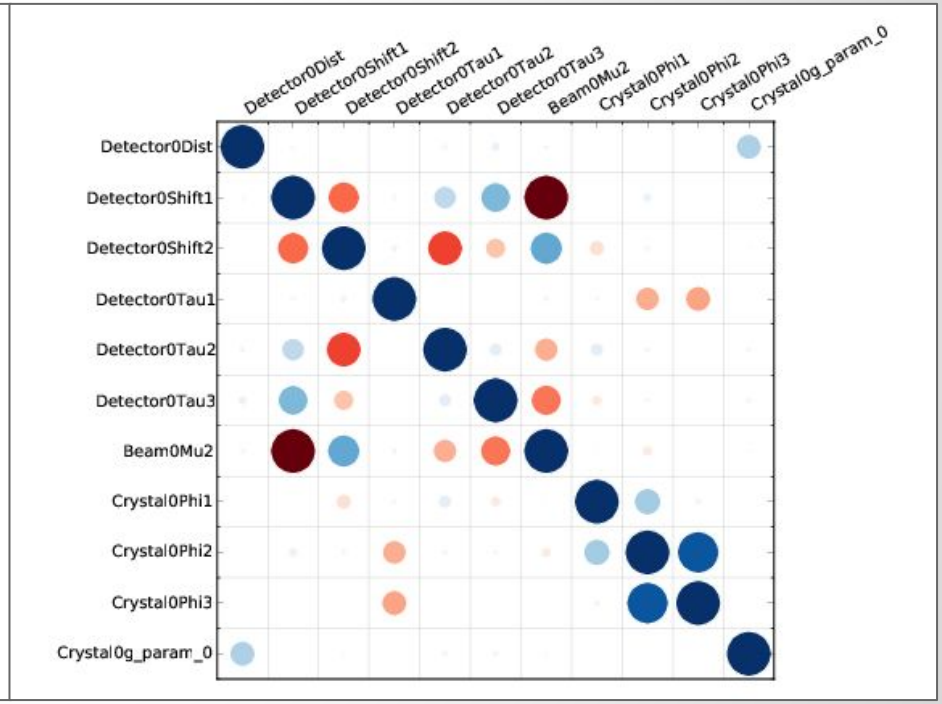
Multiple Experiments

Cubic polyhedrin crystals, 1° scans

One lattice



5 sweeps (16 lattices)

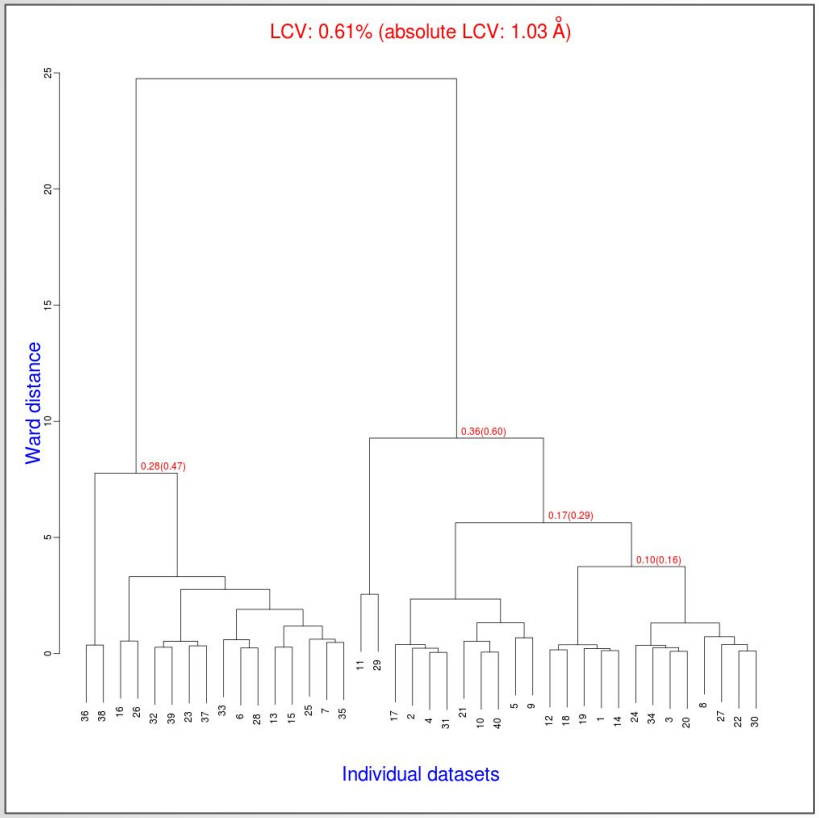


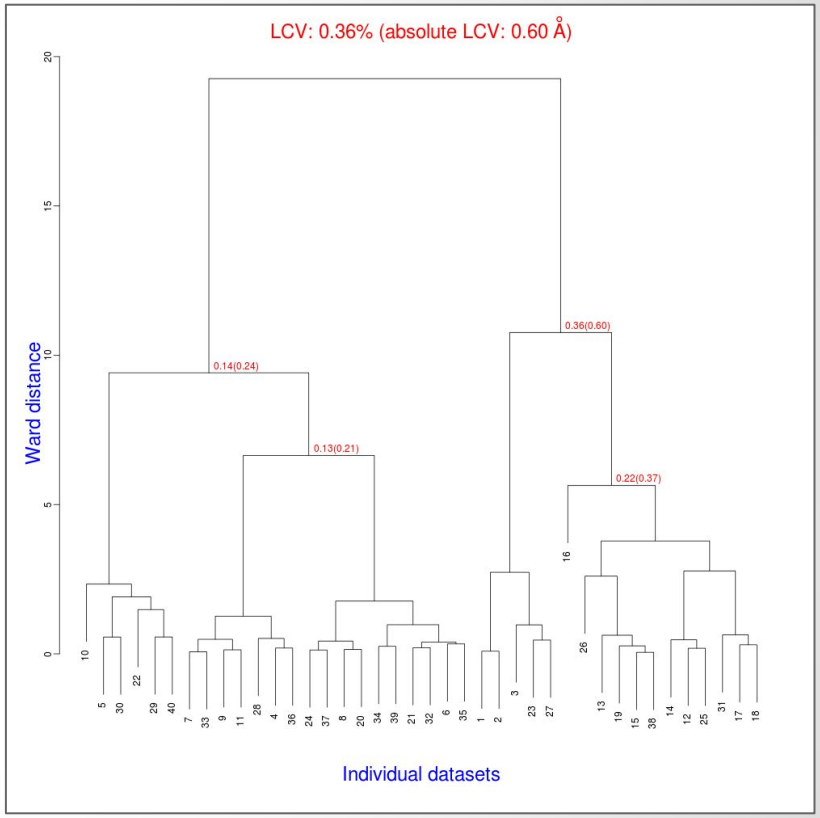
CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Multiple Experiments

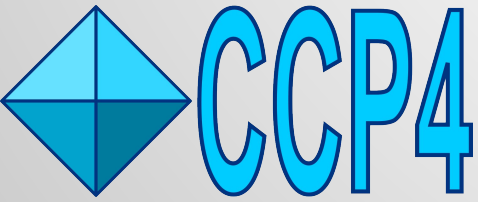
- We are investigating the use of joint refinement as a preparatory step for BLEND

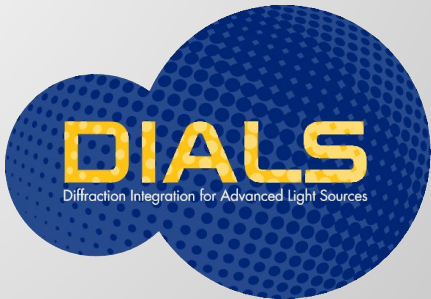




TehA membrane protein
Axford *et al. Acta Cryst.* (2015) *D***71**, 1228-1237 for original analysis

INTEGRATION





Tasks in dials.integrate

Calculate the bounding box parameters from strong reflections

Predict the positions of reflections on the images

Build reference profiles across all images

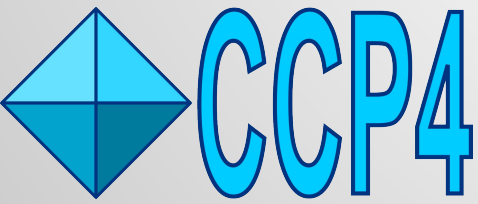
Integrate the reflections and save output

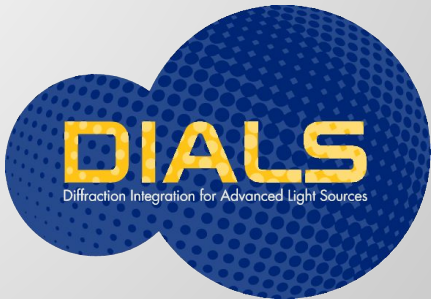
CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Computing reflection shoeboxes

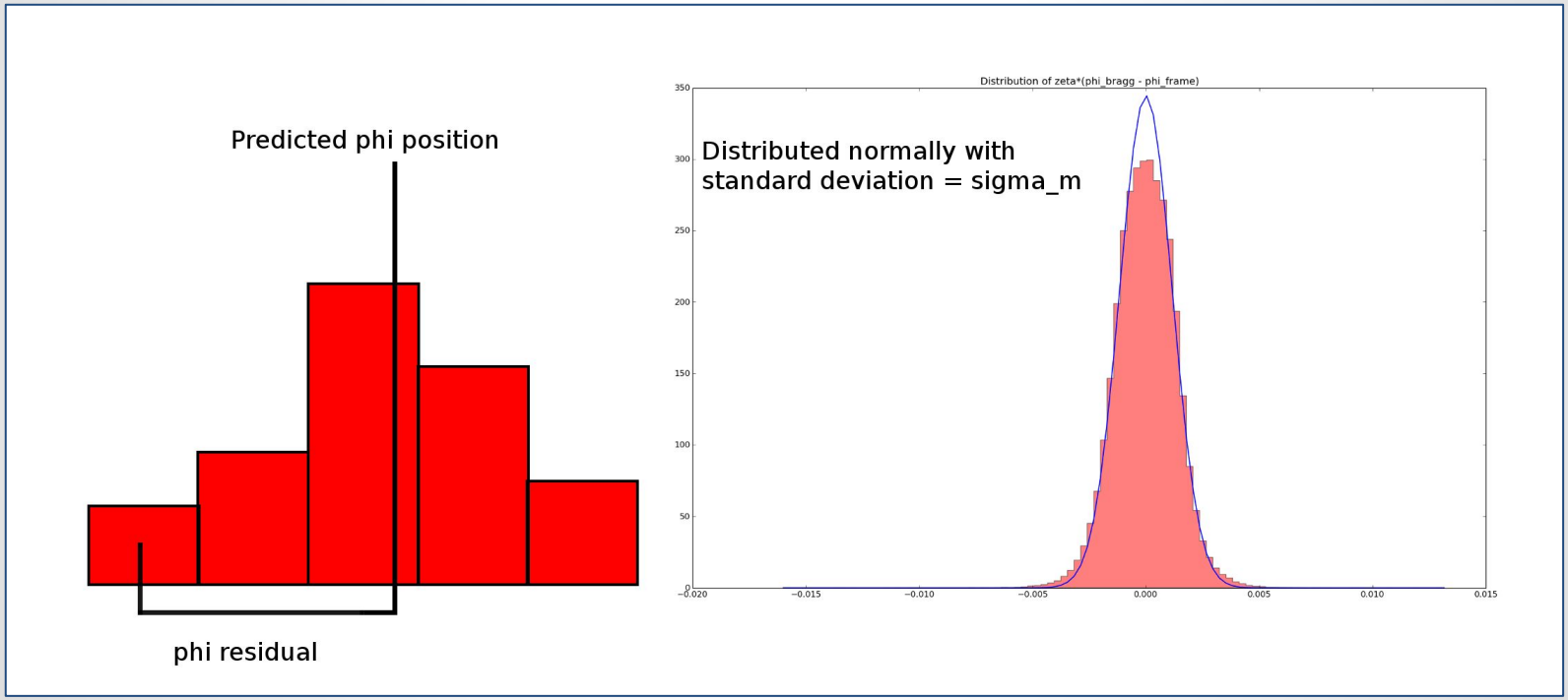
- Uses Kabsch's model of a 3D Gaussian to determine the reflection extent
- The parameter σ_D controls the size of the reflection on the surface of the Ewald sphere
- The parameter σ_M controls the size in the direction of the reflection's passage through the Ewald sphere
- The peak region is defined as within 3σ, the full extent to 6σ





Computing reflection shoeboxes

σ_M is calculated by a maximum likelihood method, assuming a normal distribution of phi residuals for each strong pixel in the spot

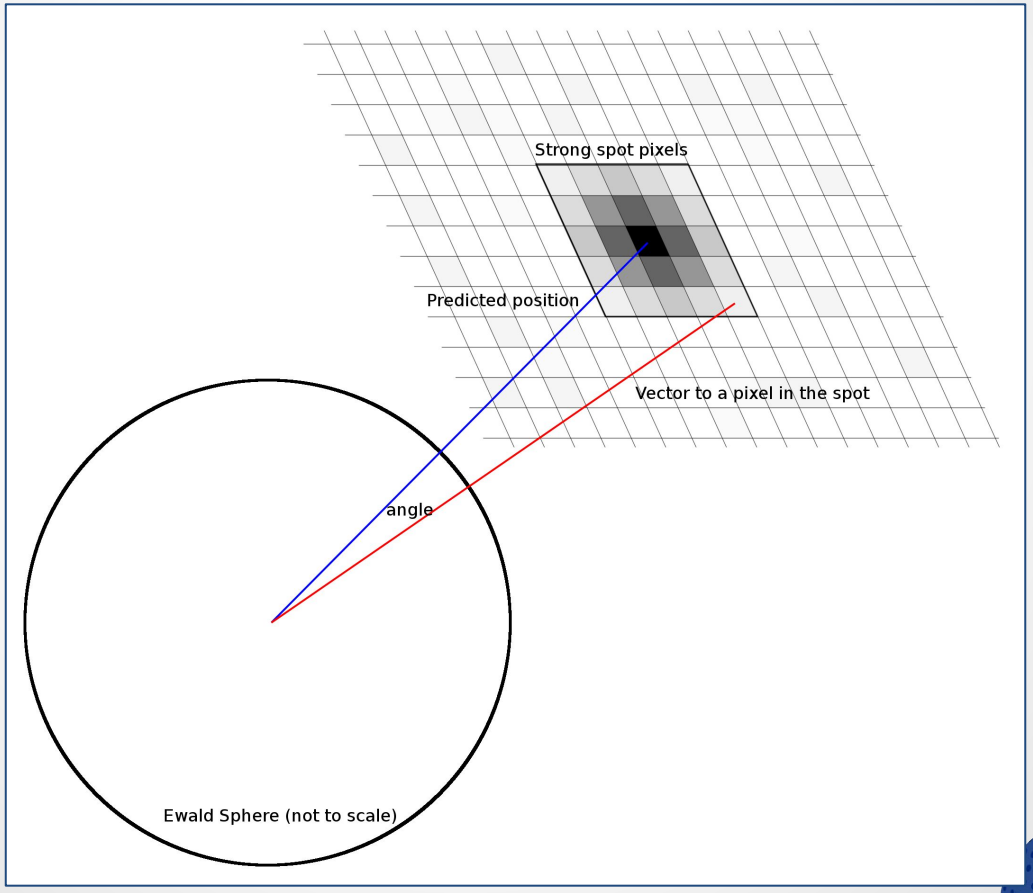


CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Computing reflection shoeboxes

σ_D is calculated from the spread of angles between the predicted centroid and each strong pixel in the spot

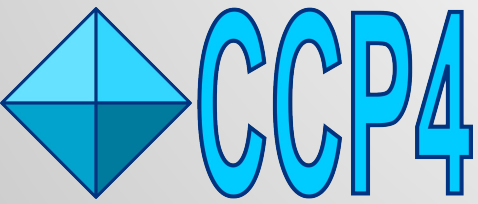


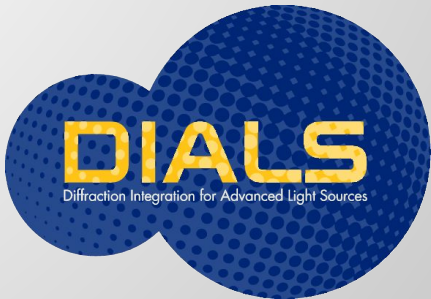
CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Background modelling

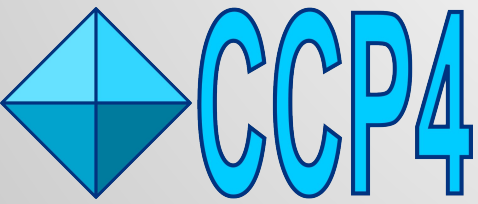
- Options to model the background under the peak as either
  - A constant across each image in the spot
  - A constant across all images in the spot - default
  - A plane across each image in the spot
  - A hyper-plane across all images in the spot
- Computed using simple linear least squares

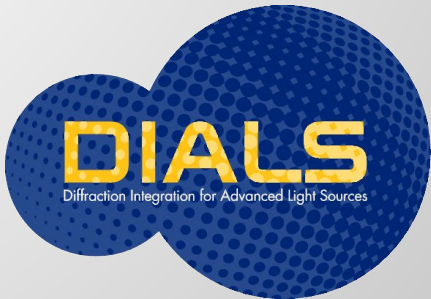




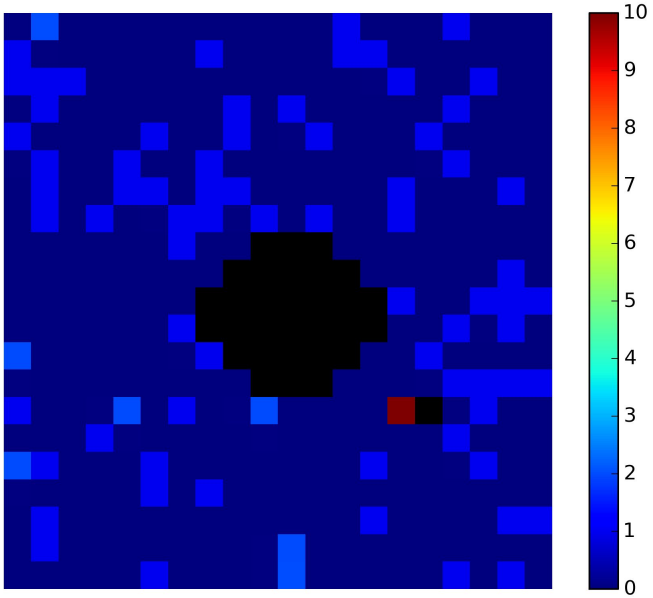
Background outliers

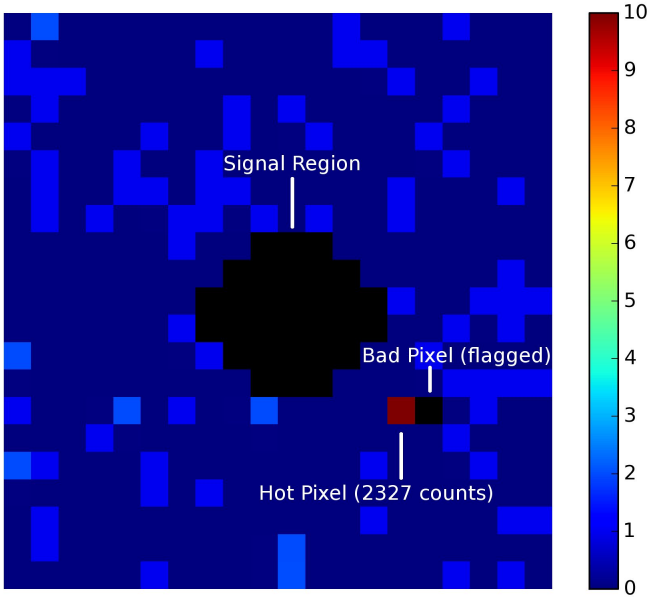
- Large-valued outliers can cause *over*-estimation of the *background*
- This then causes the reflection *intensity* to be *under*-estimated
- Outliers in the background can come from:
  - Intensity from neighbouring spots
  - Hot pixels
  - Zingers
  - Unpredicted reflections
  - Ice rings
  - …

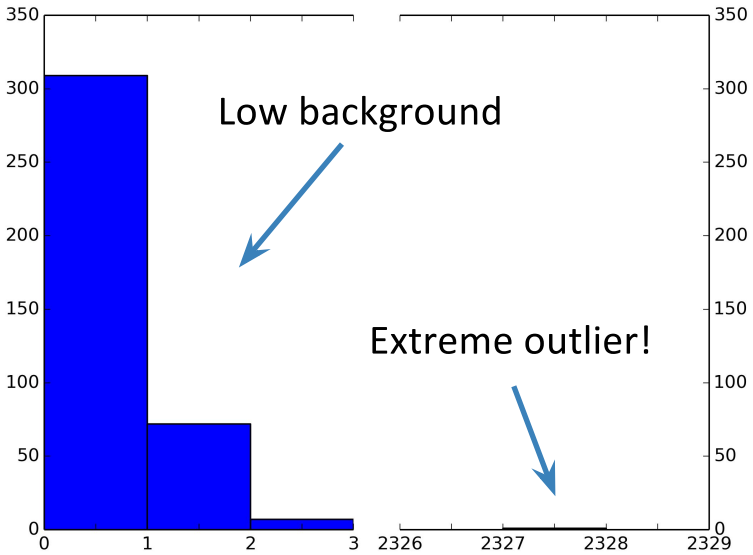




Background outliers







With Hot Pixel

Mean	6.20
Variance/Mean	2237.90

Catastrophic over-estimation!

Should be ~1 for Poisson distribution

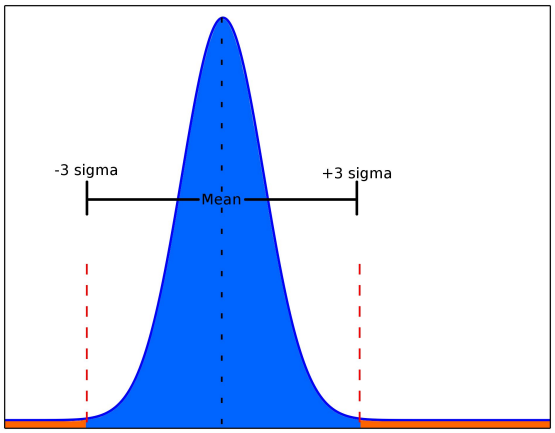
Without Hot Pixel

Mean	0.22
Variance/Mean	0.926

DIALS
Diffraction Integration for Advanced Light Sources

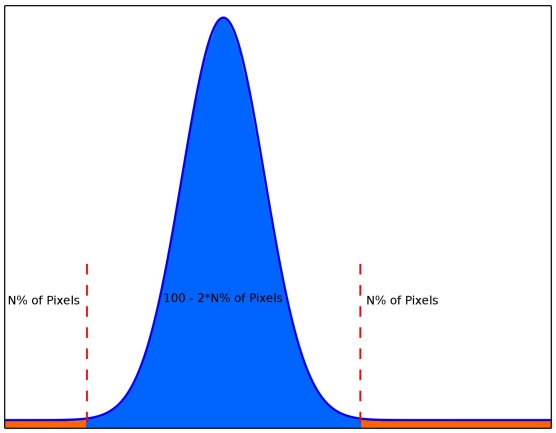
Simple outlier rejection

`outlier.algorithm=nsigma`



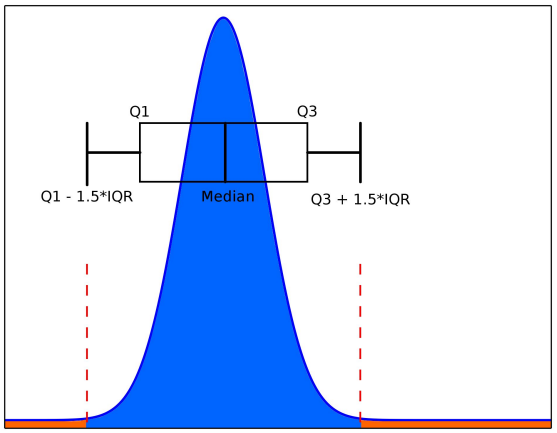
Reject pixels N sigma from the mean

`outlier.algorithm=truncated`



Reject N% of the highest and lowest valued pixels

`outlier.algorithm=tukey`



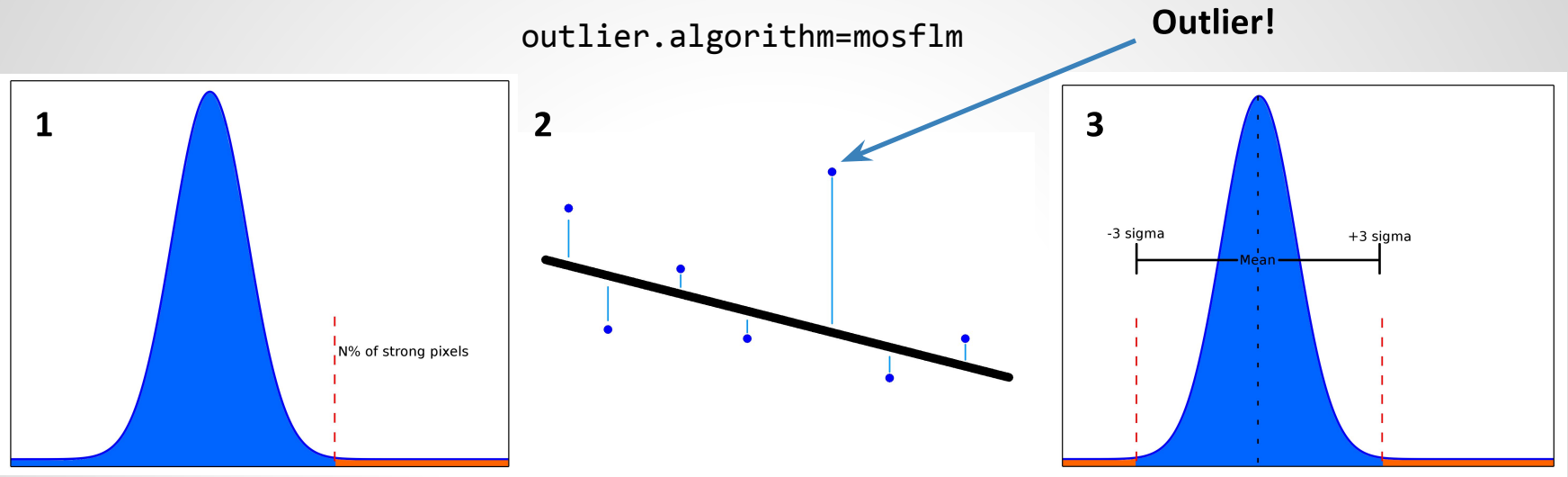
Reject pixels based on the interquartile range

CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Mosflm-like outlier rejection

`outlier.algorithm=mosflm`



Remove N% of strongest pixels and compute the background plane

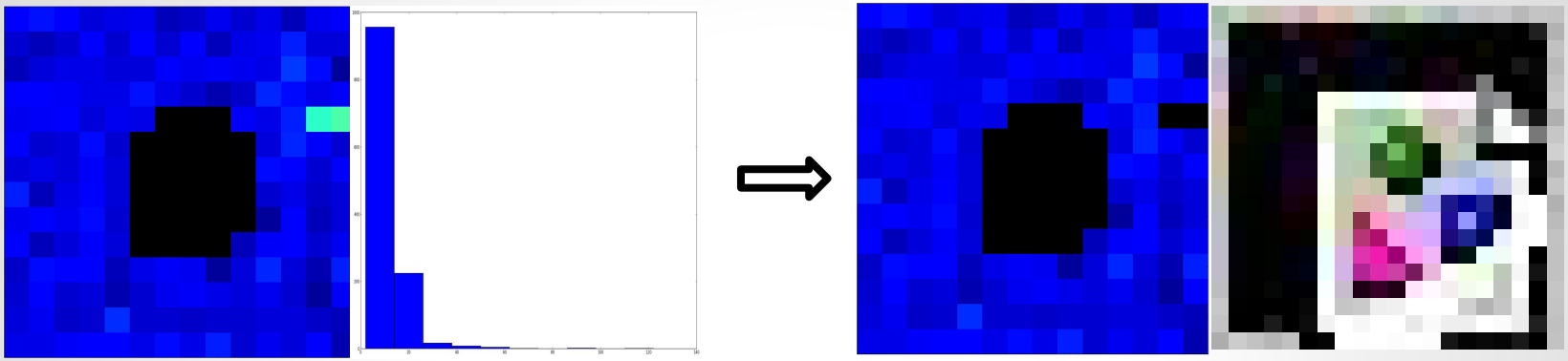
Compute the residuals of all background pixels to the plane

Remove pixels whose residuals are greater than N sigma from the plane

CCP4

DIALS
Diffraction Integration for Advanced Light Sources

XDS-like outlier rejection



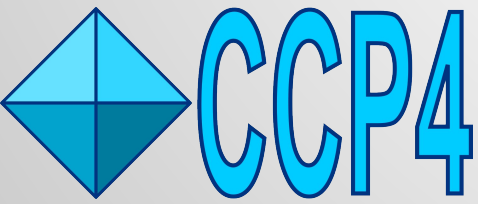
Iteratively remove high valued pixels until the distribution of pixel counts resembles a normal distribution

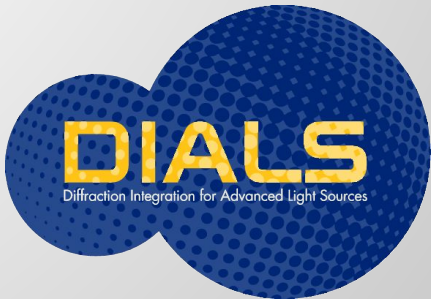
CCP4

DIALS
Diffraction Integration for Advanced Light Sources

GLM background estimation

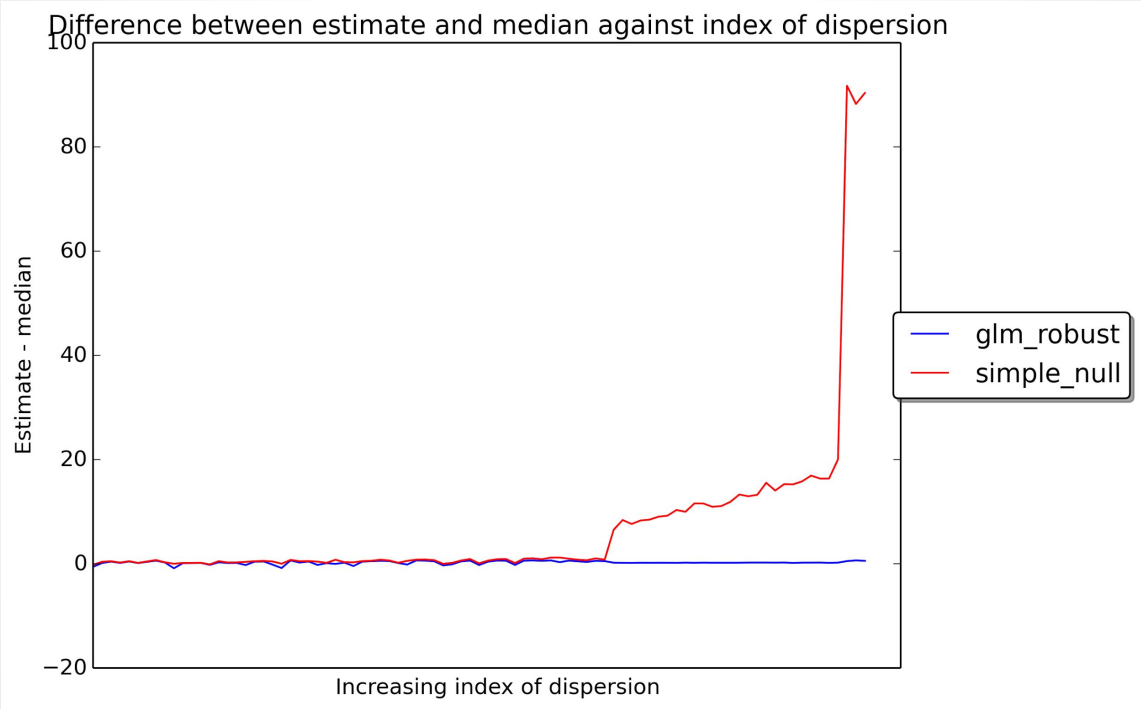
- Most of our methods assume a normal distribution of counts – *not a good approximation for data with low background*
- Instead of rejecting outliers could we use a robust estimation method?
- Most methods also focus on normally distributed data – leads to under-estimation
- Try a robust generalized linear model approach, assuming a Poisson distribution





GLM background estimation

- The more sophisticated GLM method looks promising
- This is now the default in DIALS



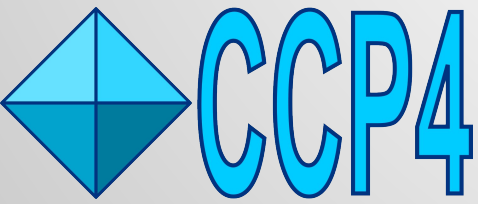
Looks like we're handling outliers ok

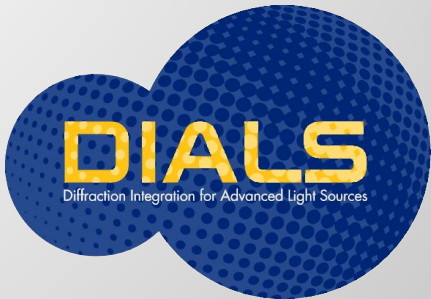
CCP4

DIALS
Diffraction Integration for Advanced Light Sources

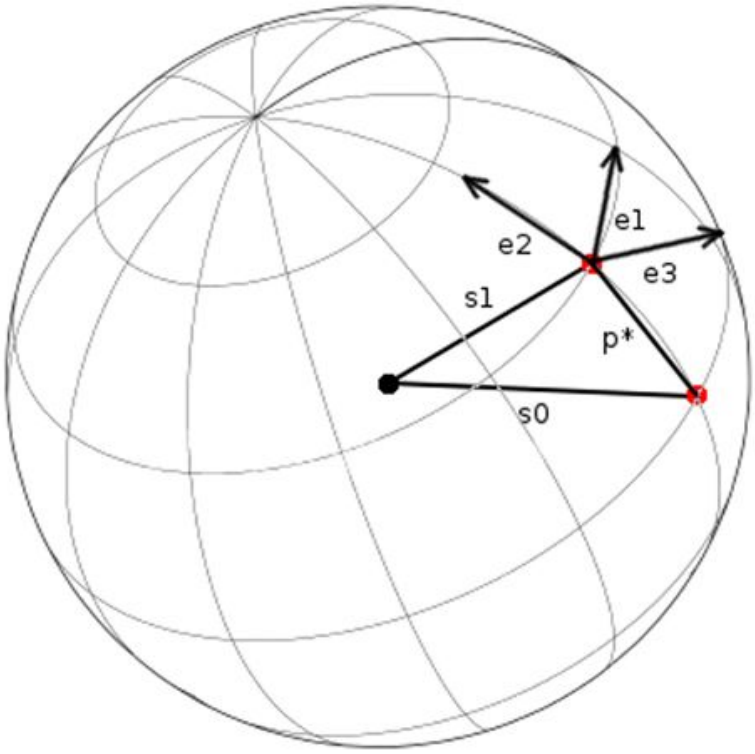
Signal integration

- Integration algorithm options:
  - Summation
  - 3D profile fitting (as in XDS)
  - 2D profile fitting (future)





3D profile fitting coordinate system



Profile coordinate system

Use Kabsch coordinate system

- Corrects for geometrical distortions
- Makes spots appear to have taken shortest path through Ewald sphere
- Model assumes a Gaussian profile in Kabsch coordinate system

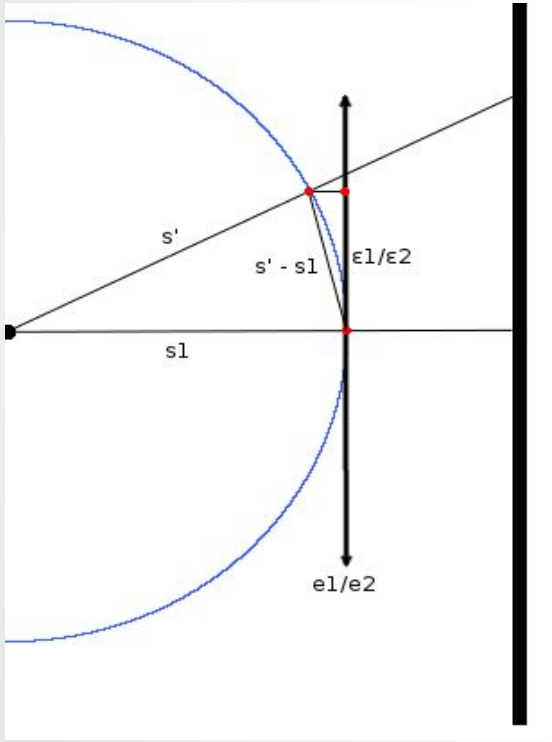
$$\boldsymbol{e}_1 = \boldsymbol{S}_1 \times \boldsymbol{S}_0 / |\boldsymbol{S}_1 \times \boldsymbol{S}_0|$$
$$\boldsymbol{e}_2 = \boldsymbol{S}_1 \times \boldsymbol{e}_1 / |\boldsymbol{S}_1 \times \boldsymbol{e}_1|$$
$$\boldsymbol{e}_3 = (\boldsymbol{S}_1 + \boldsymbol{S}_0) / |\boldsymbol{S}_1 + \boldsymbol{S}_0|$$

CCP4

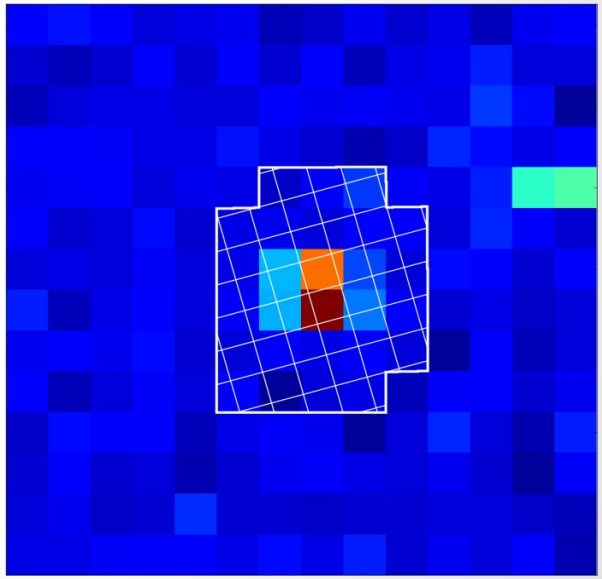
DIALS
Diffraction Integration for Advanced Light Sources

3D profile fitting
pixel gridding

Pixels are mapped back to the Ewald sphere



Counts are redistributed to Ewald sphere grid according to fractional overlap of each pixel and Ewald sphere grid point

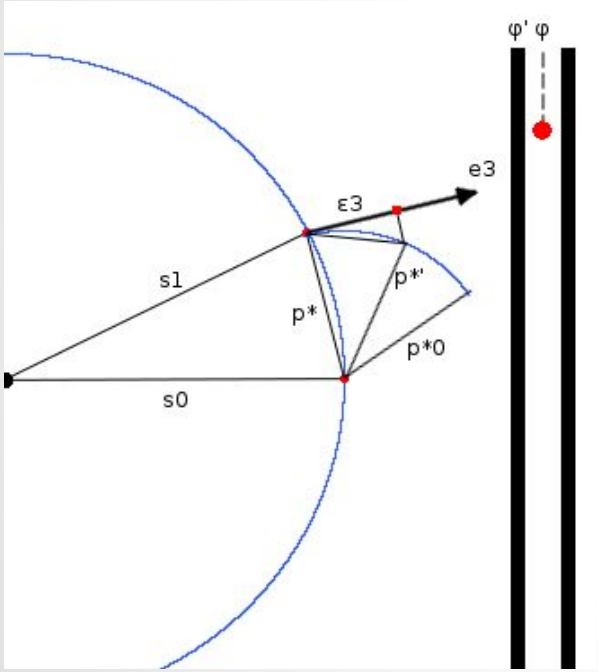


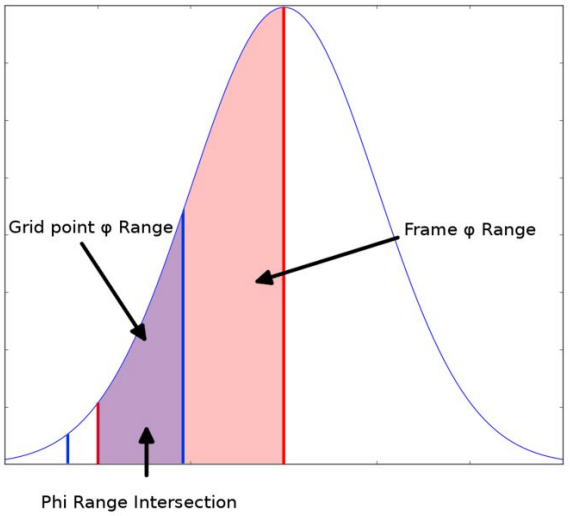
CCP4

DIALS
Diffraction Integration for Advanced Light Sources

3D profile fitting
rotation gridding

Transformation makes reflections appear to have taken shortest path through the Ewald sphere





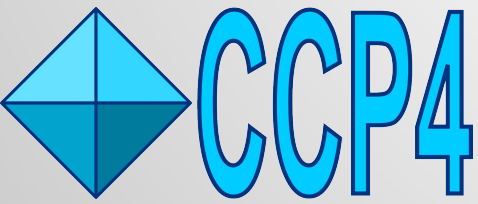
Counts are redistributed according to angular overlap between each grid point and each image

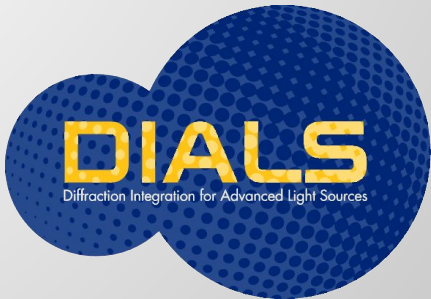
CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Building reference profiles

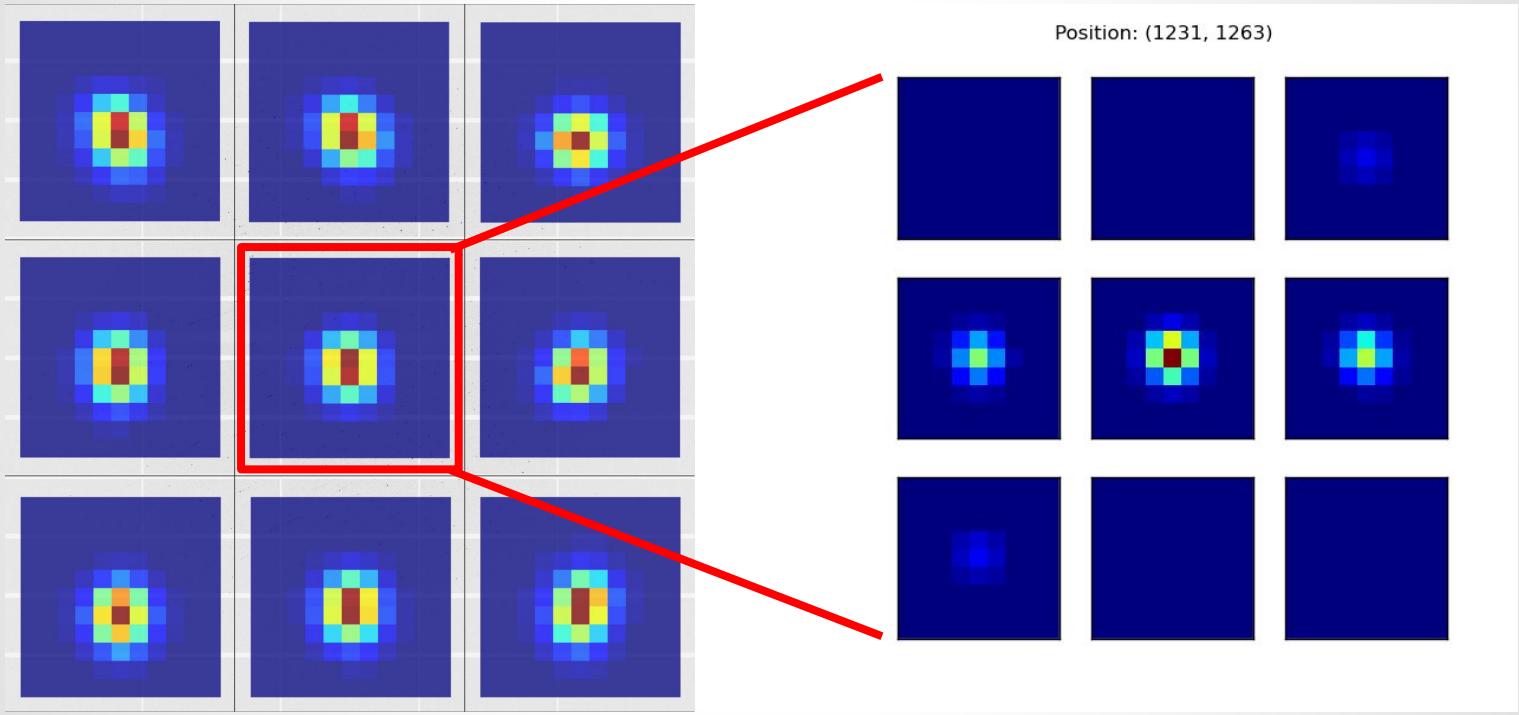
- Reference profiles are formed on a grid covering a given angular range
- Grid options include:
  - Rectangular grid (as in Mosflm) - default
  - Circular grid (as in XDS)
  - Single reflection (currently for multi-panel detectors; rare)





Building reference profiles

- Rectangular grid version:

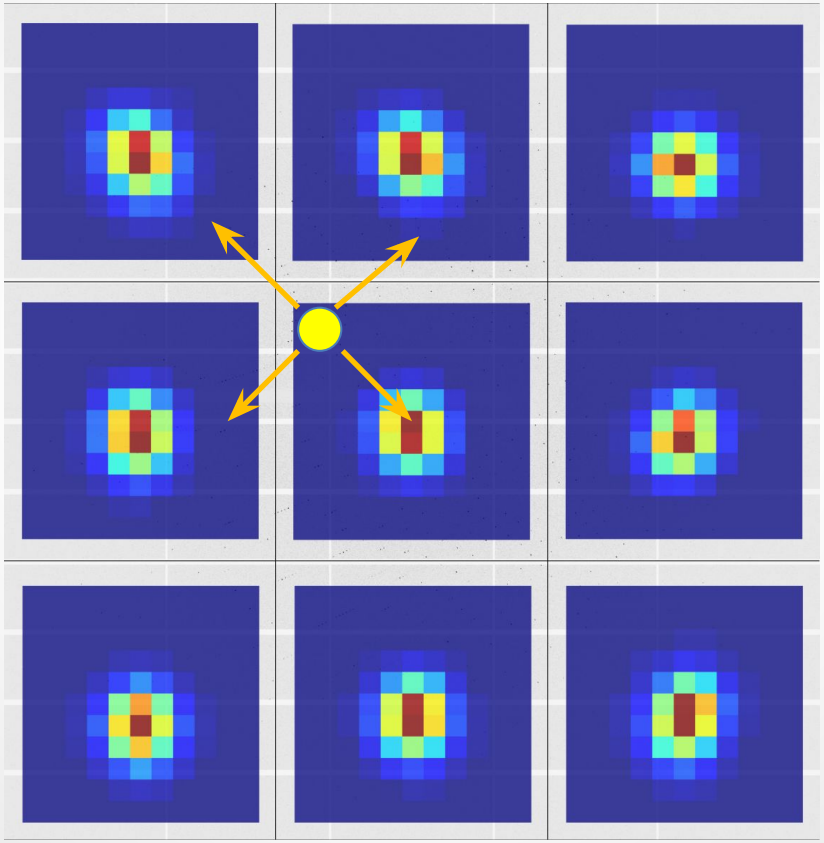


CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Building reference profiles

Each strong spot contributes to building the profile at adjacent grid points

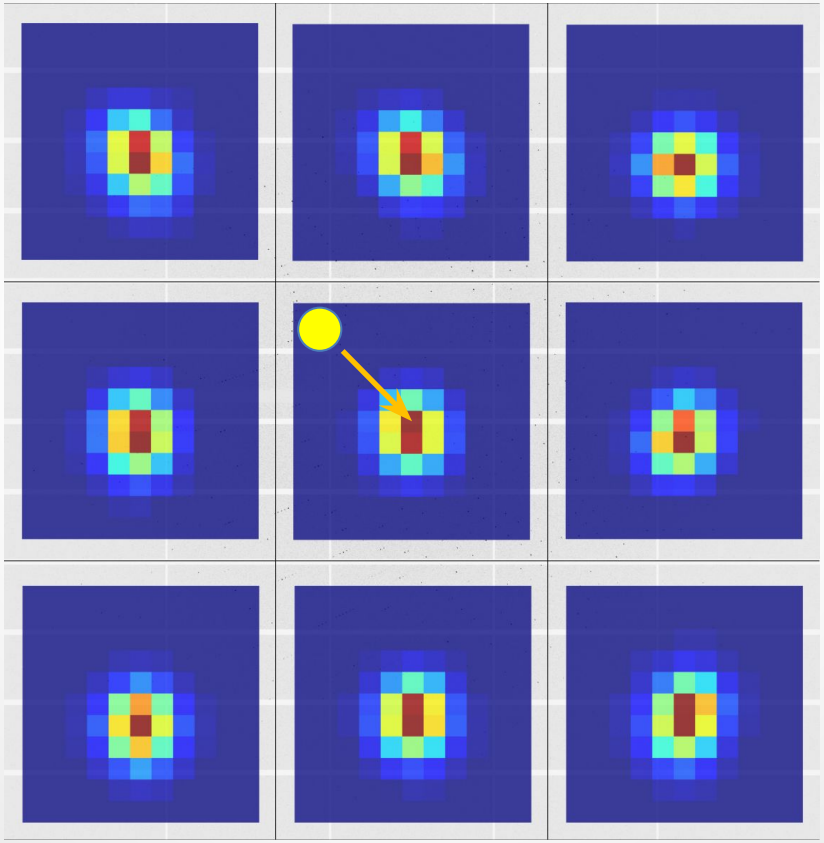


CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Fitting reference profiles

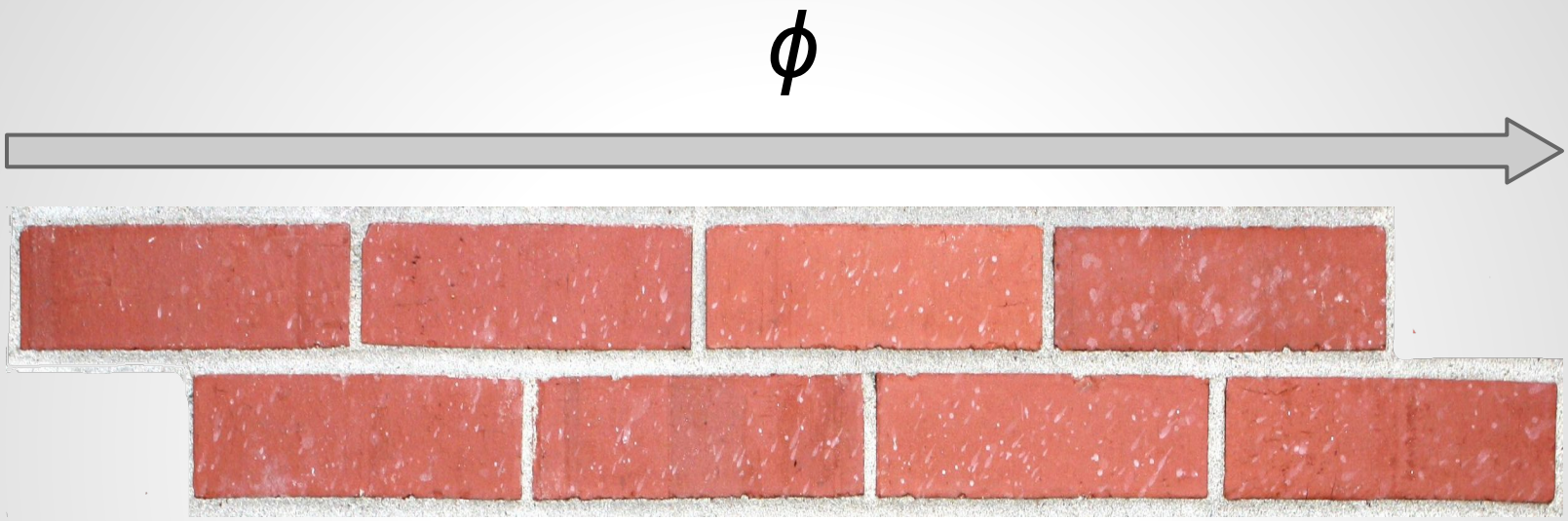
Each reflection is fitted against its closest reference profile



CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Fitting reference profiles



Profile for reflection at position x derived from average of strong reflections in block with centre nearest x

CCP4

DIALS
Diffraction Integration for Advanced Light Sources

Tutorials

http://dials.github.io/documentation/tutorials/

DIALS
Diffraction Integration for Advanced Light Sources

Processing in Detail

Introduction

DIALS processing may be performed by either running the individual tools (spot finding, indexing, refinement, integration, exporting to MTZ) or you can run `xia2 -dials`, which makes informed choices for you at each stage. In this tutorial we will run through each of the steps in turn, checking the output as we go. We will also enforce the correct lattice symmetry.

Tutorial data

The following example uses a Thaumatin dataset collected using beamline I04 at Diamond Light Source which is available for download from DOI 10.5281/zenodo.10271.

Import

The first stage of step-by-step DIALS processing is to import the data - all that happens here is that the image headers are read, and a file describing their contents (*datablock.json*) is written. It's worth noting that if this file is changed subsequent processing can use this.

```
dials.import data/th_8_2_0*cbf
```

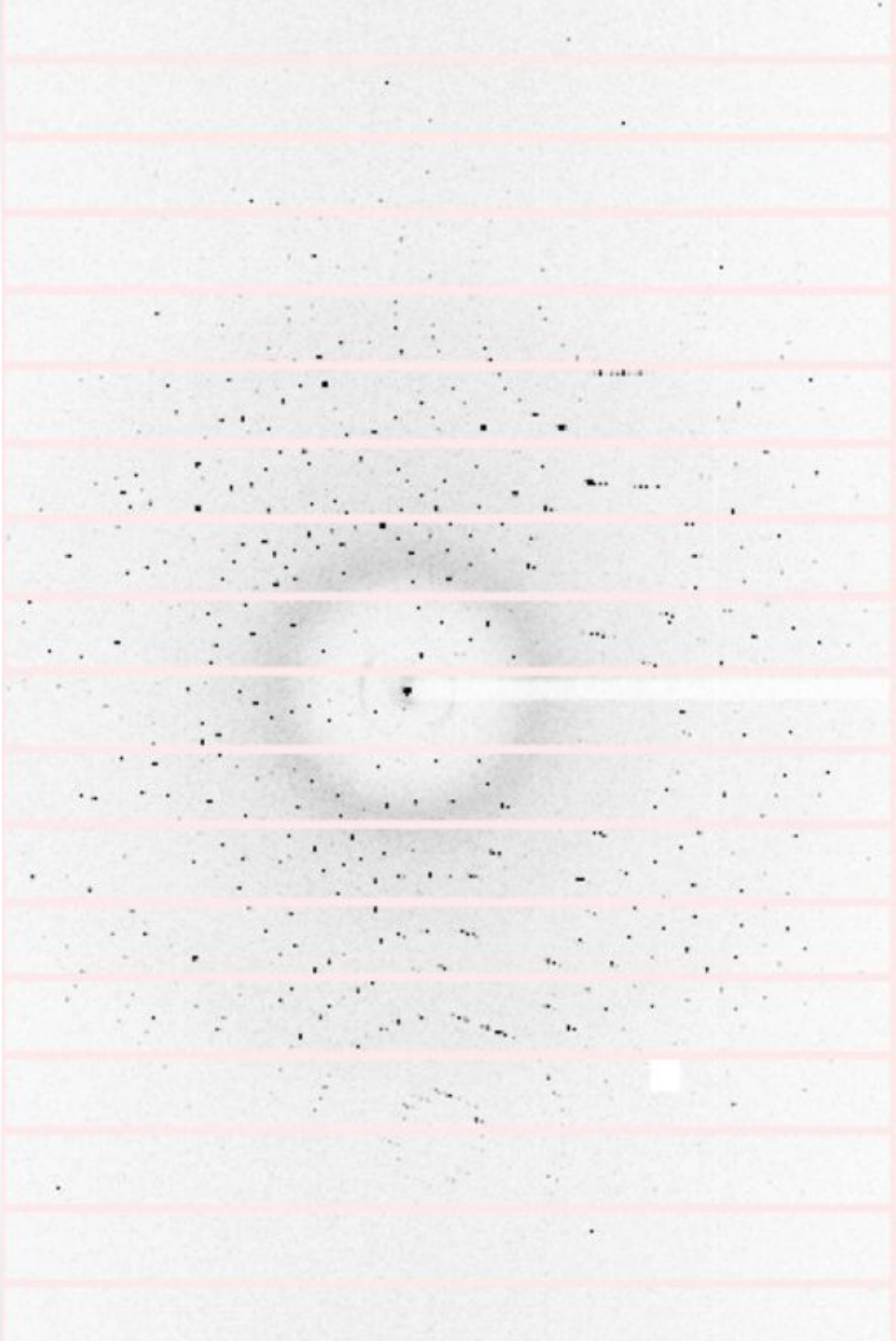
The output just describes what the software understands of the images it was passed, in this case one sweep of data containing 540 images.

Non-planar detectors

- Long wavelength beamline I23 @ Diamond Light Source
- Custom P12M
- Excellent data - experiment in vacuum
- Support simple via dxtbx

Data

- Very low background (in vacuum)
- Very sharp spots despite lack of beam focussing at the time
- Data collection at 9 keV
- Two theta to around 75º hence design of detector
- Currently single axis goniometer

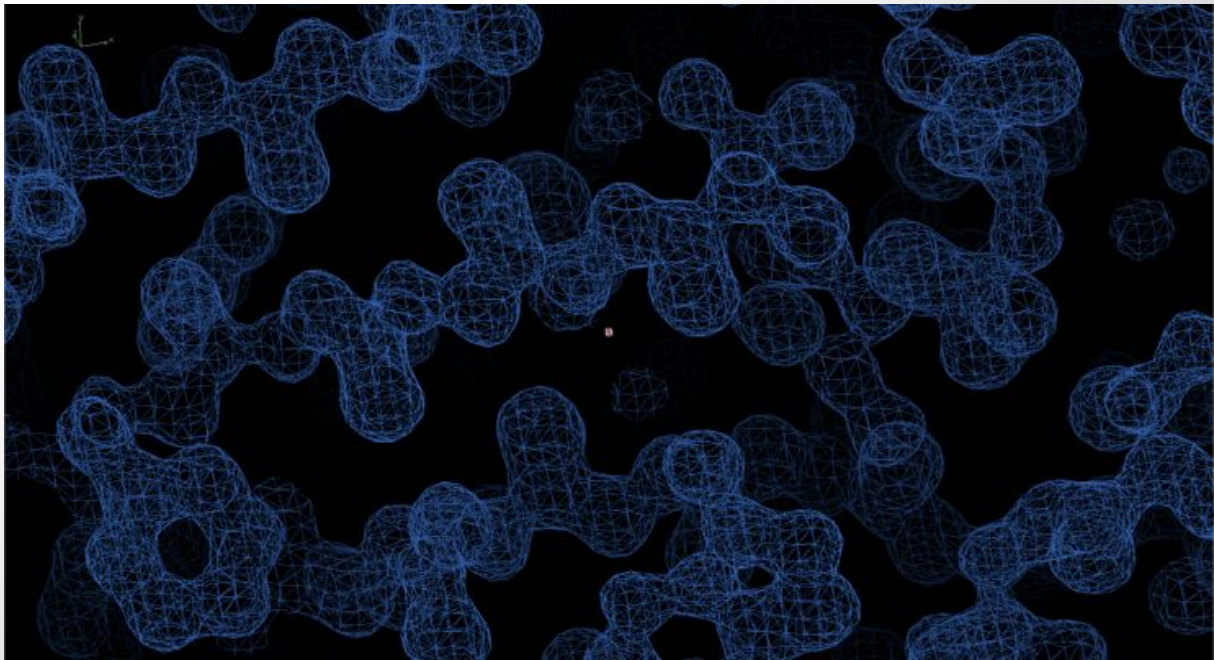


Data at low resolution

Results

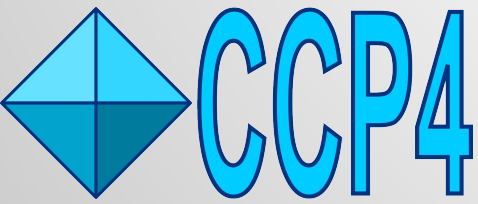
Processed using “standard” script
Scaled with AIMLESS again using standard commands
Phased with shelxc/d/e - gives excellent maps

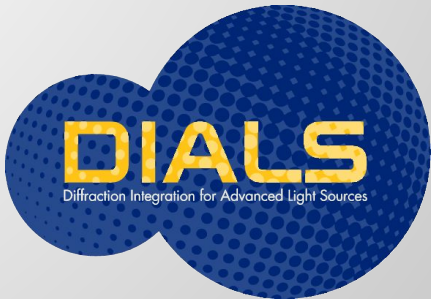
	Overall	InnerShell	OuterShell
Low resolution limit	28.93	28.93	1.12
High resolution limit	1.10	6.02	1.10
Rmerge (within I+/I-)	0.047	0.022	0.715
Rmerge (all I+ and I-)	0.050	0.025	0.766
Rmeas (within I+/I-)	0.052	0.024	0.820
Rmeas (all I+ & I-)	0.053	0.027	0.821
Rpim (within I+/I-)	0.022	0.009	0.392
Rpim (all I+ & I-)	0.016	0.008	0.288
Rmerge in top intensity bin	0.023	-	-
Total number of observations	910205	7848	30945
Total number unique	93738	781	4010
Mean((I)/sd(I))	15.3	37.0	2.5
Mn(I) half-set correlation CC(1/2)	1.000	1.000	0.634
Completeness	90.2	99.3	79.0
Multiplicity	9.7	10.0	7.7
Anomalous completeness	88.6	100.0	77.6
Anomalous multiplicity	4.9	6.2	3.9
DelAnom correlation between half-sets	0.109	0.467	0.024
Mid-Slope of Anom Normal Probability	0.917	-	-



CCP4

Not just data processing

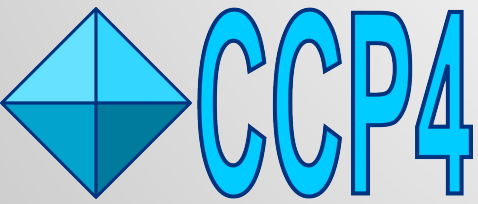


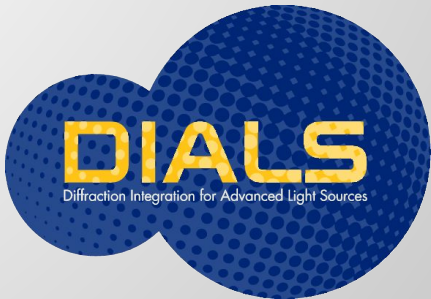


DIALS TOOLBOX

FULL TOOLBOX

```
dials.analyse_output
dials.apply_mask
dials.background_lookup
dials.check_indexing_symmetry
dials.combine_experiments
dials.compare_mosflm_dials
dials.compare_mosflm_xds
dials.compare_orientation_matrices
dials.compare_reflections
dials.compare_xds_dials
dials.compare_xds_dials2
dials.create_profile_model
dials.detector_max_resolution
dials.discover_better_experimental_model
dials.display_reference_profiles
dials.estimate_gain
dials.estimate_resolution_limit
dials.export
dials.export_mosflm
dials.export_mtz
dials.export_nxmx
dials.export_nxmx_to_mtz
dials.export_spot_xds
dials.export_text
dials.export_xds
dials.extract_shoeboxes
dials.filter_good_intensities
dials.filter_reflections
dials.find_hot_pixels
dials.find_overlaps
dials.find_spots
dials.find_spots_client
dials.find_spots_server
dials.generate_mask
dials.generate_process_test_reflections
dials.generate_test_reflections
dials.image_viewer
dials.import
dials.import_xds
dials.index
dials.integrate
dials.merge_cbf
dials.merge_reflection_lists
dials.plot_reflections
dials.plot_scan_varying_crystal
dials.predict
dials.print_test_reflections
dials.python
dials.reciprocal_lattice_viewer
dials.refine
dials.refine_bravais_settings
dials.reflection_viewer
dials.reindex
dials.remove_absent_reflections
dials.rs_mapper
dials.scale_down_image
dials.scale_down_images
dials.show_build_path
dials.show_dist_paths
dials.show_extensions
dials.show_indexed_strong
dials.show_isig_rmsd
dials.show_models
dials.show_profiles
dials.show_spots
dials.simulate
dials.simulate_rs
dials.slice_sweep
dials.sort_reflections
dials.split_experiments
dials.spot_counts_per_image
dials.stereographic_projection
dials.stills_detector_hybrid_refine
dials.version
```

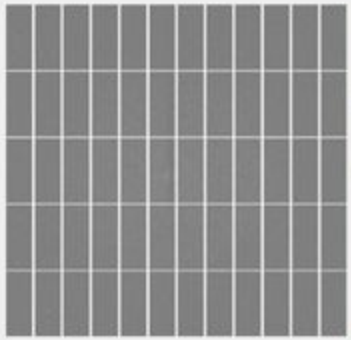


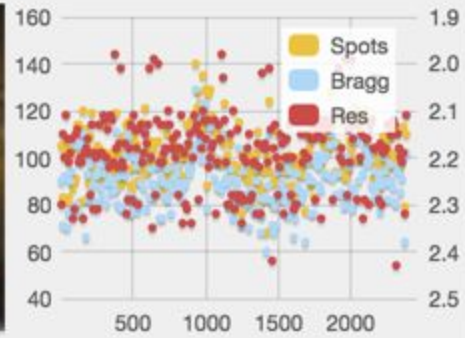


User feedback: per-image analysis

15-05-2015 15:50:35 - GlIso/GlIso_M6S7_1_####.cbf

Sample: GI_7	Ω Start: 0.0°
Ω Osc: 0.15°	Ω Overlap: 0°
No. Images: 2400	Resolution: 2.27Å
Wavelength: 1.8903Å	Exposure: 0.040s
Transmission: 0.50%	Beamsize: 11x5µm
Type: Data Collection	
Comment: (900,189,-14) EDNAStrategy3: subWedge:1Aperture: Large	





Auto Processing — Fast DP: ✓ Xia2: ✓ ✗ ✗

Fast DP | DIALS

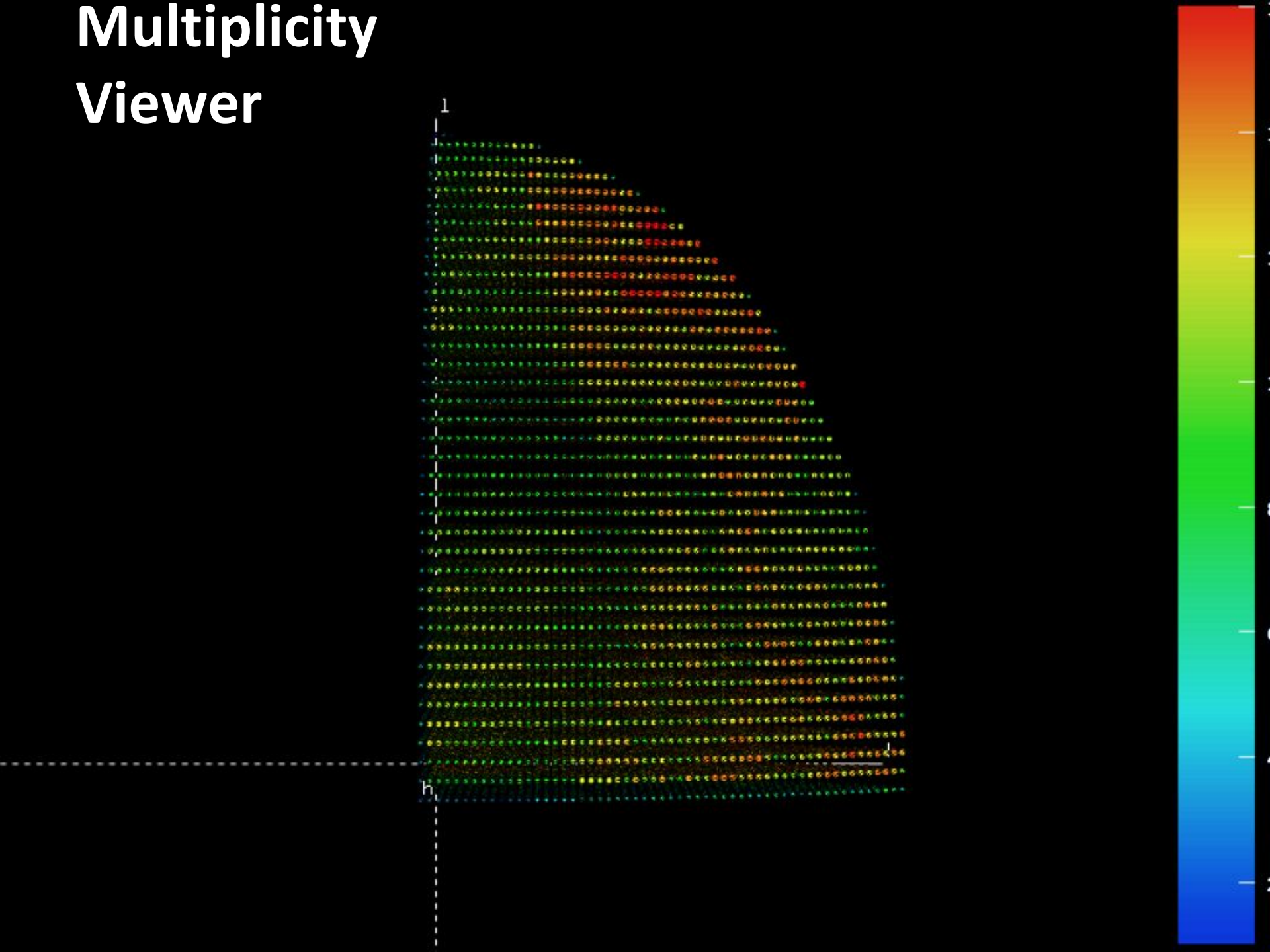
Space Group	A	B	C	α	β	γ
I 2 2 2	93.04	98.22	102.79	90.00	90.00	90.00

MTZ file | Log file | Lookup Cell

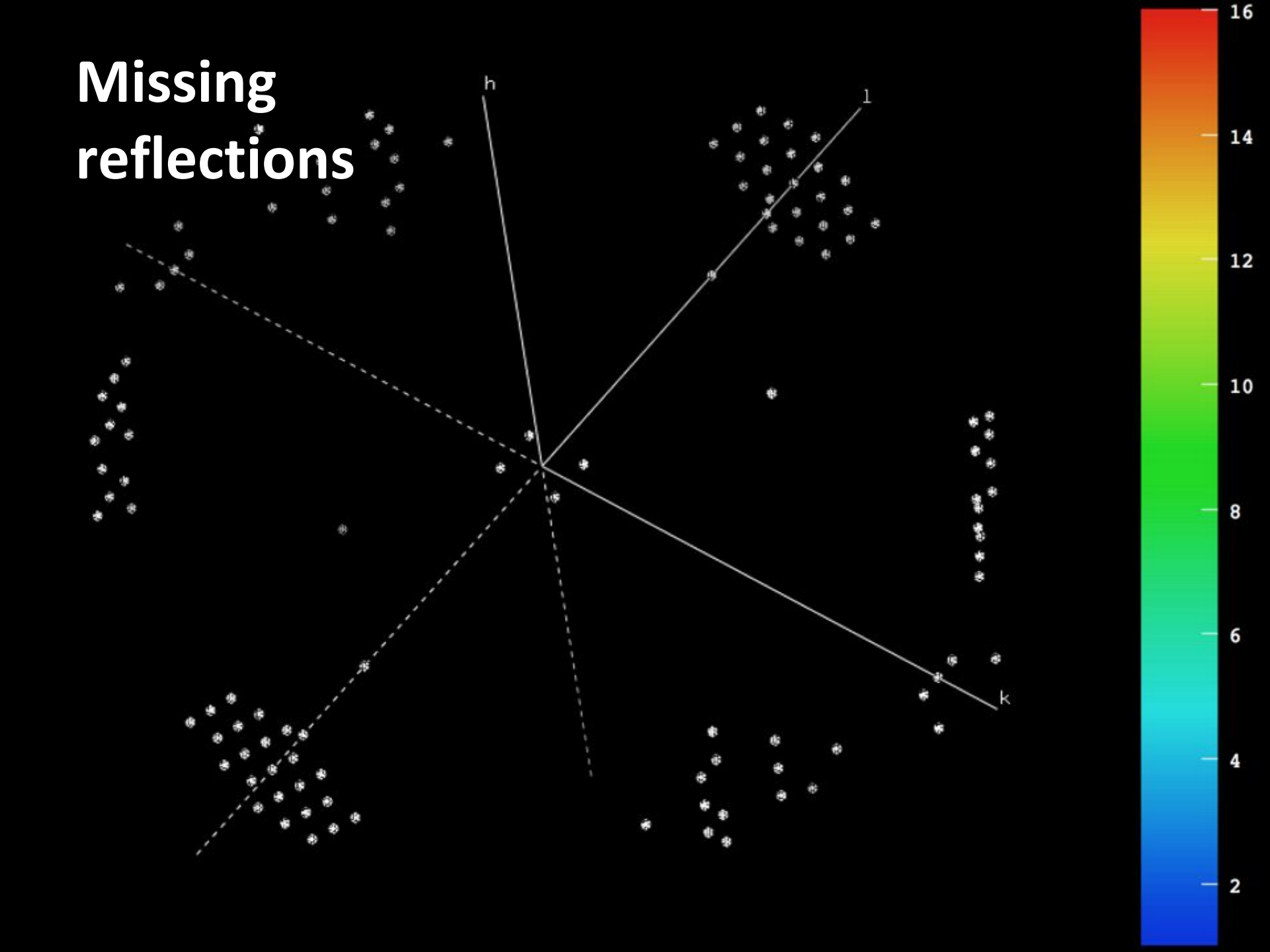
Shell	Observations	Unique	Resolution	Rmeas	I/sig(I)	Completeness	Multiplicity	Anom Completeness	Anom Multiplicity
outerShell	947	519	1.90 - 1.95	0.571	2.0	18.5	1.8	4.3	1.5
innerShell	6208	521	8.29 - 71.01	0.082	17.8	99.9	11.9	100.0	7.2
overall	341484	30333	1.90 - 71.01	0.148	11.8	81.4	11.3	78.7	5.8

Downstream Processing — Fast EP: ✓ Dimple: ✓

Multiplicity Viewer

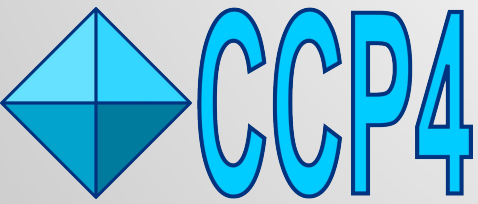


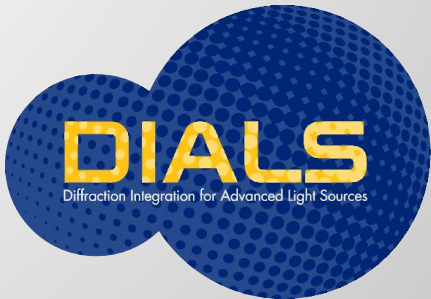
Missing reflections



Summary

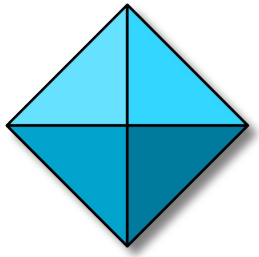
- DIALS now used routinely for automated data processing with xia2
- xia2 is the “friendly” DIALS user interface for synchrotron data (and is bundled with DIALS)
- Software available from http://dials.github.io/ under BSD license
- Binary builds available for Mac and Linux from http://dials.github.io/installation.html
- DIALS 1.1 now in CCP4 7.0 (including Windows)





diamond

Research Complex
at Harwell



CCP4

Science & Technology
Facilities Council

BioStructX