DIALS and BLEND

James Parkhurst

NJUST / CCP4 crystallography school

September 2015

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

The usage and control of recent modifications
package *XDS* for the processing of rotati
described in the context of previous versions
include automatic determination of spot size
range and recognition and assignment of cry
Moreover, the limitations of earlier package
number of correction/scaling factors and the re
pixel contents have been removed. Large prog
been restructured for parallel processing so t
and completeness of collected data can be asse
measurement.

1. Functional specification

The program package *XDS* (Kabsch, 1988*a*,*b*,
developed for the reduction of single-crystal
recorded on a planar detector by the rotatio
monochromatic X-rays. It includes a set of thr

XDS accepts a sequence of adjacent n
rotation images from a variety of imaging-pla
and multiwire area detectors, infers crystal
metrics and produces a list of corrected integr
of the reflections occurring in the images in a n
way. The program assumes that each image c
positive amount of crystal rotation and that th
incident beam and crystal intersect at one poin
imposes no limitations on the detector pc
directions of the rotation axis and incident b
oscillation range covered by each image.

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

X-ray diffraction images from two-dimen
sensitive detectors can be characterized as
depending on whether the rotation-angle incr
is greater than or less than the crystal mosaic
The expectations and consequences of the pr
and thin images in terms of spatial overlap,
X-ray background and $I/\sigma(I)$ are discussed
software suite for processing diffraction i
introduced, and results from *d*TREK* are
those from another popular package.

1. Introduction

Two-dimensional position-sensitive detectors
for many years in X-ray diffraction data coll
cular, data from crystals of macromolecules
oligonucleotides and their complexes are
acquired with an area detector such as film
obsolete), a multi-wire system, an imagin
recently commercialized charge-coupled
coupled to a phosphor-coated fiber-optic tape
detectors, the crystal, centered in the X-ray be
oscillated around a single axis through a sma
~2.0°, while counts from diffracted photons
for a specified time. At the end of the sm
detector is read out and the counts are store
two-dimensional array with each array elemer
to a distinct position on the detector and

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*.

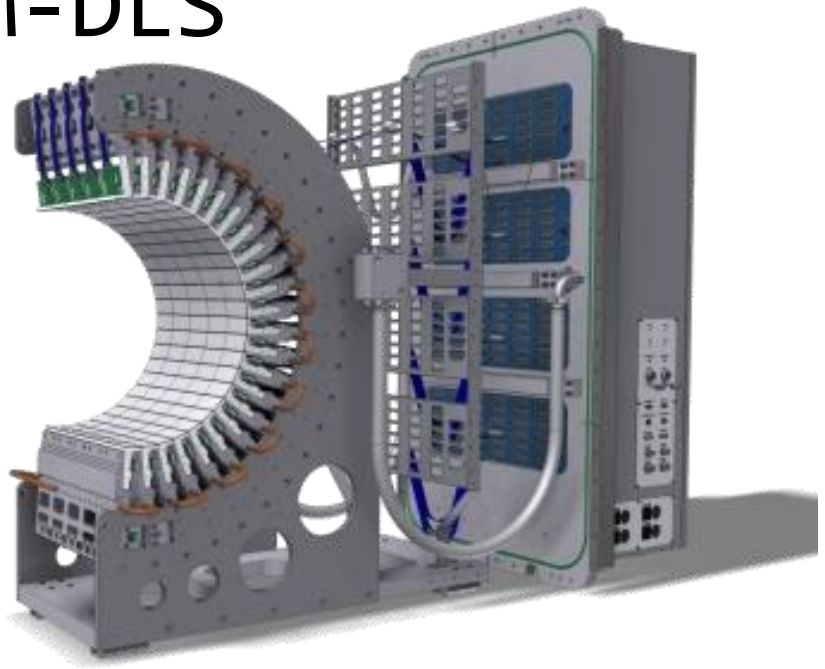
1. Introduction

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

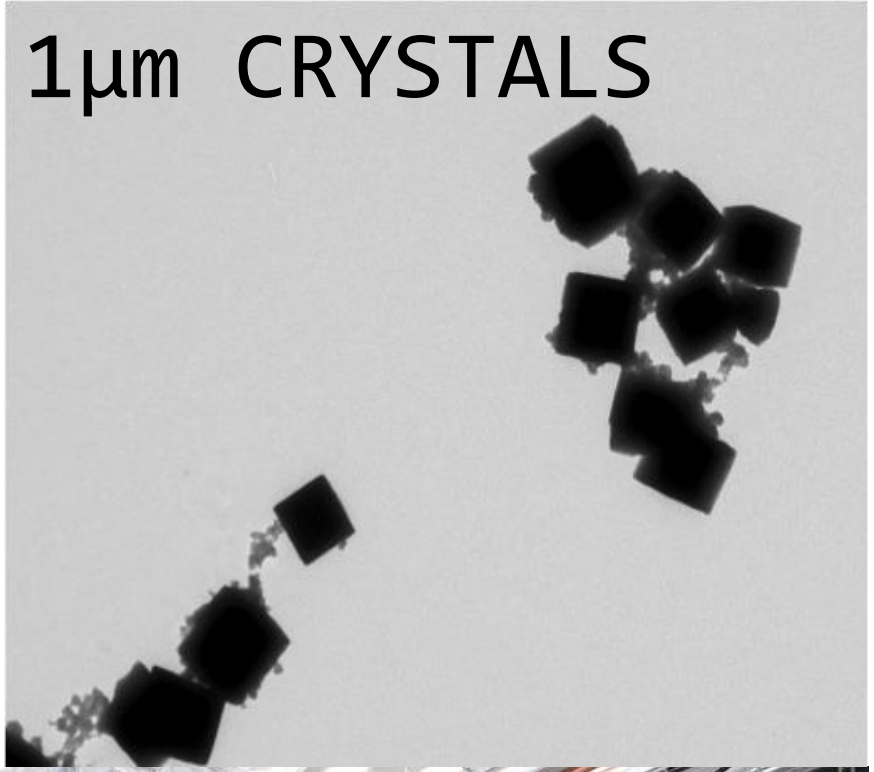
DIALS

Diffraction Integration for Advanced Light Sources

P12M-DLS

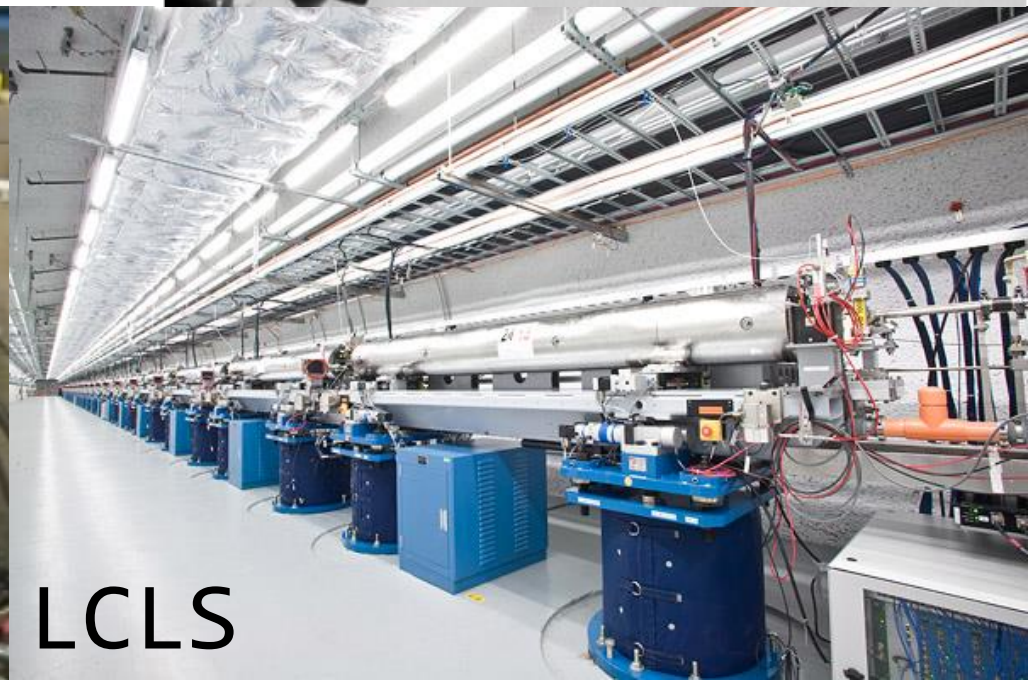


1µm CRYSTALS

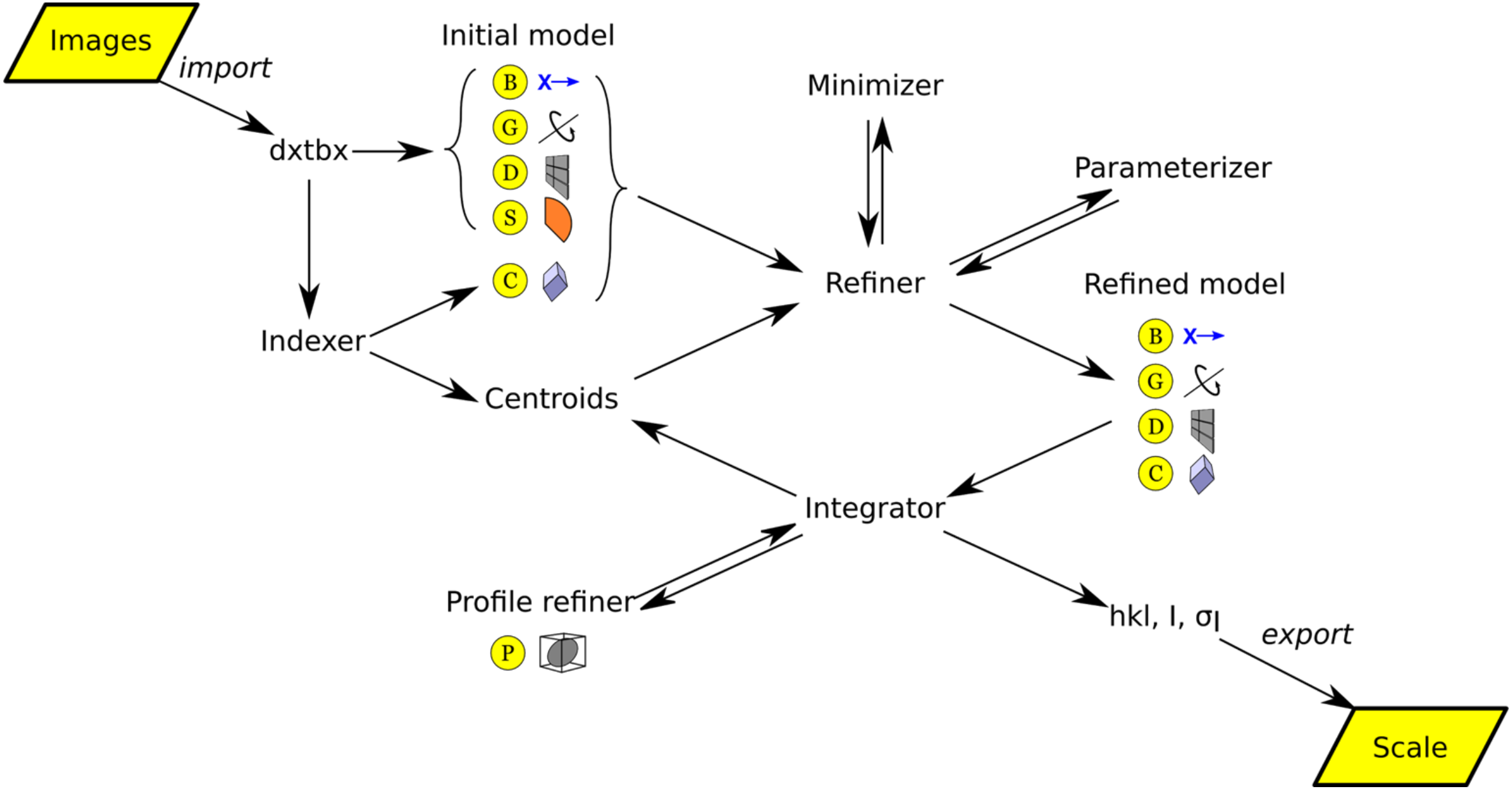


New Challenges

DIAMOND I24



LCLS


Images
import
dxtbx
Initial model
B
G
D
S
C
Indexer
Centroids
Minimizer
Refiner
Parameterizer
Refined model
B
G
D
C
Integrator
Profile refiner
P
hkl, I, σI
export
Scale

Main DIALS programs

- dials.import
- dials.find_spots
- dials.index
- dials.refine_bravais_settings
- dials.refine
- dials.integrate
- dials.export_mtz
- (POINTLESS/AIMLESS)

DIALS on the command line

```
$ 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
  outlier.algorithm=tukey use_all_reflections=true
  scan_varying=true
  output.experiments=refined_experiments.json

$ dials.integrate refined_experiments.json
  refined.pickle outlier.algorithm=null nproc=4
```

```
$ dials.export_mtz integrated.pickle
  refined_experiments.json hklout=integrated.mtz

$ pointless hklin integrated.mtz hklout sorted.mtz >
  pointless.log

$ aimless hklin sorted.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?

Useful for automation

- Well defined step-by-step analysis process
- Cleanly defined command-line interface
- Comprehensive range of capabilities
- Python-compatible data files for results etc.
- Robust, tested

Less useful for automation

- Step-by-step analysis process => lots of work to integrate
- Does not include “intelligent decision making”
- Does not (really) include scaling

For more friendly user interface we can use XIA2

Phase — gw56@cs03r-sc-serv-16:/dls/i04/data/2014/nt5073-4/LIC/2/5 — ...

```
Graemes-MacBook-Pro-3:DIALS-6 graeme$ cd LIC25_32/
Graemes-MacBook-Pro-3:LIC25_32 graeme$ xia2 -atom Zn -dials /Volumes/GraemeData/
data/i04-soleil/Lic25_strong/
Environment configuration...
XIA2_ROOT => /Users/graeme/svn/cctbx/modules/xia2
XIA2CORE_ROOT => /Users/graeme/svn/cctbx/modules/xia2/core
Python => /Users/graeme/svn/cctbx/build/../base/Python.framework/Versions/2.7/Re
sources/Python.app/Contents/MacOS/Python
CCTBX => /Users/graeme/svn/cctbx/modules
CCP4 => /Applications/ccp4-6.5
CLIBD => /Applications/ccp4-6.5/lib/data
CCP4_SCR => /var/folders/hs/gvssfcd910s0jczjhjczj6h80000gn/T/tmpreQPsS
Working directory: /Users/graeme/Projects/DIALS-6/LIC25_32
Free space:        218.08 GB
Build: 5267
Contact: xia2.support@gmail.com
XIA2 0.3.8.0
Command line: xia2 -atom Zn -dials /Volumes/GraemeData/data/i04-soleil/Lic25_str
ong/
------------------- Autoindexing SWEEP1 --------------------
All possible indexing solutions:
tP 121.50 121.50  57.03  90.00  90.00  90.00
oC 171.82 171.84  57.03  90.00  90.00  90.00
oP  57.01 121.46 121.50  90.00  90.00  90.00
```

Phase — gw56@cs03r-sc-serv-16:/dls/i04/data/2014/nt5073-4/LIC/2/5 —...

For AUTOMATIC/DEFAULT/SAD

High resolution limit	1.54	6.89	1.54
Low resolution limit	121.50	121.50	1.58
Completeness	100.0	99.9	100.0
Multiplicity	10.7	9.6	10.4
I/sigma	17.0	39.3	3.4
Rmerge	0.065	0.037	0.609
Rmeas(I)	0.077	0.048	0.683
Rmeas(I+/-)	0.072	0.040	0.675
Rpim(I)	0.023	0.016	0.210
Rpim(I+/-)	0.030	0.017	0.289
CC half	0.999	0.998	0.893
Wilson B factor	14.071		
Anomalous completeness	100.0	100.0	100.0
Anomalous multiplicity	5.5	5.8	5.3
Anomalous correlation	0.498	0.735	0.049
Anomalous slope	1.251	0.000	0.000
dF/F	0.081		
dI/s(dI)	1.260		
Total observations	682919	8089	48088
Total unique	63570	845	4625

Assuming spacegroup: P 42 21 2

Unit cell:

121.505 121.505 57.029

dials.find_spots

DIALS: Diffraction Integration for Advanced Light Sources

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
```

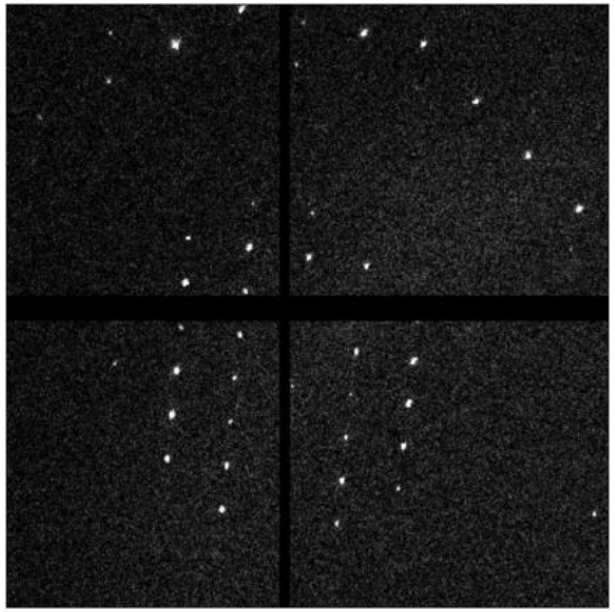
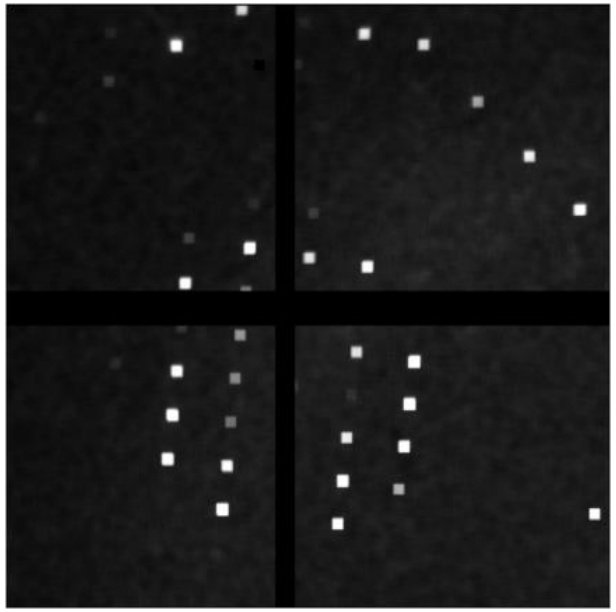
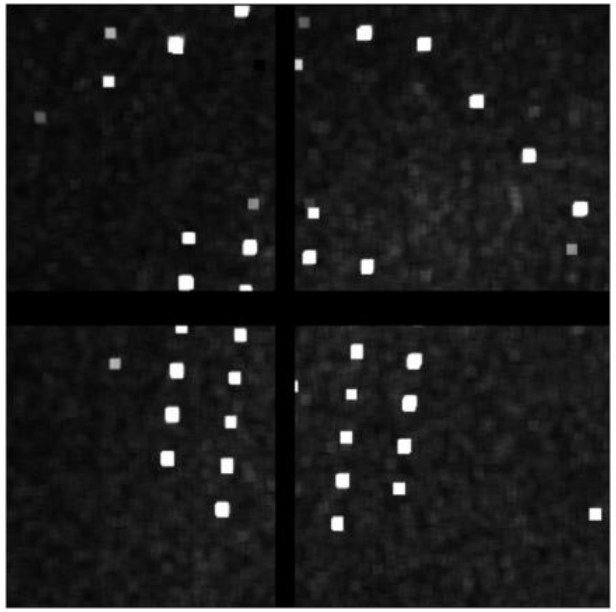
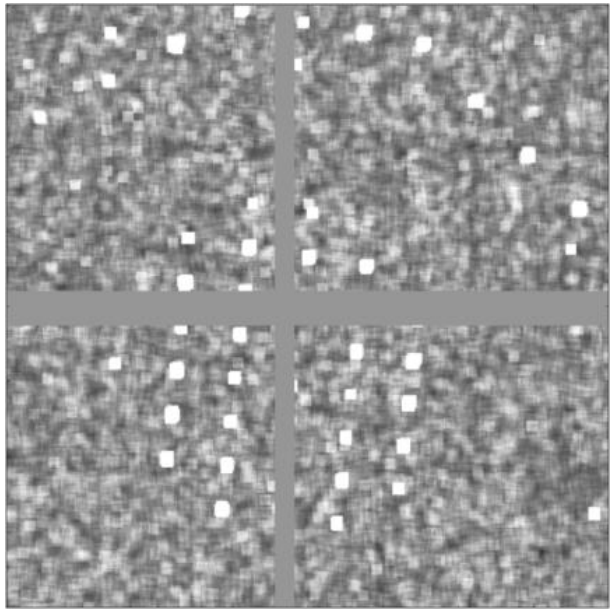
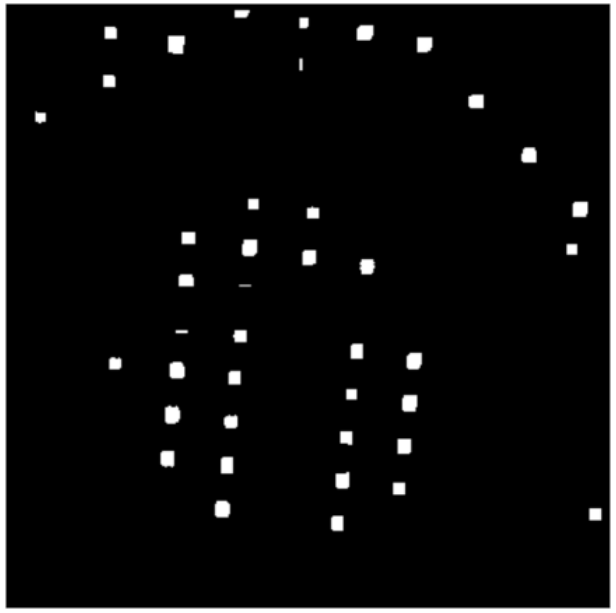
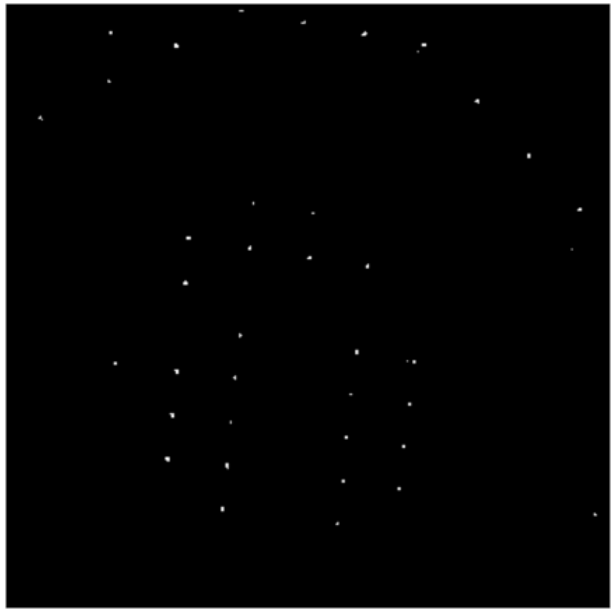
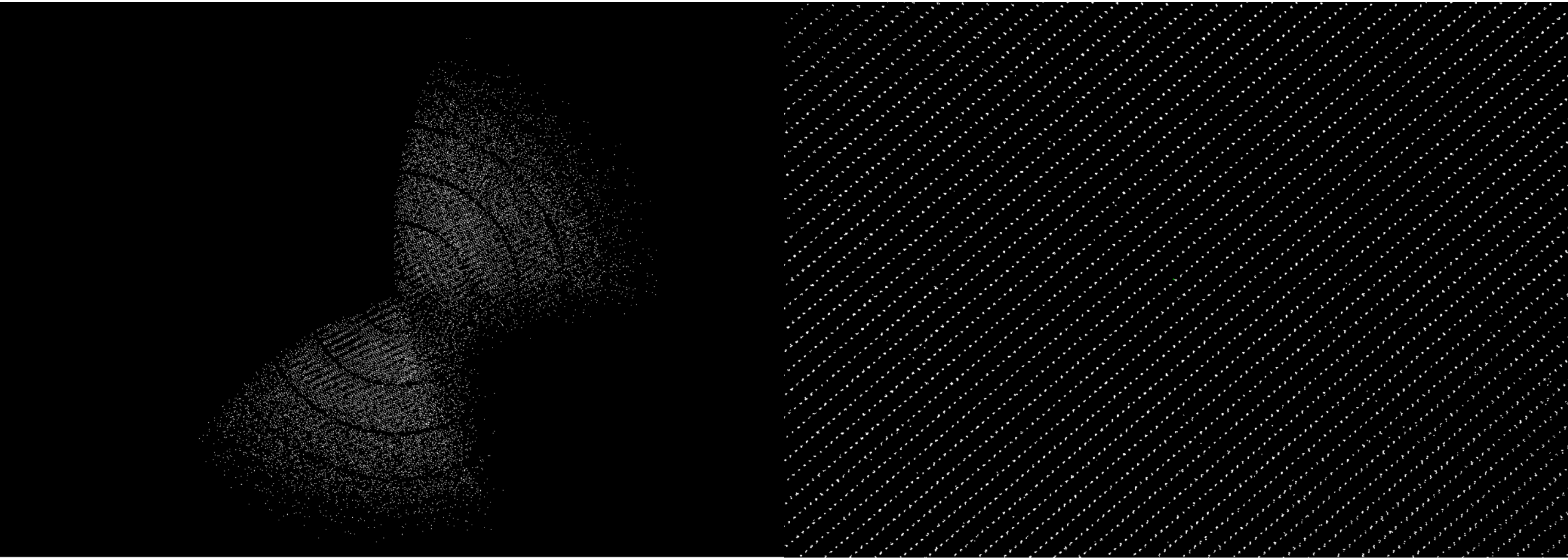







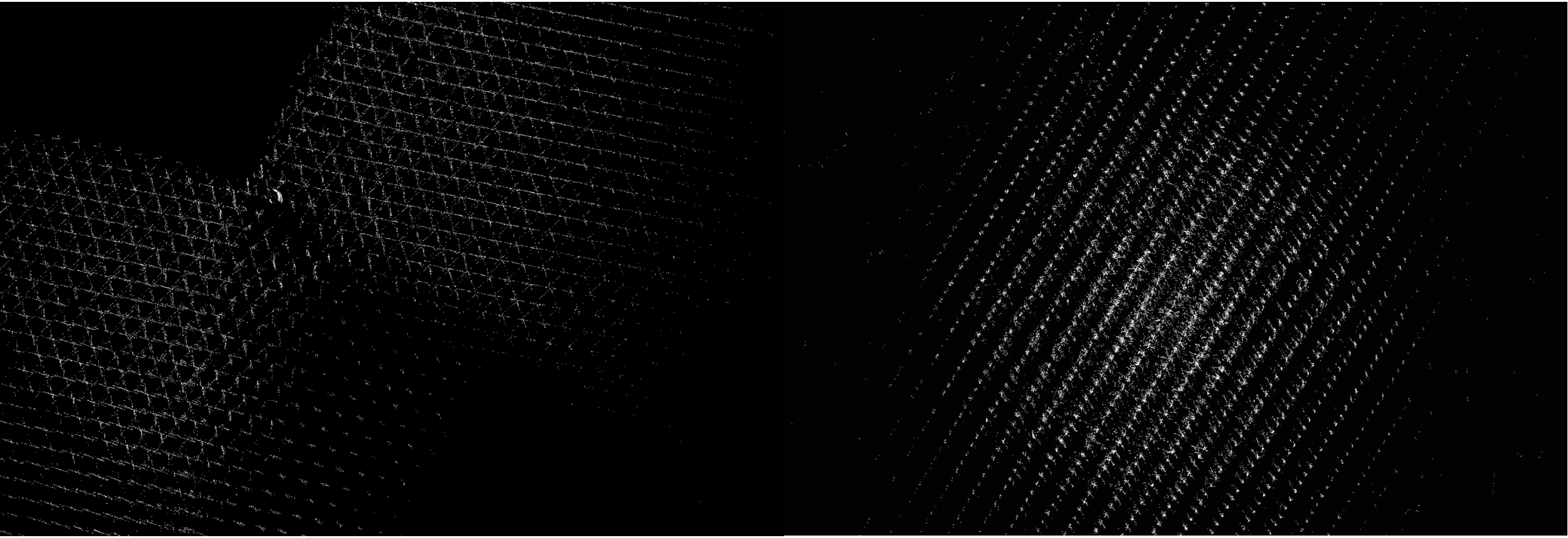
Image processing steps

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

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



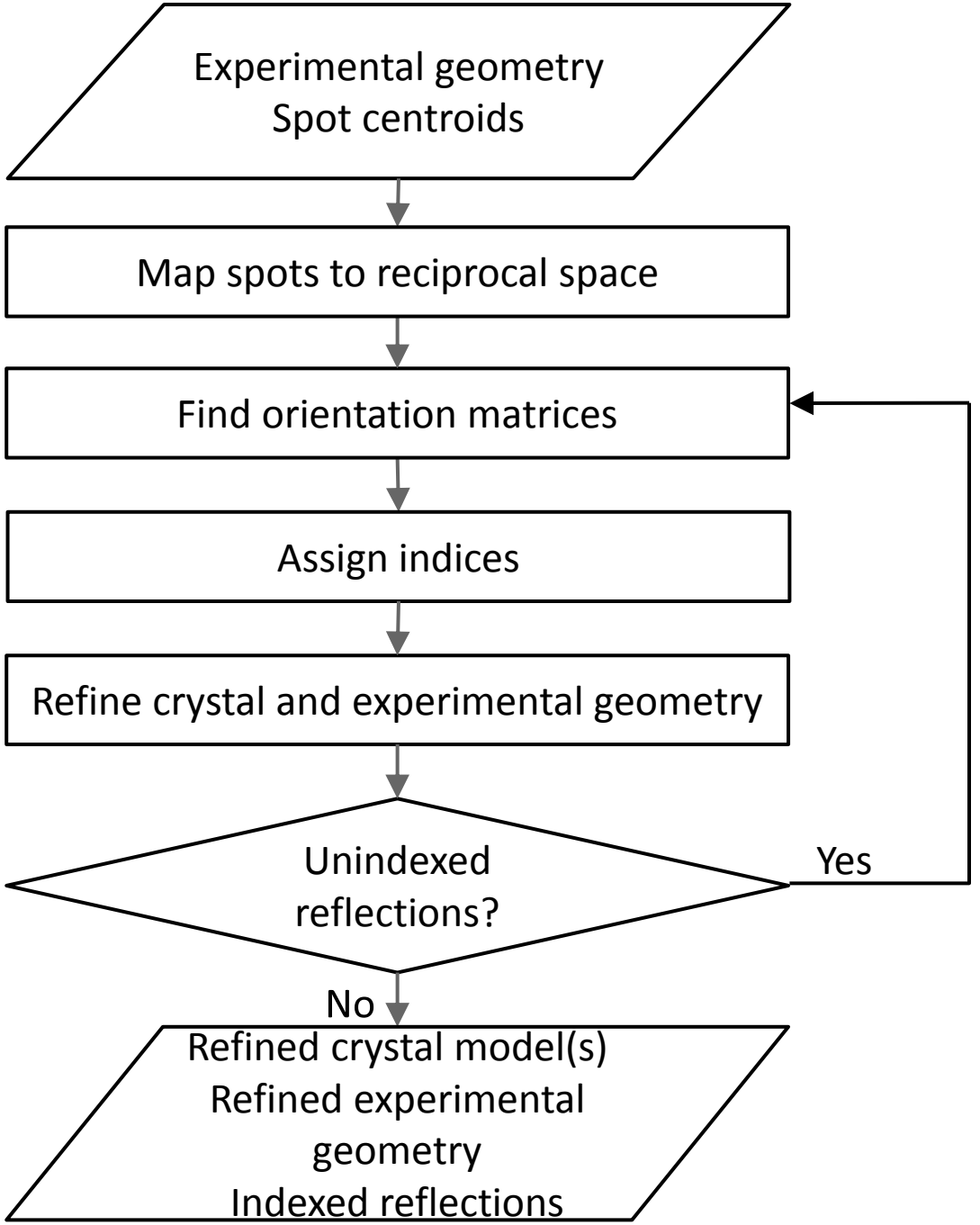
```
$ dials.reciprocal_lattice_viewer datablock.json strong.pickle
```



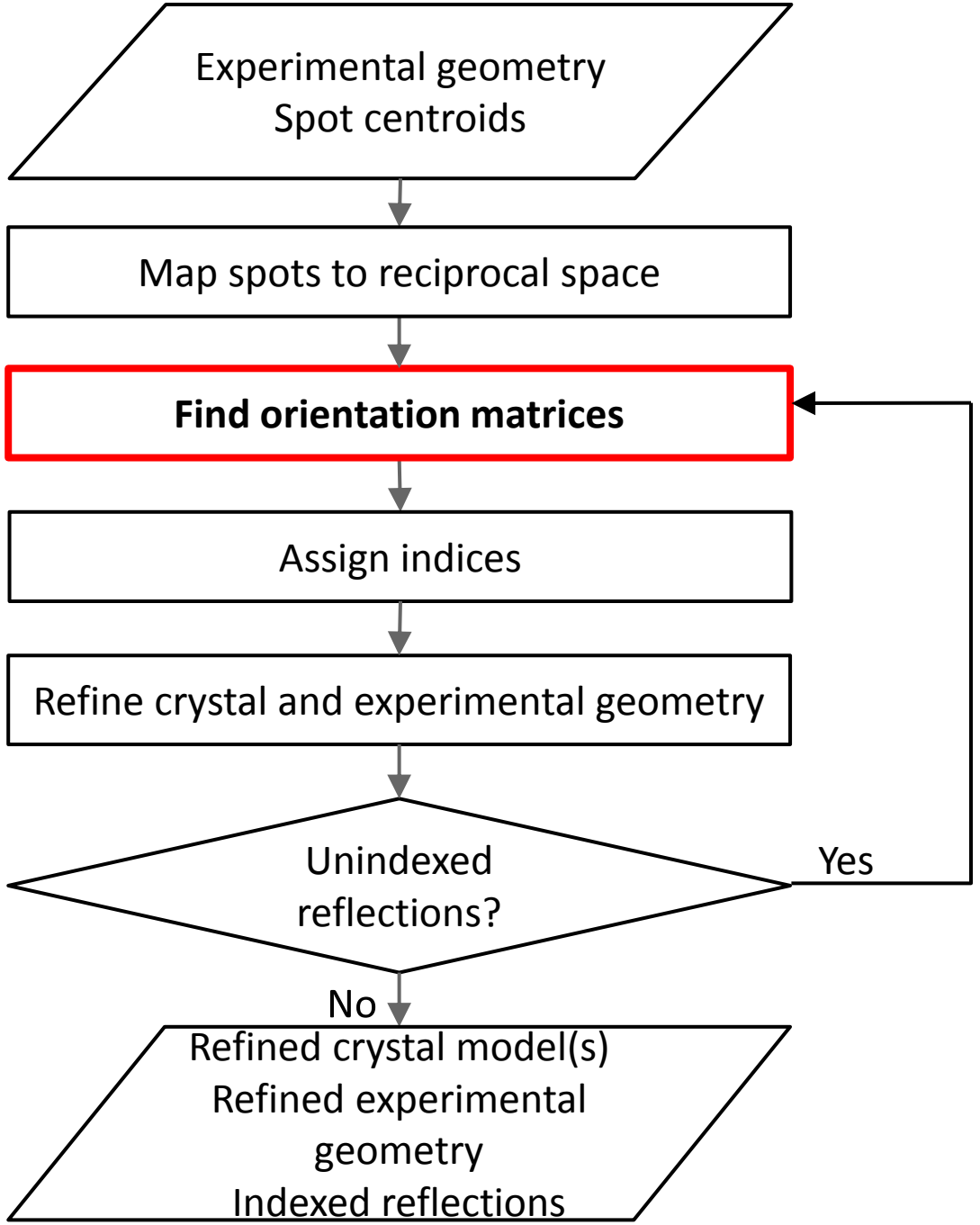
dials.index

DIALS: Diffraction Integration for Advanced Light Sources

dials.index


Experimental geometry
Spot centroids
Map spots to reciprocal space
Find orientation matrices
Assign indices
Refine crystal and experimental geometry
Unindexed reflections?
Yes
No
Refined crystal model(s)
Refined experimental geometry
Indexed reflections

dials.index


Experimental geometry
Spot centroids
Map spots to reciprocal space
Find orientation matrices
Assign indices
Refine crystal and experimental geometry
Unindexed reflections?
Yes
No
Refined crystal model(s)
Refined experimental geometry
Indexed reflections

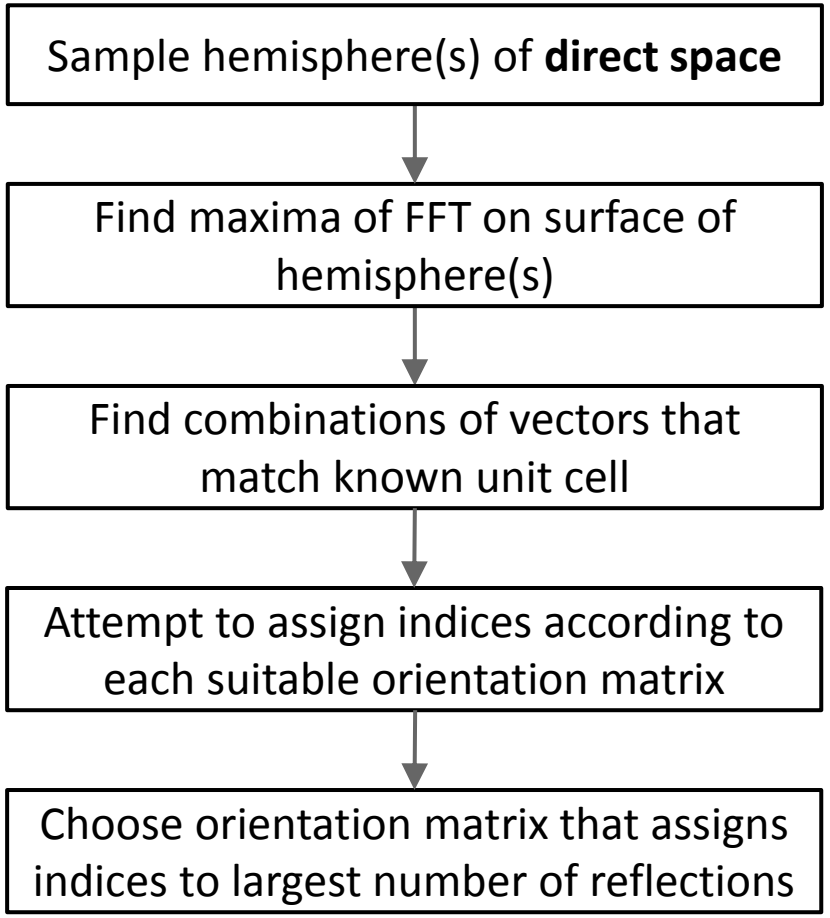
dials.index

- Choice of 1D & 3D FFT methods or new real space grid search algorithm
- 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
```

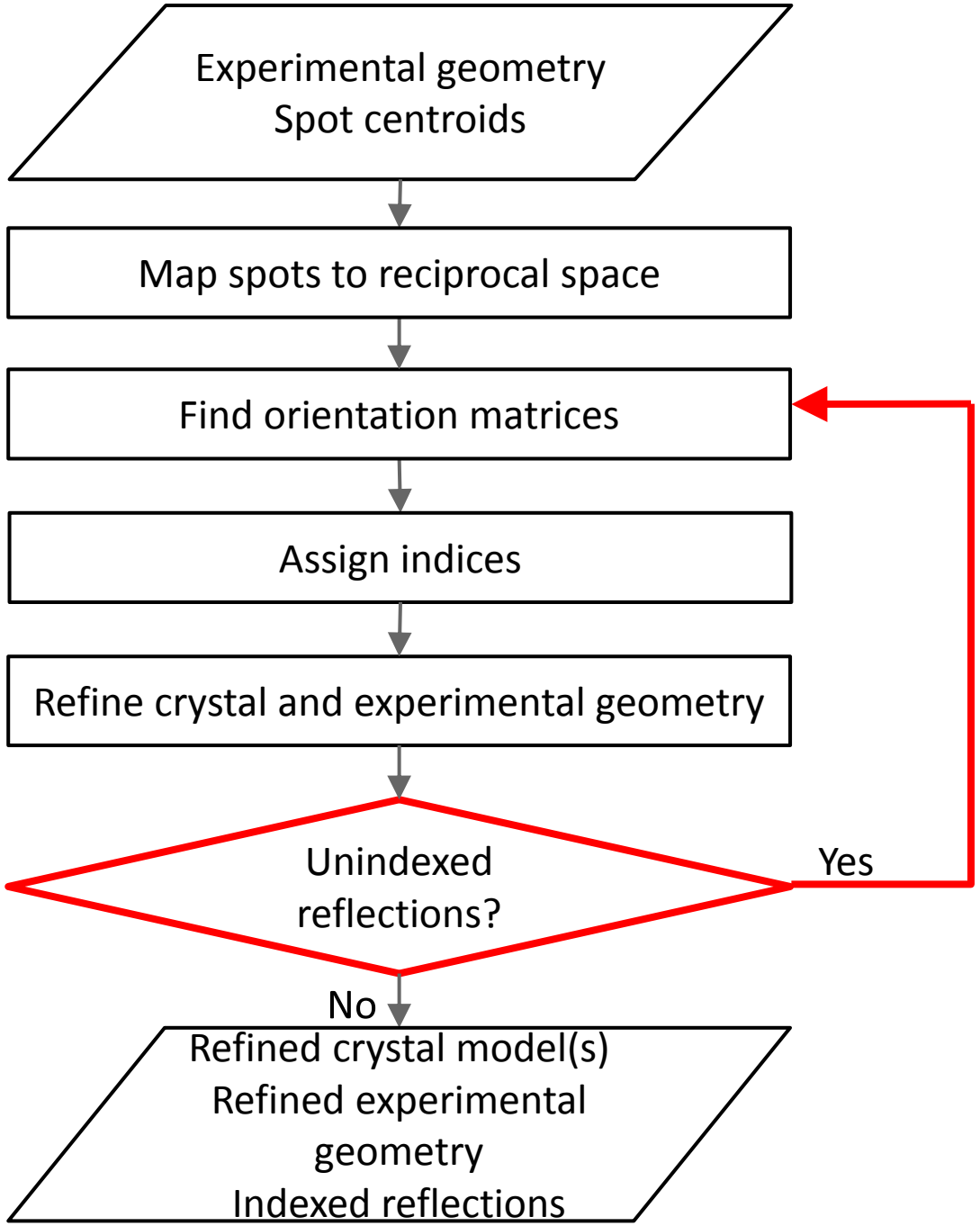
Real space grid search

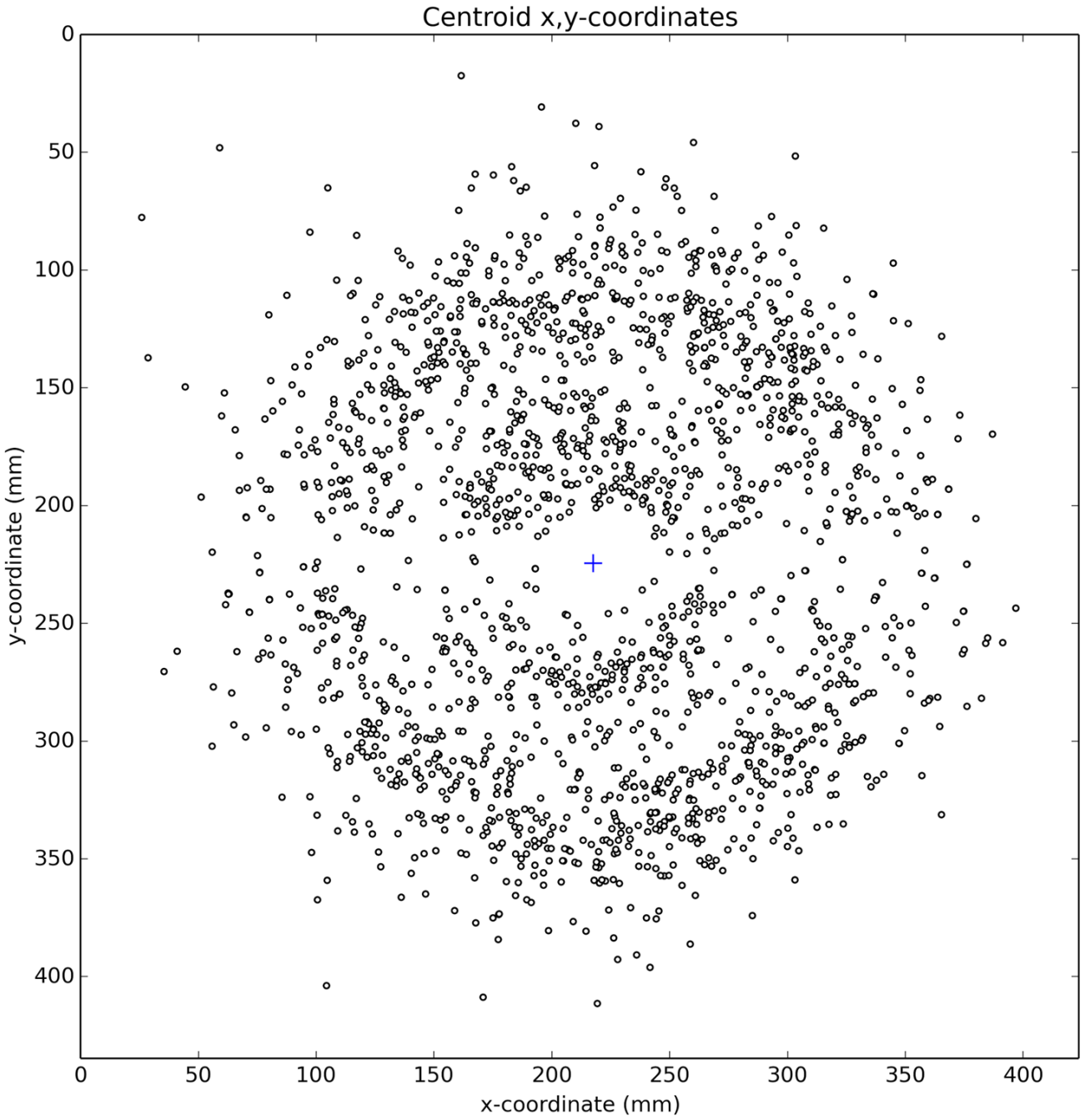
- In many cases the unit cell is known - why not make use of this information?
- 1D FFT (DPS) and 3D FFT algorithms try to determine the magnitude and direction of the basis vectors simultaneously
- If the unit cell is already known, then we only need to determine the direction of the basis vectors



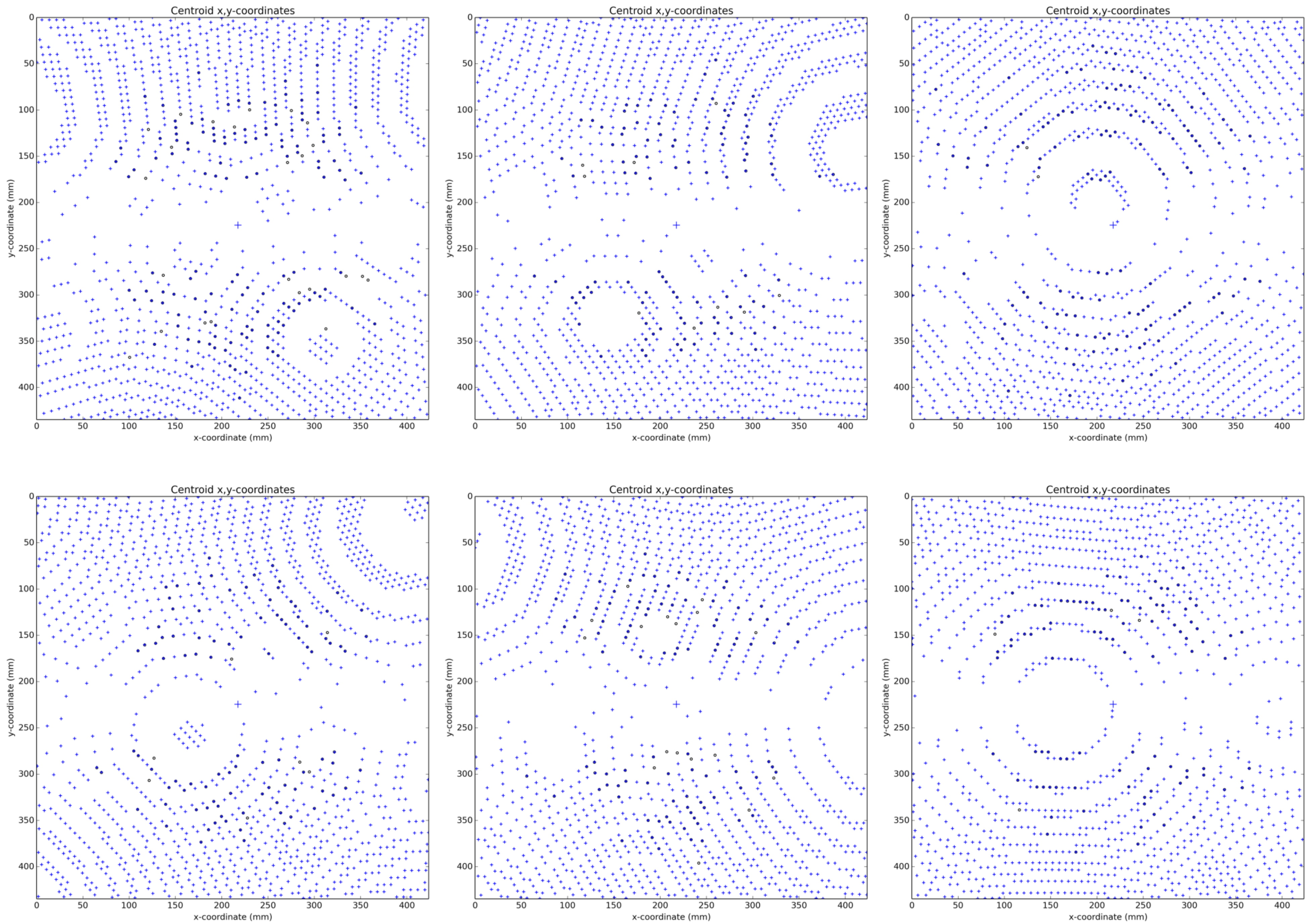
New methods for indexing multi-lattice diffraction data. Gildea *et al.*, *Acta Cryst.* (2014) **D**70, 2652-2666.

dials.index


Experimental geometry
Spot centroids
Map spots to reciprocal space
Find orientation matrices
Assign indices
Refine crystal and experimental geometry
Unindexed reflections?
Yes
No
Refined crystal model(s)
Refined experimental geometry
Indexed reflections



1° wedge of data
1858 spots



6 lattices identified

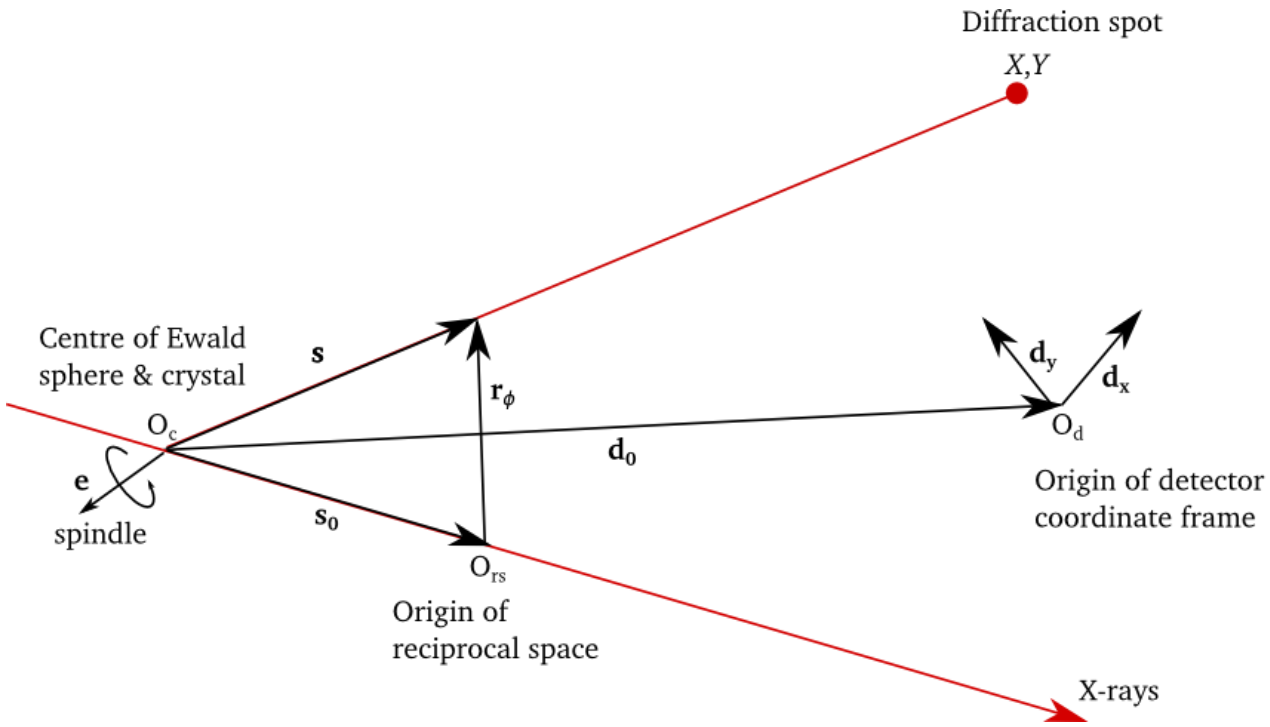
dials.refine_bravais_settings

```
$ 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
-----------------------------------------------------------------------------------------------------------

$ ls bravais_setting_*.json
bravais_setting_1.json bravais_setting_2.json bravais_setting_3.json bravais_setting_4.json bravais_setting_5.json
   bravais_setting_6.json bravais_setting_7.json bravais_setting_8.json bravais_setting_9.json
```

dials.refine

DIALS: Diffraction Integration for Advanced Light Sources

Centroid refinement



$$\mathbf{d} = (d_x | d_y | d_0)$$

$$\mathbf{D} = \mathbf{d}^{-1}$$

$$\mathbf{v} = \mathbf{Ds} = \begin{pmatrix} \alpha X \\ \alpha Y \\ 1/\alpha \end{pmatrix}$$

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

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

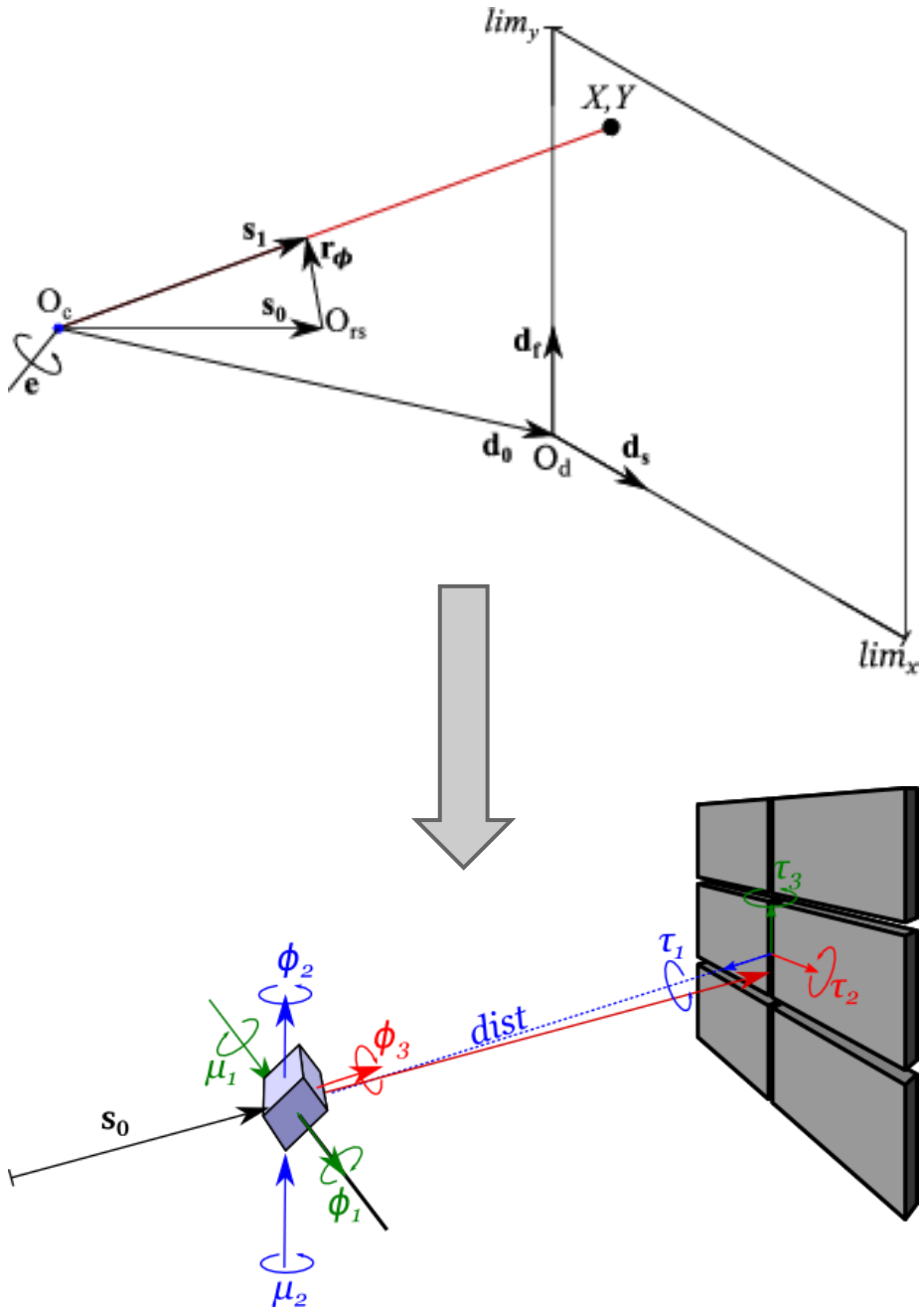
Parameterisation

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



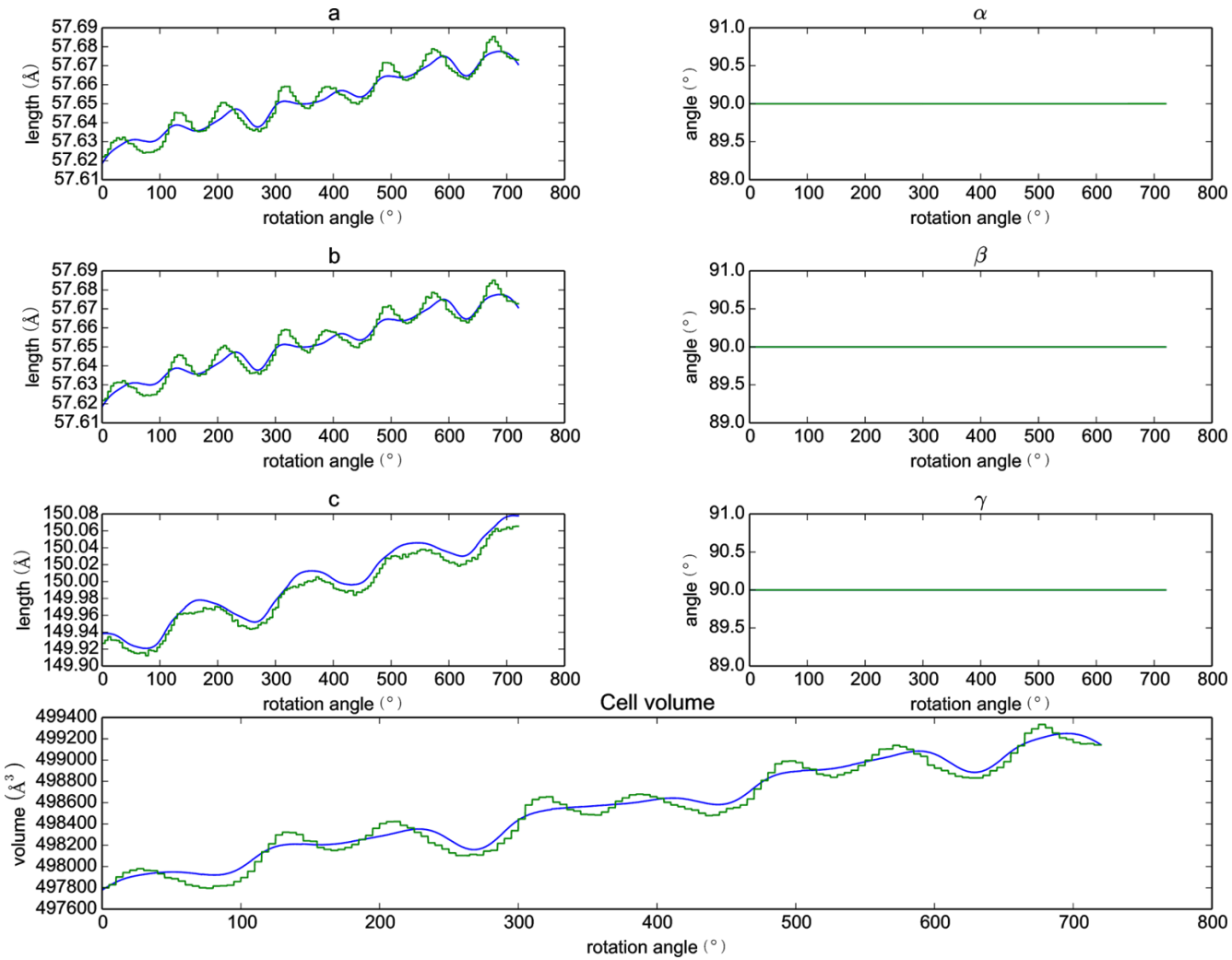
Scan-varying refinement

- 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:

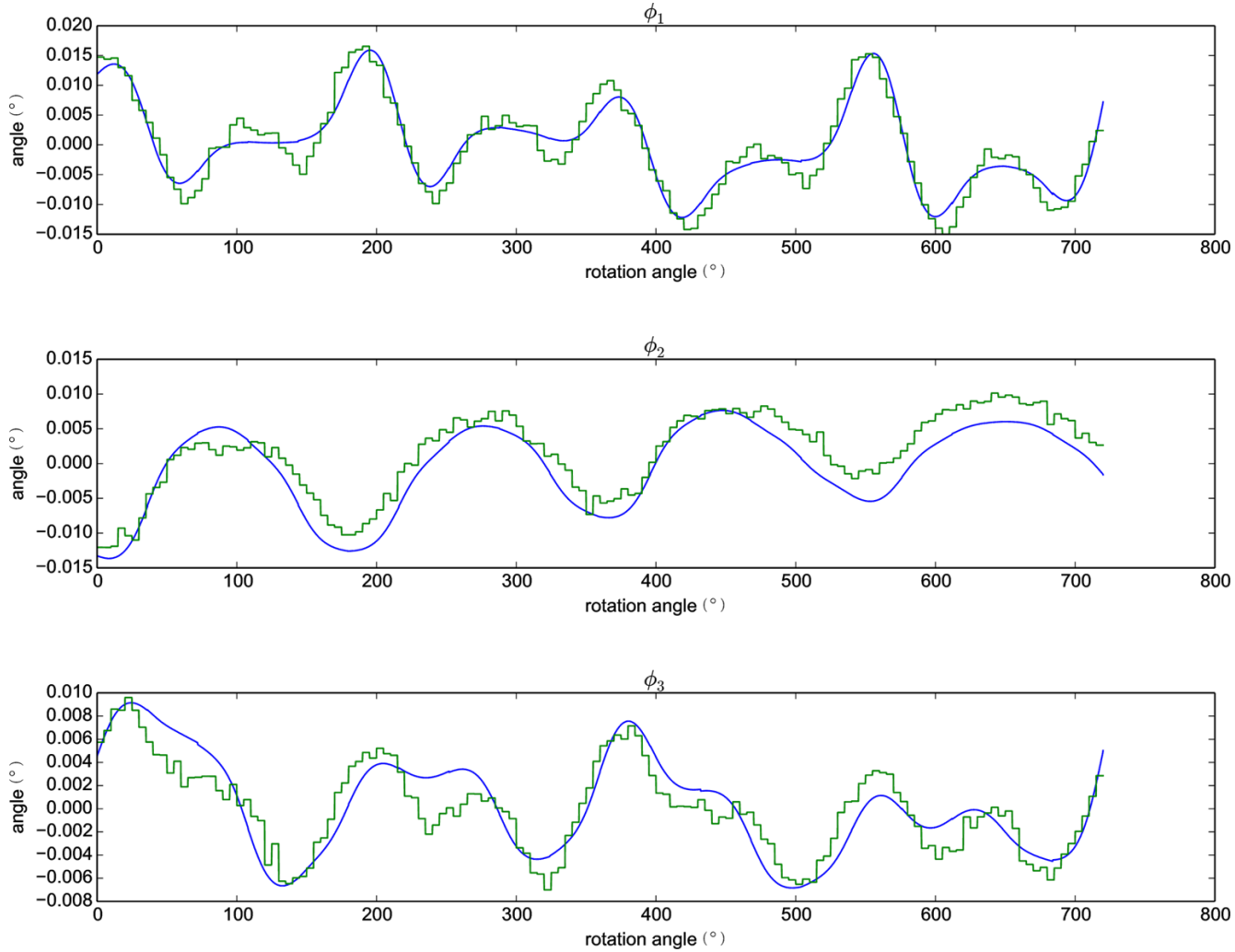
- 6 detector
- 1 beam
- 3 crystal orientation × 22 "samples"
- 2 unit cell parameters × 22 "samples"



Scan-varying refinement

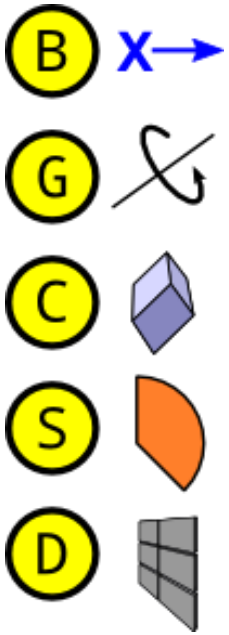
117 parameters:

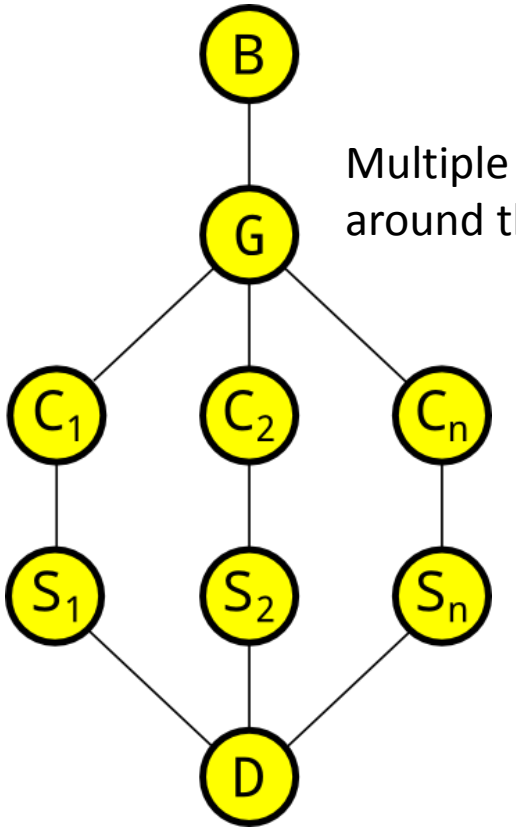
- 6 detector
- 1 beam
- 3 crystal orientation × 22 "samples"
- 2 unit cell parameters × 22 "samples”



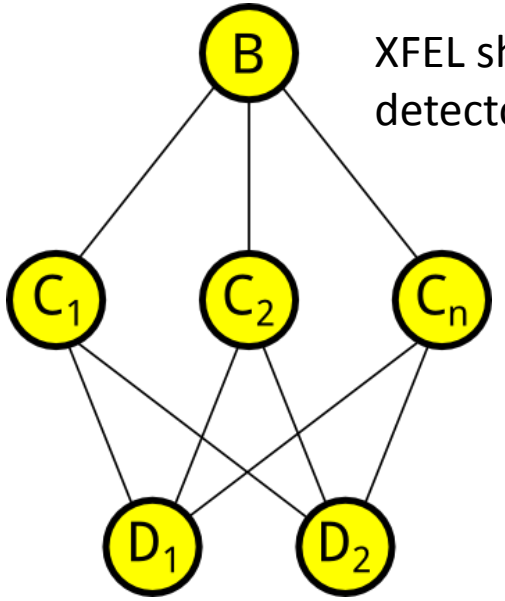
Multiple experiments

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





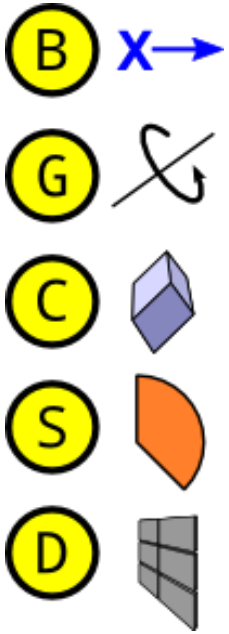
Multiple rotation scans around the same axis

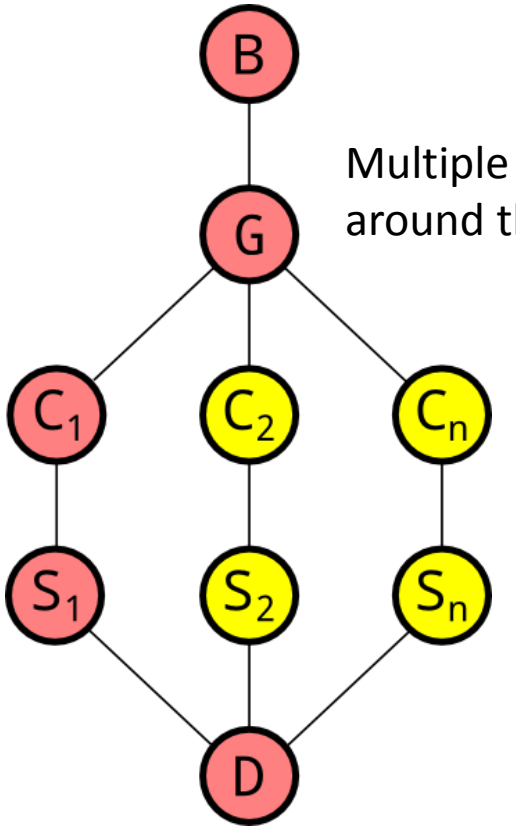


XFEL shots with two detectors

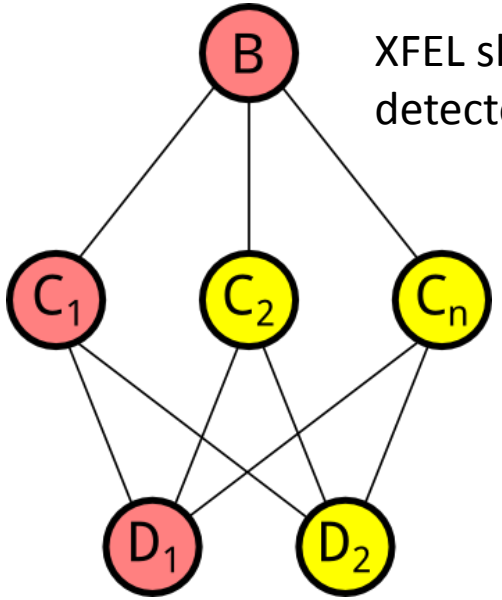
Multiple experiments

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





Multiple rotation scans around the same axis

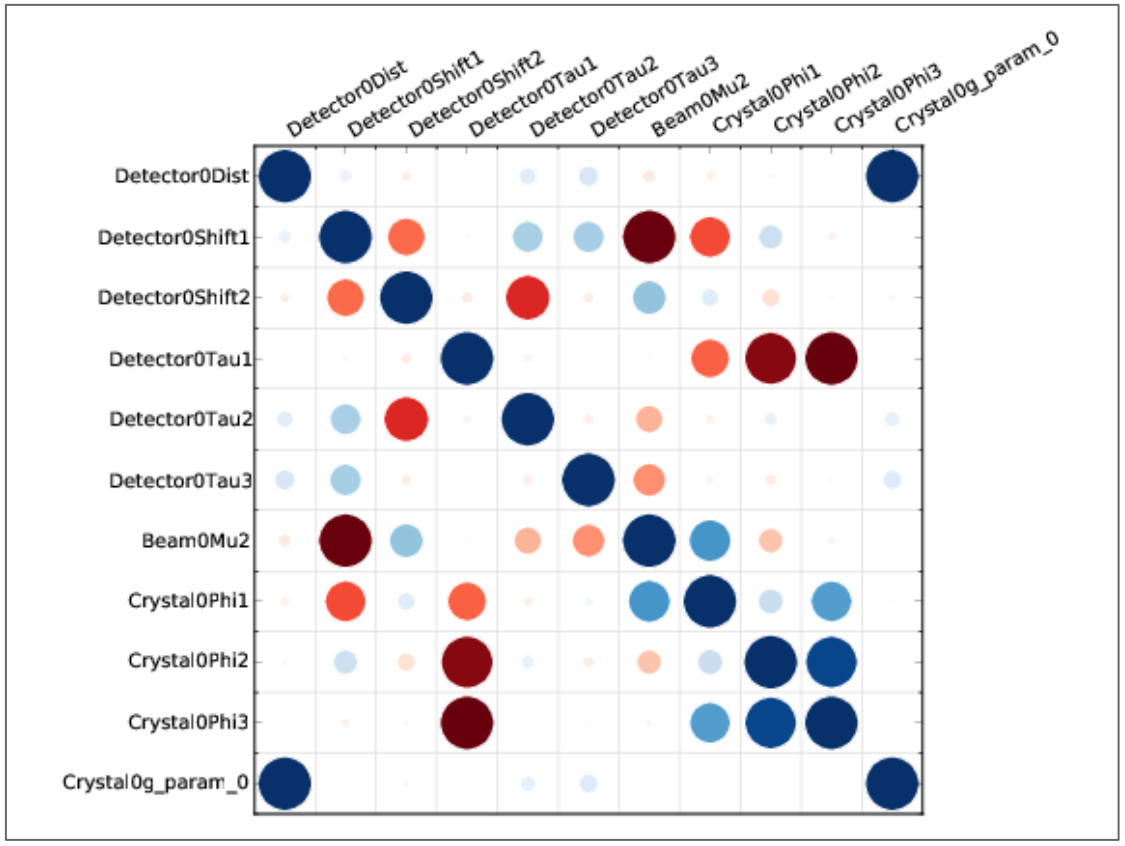


XFEL shots with two detectors

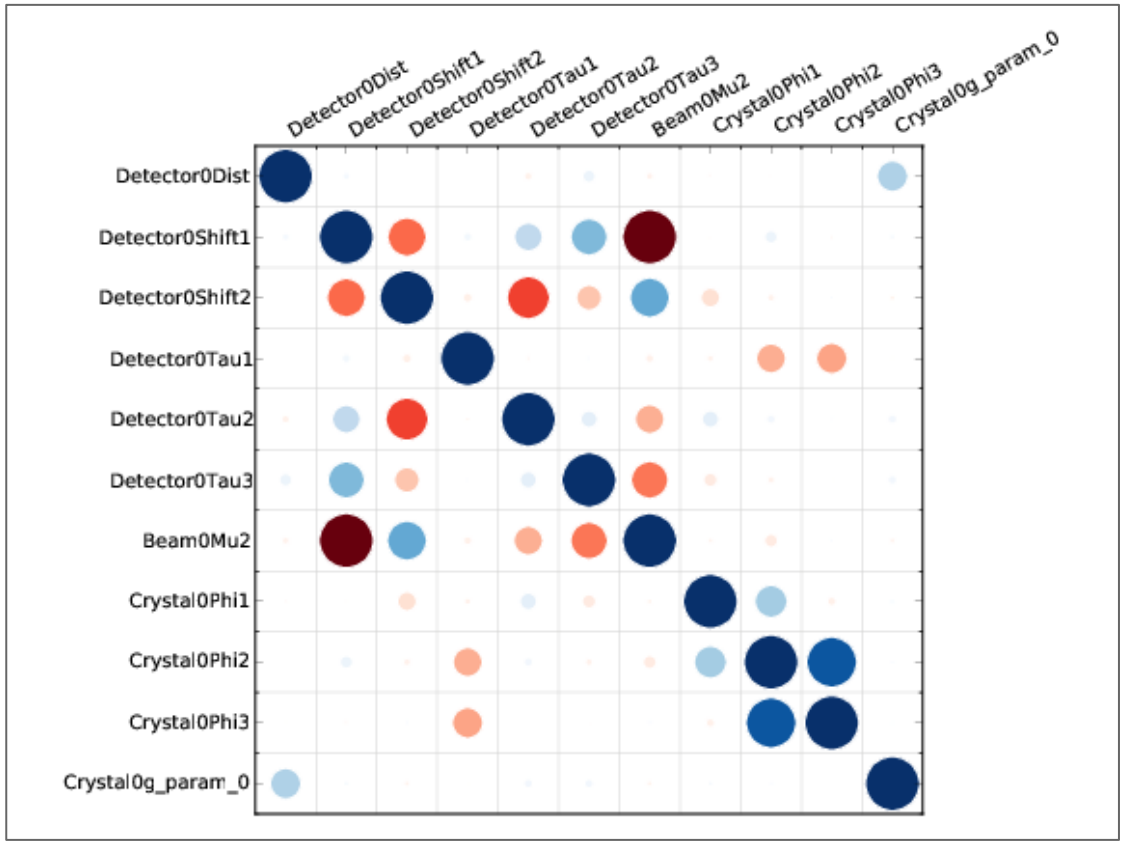
Multiple experiments

Cubic polyhedrin crystals, 1° scans

One lattice



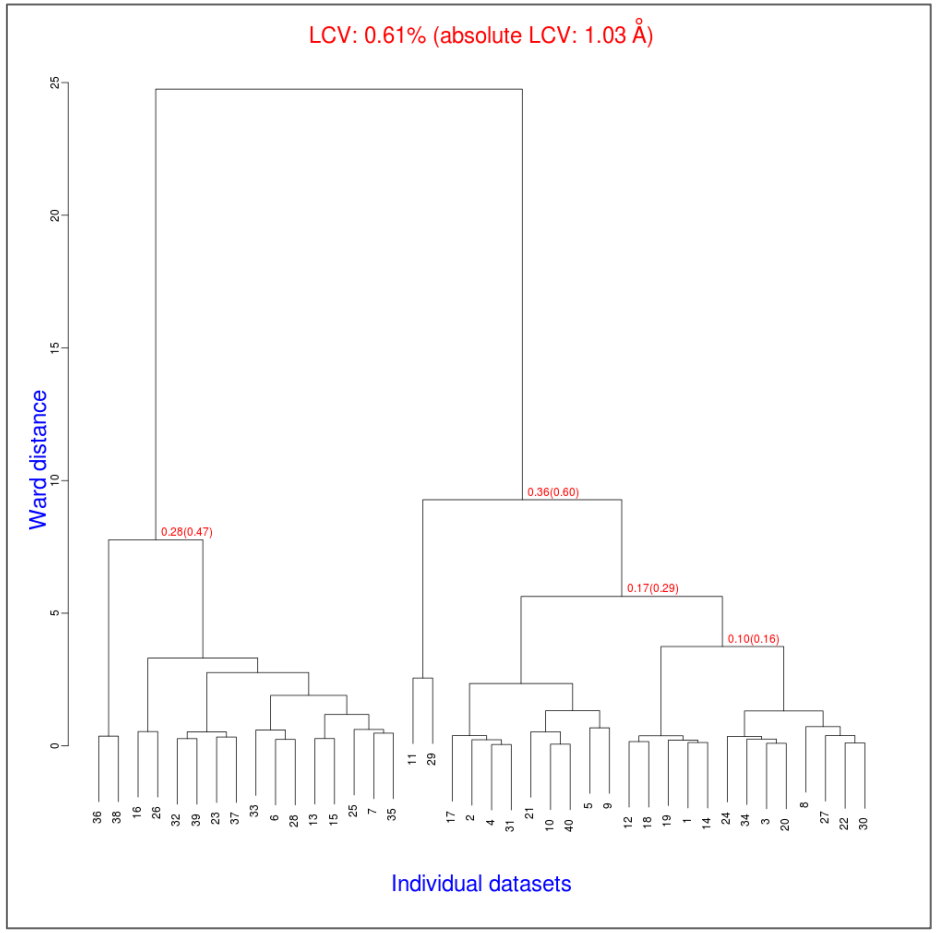
5 sweeps (16 lattices)

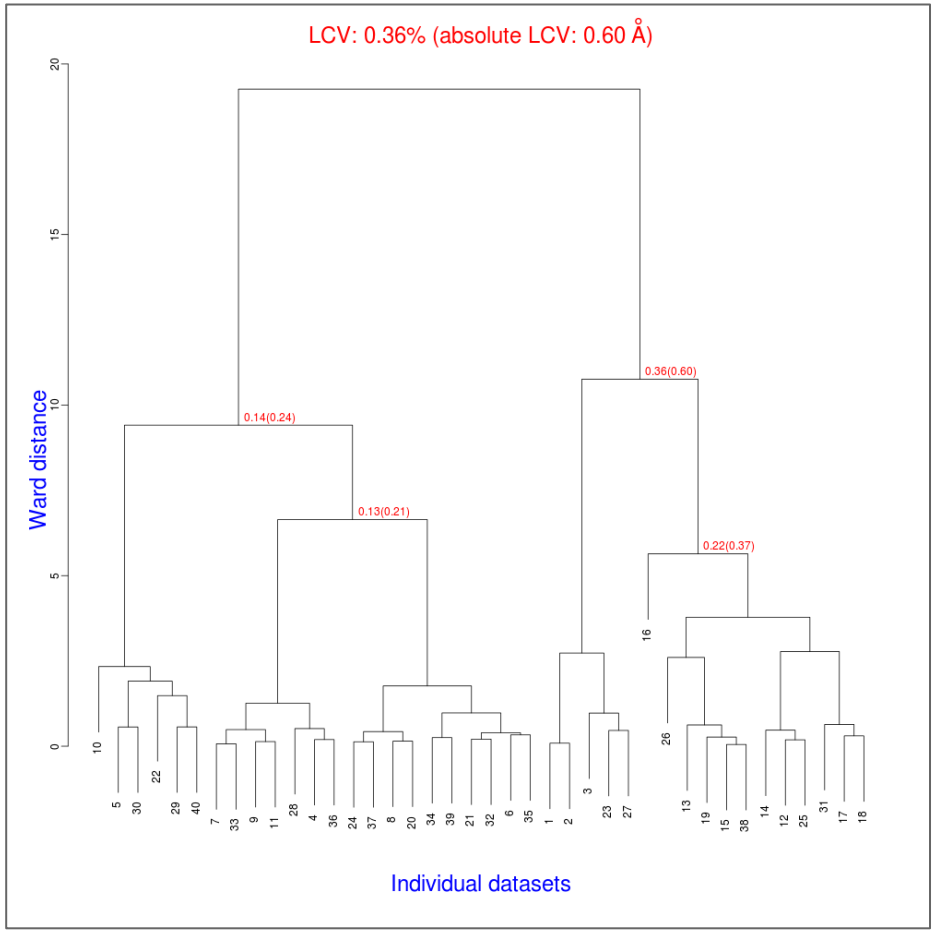


Joint refinement

Multiple experiments

Use joint refinement as a preparatory step for BLEND





TehA data. See forthcoming *Acta Cryst. D***71** (June 2015) for original analysis

dials.integrate

DIALS: Diffraction Integration for Advanced Light Sources

Tasks in dials.integrate

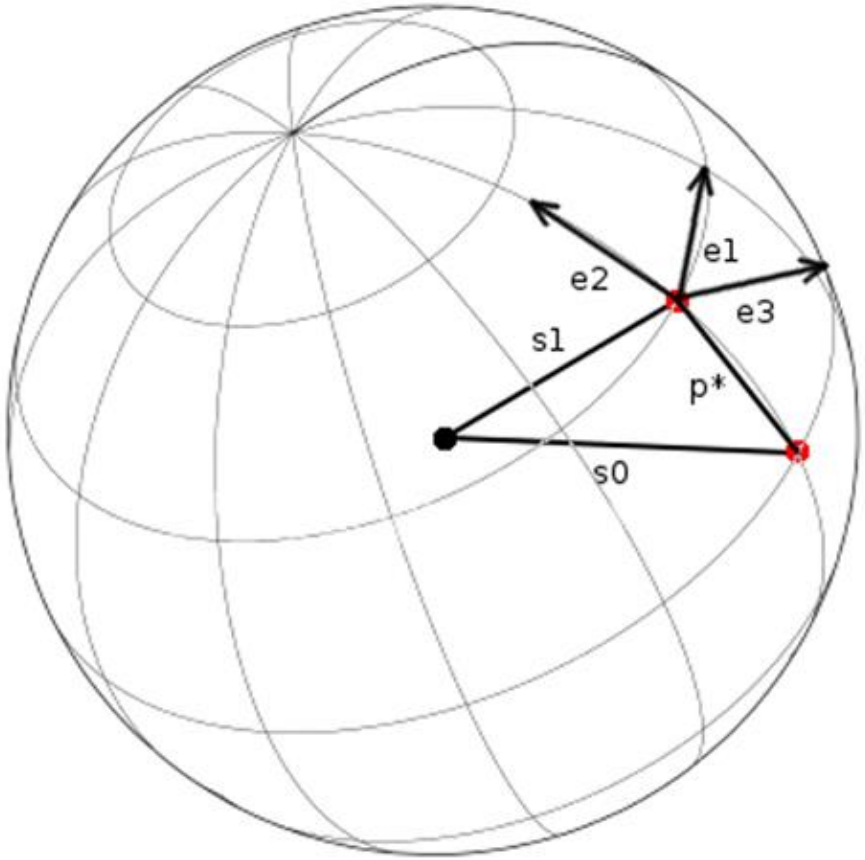
Calculate the bounding box parameters from strong reflections

Predict the positions of reflections on the images

Build reference profiles across all images

Integrate the reflections and save output

Computing reflection shoeboxes



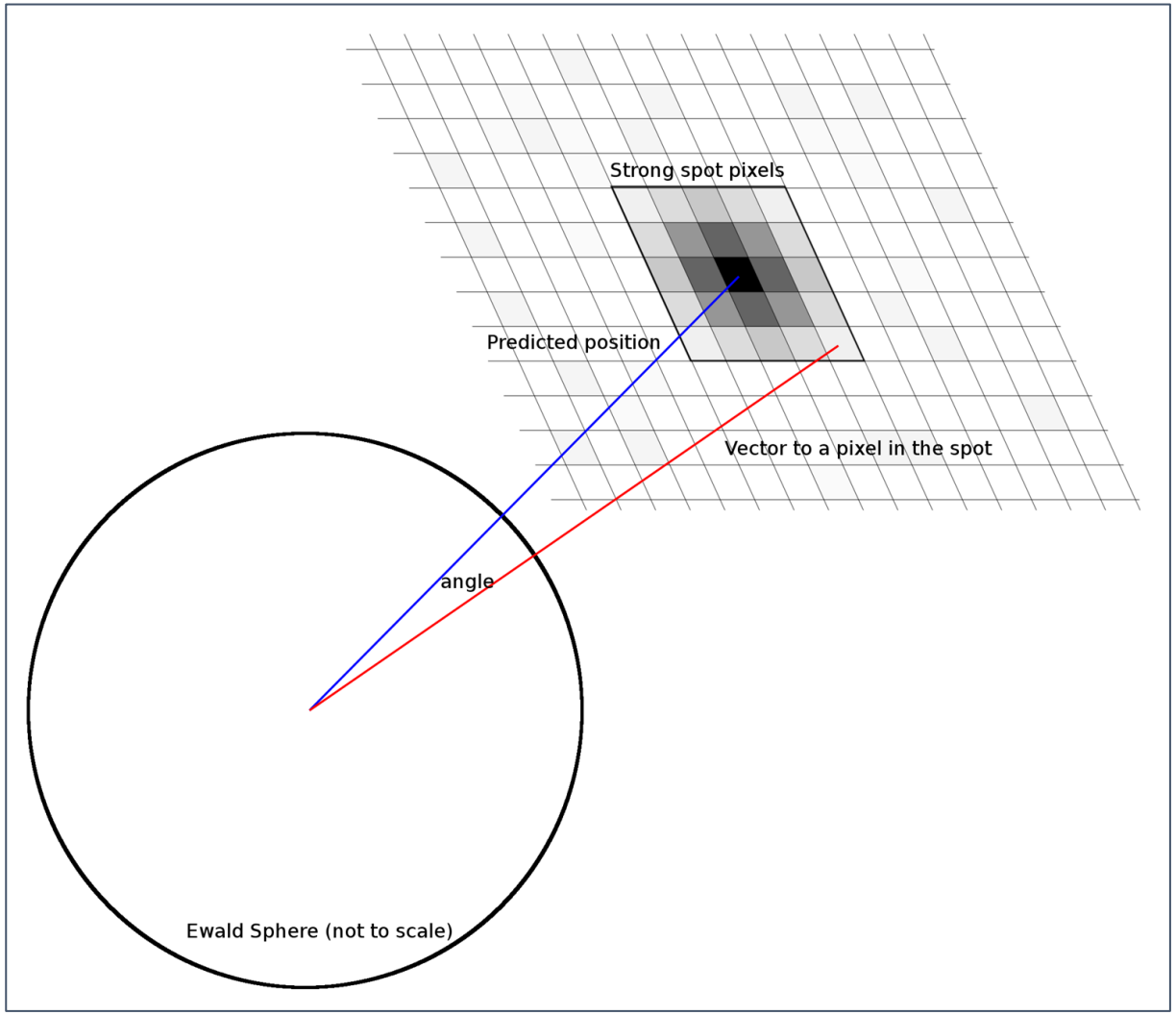
Profile coordinate system

Use the kabsch model of a normal distribution on the surface of the Ewald sphere

$$\exp(\frac{-\epsilon_1^2}{2\sigma_D^2})\ \exp(\frac{-\epsilon_2^2}{2\sigma_D^2})\exp(\frac{-\epsilon_3^2}{2\sigma_M^2})$$

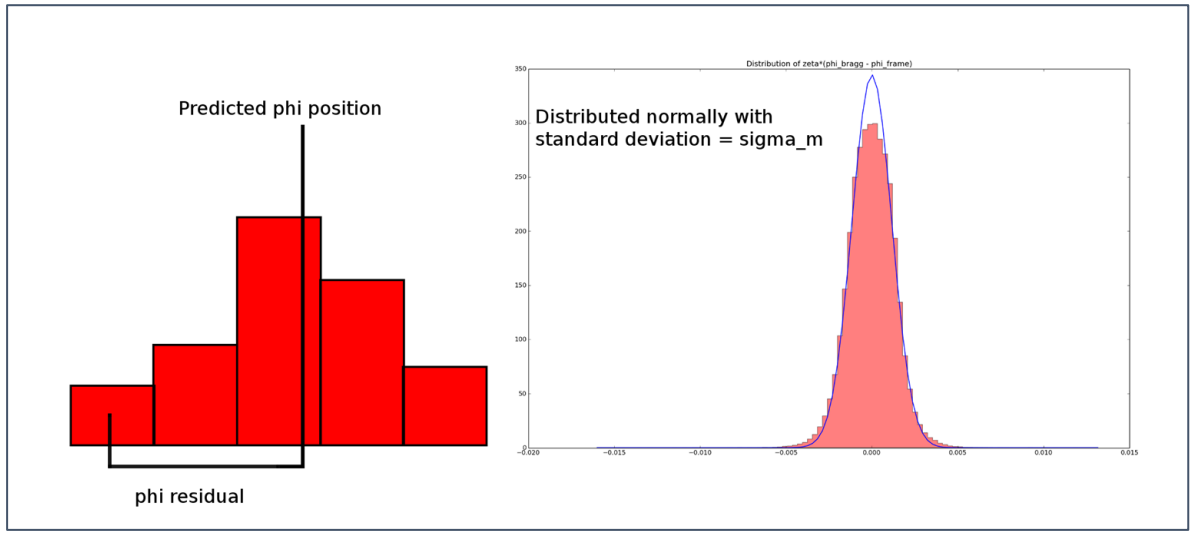
$$\boldsymbol{e}_1 = \boldsymbol{S}_1 \times \boldsymbol{S}_0 / |\boldsymbol{S}_1 \times \boldsymbol{S}_0|$$
$$\boldsymbol{e}_2 = \boldsymbol{S}_1 \times \boldsymbol{e}_1 / |\boldsymbol{S}_1 \times \boldsymbol{e}_1|$$
$$\boldsymbol{e}_3 = (\boldsymbol{S}_1 + \boldsymbol{S}_0) / |\boldsymbol{S}_1 + \boldsymbol{S}_0|$$

Computing reflection shoeboxes

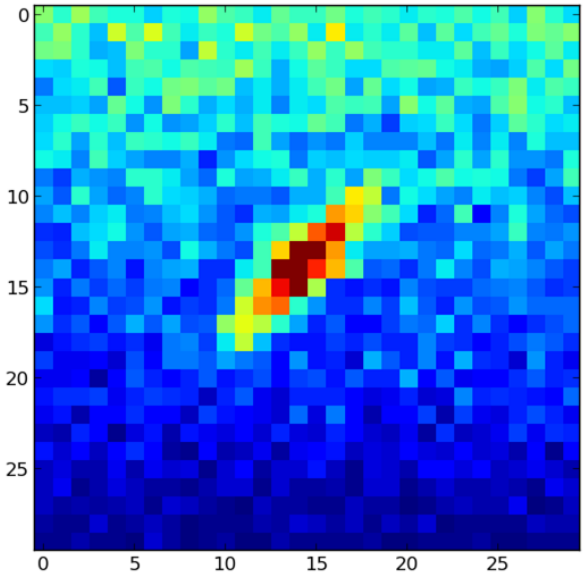


σ_D is calculated from the spread of angles between the predicted diffracted beam vector and the vector for each strong pixel in the spot

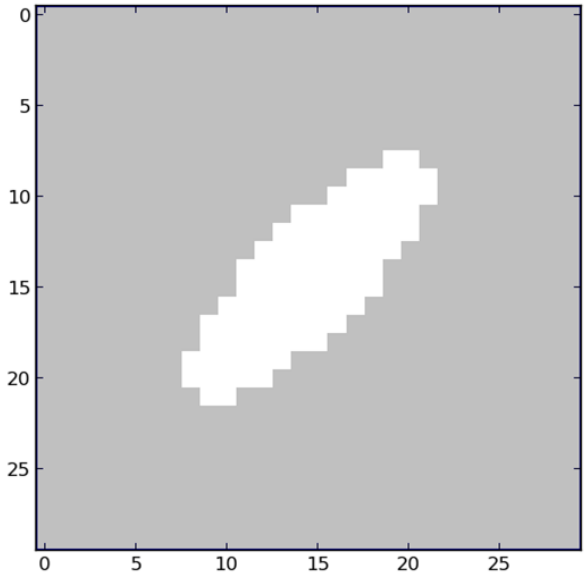
σ_M is calculated by maximum likelihood method assuming a normal distribution of phi residuals for each strong pixel in the spot



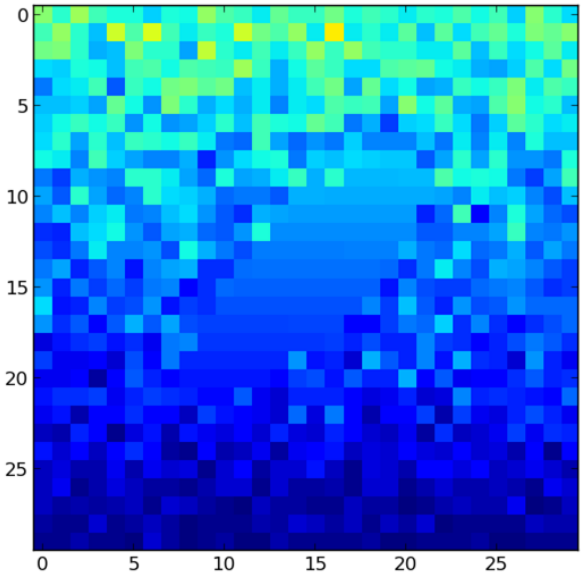
Background modelling



Reflection data



Reflection mask



Calculated background

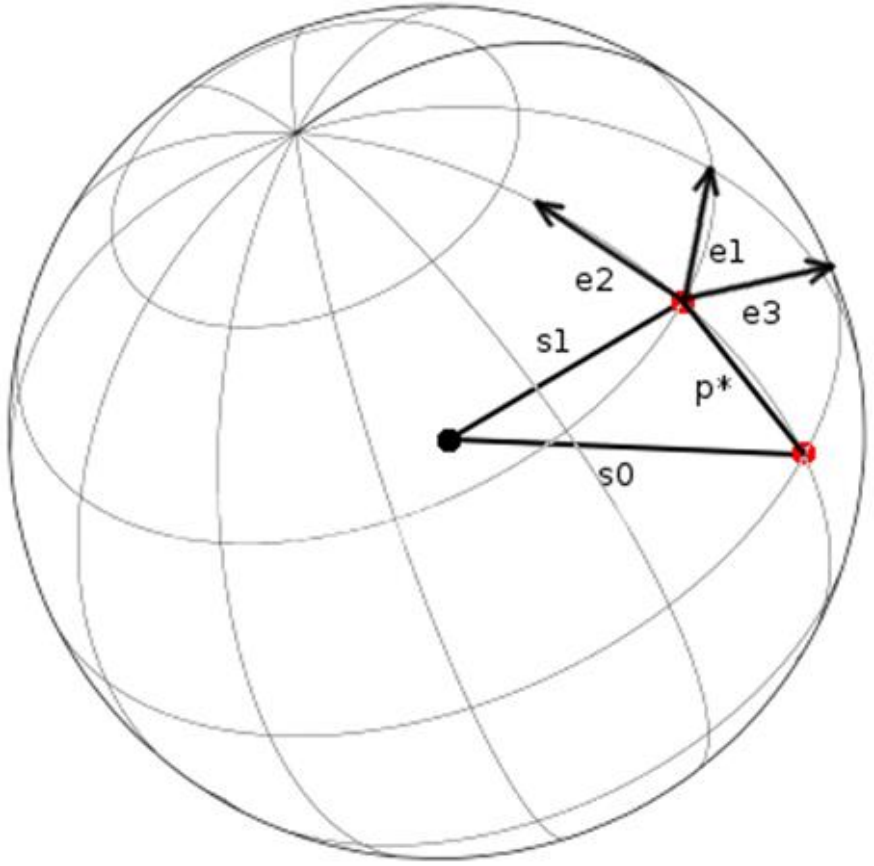
Background models

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

Integration

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

3D profile fitting coordinate system



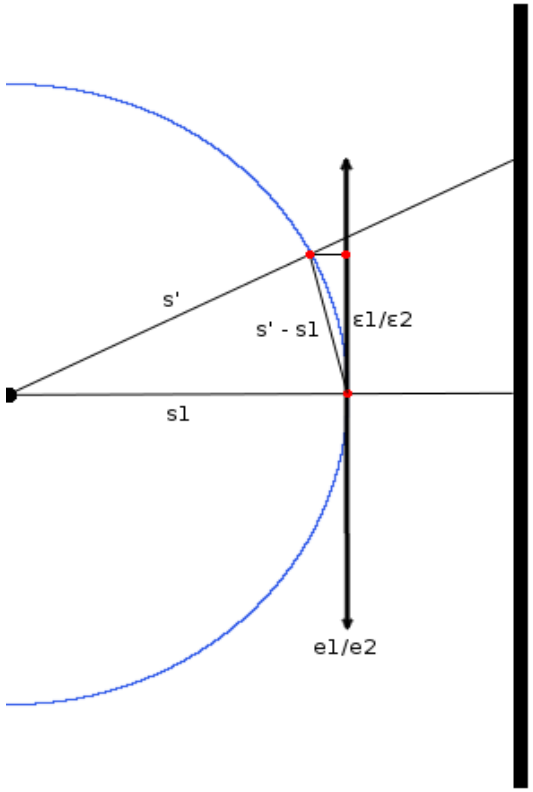
Profile coordinate system

Use Kabsch coordinate system

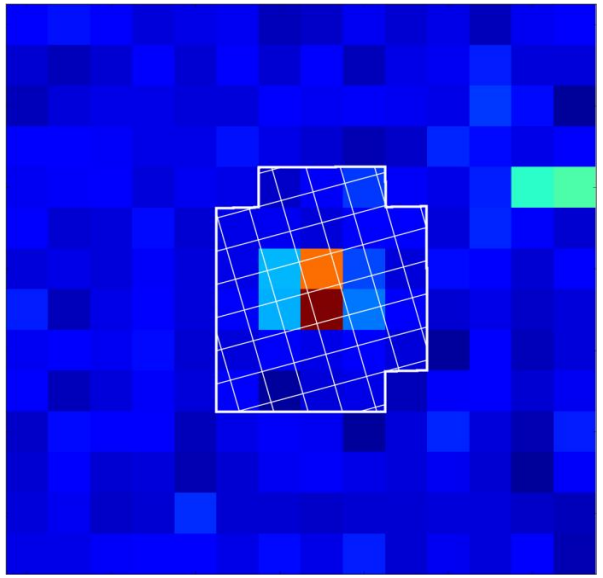
- Corrects for geometrical distortions
- Makes spots appear to have taken shortest path through Ewald sphere
- Model assumes a Gaussian profile in Kabsch coordinate system

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

3D profile fitting pixel gridding

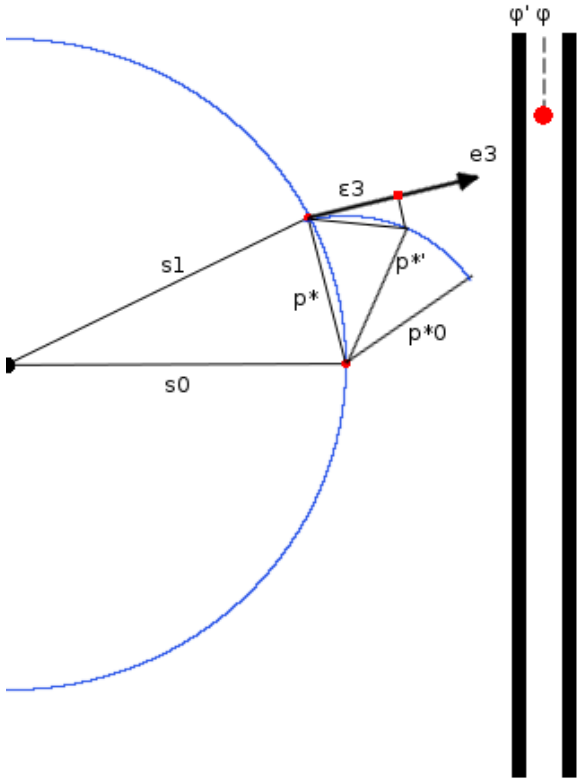


Pixels are mapped to the Ewald sphere

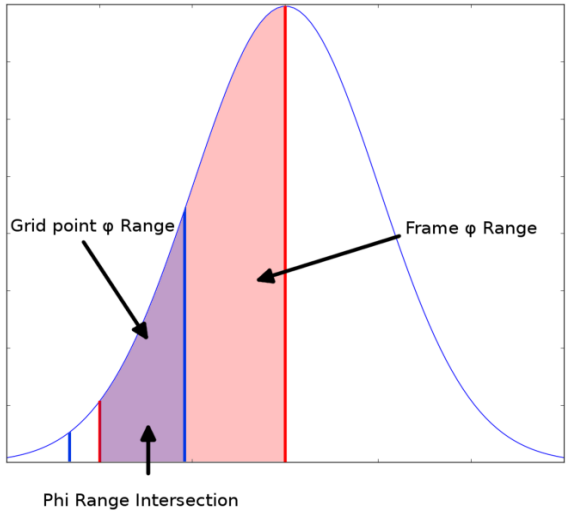


Counts are redistributed to Ewald sphere grid by computing fractional overlap of each pixel and Ewald sphere grid point

3D profile fitting phi gridding

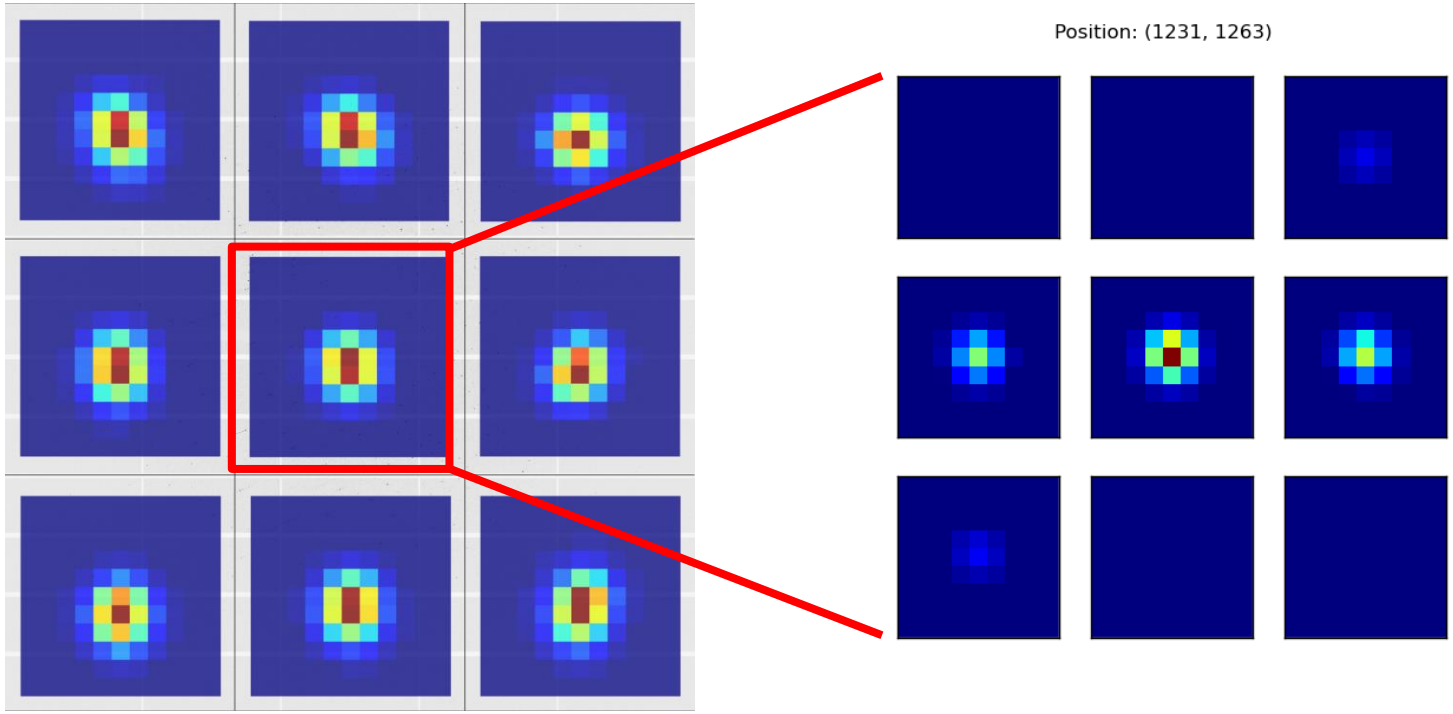


Frames are transformed to make reflection appear as if it took the shortest path through the Ewald sphere



Counts on each image are distributed by finding the angular overlap between each grid point and each image and integrating over the intersection

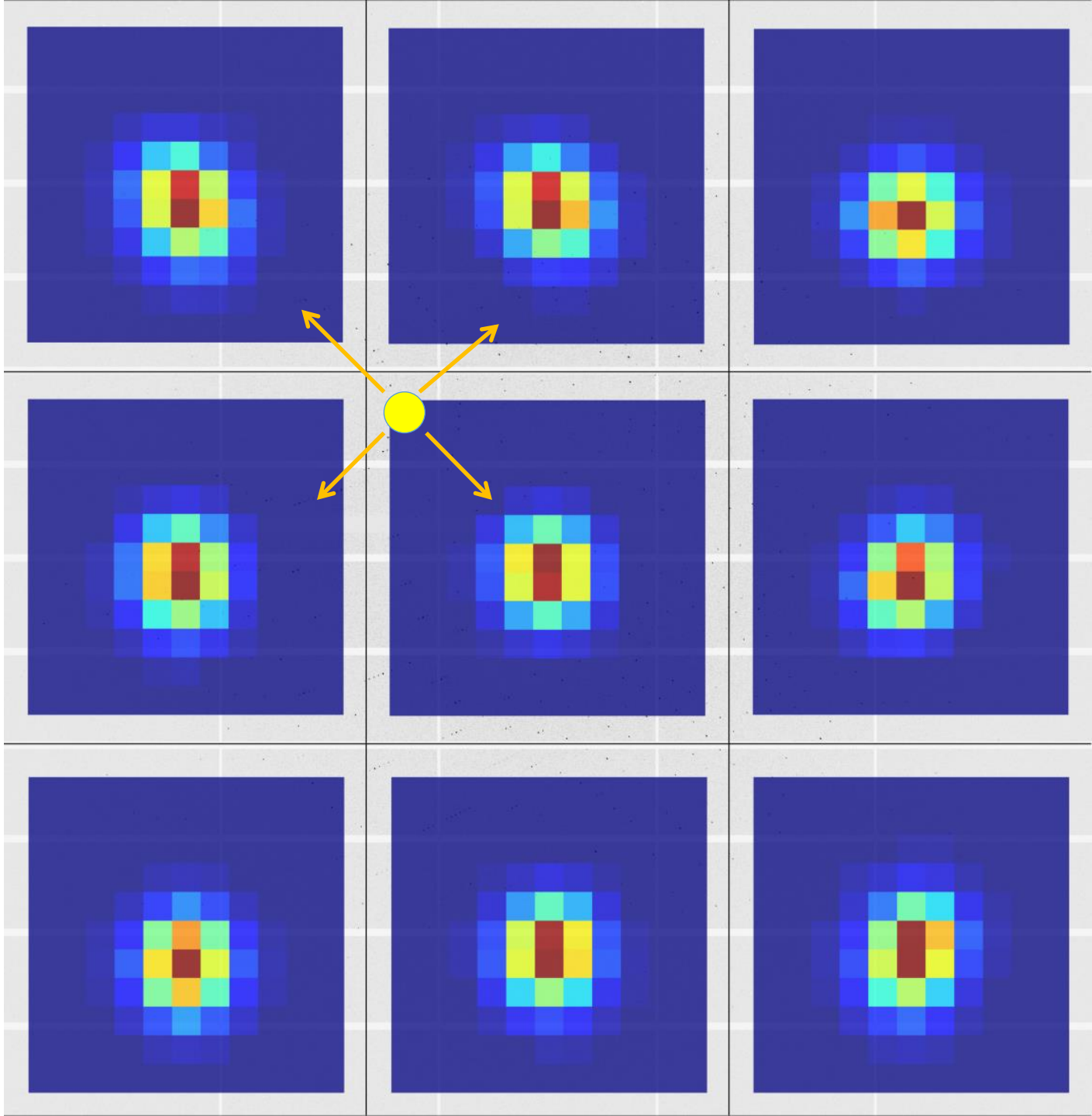
Building reference profiles



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

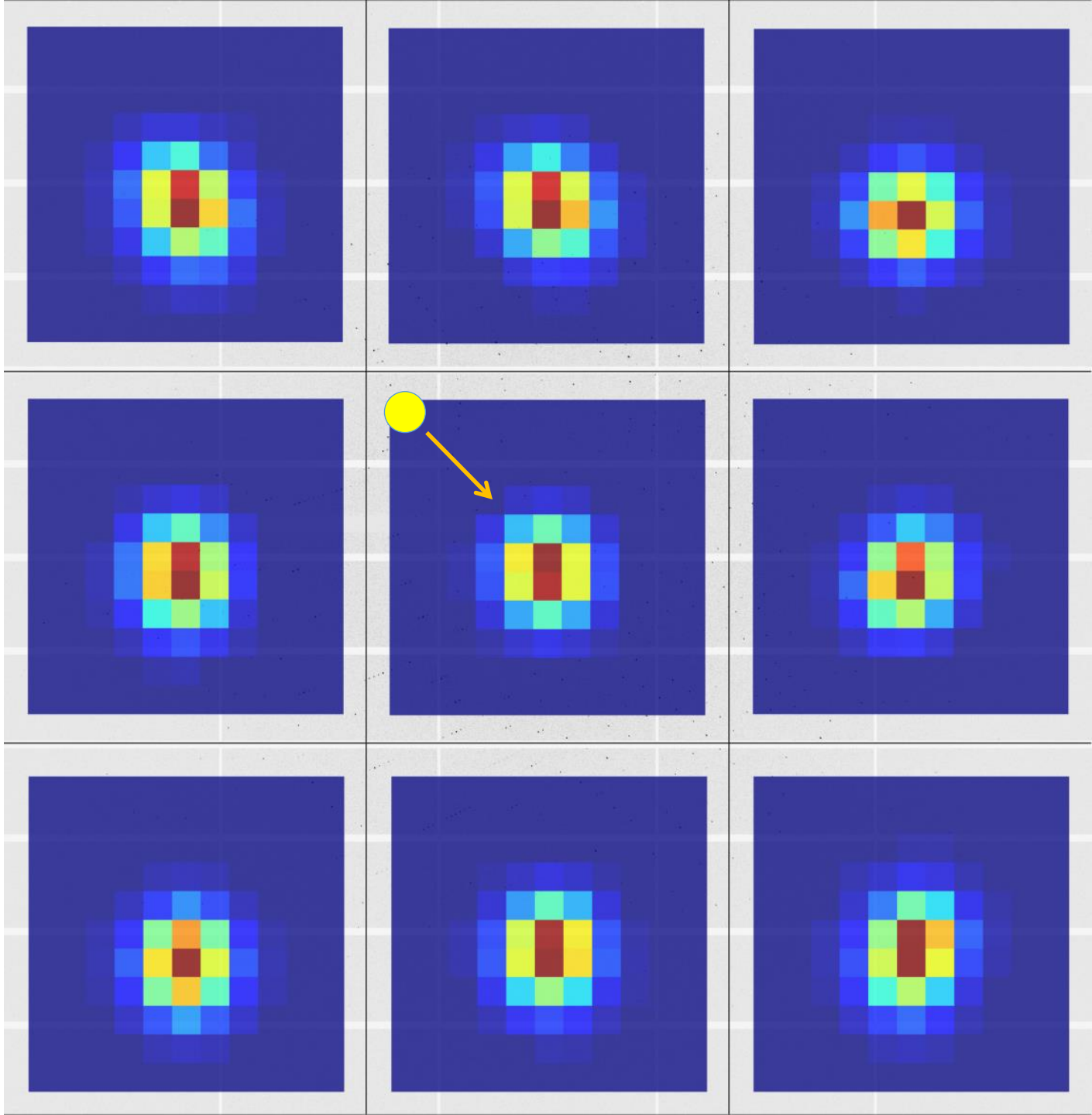
Building reference profiles

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



Fitting reference profiles

Each reflection is fitted against its closest reference profile



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.

```
The following parameters have been modified:

input {
  datablock = <image files>
}

--------------------------------------------------------------------------------
DataBlock 0
  format: <class 'dxtbx.format.FormatCBFMiniPilatusDLS6MSN100.FormatCBFMiniPilatusDLS6MSN100'>
  num images: 540
  num sweeps: 1
  num stills: 0
--------------------------------------------------------------------------------
Writing datablocks to datablock.json
```

Find Spots

Quick start guide

http://xia2.sourceforge.net/quick_start.html

Navigation

If you don't like reading manuals and just want to get started, try:

```
xia2 -2d /here/are/my/images
```

or:

```
xia2 -3d /here/are/my/images
```

or:

```
xia2 -dials /here/are/my/images
```

(remembering of course -atom X if you want anomalous pairs separating in scaling.) If this appears to do something sensible then you may well be home and dry. Some critical options:

Option	Usage
-atom X	tell xia2 to separate anomalous pairs i.e. I(+) ≠ I(−) in scaling
-2d	tell xia2 to use MOSFLM and Aimless
-3d	tell xia2 to use XDS and XSCALE
-3dii	tell xia2 to use XDS and XSCALE, indexing with peaks found from all images
-dials	tell xia2 to use DIALS and Aimless

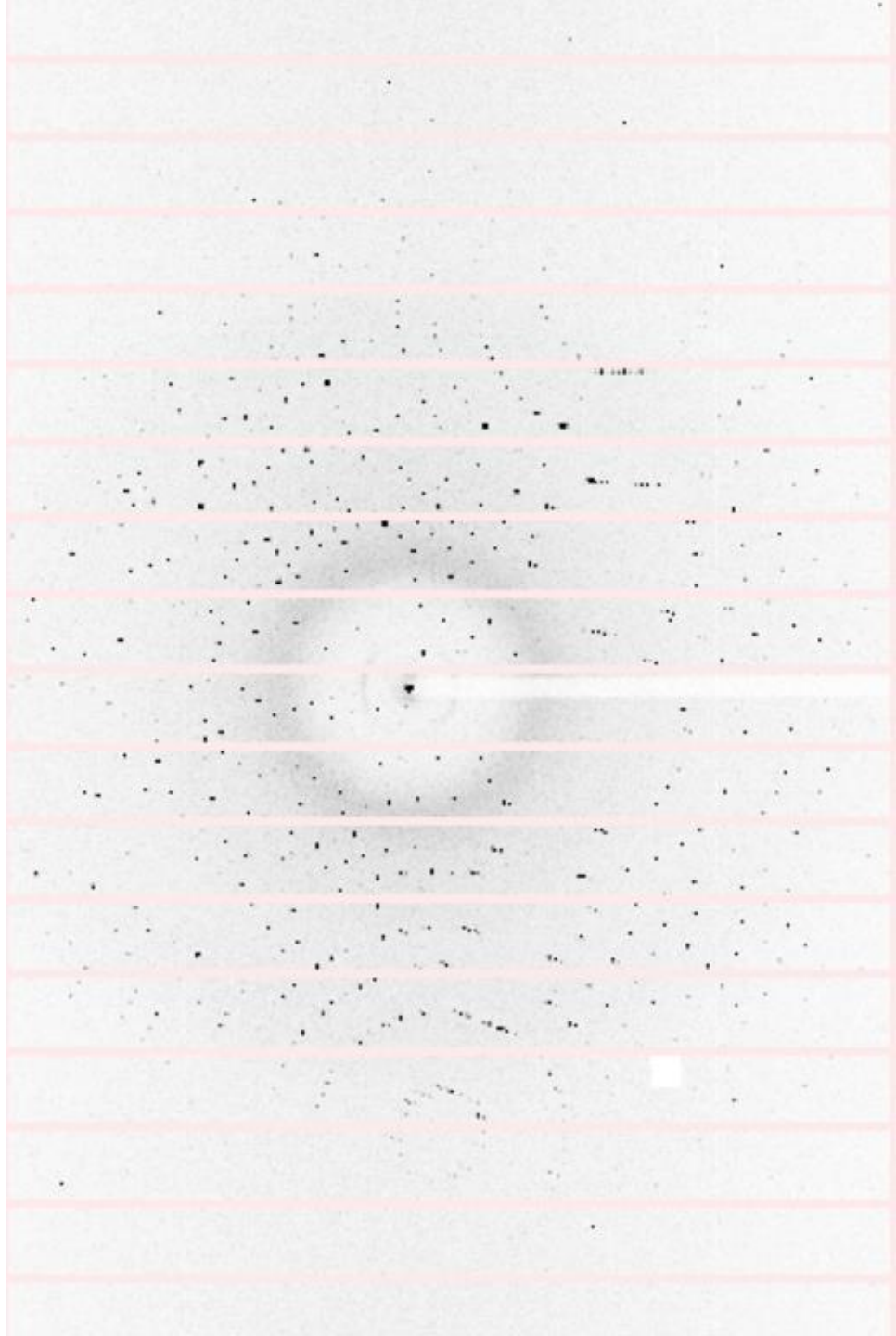
If this doesn't hit the spot, you'll need to read the rest of the documentation.

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

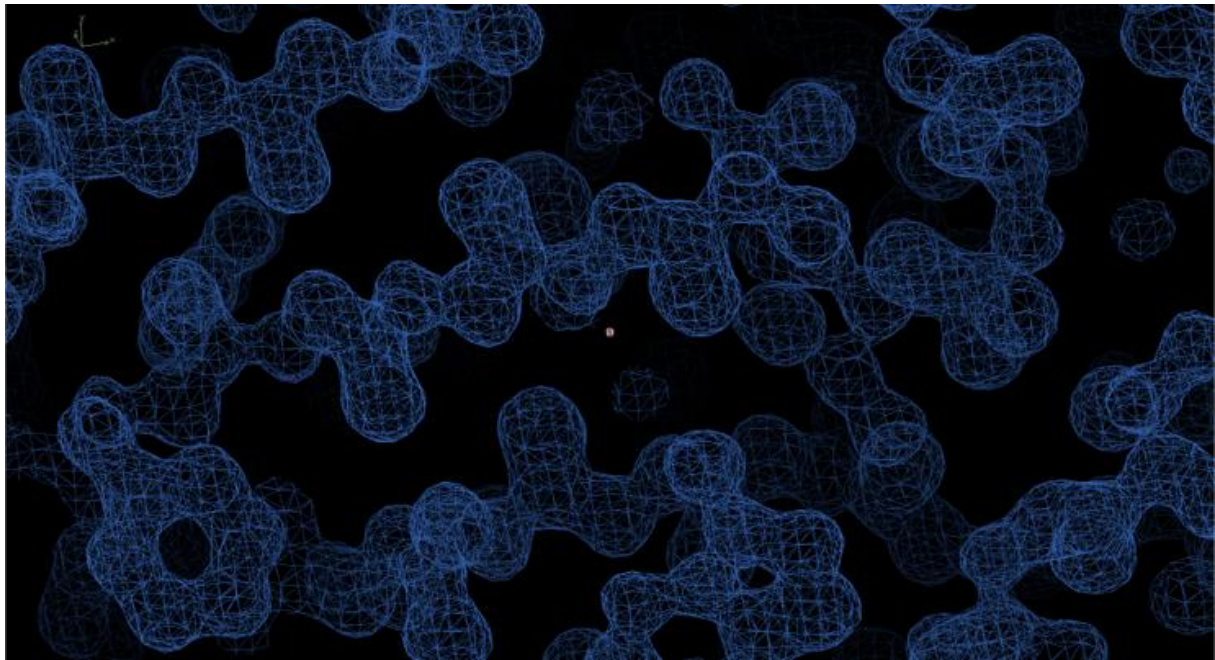


Data at low resolution

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	-	-



Not just data processing

DIALS: Diffraction Integration for Advanced Light Sources

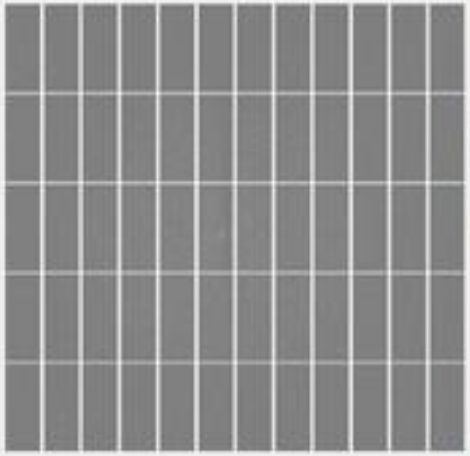
DIALS is a toolbox

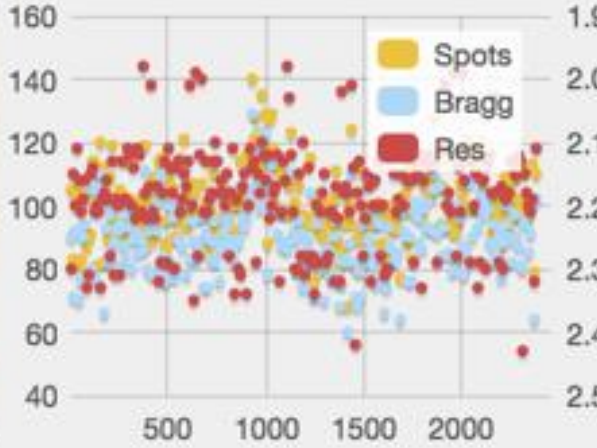
Full toolbox

```
dials.analyse_output
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.dqe
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.process
dials.python
dials.pythonw
dials.reciprocal_lattice_viewer
dials.refine
dials.refine_bravais_settings
dials.reflection_viewer
dials.reindex
dials.remove_absent_reflections
dials.rs_mapper
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
```

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

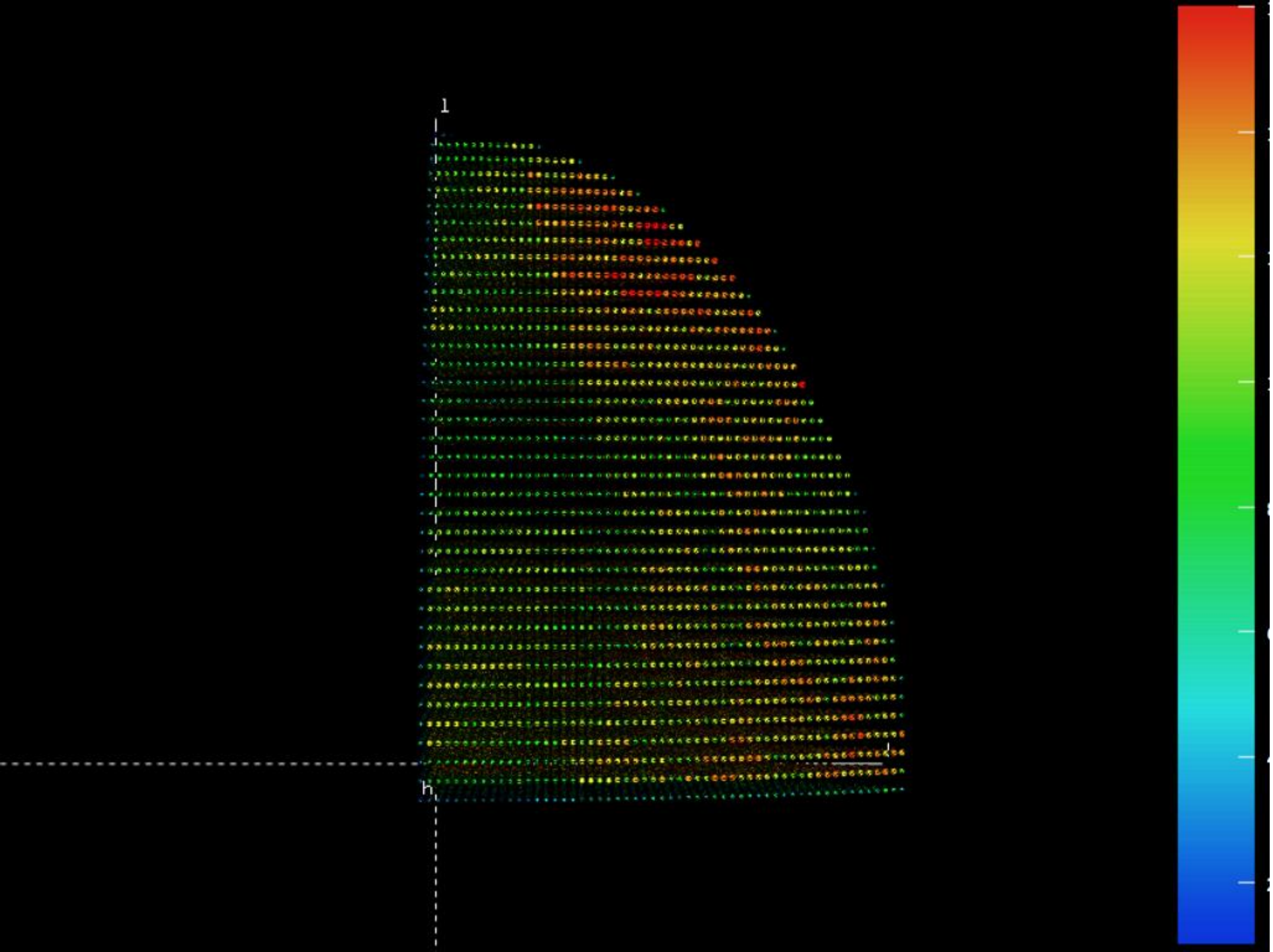
MTZ file | Log file | Lookup Cell

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

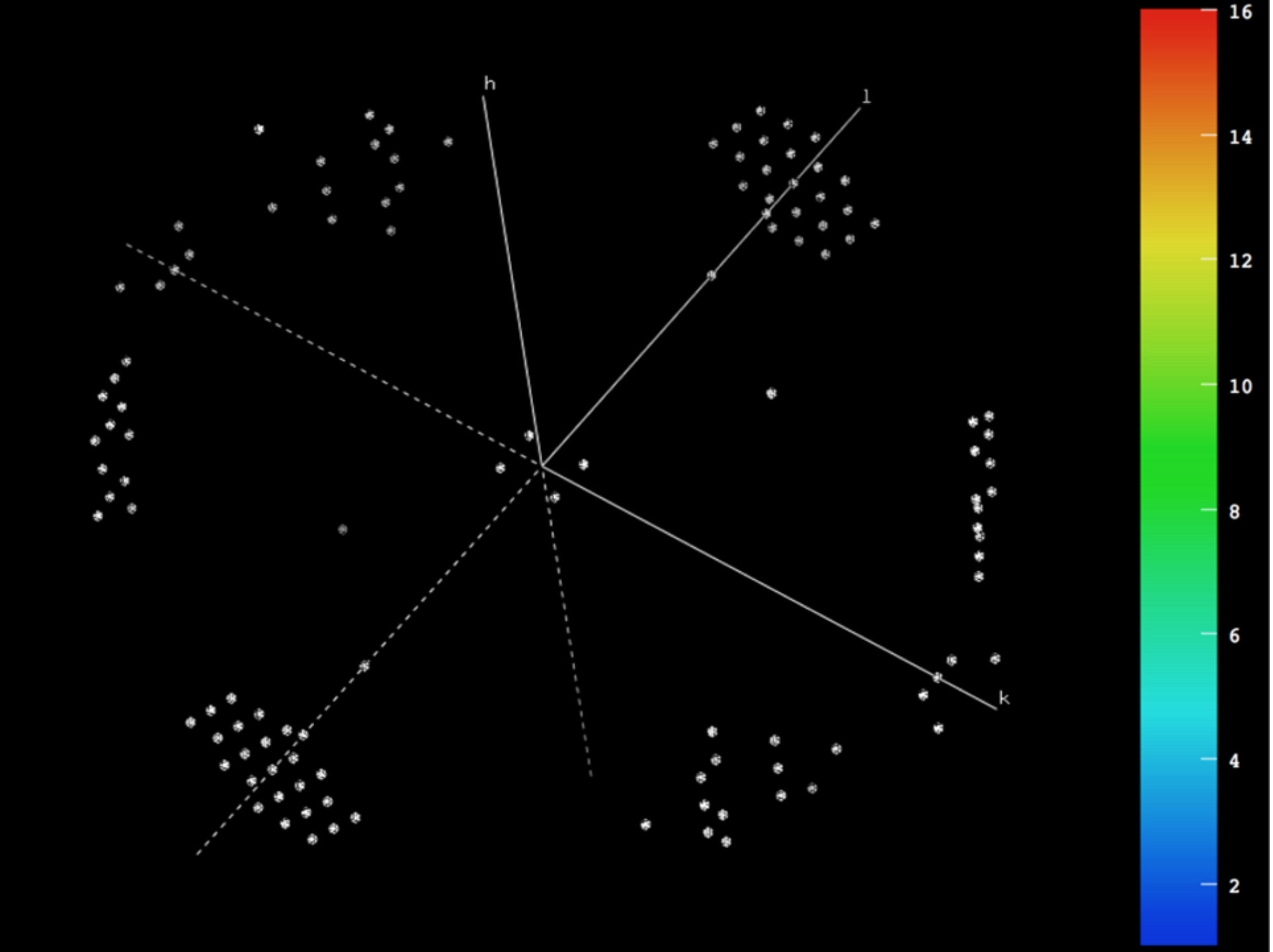
Shell	Observations	Unique	Resolution	Rmeas	I/sig(I)	Completeness	Multiplicity	Anom Completeness	Anom Multiplicity
outerShell	947	519	1.90 - 1.95	0.571	2.0	18.5	1.8	4.3	1.5
innerShell	6208	521	8.29 - 71.01	0.082	17.8	99.9	11.9	100.0	7.2
overall	341484	30333	1.90 - 71.01	0.148	11.8	81.4	11.3	78.7	5.8

Downstream Processing

Fast EP: ✓ Dimple: ✓



Multiplicity Viewer



Viewing missing reflections

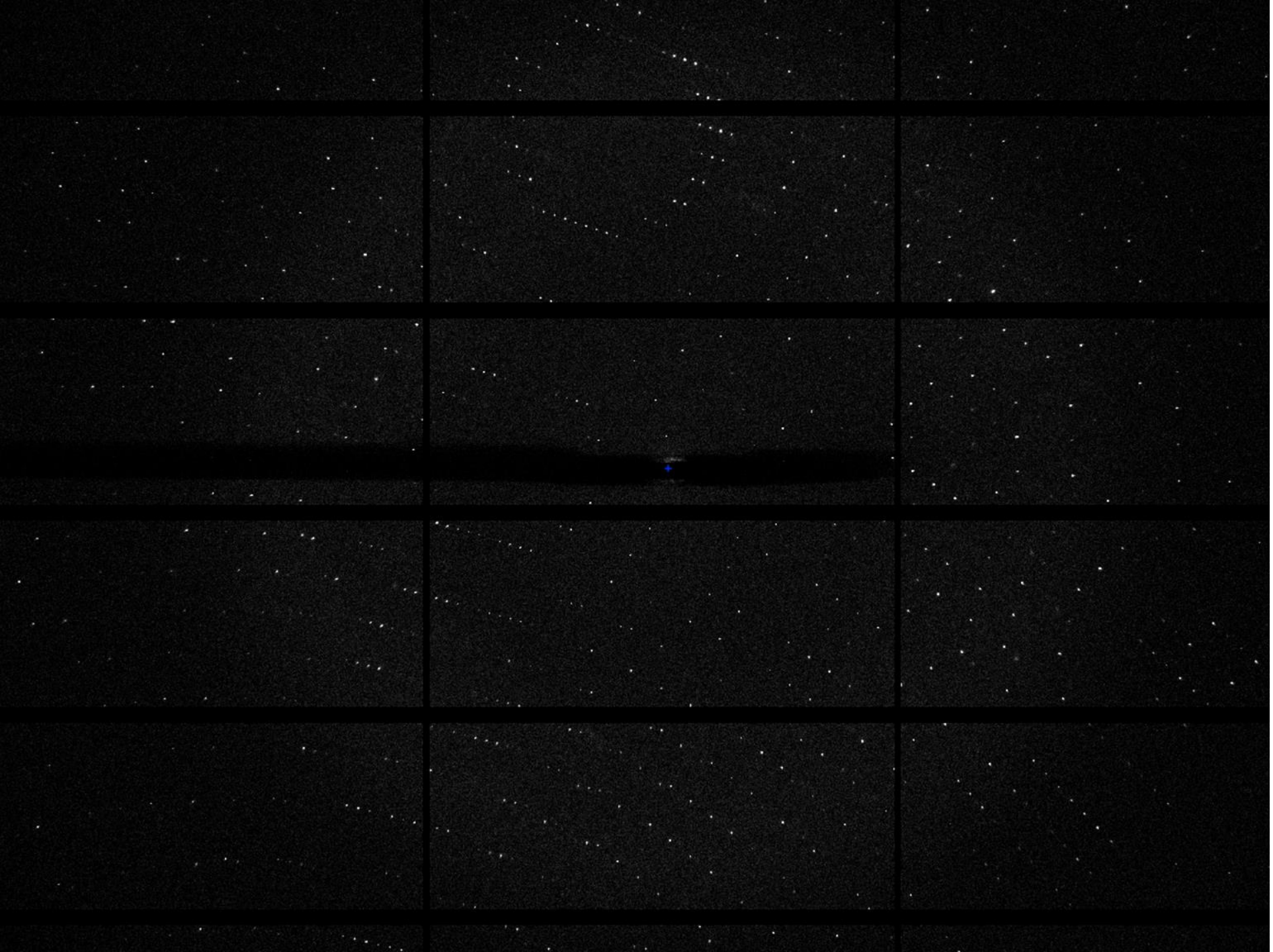


Image viewer:
Raw data

Image viewer: spots

Summary

- DIALS now used routinely for automated data processing via xia2
- xia2 is the “friendly” DIALS user interface for synchrotron data, and is bundled with DIALS
- Software available from dials.sf.net under BSD license
- Currently in “alpha” – your mileage may vary etc.
- Binary releases available for Mac and Linux (Windows coming soon)
- Will be included in CCP4

DIALS East– Diamond / CCP4

DIALS West – LBL / SLAC

Acknowledgements

- Gwyndaf Evans, Graeme Winter, James Parkhurst, Luis Fuentes-Montero, Richard Gildea, Markus Gerstel, Alun Ashton
- Andrew Leslie, Harry Powell, Phil Evans, Garib Murshudov
- Nick Sauter, Aaron Brewster, Nat Echols, Ian Rees, Nigel Moriarty, Ralf Grosse-Kunstleve
- Software users & Diamond scientists


diamond
Research Complex
at Harwell
CCP4
Science & Technology
Facilities Council
BERKELEY LAB
Lawrence Berkeley National Laboratory
DIALS
Diffraction Integration for Advanced Light Sources
BioStructX

BLEND

Management of Data from Multiple Crystals

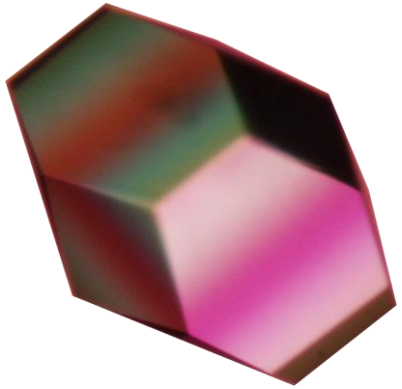
James Foadi
(author)

diamond

MPL

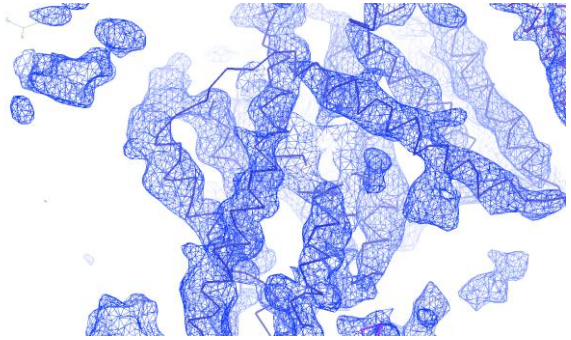
Imperial College
London

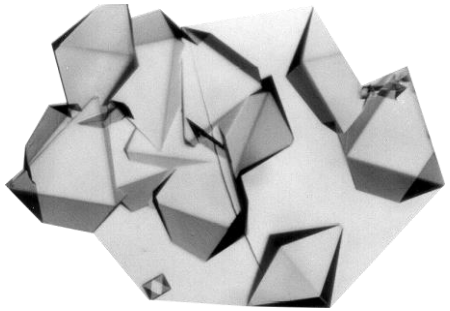
In a nutshell



Single Crystal

Established Methods and Software

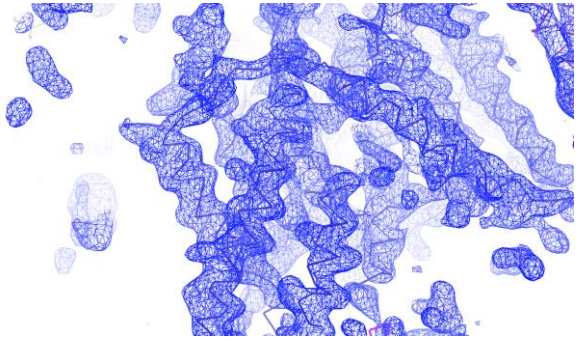




Multiple Crystals

New Methods and Software

BLEND



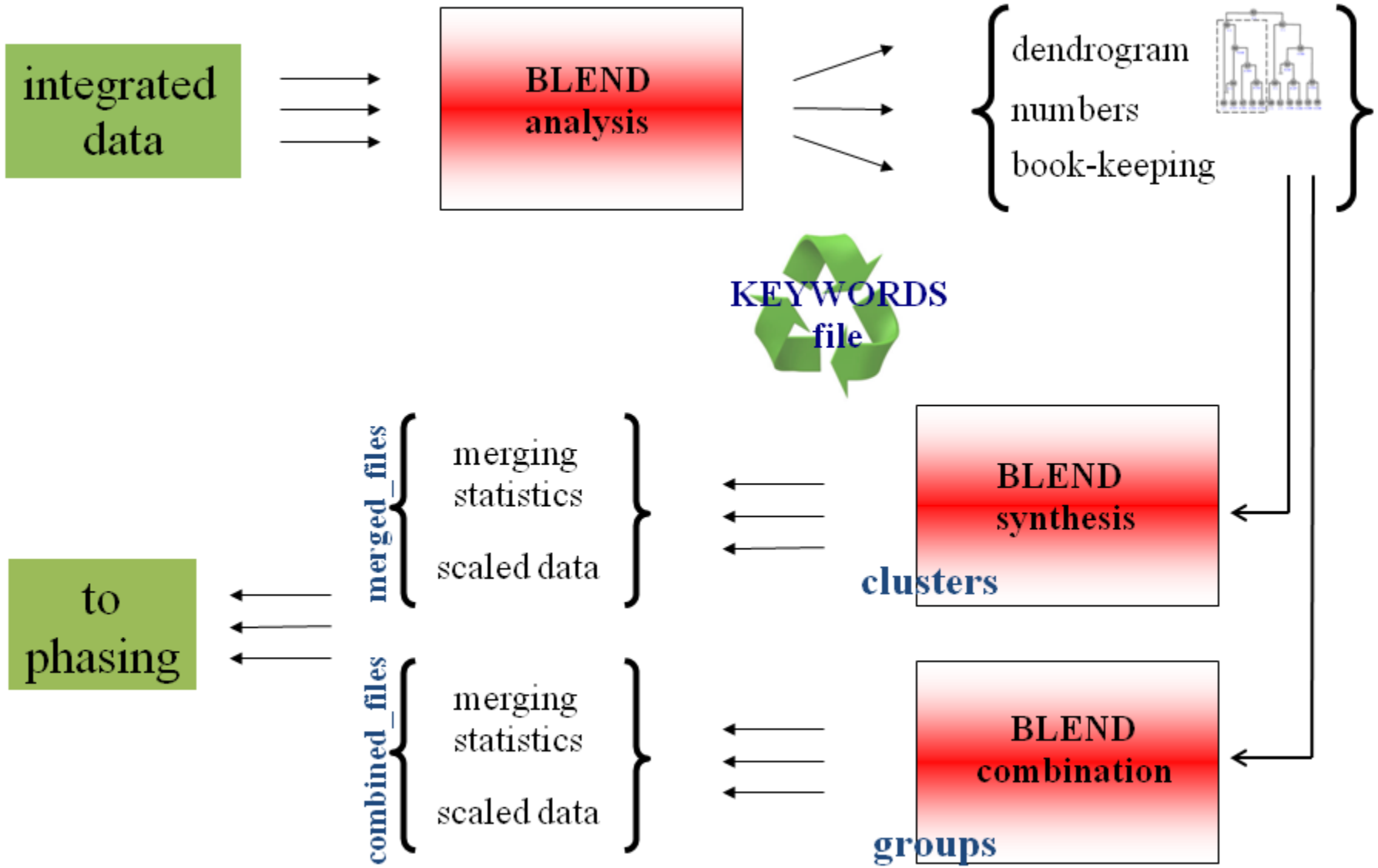
What is BLEND

Tool to manage/process multiple data sets from multiple crystals (scaling)

Help to reduce the large number of data implied by the combination of multiple data sets:

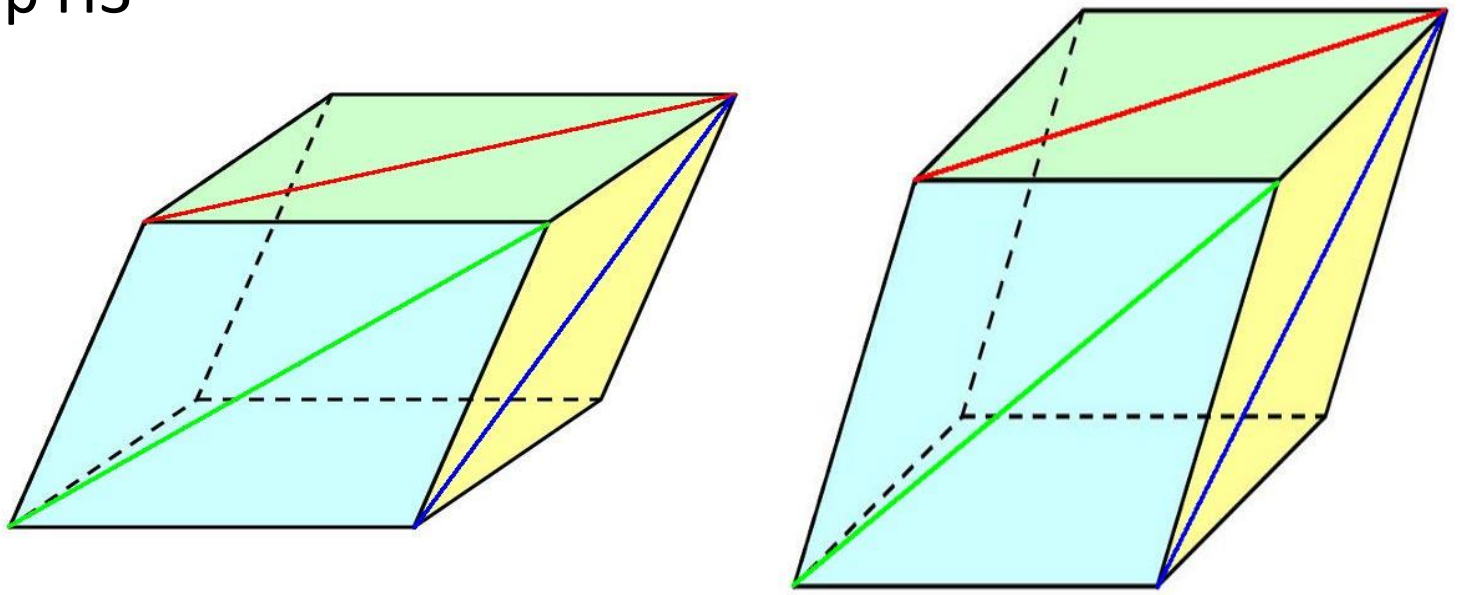
a) cluster analysis on cell parameters;
b) scaling all cluster produced
c) filtering out data sets

“Environment” to help knowing better the quality of your data


integrated data
BLEND analysis
dendrogram
numbers
book-keeping
KEYWORDS file
merged_files
merging statistics
scaled data
BLEND synthesis
clusters
to phasing
combined_files
merging statistics
scaled data
BLEND combination
groups

Example

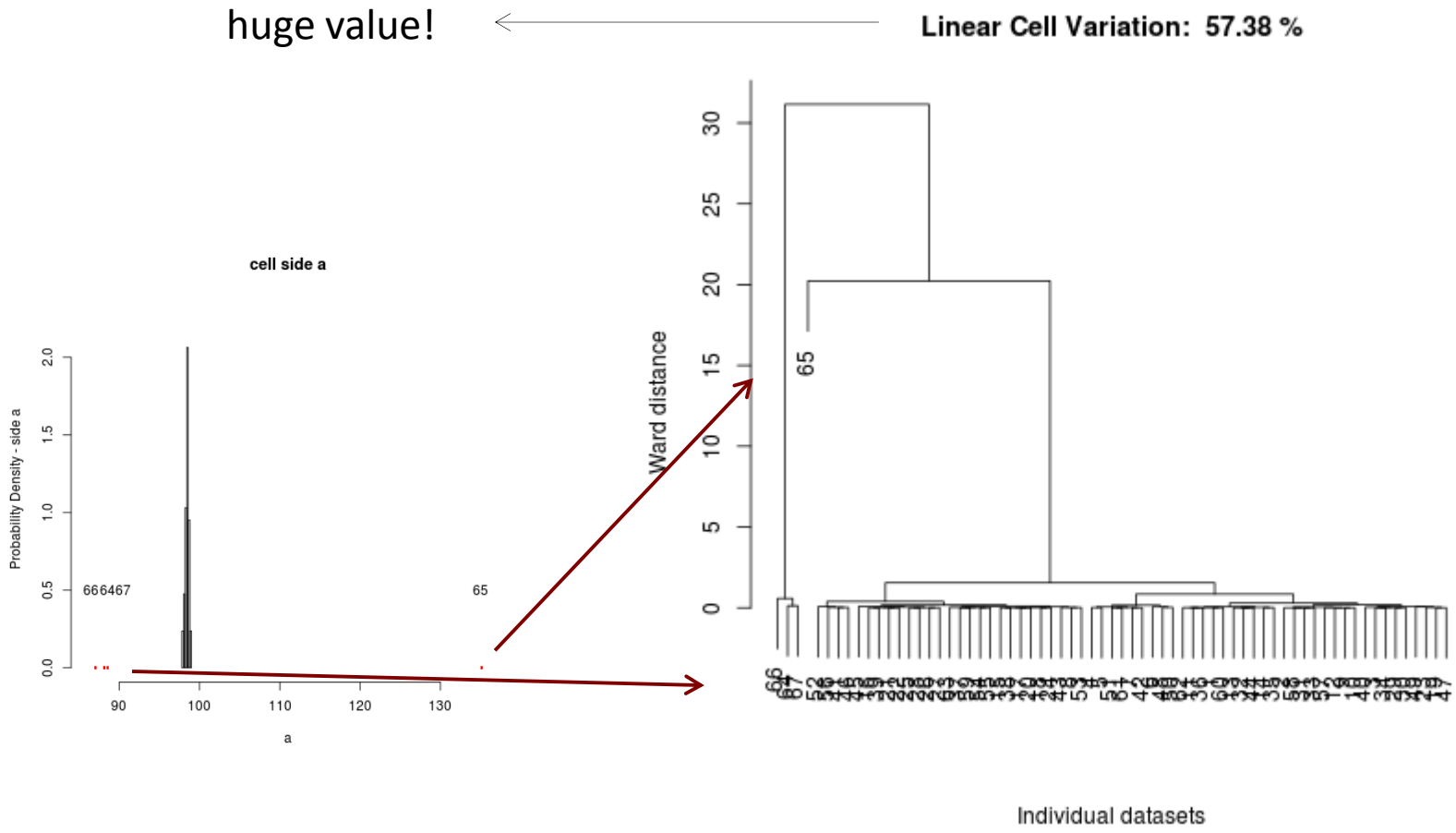
- *in situ* Tellurite resistance protein homolog
- 67 wedges / 56 crystals ; 6 to 10 degrees oscillation
- Space group H3



Meaning of **LCV**

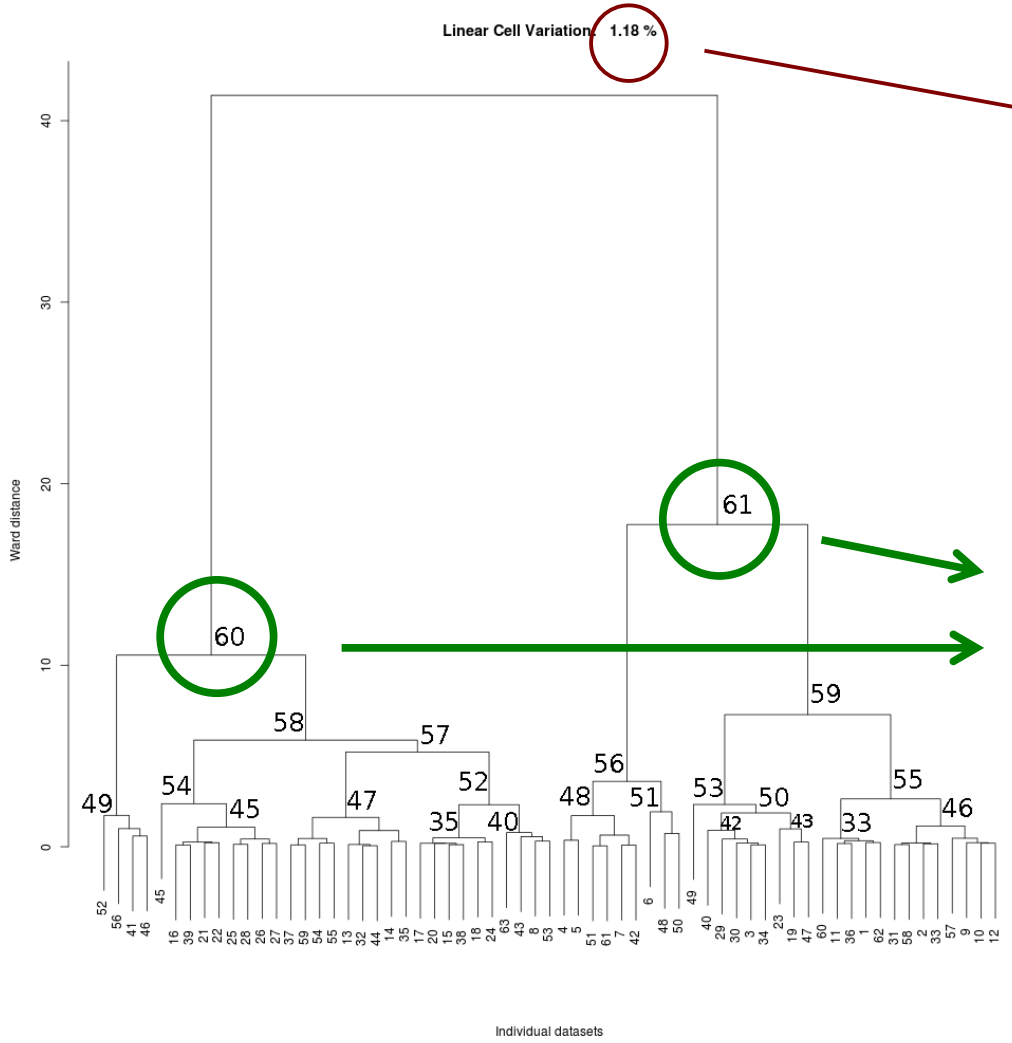
Linear Cell Variation

Dendogram



64, 65, 66 and 67 can be filtered out

Dendogram

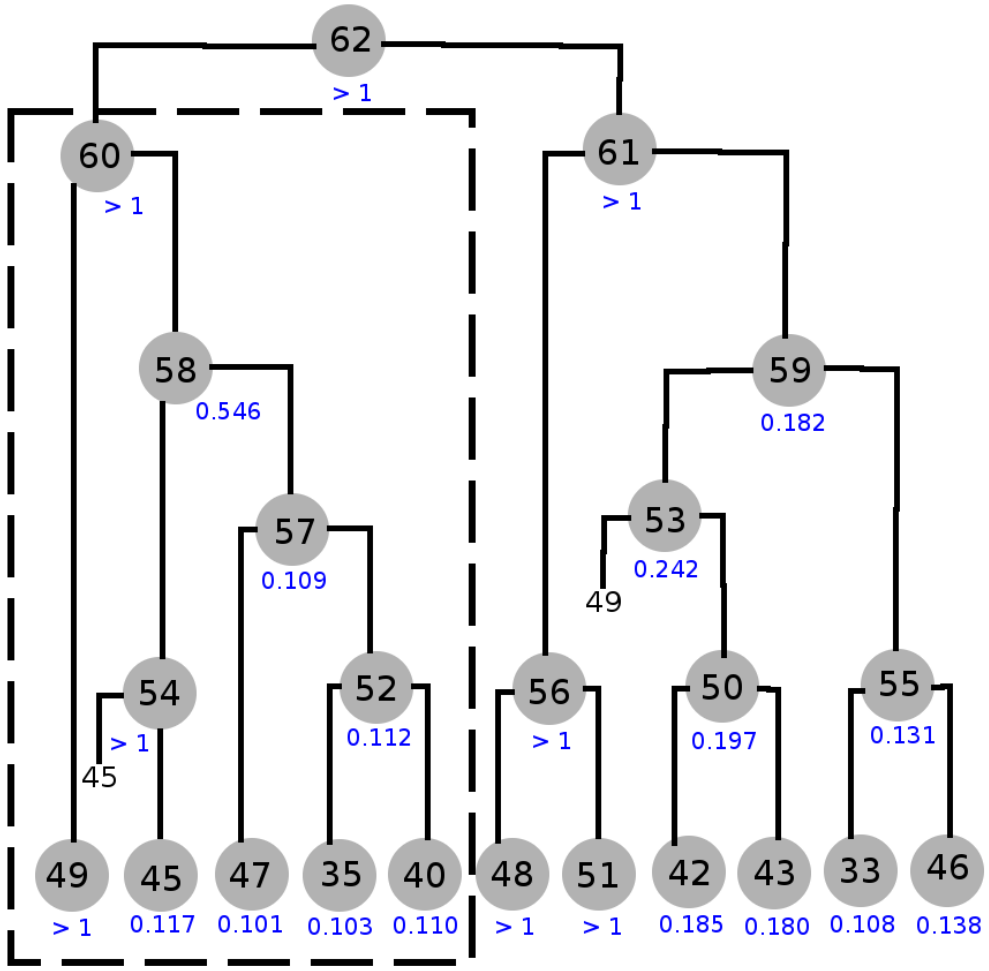


LCV = 1.18%
a reasonable value

The two different branches might point to some mild non-isomorphism.
But …
Cluster 60 is 89.7% complete
Cluster 61 is 70.7% complete

We select cluster 60

Annotated dendogram



Cluster 57 shows a good value of Rmeas.

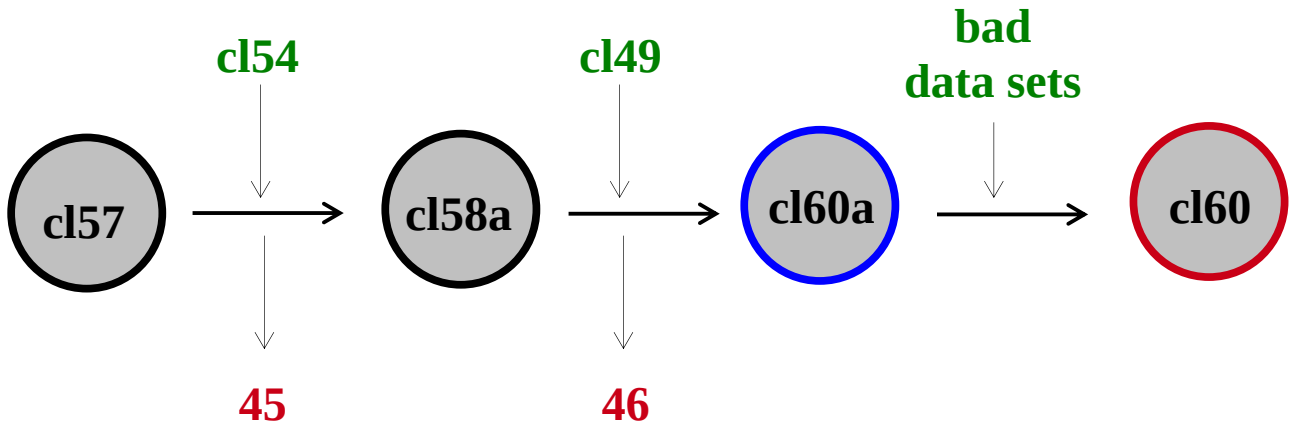
Cluster 58 shows a bad value of Rmeas.

Cluster 54 is “polluting” the good quality of cluster 57.

Cluster 60 is “polluted” by cluster 58, but could also be “polluted” by cluster 49.

→ **We need to combine cluster 57 with groups of data sets or individual data sets of cluster 54, in order to form a “cleaned” cluster 58. Then we combine with data sets from cluster 49, in ordert o form a “cleaned” cluster 60.**

Filtered dendogram



Rmeas	0.109	0.113	0.121	> 1
Rpim	0.055	0.050	0.052	> 1
Compl.	81.9%	84.1%	89.1%	89.7%

Further reading

First published paper:

J. Foadi, P. Aller, Y. Alguel, A. Cameron, D. Axford,
R. L. Owen, W. Armour, D. G. Waterman, S. Iwata and G. Evans
Clustering procedures for the optimal selection of data sets from multiple crystals in macromolecular crystallography
Acta Cryst. (2013), D**69**, 1617-1632

See also:

BLEND user guide
Diamond web site and in BLEND installation package
(or on request)

Tutorials
In BLEND installation package
(or on request)

Acknowledgements

BLEND RESEARCH

James Foadi

Gwyndaf Evans (PI)

Pierre Aller

IN-SITU MEMBRANE PROTEIN

Yilmaz Alguel

Nien-Jen Hu

Danny Axford

Gwyndaf Evans

Kostantinos Beis

Hassanul Choudhury

FUNDING

Wellcome Trust

MPL SUPPORT

Isabel DeMoraes

Momi Iwata

So Iwata

I24 SUPPORT

Danny Axford

DIAMOND SOFTWARE SUPPORT

Alun Ashton

Karl Levik

Thanks for listening!

http://dials.sf.net