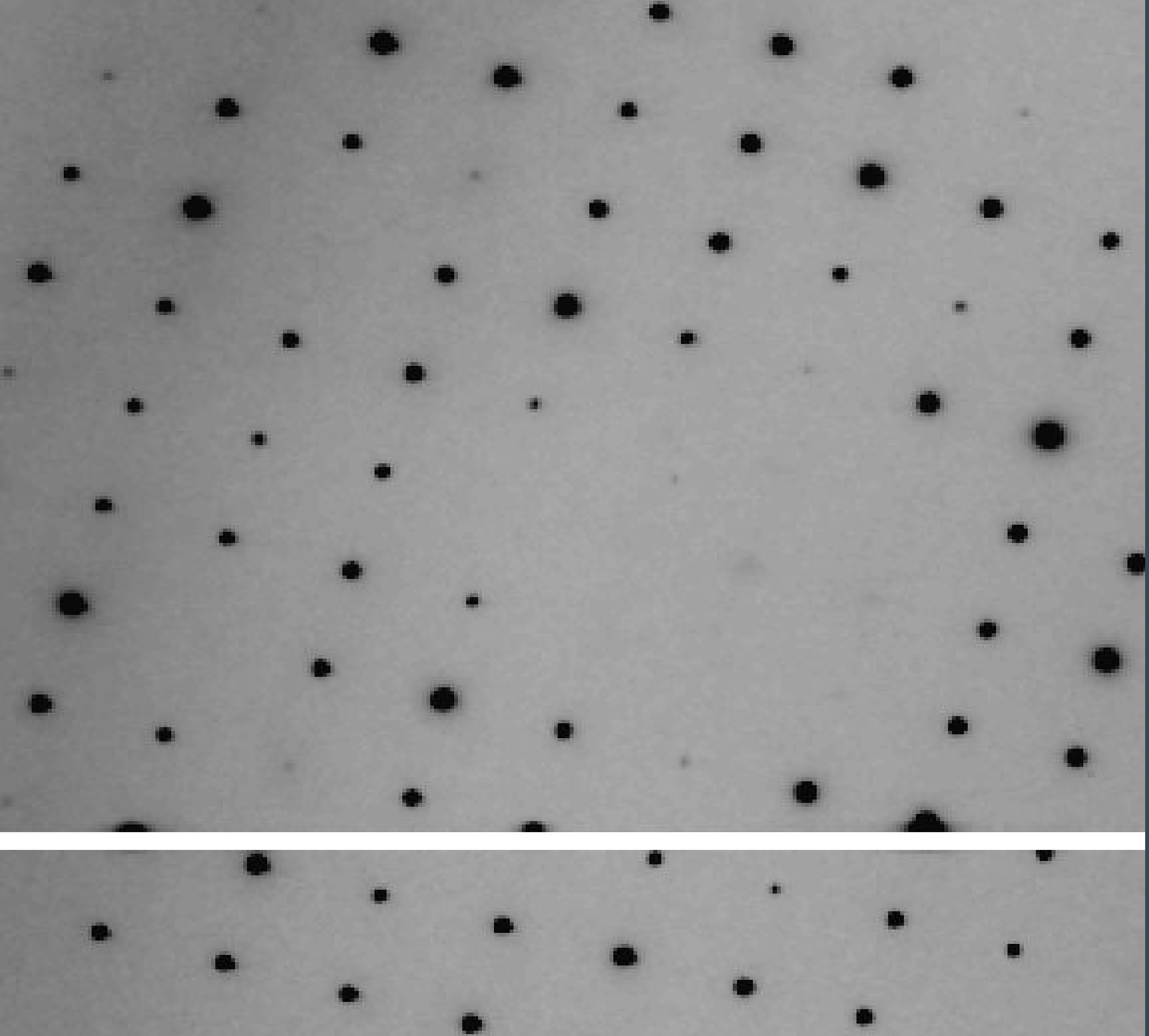
Data processing with DIALS

•••

James Parkhurst

CCP4 Spring-8 workshop, January 2017

What are we doing and why are we doing it?



Compute the intensity of each Bragg spot in a set of diffraction images

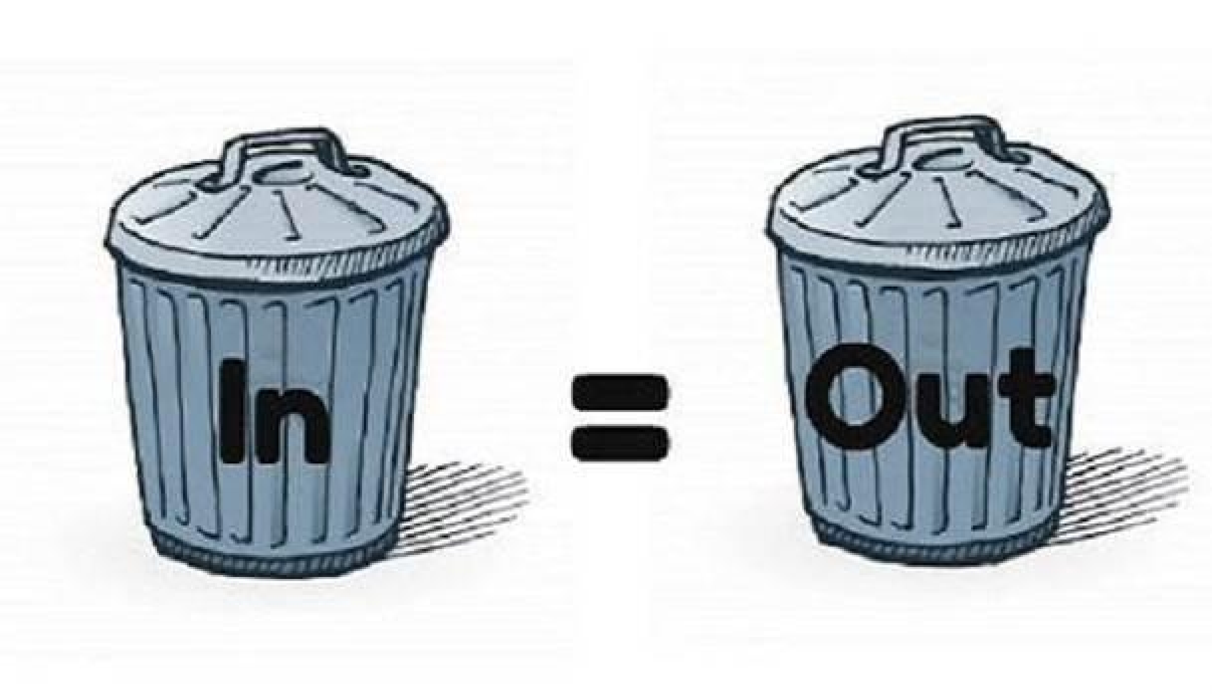
$$|F_{hkl}| = \sqrt{\frac{K I_{hkl}}{Lp}}$$

K = constant for given crystal
L = Lorentz factor
p = polarization factor

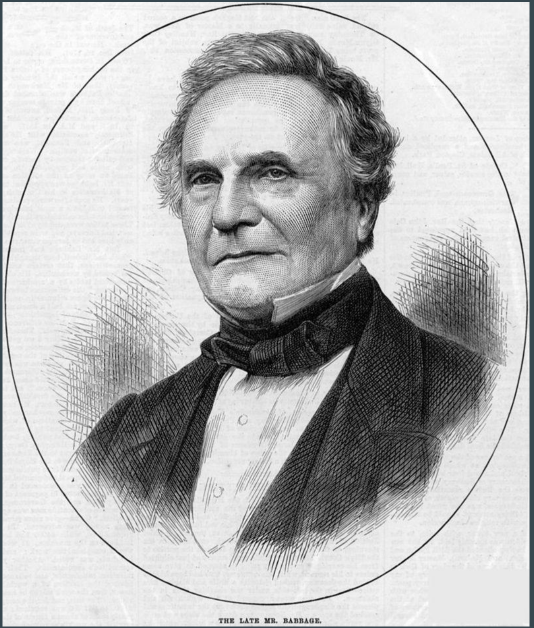
$$\rho(x, y, z) = \frac{1}{V} \sum_{h} \sum_{k} \sum_{l} F_{hkl} e^{-2\pi i(hx+ky+lz)}$$

Electron density at every point in the cell depends on the intensity of every reflection. We need to measure our intensities as well as possible!

Warning: garbage in, garbage out



Data collection is the last experimental stage; if you collect bad data you are stuck with it! Data processing programs won't be able to rescue you!



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

DIALS overview

Acknowledgements

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 contro
package *XDS* for t
described in the con
include automatic de
range and recognitic
Moreover, the limita
number of correction
pixel contents have b
been restructured fo
and completeness of
measurement.

1. Functional speci

The program packag
developed for the re
recorded on a plana
monochromatic X-ra
XDS accepts a
rotation images from
and multiwire area
metrics and produces
of the reflections occ
way. The program as
positive amount of c
incident beam and cr
imposes no limitati
directions of the rot
oscillation range cov

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The finer things in

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
sensitive detectors can be cha
depending on whether the rotatic
is greater than or less than the c
The expectations and consequen
and thin images in terms of spa
X-ray background and $I/\sigma(I)$ a
software suite for processing
introduced, and results from *d*
those from another popular pack

1. Introduction

Two-dimensional position-sensiti
for many years in X-ray diffract
cular, data from crystals of macr
oligonucleotides and their cor
acquired with an area detector
obsolete), a multi-wire system
recently commercialized char
coupled to a phosphor-coated fib
detectors, the crystal, centered in
oscillated around a single axis th
~2.0°, while counts from diffract
for a specified time. At the en
detector is read out and the cou
two-dimensional array with each
to a distinct position on the d

Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

The integrati

Andrew G. W. Leslie

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

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

The objective of any
produce from a set of
with their associated
uncertainties), togeth
crystal unit-cell param
reliable, but should i
intervention. The pro
three stages. The first
parameters and the o
parameters may indica
The second step is to r
parameters and also t
known as post-refinem
images, which consists
reflections on each in
intensity of each refle
out while simultaneou
parameters. Basic fea
each of these three s
with reference to the

1. Introduction

The collection of mac
gone dramatic advan
advent of two-dimensi
and CCDs, crystal cry
monochromatic and

Centre National de la Recherche Scientifique
Université Paris-Sud

Laboratoire pour l'Utilisation du Rayonnement Electromagnétique

Proceedings

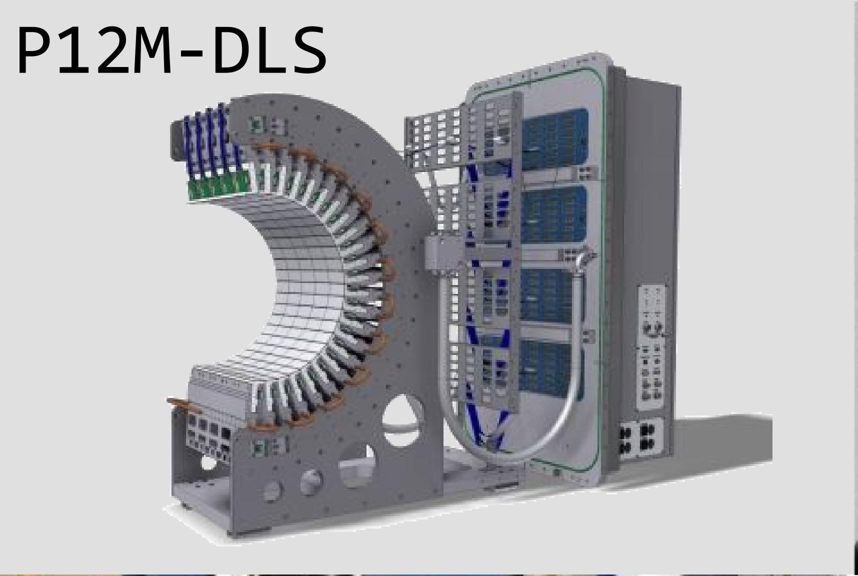
of the EEC Cooperative Workshop

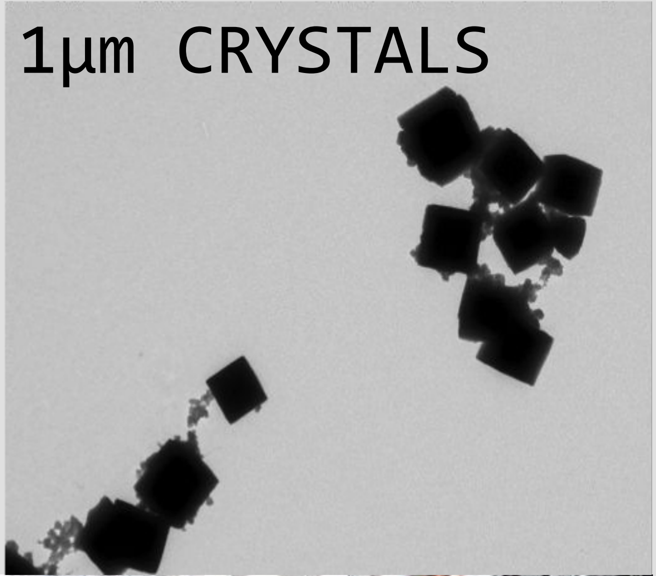
on Position-Sensitive Detector Software

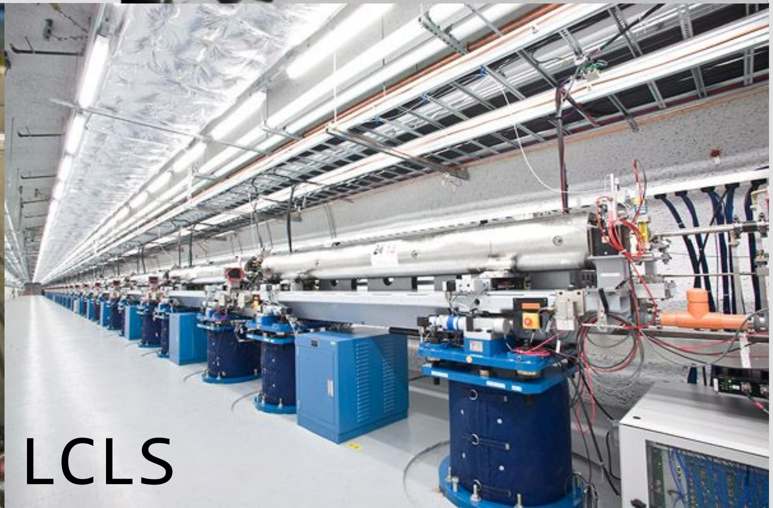
(Phases I & II)

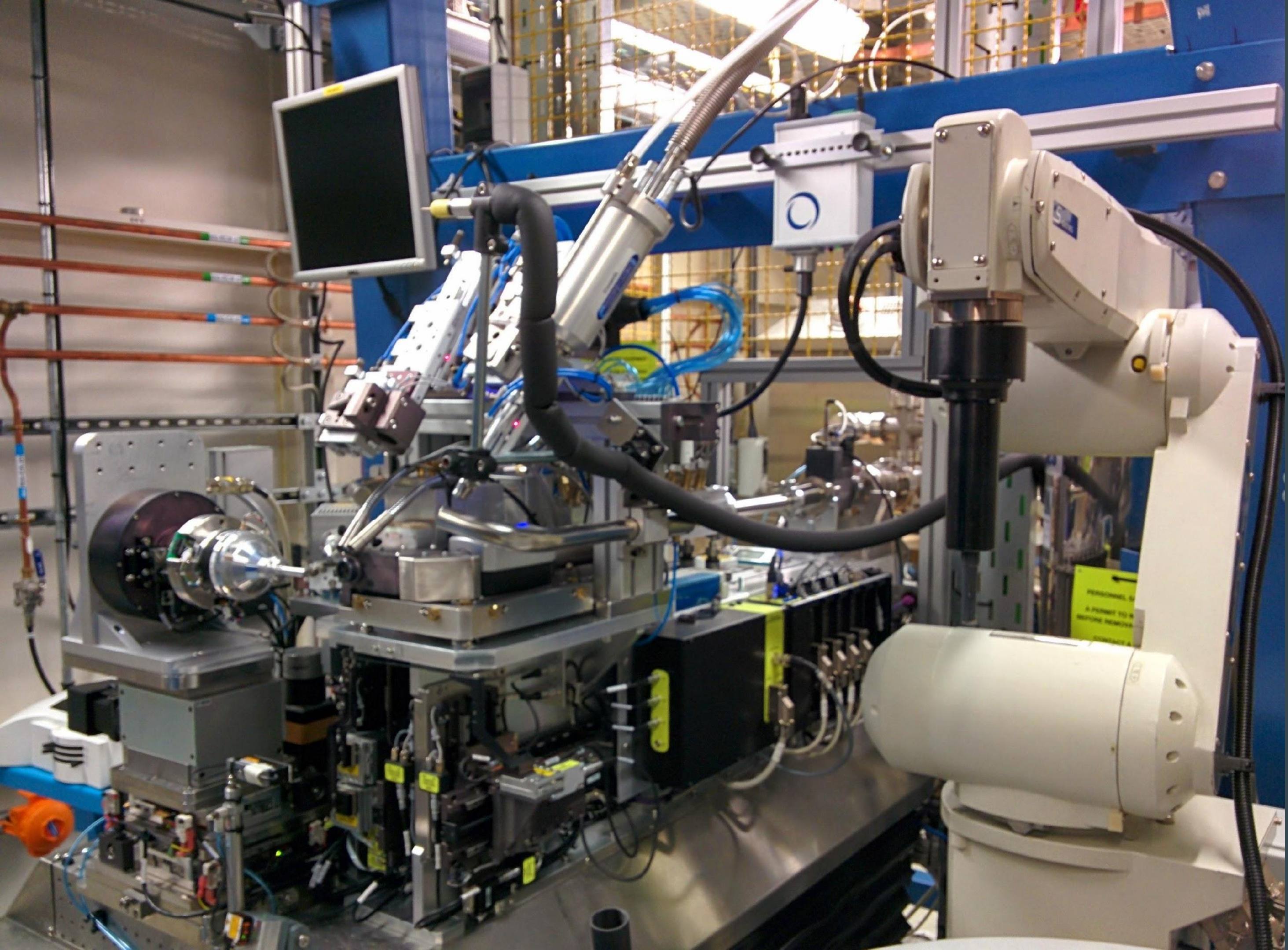
held at L.U.R.E. from May 26 to June 7, 1986.

New Challenges

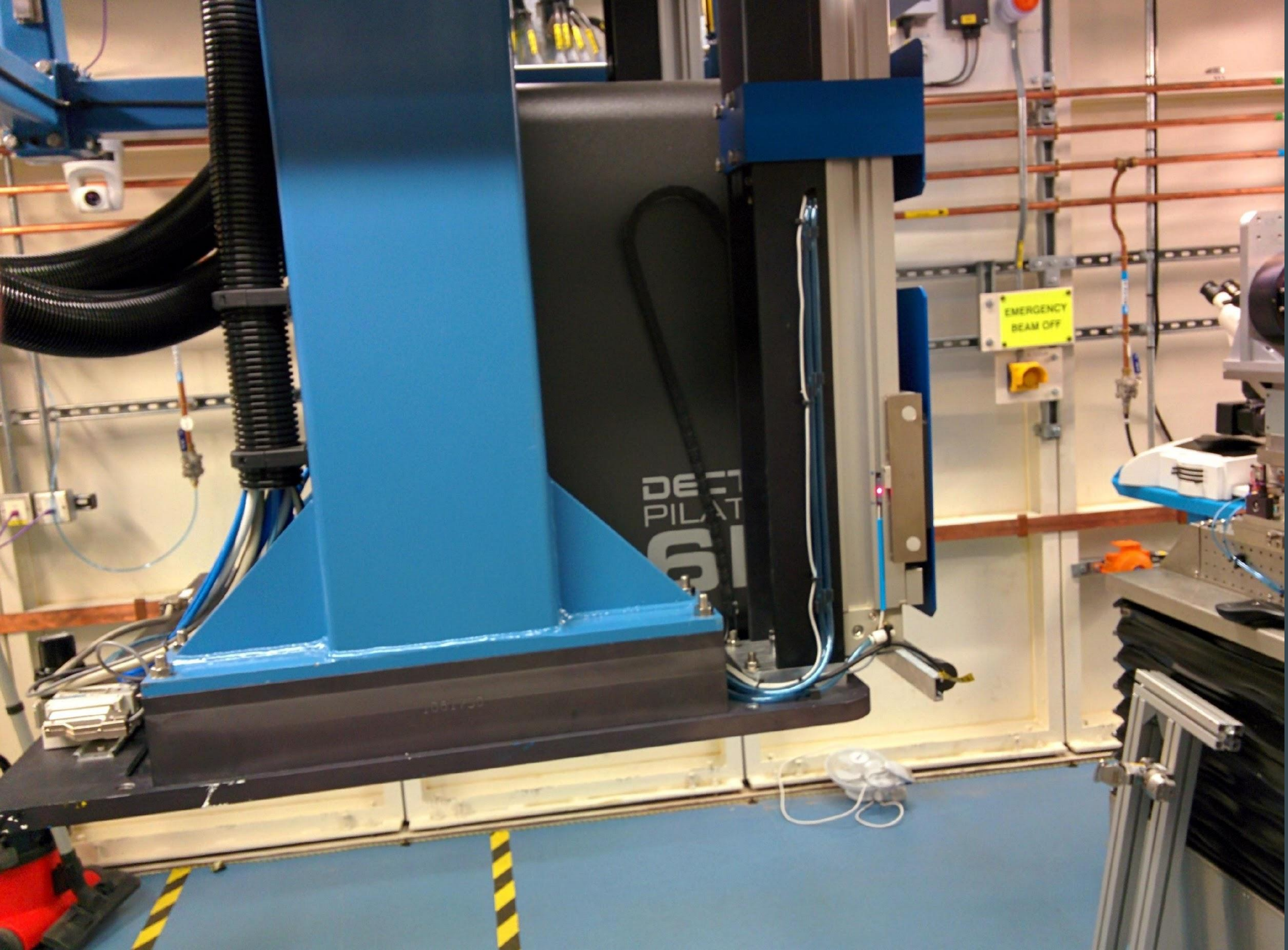




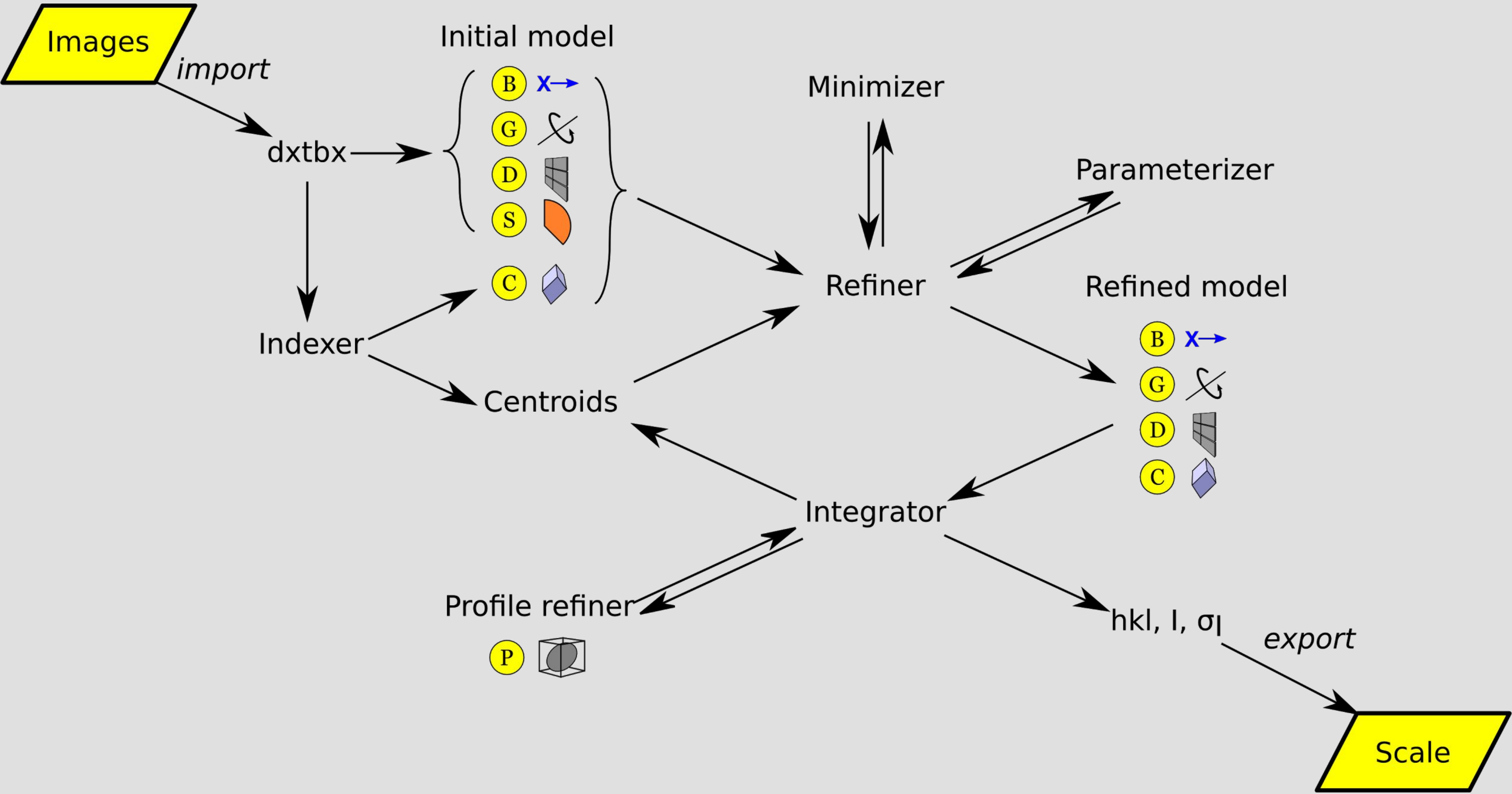




Current Challenges



Current Challenges



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

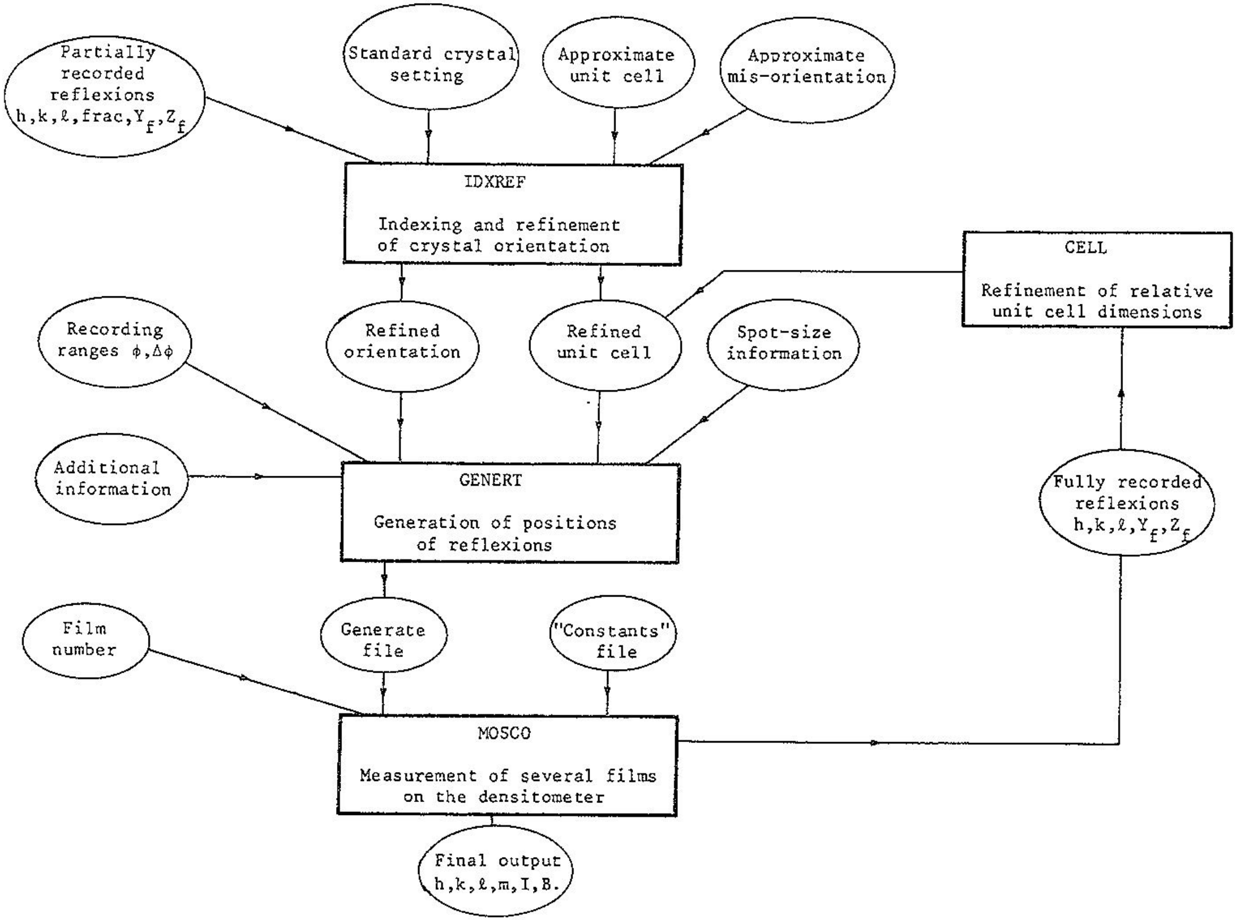


Fig. 10.1. Simplified flow-diagram of the Cambridge system, showing the inter-relation of the component programs, IDXREF, GENERT, MOSCO and CELL.

Main DIALS programs

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

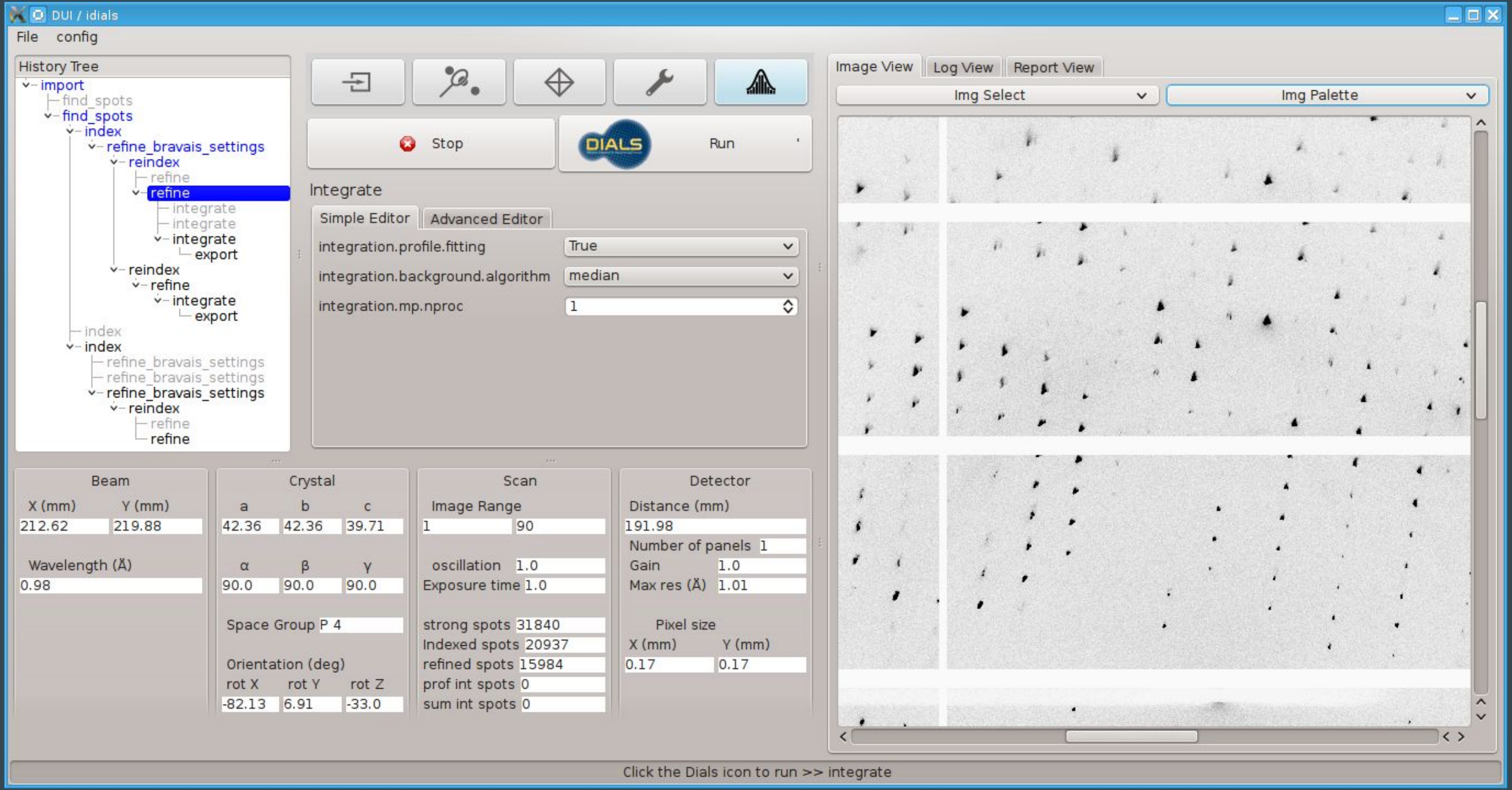
Then onwards into the CCP4 data processing pipeline:
`POINTLESS` → `AIMLESS` → `CTRUNCATE...`

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?

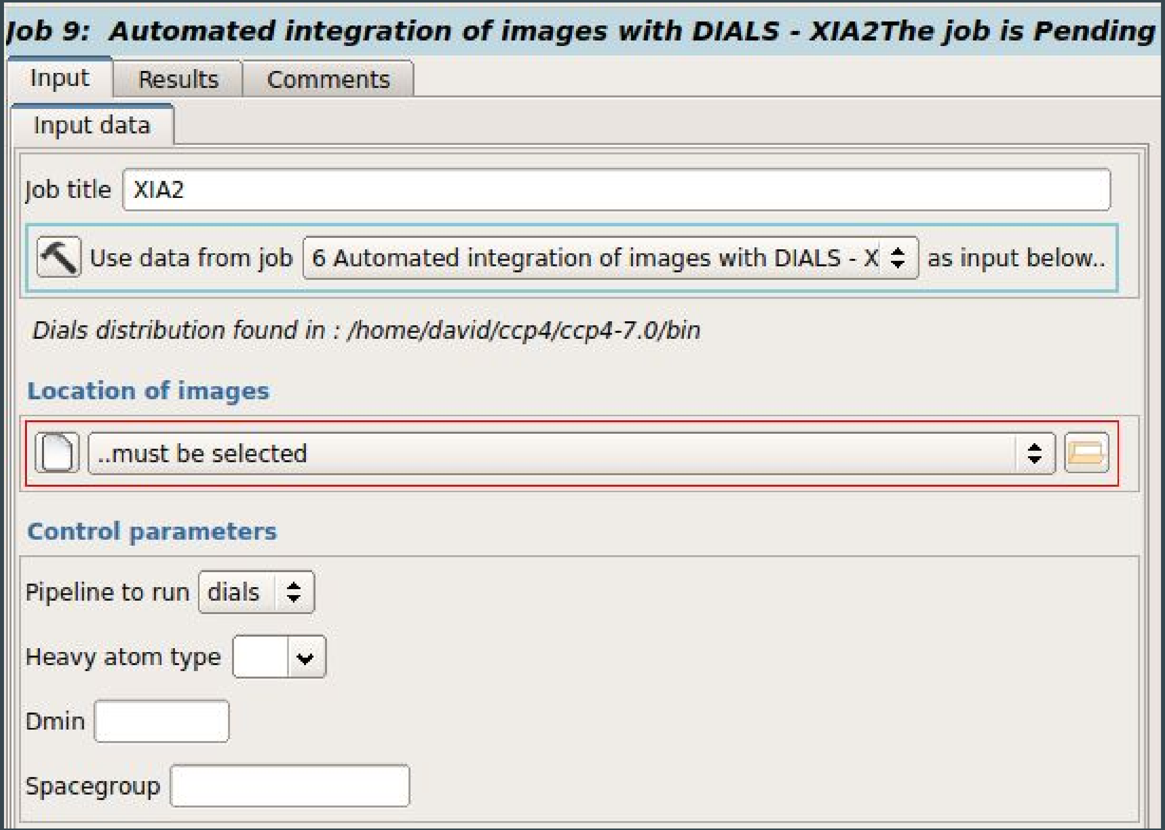
Future: DIALS GUI (currently in development)



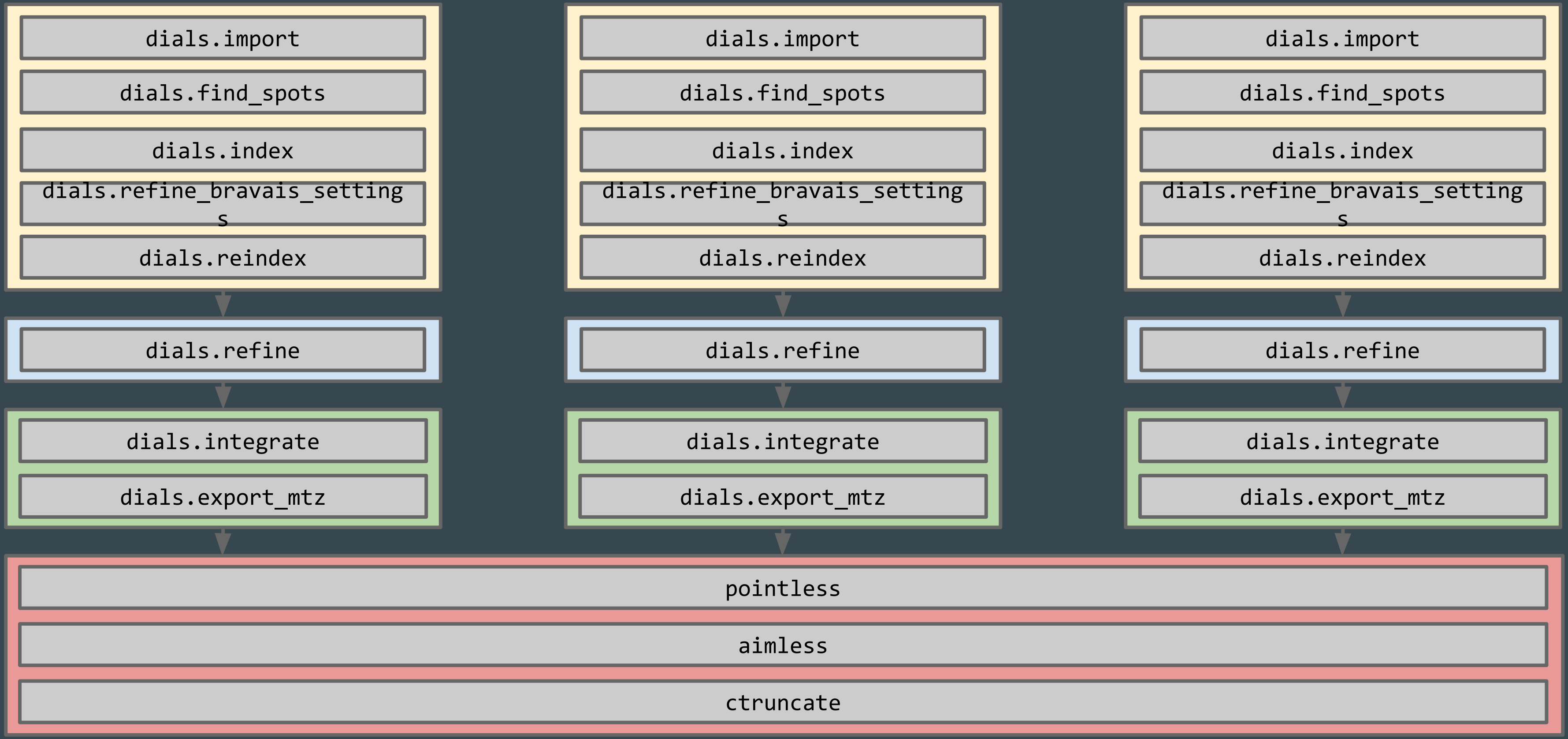
xia2 -dials ${data_directory}

```
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
```

XIA2 in CCP4 I2



XIA2 DIALS pipeline



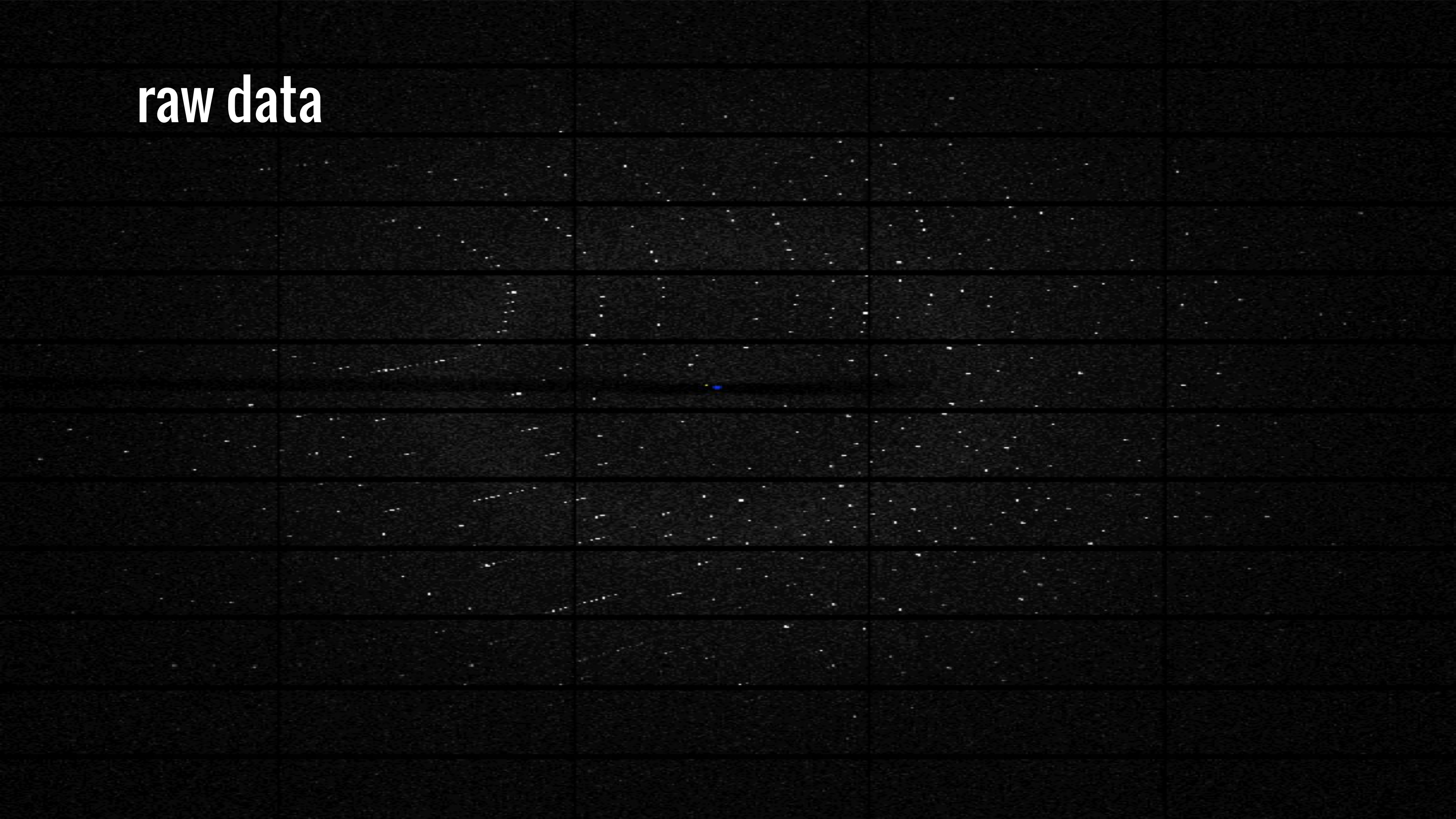
Spot finding

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
```

raw data



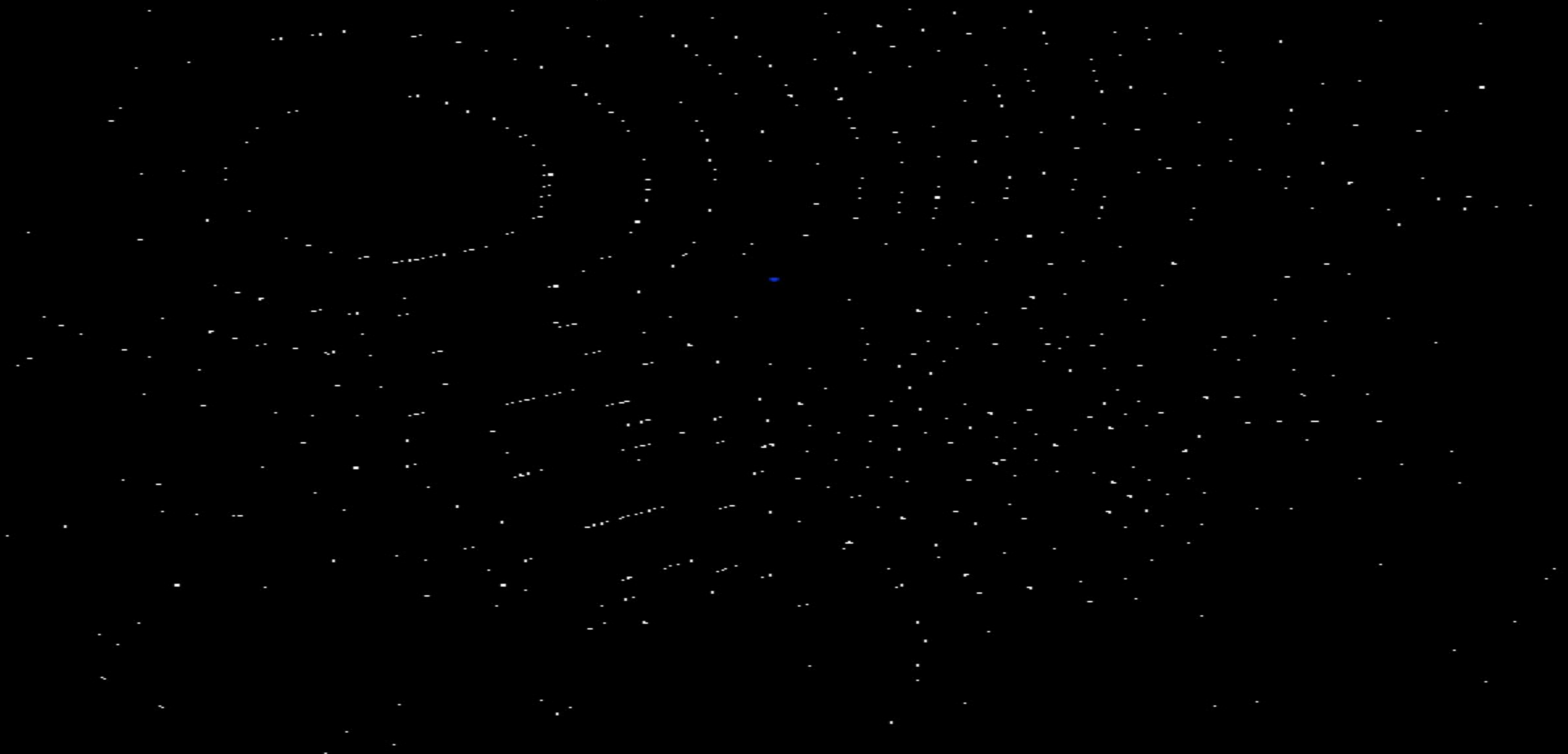
mean

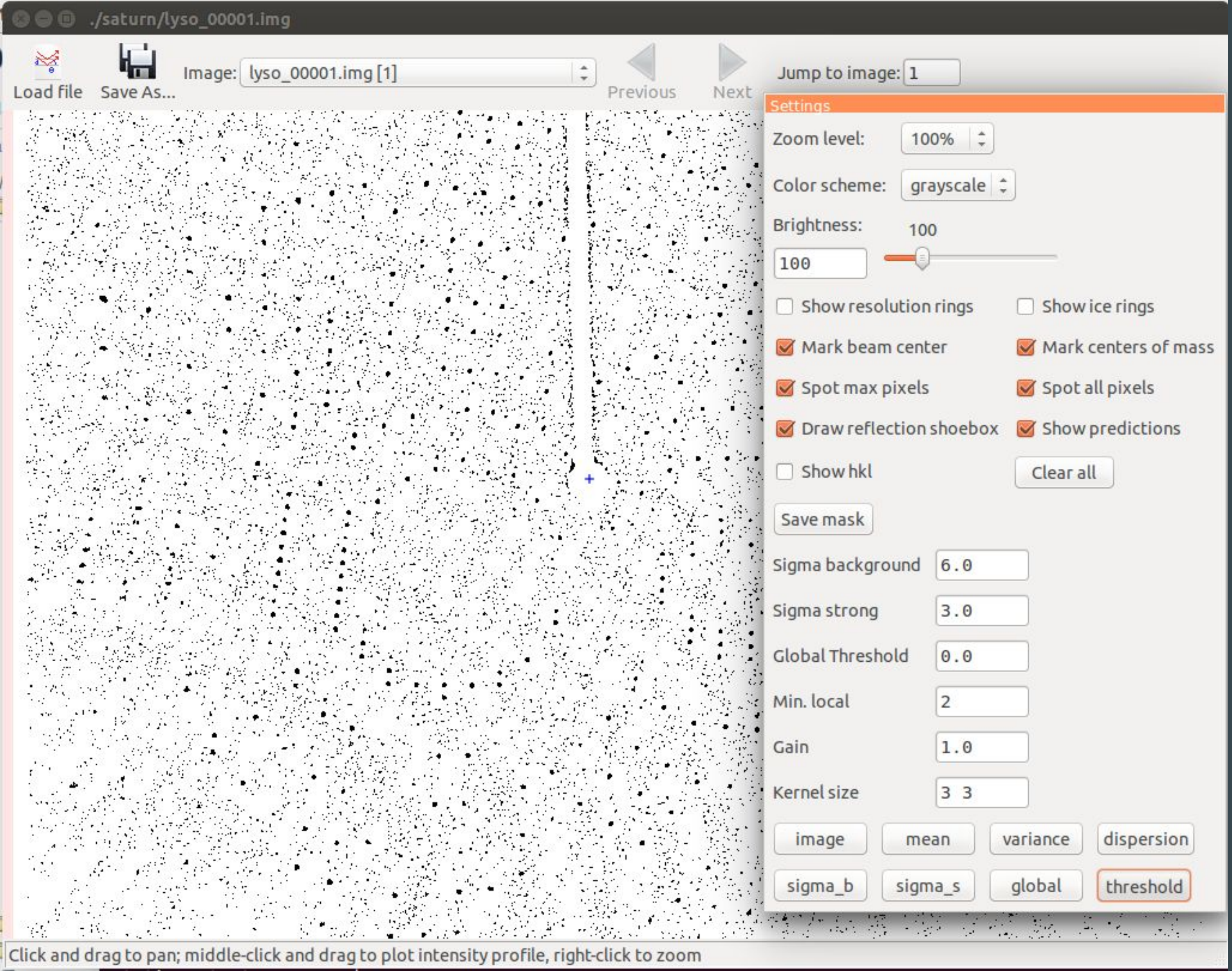
variance

dispersion = variance / mean

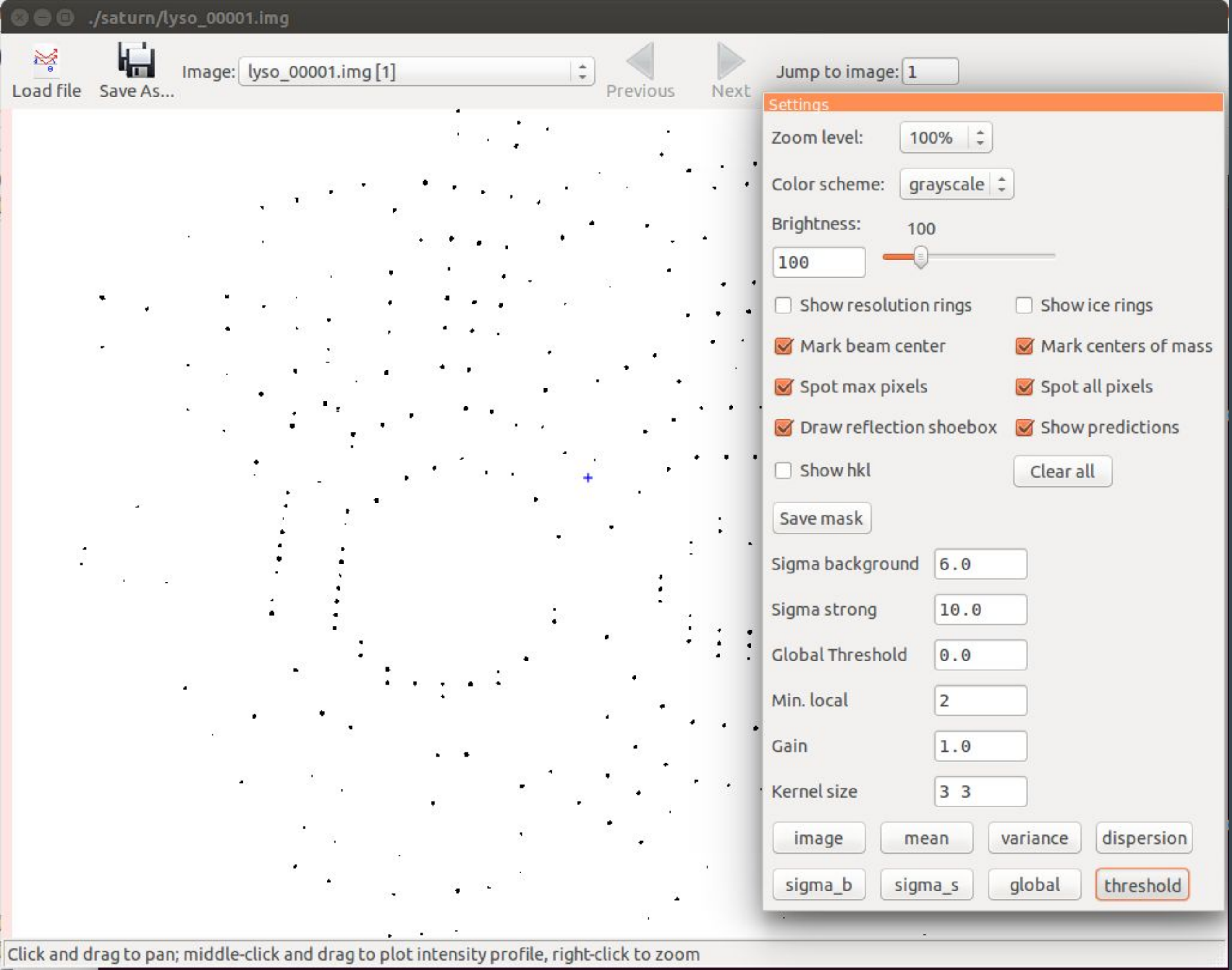
$$\text{dispersion} > 1 + \text{sigma\_s} * \text{sqrt}(2/(m-1))$$

raw data > mean + sigma_b * sqrt(variance)



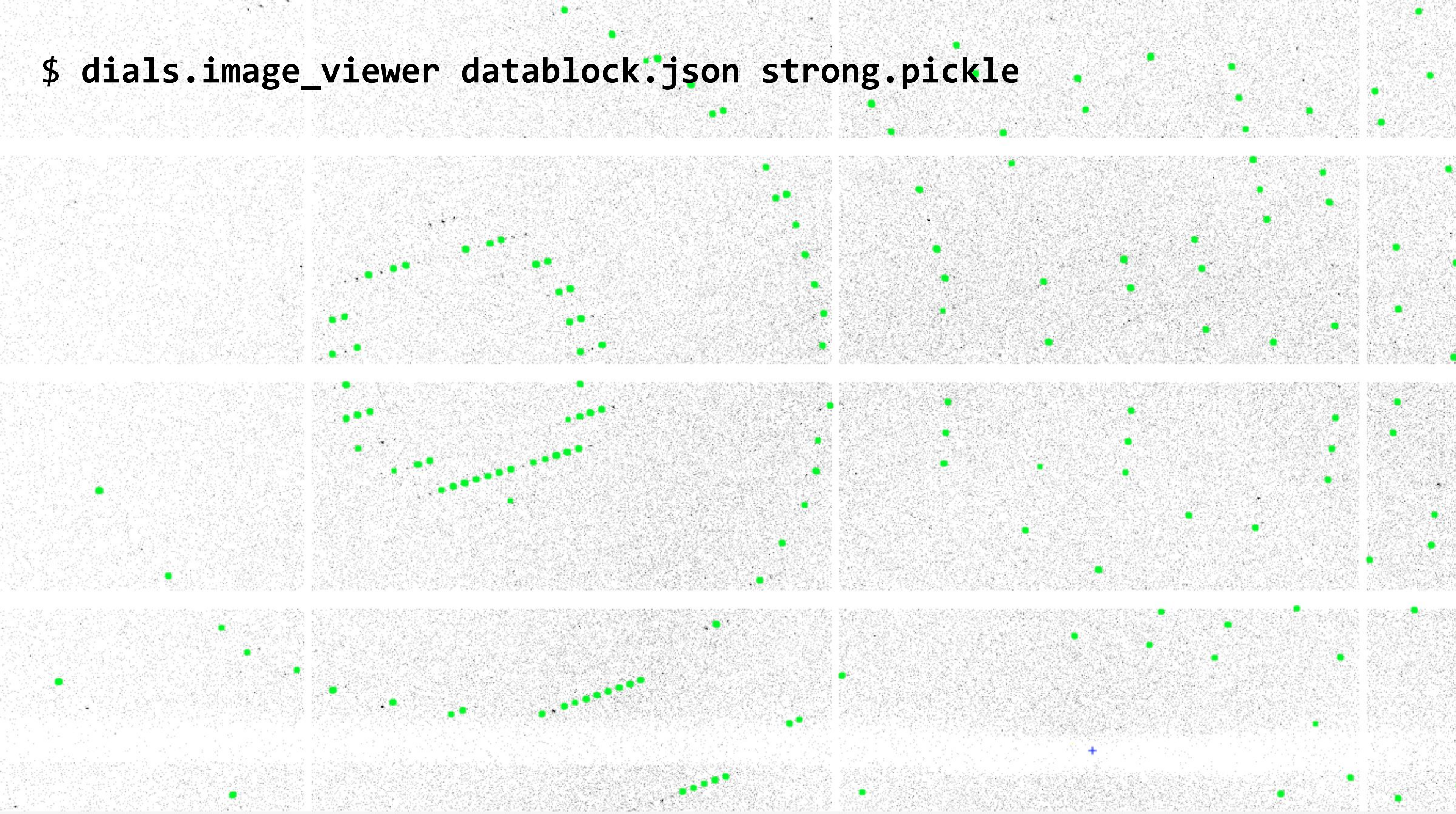


Default spot finding parameters are often not suitable for CCD images

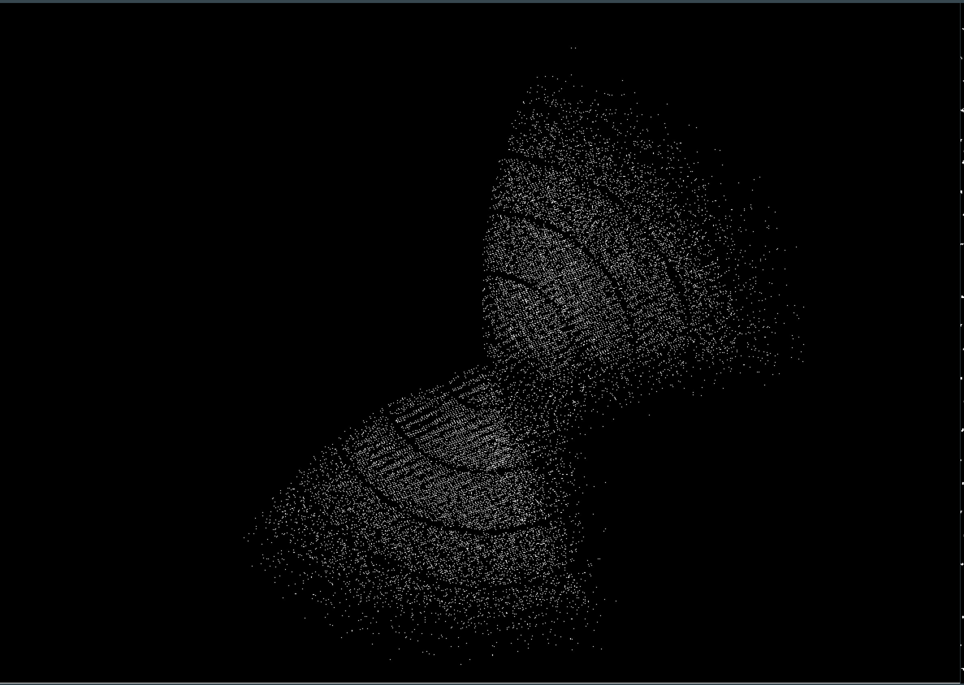


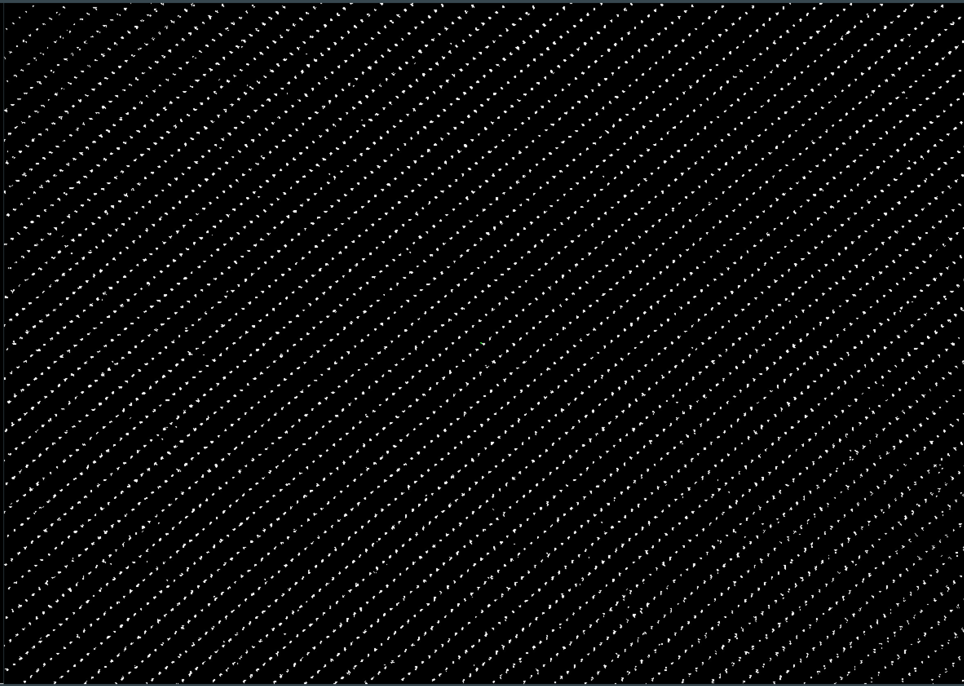
Default spot finding parameters are often not suitable for CCD images

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

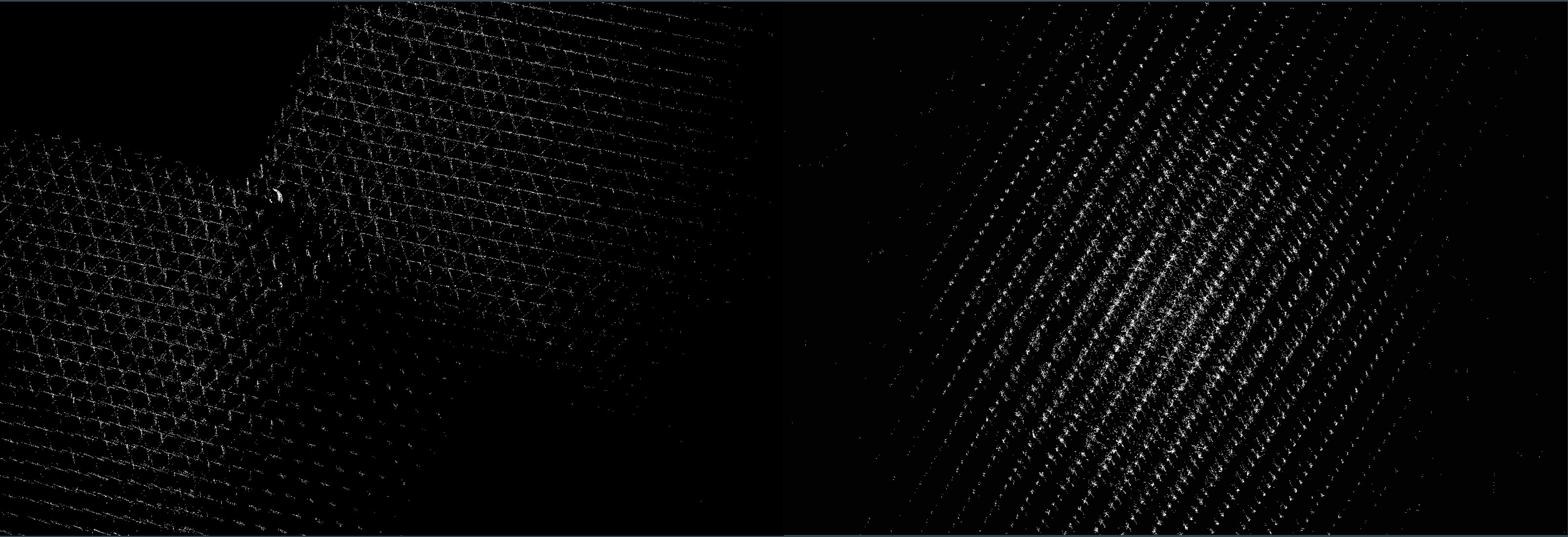


$ dials.reciprocal_lattice_viewer datablock.json strong.pickle



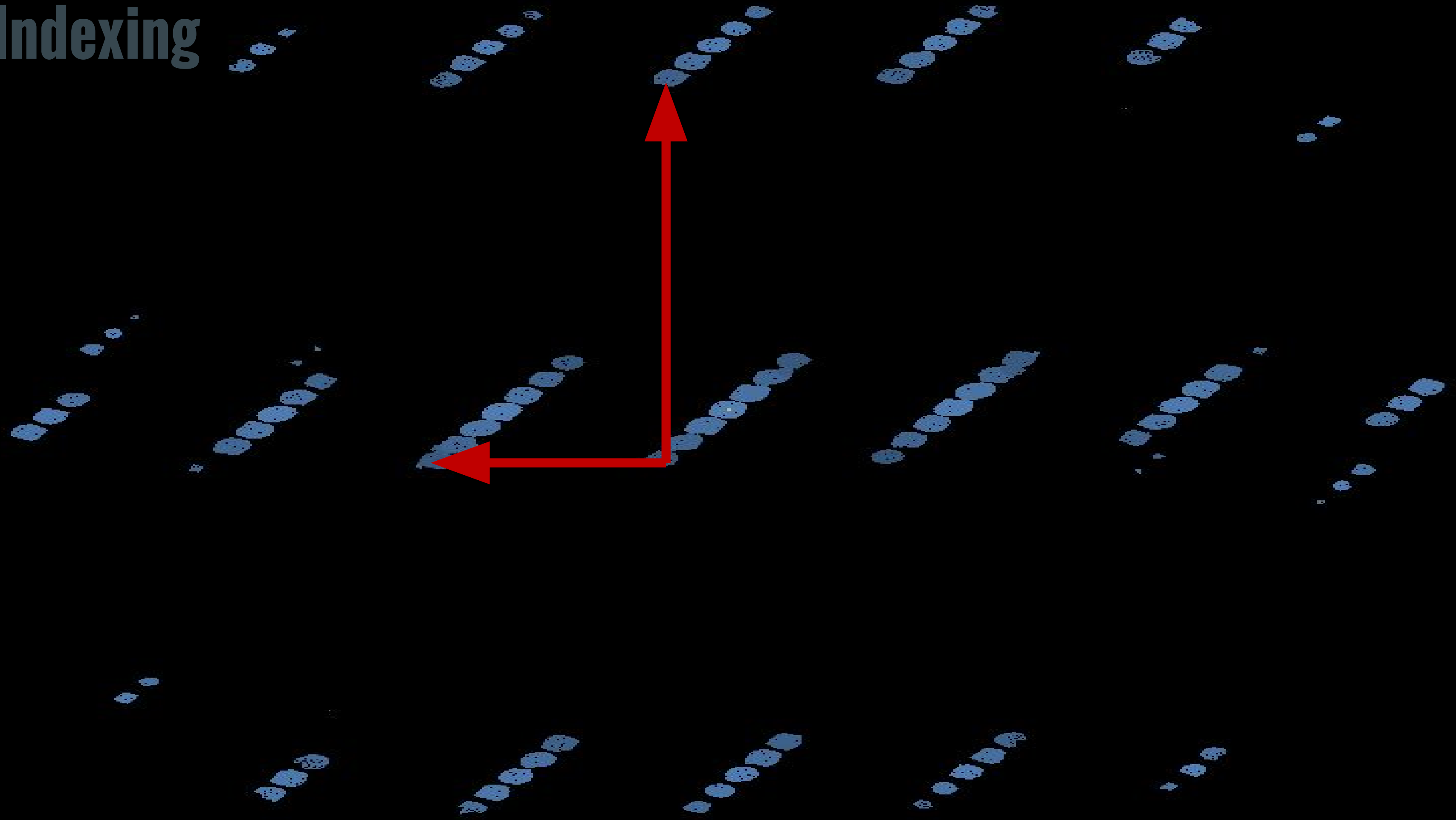


```
$ dials.reciprocal_lattice_viewer datablock.json strong.pickle
```

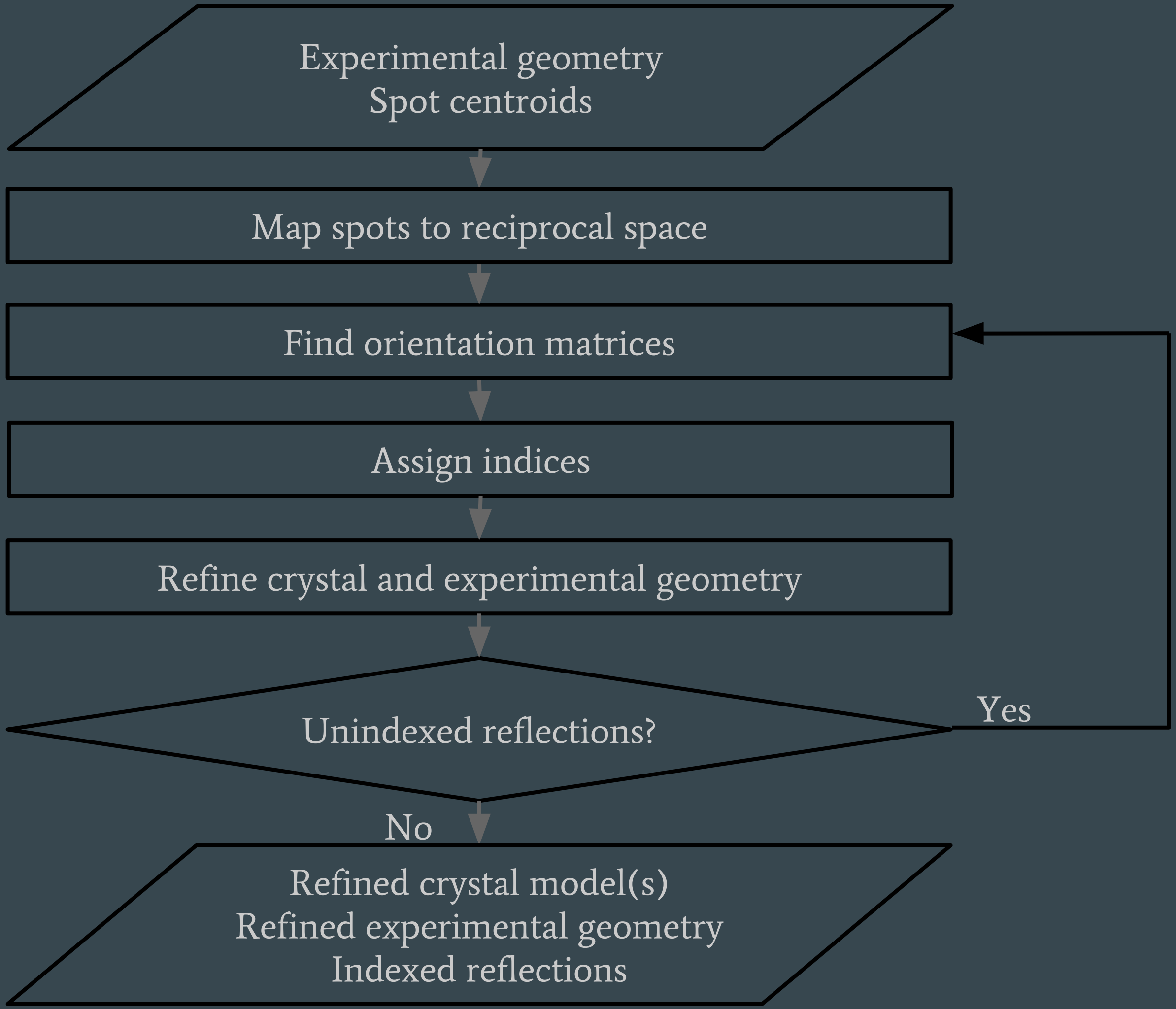


Indexing

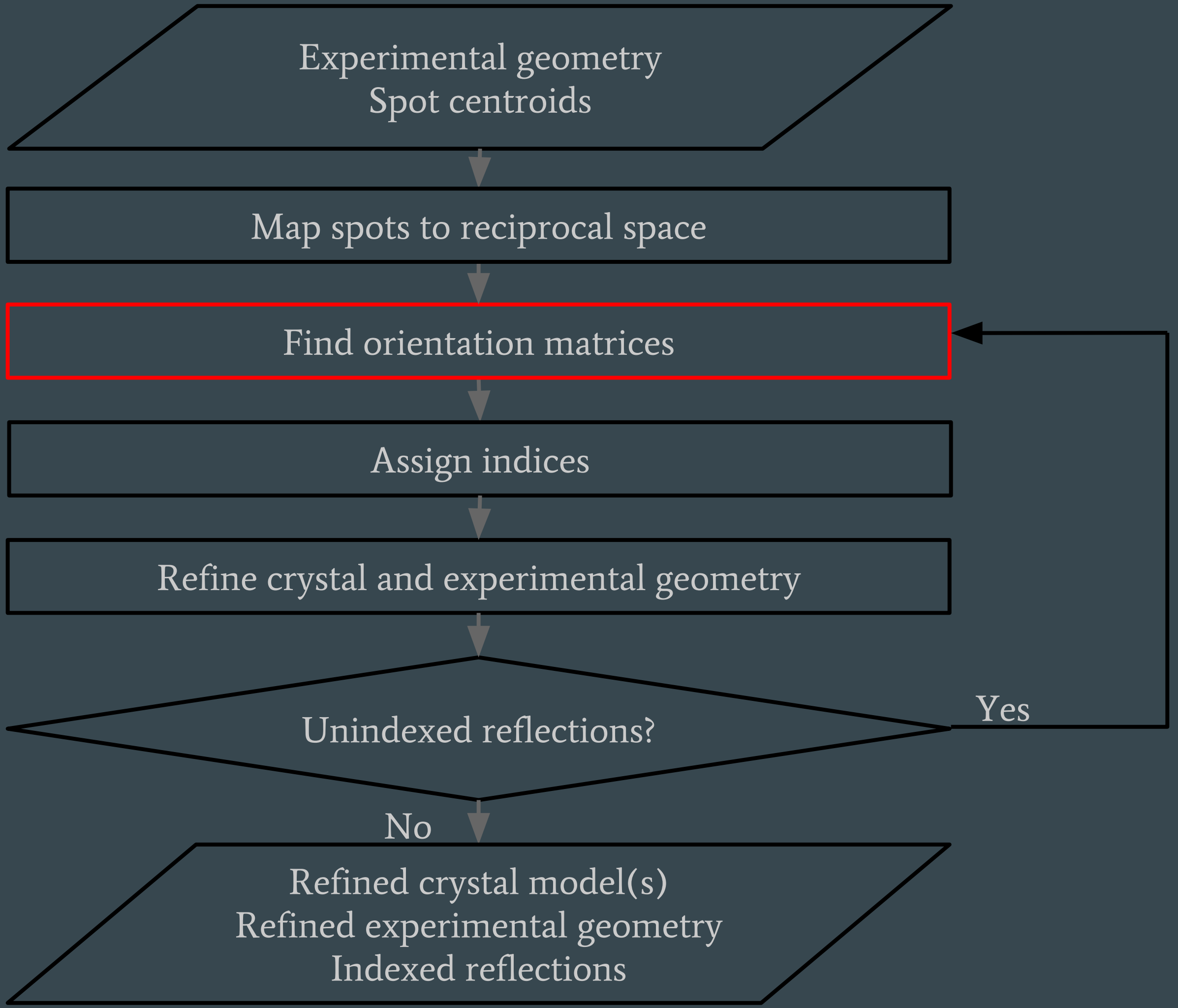
Indexing



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

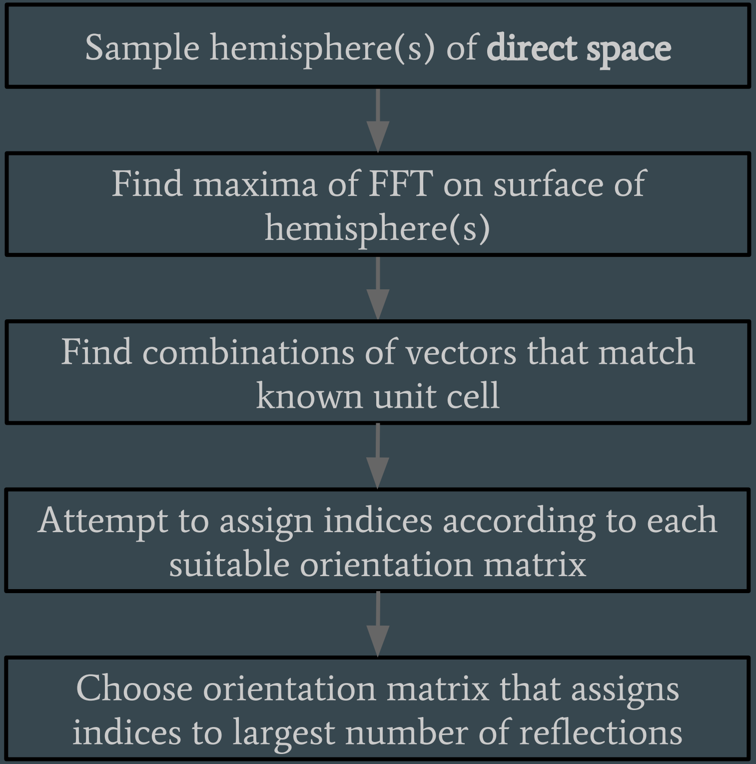
dials.index

- Choice of method:
  - 1D FFT (DPS)
  - 3D FFT - default
  - 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
```

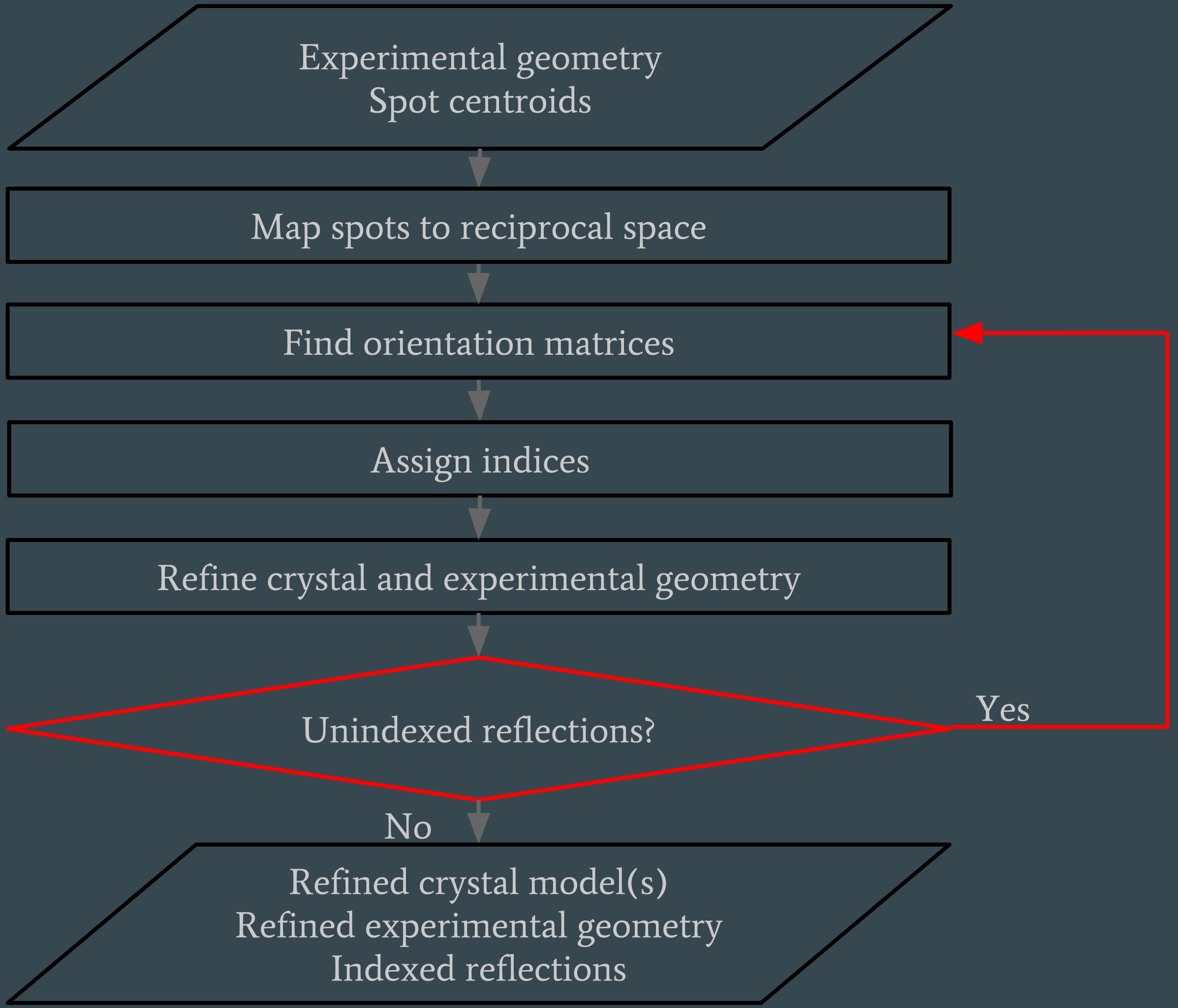
Real space grid search

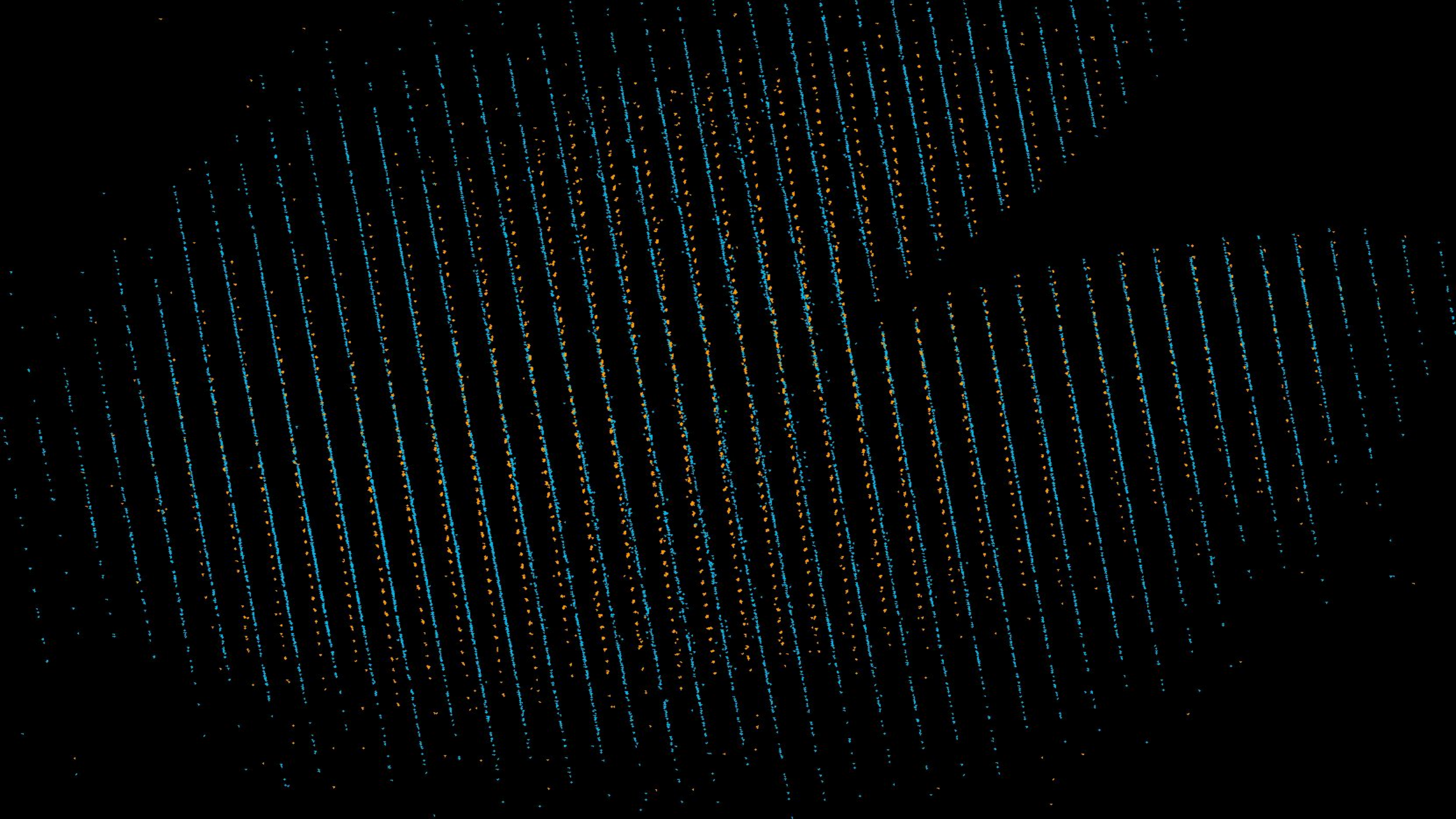
- In many cases the unit cell is known - why not make use of this information?
- 1D FFT (DPS) and 3D FFT algorithms try to determine the magnitude and direction of the basis vectors simultaneously
- If the unit cell is already known, then we only need to determine the direction of the basis vectors

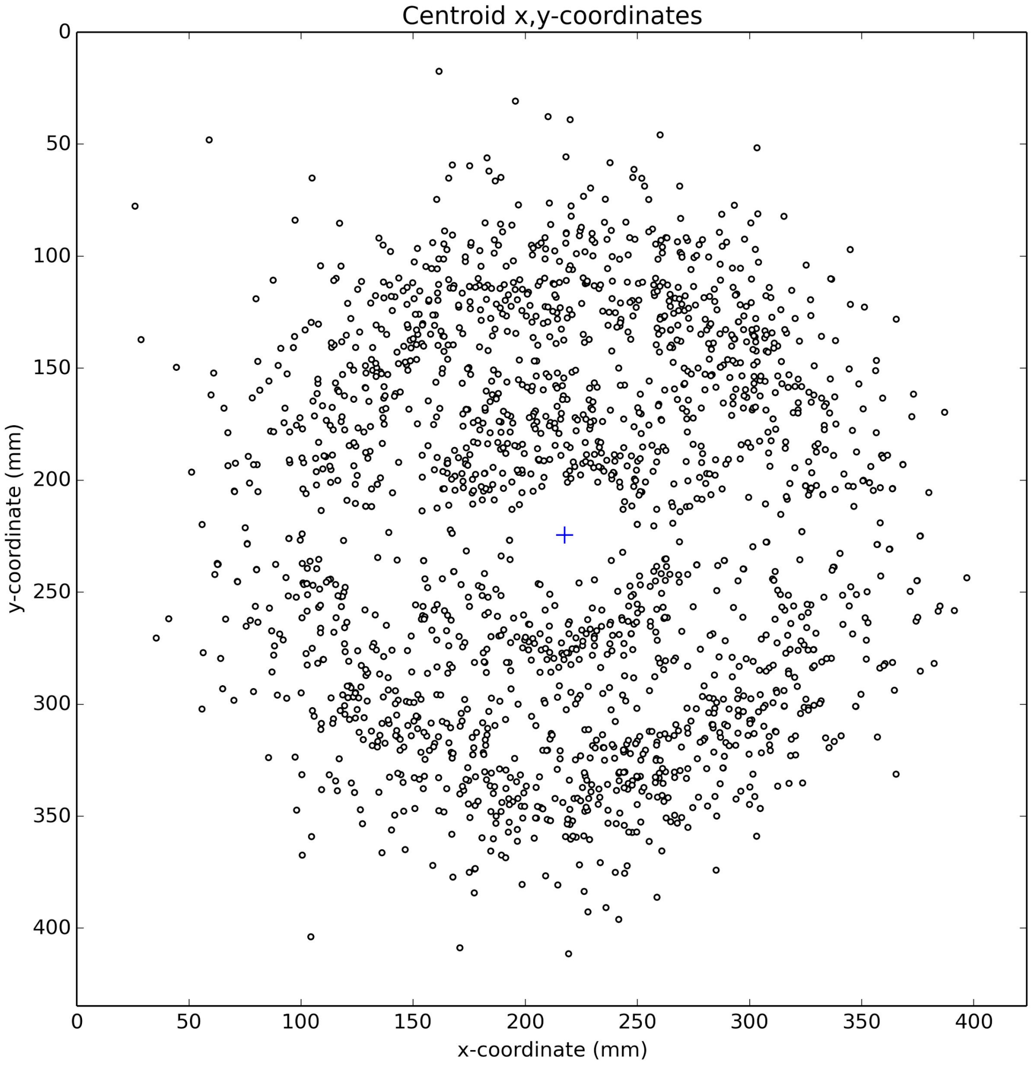


New methods for indexing multi-lattice diffraction data. Gildea *et al.*, *Acta Cryst.* (2014) **D70**, 2652-2666.

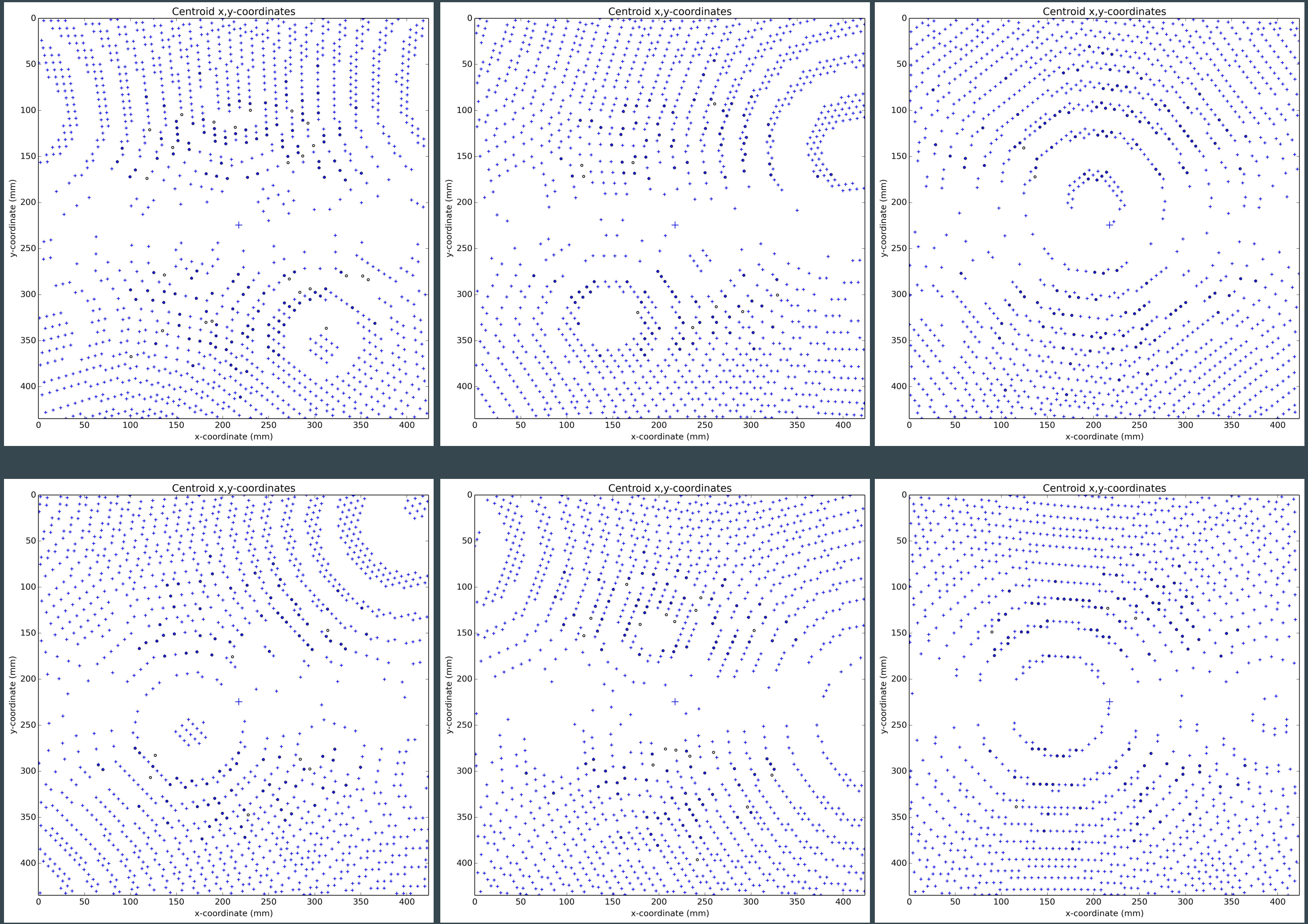
dials.index


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

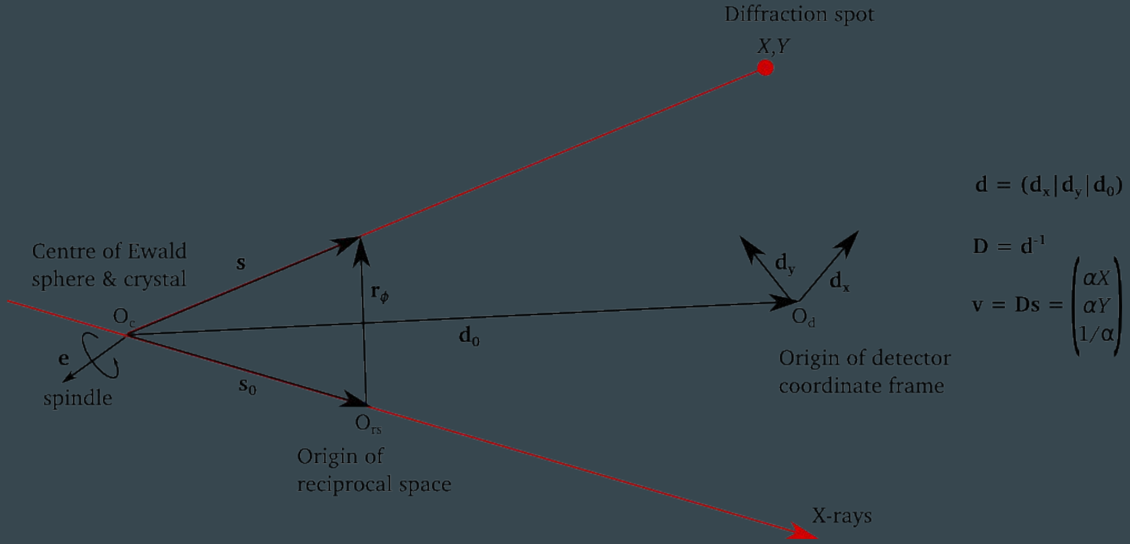
dials.refine_bravais_settings

- After indexing, 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 solution to take further

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

-----------------------------------------------------------------------------------------------------------
Solution Metric fit  rmsd  min/max cc #spots lattice                               unit_cell  volume      cb_op
-----------------------------------------------------------------------------------------------------------

*      9      0.0311 0.063 0.800/0.857   8099      tP  57.78   57.78 150.00  90.00  90.00  90.00  500867      a,b,c
*      8      0.0311 0.063 0.800/0.969   8099      oC  81.72   81.73 150.01  90.00  90.00  90.00 1002008  a-b,a+b,c
*      7      0.0272 0.061 0.969/0.969   8099      mC  81.73   81.74 150.03  90.00  89.99  90.00 1002365  a-b,a+b,c
*      6      0.0311 0.062 0.805/0.805   8099      mC  81.73   81.72 150.02  90.00  89.99  90.00 1002012 a+b,-a+b,c
*      5      0.0154 0.061 0.800/0.906   8099      oP  57.79   57.76 149.99  90.00  90.00  90.00  500672      a,b,c
*      4      0.0147 0.060 0.821/0.821   8099      mP  57.77   57.80 150.01  90.00  90.02  90.00  500853   -b,-a,-c
*      3      0.0154 0.060 0.906/0.906   8099      mP  57.80   57.78 150.02  90.00  89.98  90.00  500945      a,b,c
*      2      0.0152 0.061 0.800/0.800   8099      mP  57.78  150.01  57.80  90.00  89.99  90.00  500925      b,c,a
*      1      0.0000 0.060         -/-   8099      aP  57.80   57.78 150.03  90.01  89.99  89.99  501086      a,b,c
-----------------------------------------------------------------------------------------------------------
* = recommended solution
```

Refinement

Centroid refinement



- 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

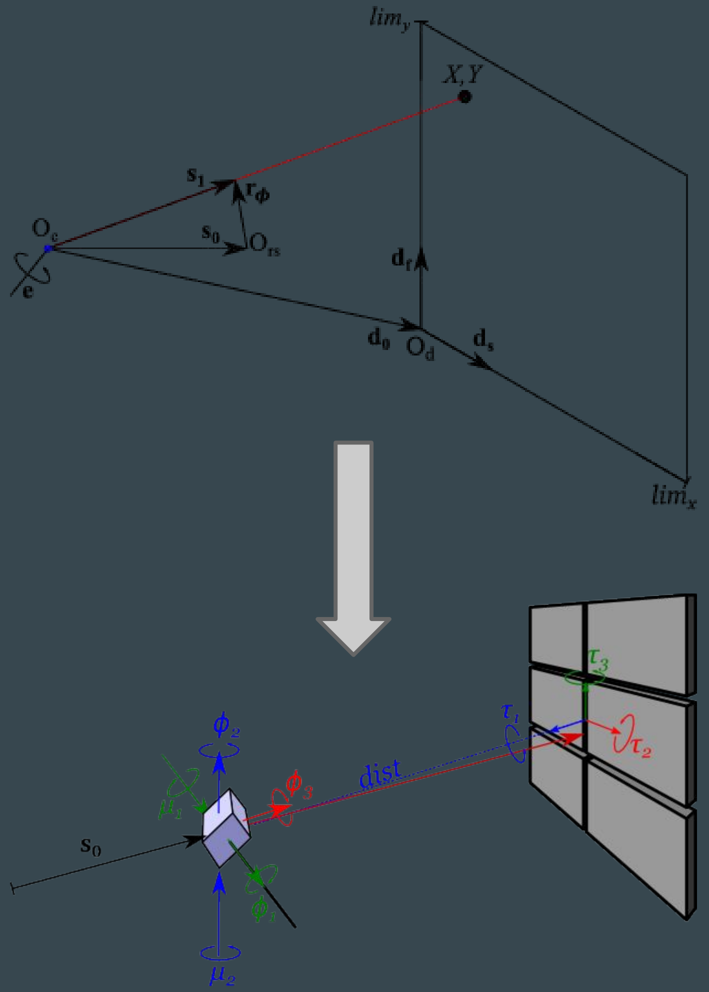
Parameterisation

There are 18 parameters in the $P1$ 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	**U**	ϕ_1	rotation about laboratory X
		ϕ_2	rotation about laboratory Y
		ϕ_3	rotation about laboratory Z
Crystal unit cell	**B**	g^*_{11}	set metrical matrix elements
		g^*_{22}	
		g^*_{33}	
		g^*_{12}	
		g^*_{13}	
		g^*_{23}	
Detector	**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



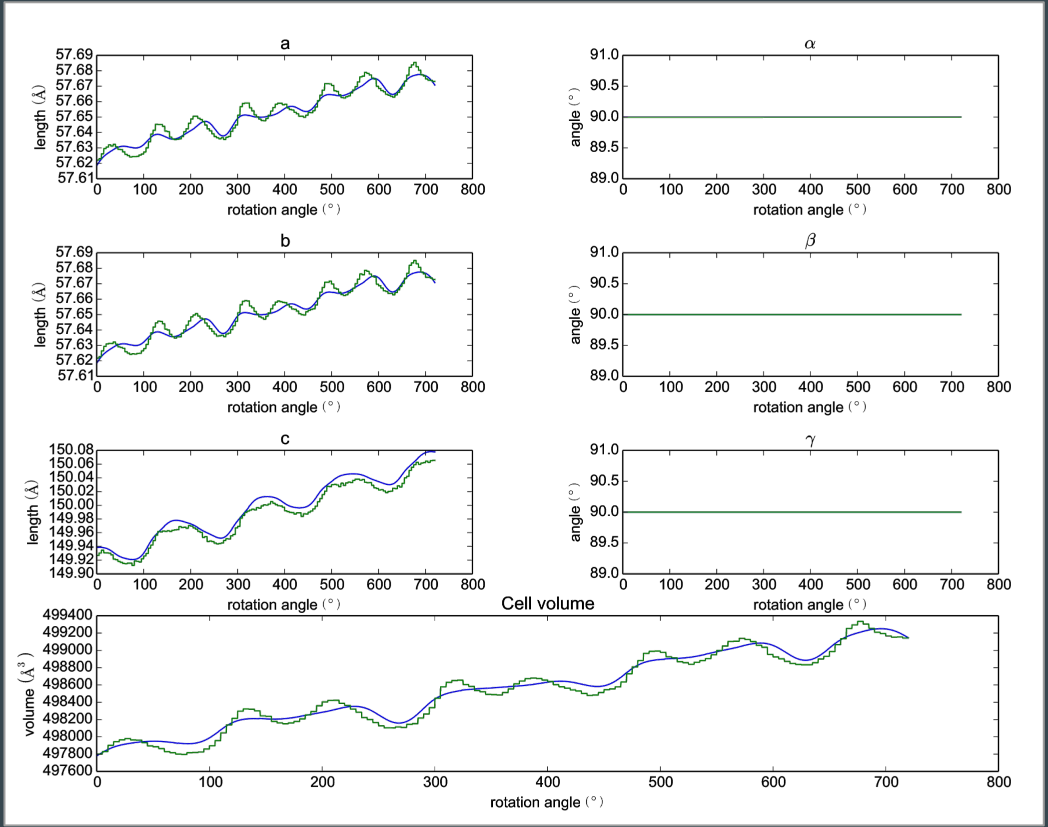
Scan-varying refinement

- 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:

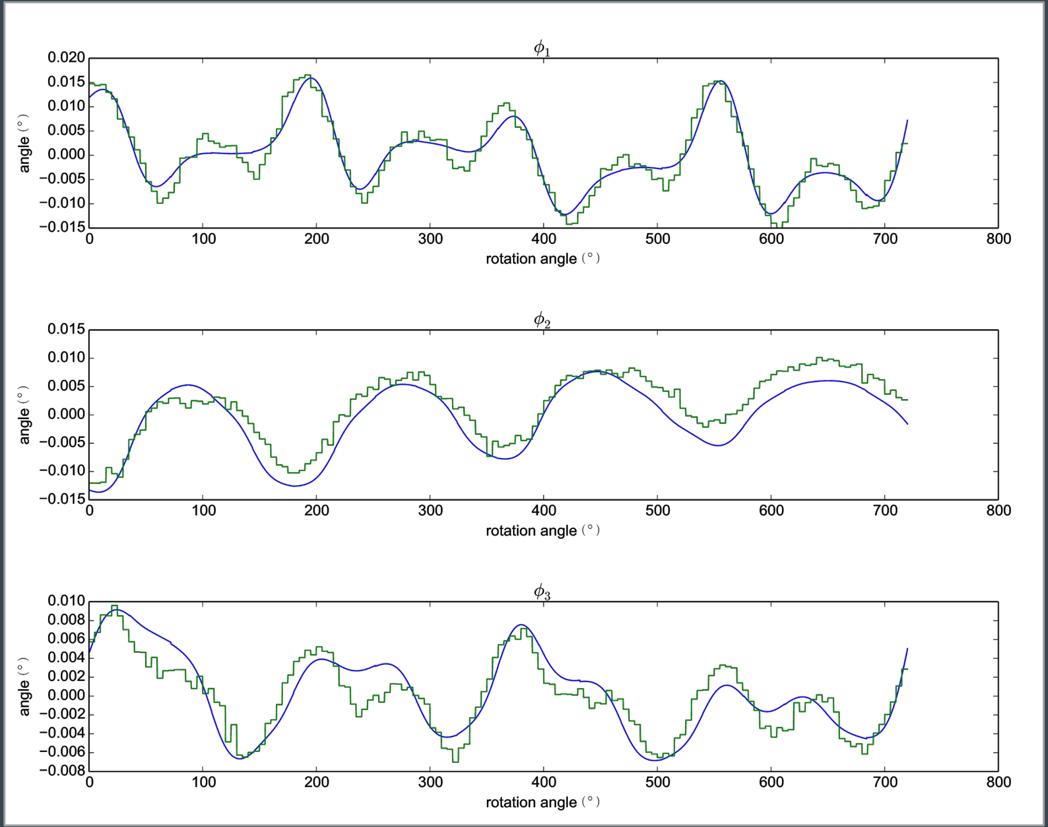
- 6 detector
- 1 beam
- 3 crystal orientation × 22 "samples"
- 2 unit cell parameters × 22 "samples”



Scan-varying refinement

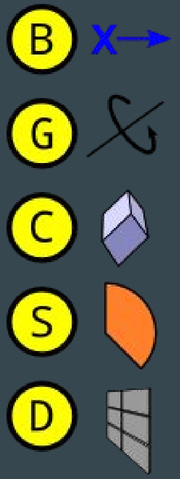
117 parameters:

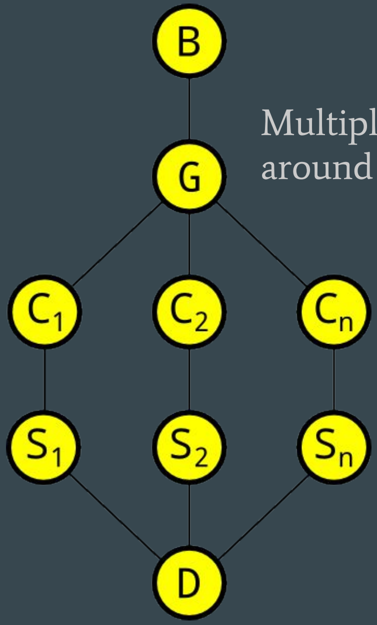
- 6 detector
- 1 beam
- 3 crystal orientation × 22 "samples"
- 2 unit cell parameters × 22 "samples”



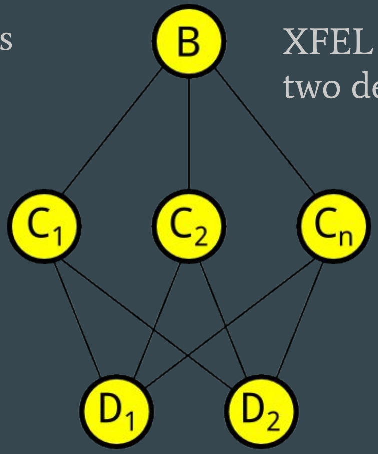
Multiple experiments

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





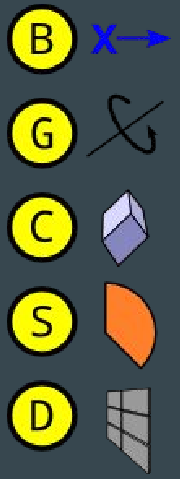
Multiple rotation scans around the same axis

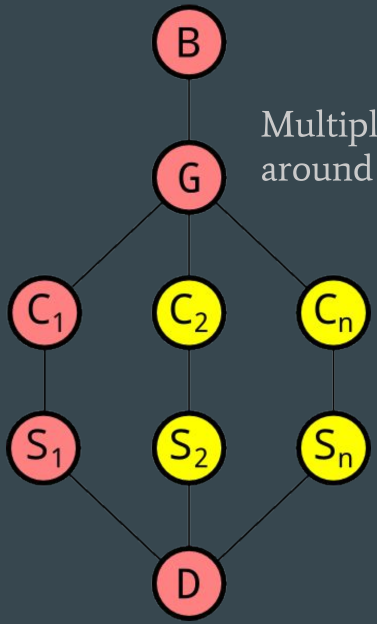


XFEL shots with two detectors

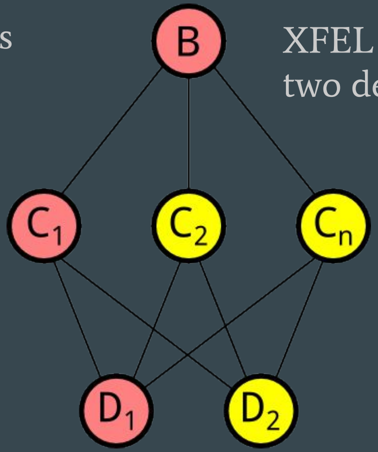
Multiple experiments

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





Multiple rotation scans around the same axis

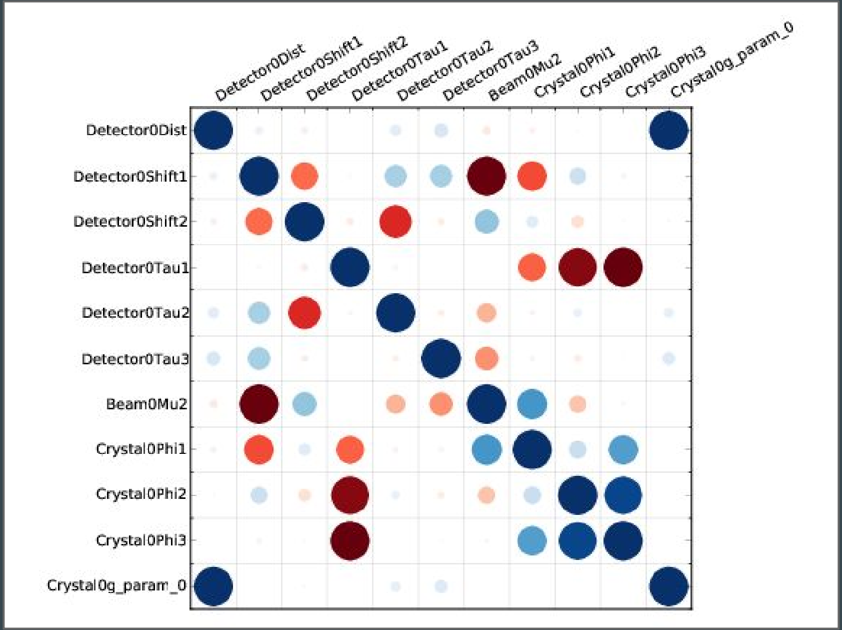


XFEL shots with two detectors

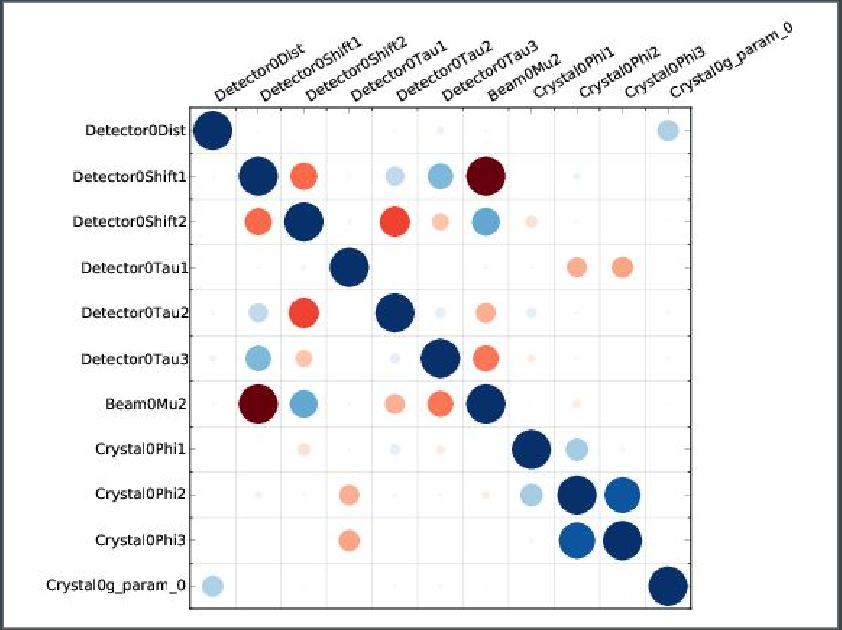
Multiple experiments

Cubic polyhedrin crystals, 1° scans

One lattice



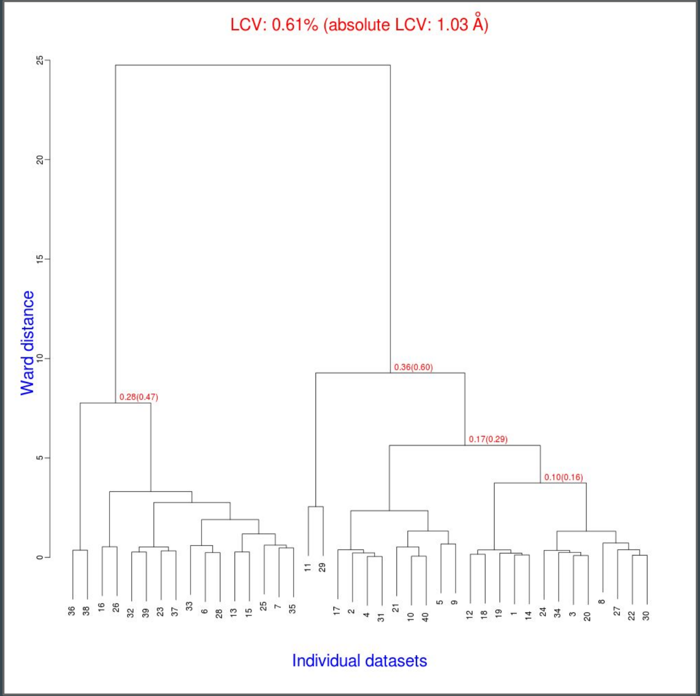
5 sweeps (16 lattices)

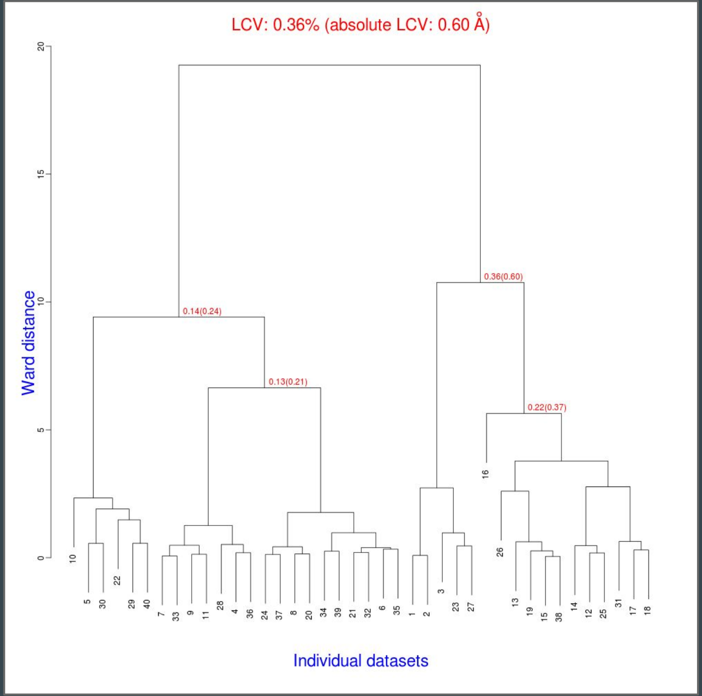


Joint refinement

Multiple experiments

Use joint refinement as a preparatory step for BLEND





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

Integration

Tasks in dials.integrate

Calculate the bounding box parameters from strong reflections

↓

Predict the positions of reflections on the images

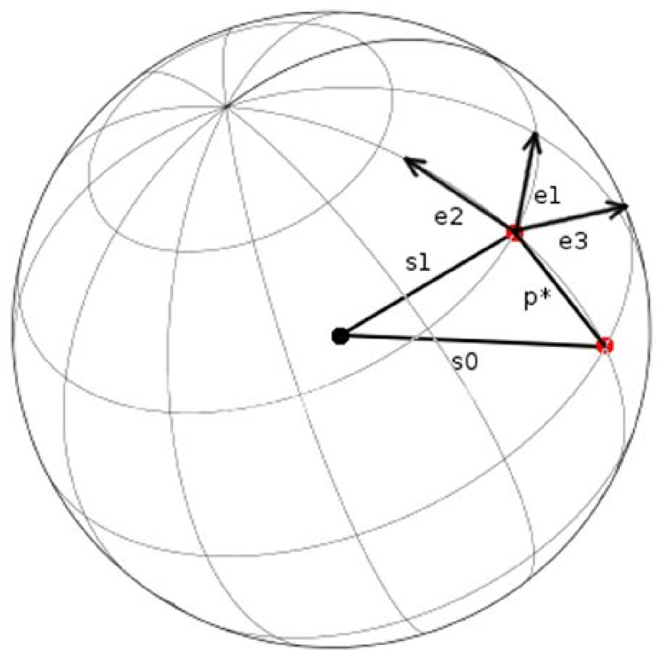
↓

Build reference profiles across all images

↓

Integrate the reflections and save output

Computing reflection shoeboxes



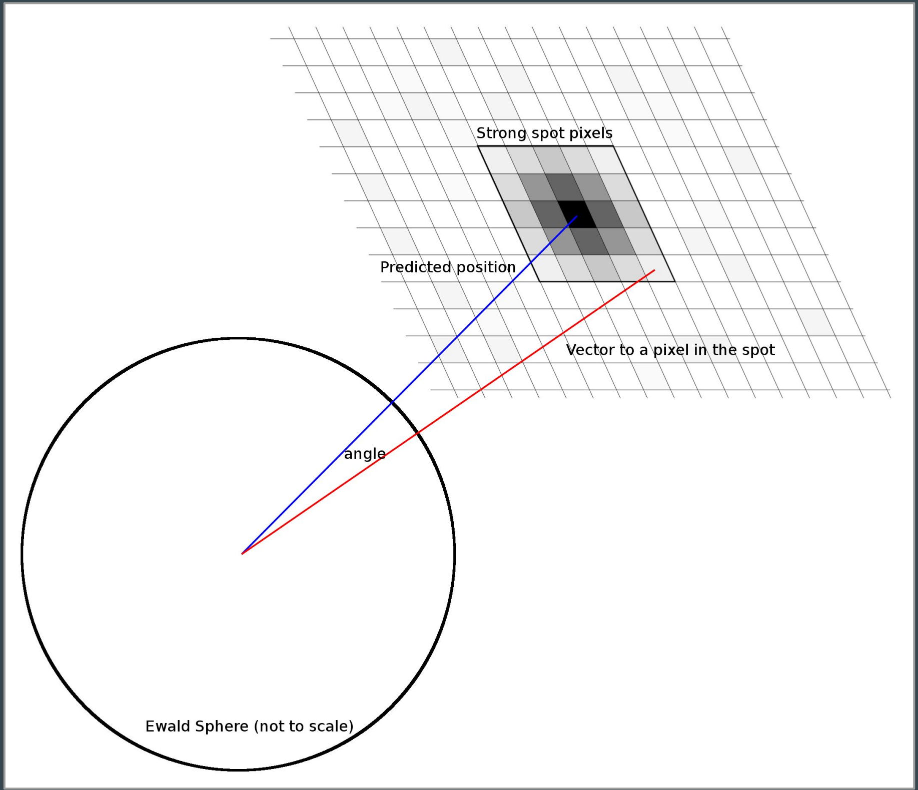
Profile coordinate system

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

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

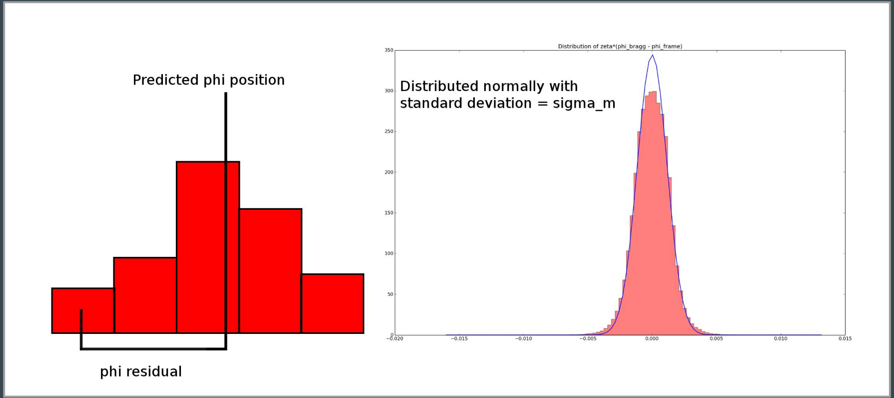
$$\boldsymbol{e}_1 = \boldsymbol{S}_1 \times \boldsymbol{S}_0 / |\boldsymbol{S}_1 \times \boldsymbol{S}_0|$$
$$\boldsymbol{e}_2 = \boldsymbol{S}_1 \times \boldsymbol{e}_1 / |\boldsymbol{S}_1 \times \boldsymbol{e}_1|$$
$$\boldsymbol{e}_3 = (\boldsymbol{S}_1 + \boldsymbol{S}_0) / |\boldsymbol{S}_1 + \boldsymbol{S}_0|$$

Computing reflection shoeboxes

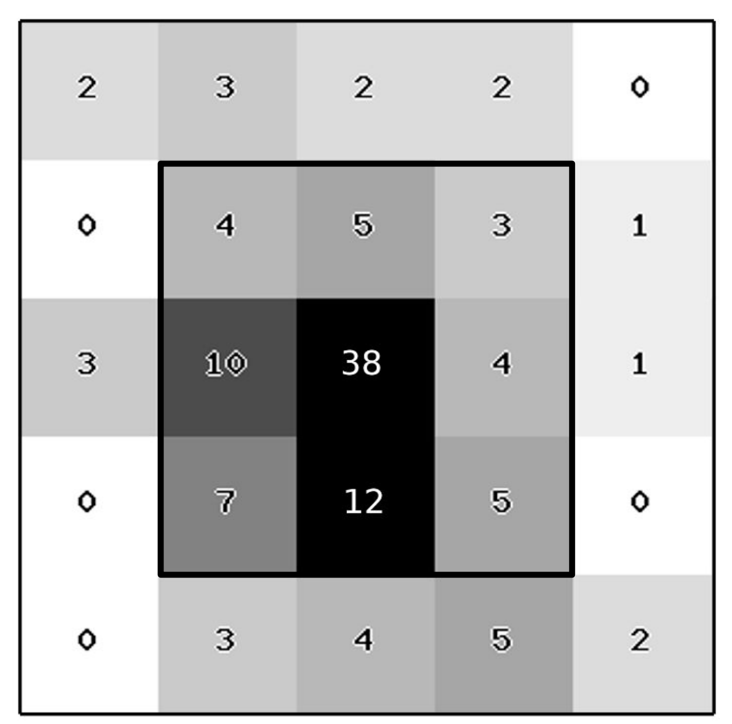


σ_D is calculated from the spread of angles between the predicted diffracted beam vector and the vector for each strong pixel in the spot

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



Integration



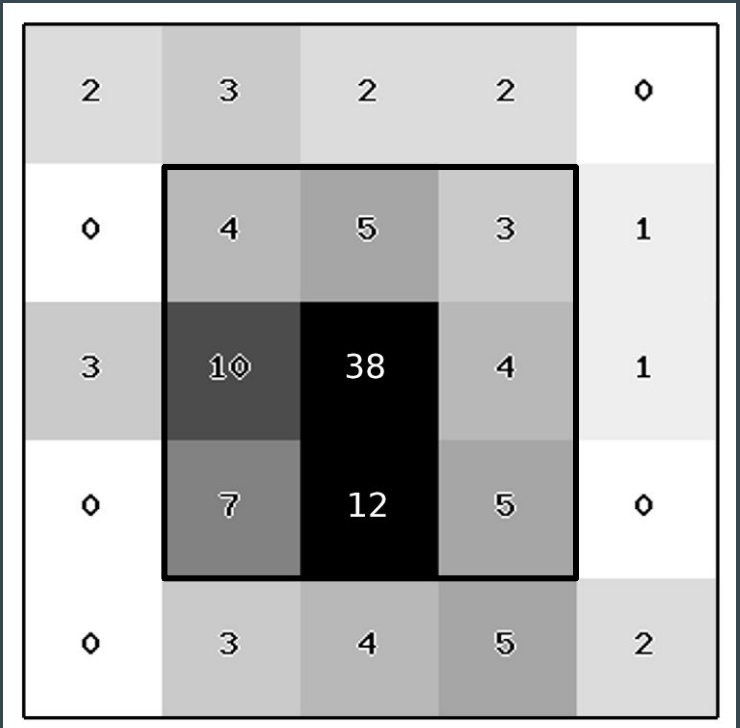
Summation integration: estimate the reflection intensity by summing the counts contributing to the reflection and subtracting the background

I = SUM(Counts - Background)

Profile fitting: fit a known profile shape to the reflection to estimate the intensity

Need to estimate background under reflection peak

Background determination

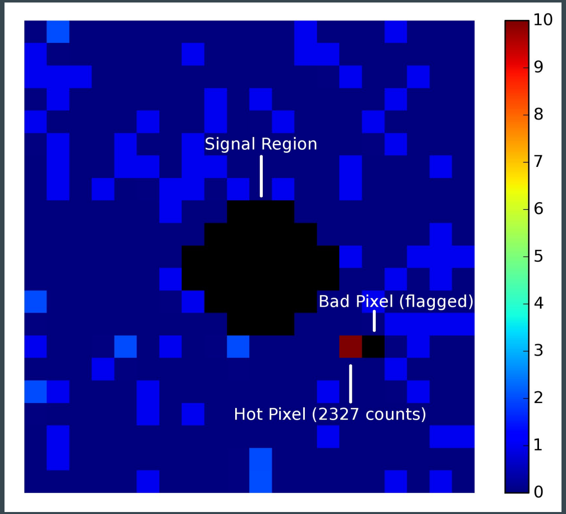


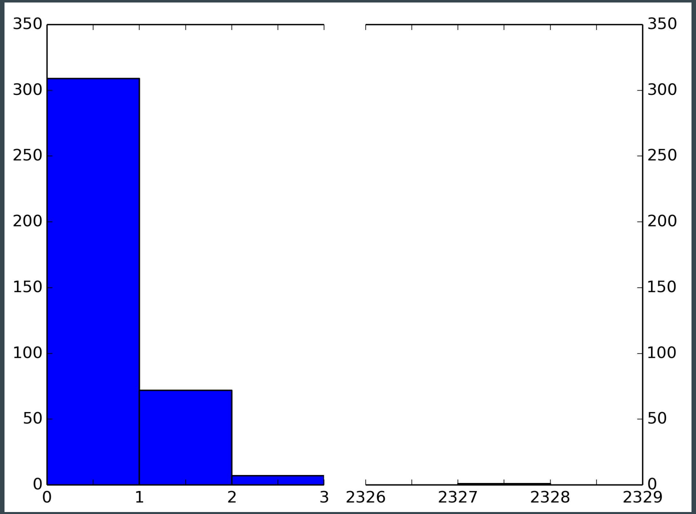
Don’t know background in signal region so estimate from the surrounding pixels

Background = MEAN(Background Counts)
Background = (2+3+2+2+0+0+1+3+1+0+0+0+3+4+5+2) / 16
Background = 2.5

I = SUM(Counts - Background)
I = 4+5+3+10+38+4+7+12+5 - 9*2.5
I = 88 - 22.5
I = 65.5

Background outlier pixels





	With Hot Pixel	Without Hot Pixel
Mean	6.20	0.22
Variance/Mean	2237.90	0.926

~1 for Poisson distribution

Background modelling with outlier pixels

DIALS has multiple options for outlier pixel handling

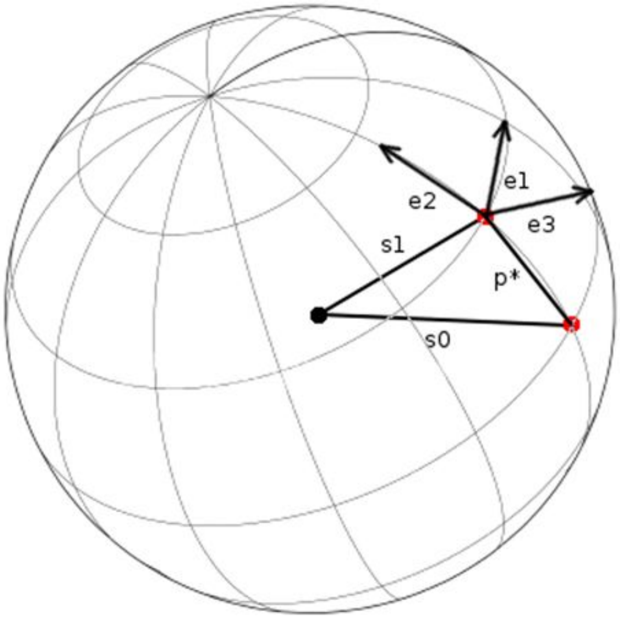
- Truncated - removed percentage of high and low valued pixels
- Normal - remove pixels above and below 3 STD around the mean
- Tukey - remove pixels based on interquartile range
- Plane - compute a plane and remove pixels based on deviation from plane (based on published mosflm algorithm)
- Normal - iteratively remove high valued pixels until they are approximately normally distributed (based on published XDS algorithm)

However, these methods assume a normal distribution and result in biases intensity estimates (particularly for low background)

Default algorithm in DIALS used a GLM algorithm

- assumes a Poisson distribution of pixel counts.
- Provides an unbiased estimate of the reflection background

3D profile fitting coordinate system



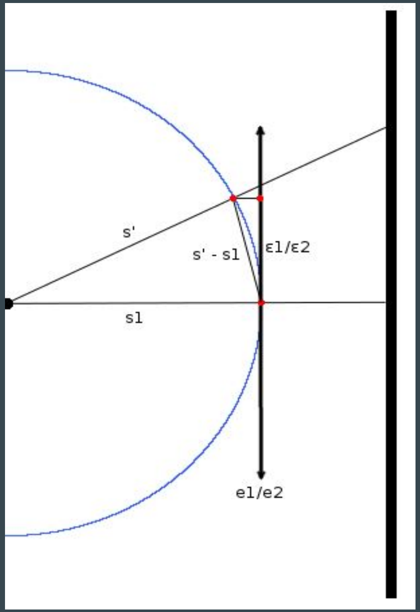
Profile coordinate system

Use Kabsch coordinate system

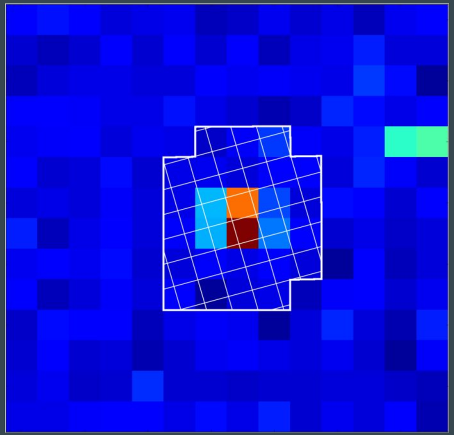
- Corrects for geometrical distortions
- Makes spots appear to have taken shortest path through Ewald sphere
- Model assumes a Gaussian profile in Kabsch coordinate system

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

3D profile fitting pixel gridding

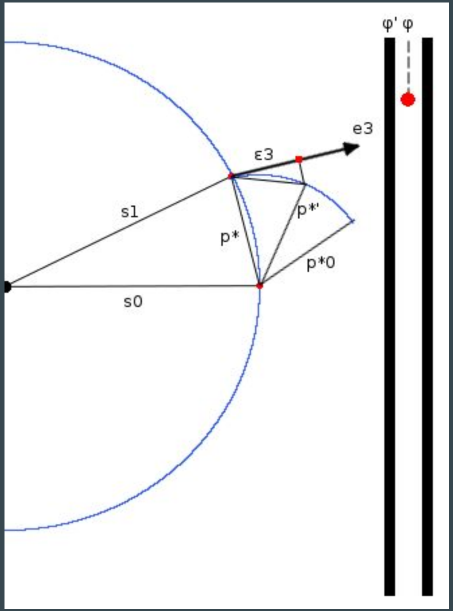


Pixels are mapped to the Ewald sphere

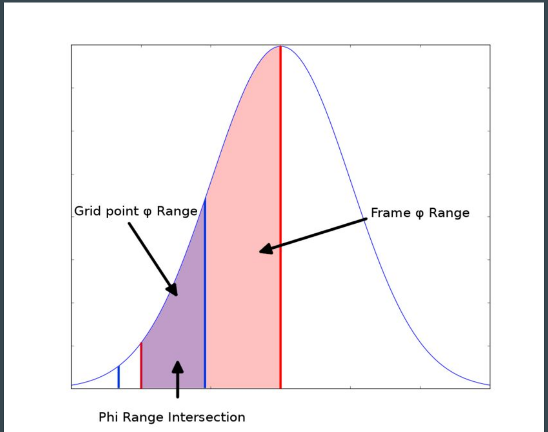


Counts are redistributed to Ewald sphere grid by computing fractional overlap of each pixel and Ewald sphere grid point

3D profile fitting phi gridding

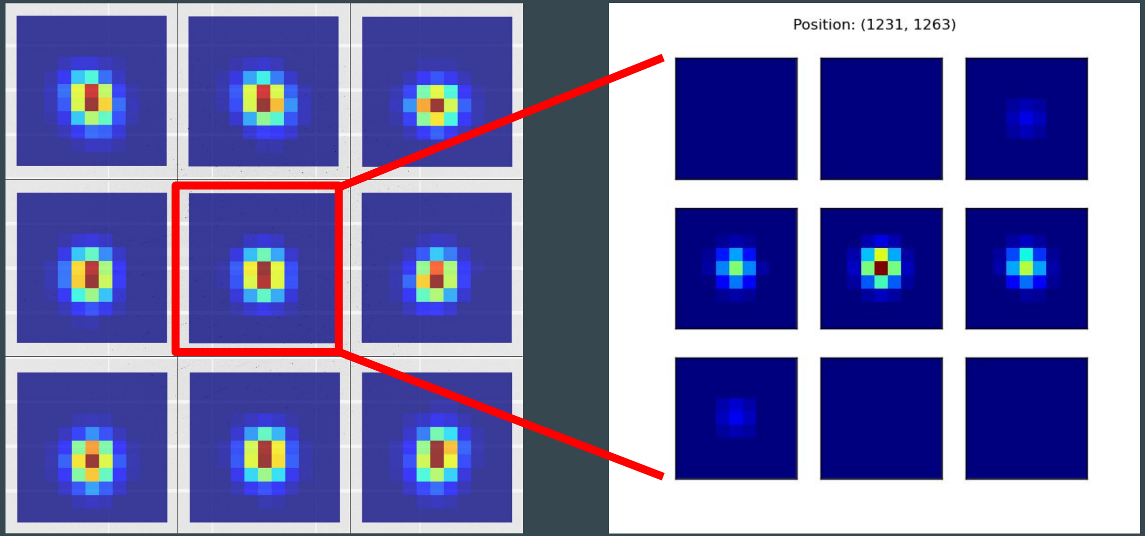


Frames are transformed to make reflection appear as if it took the shortest path through the Ewald sphere



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

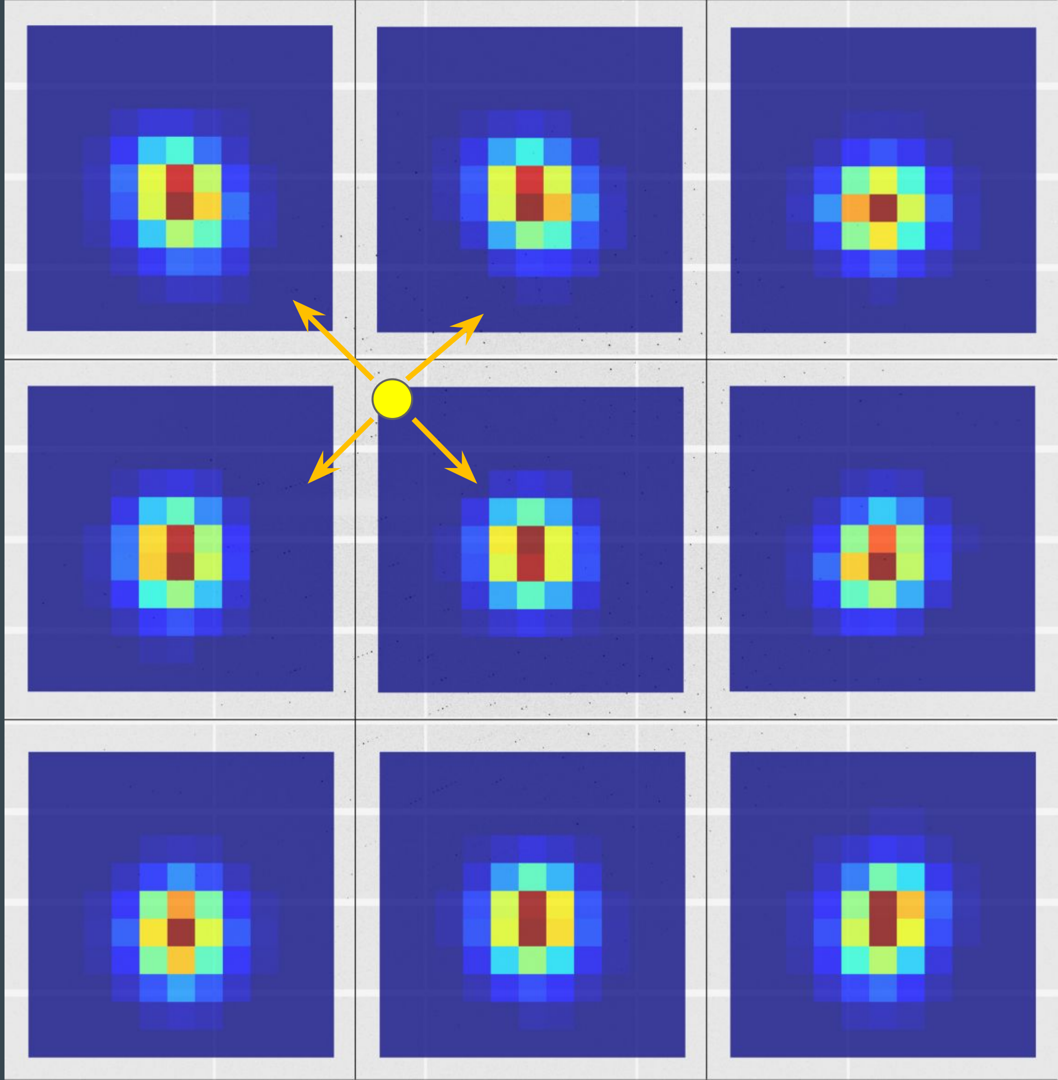
Building reference profiles



- 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)
  - Projected Ewald sphere grid (needed for multi-panel detectors)
  - Single reflection

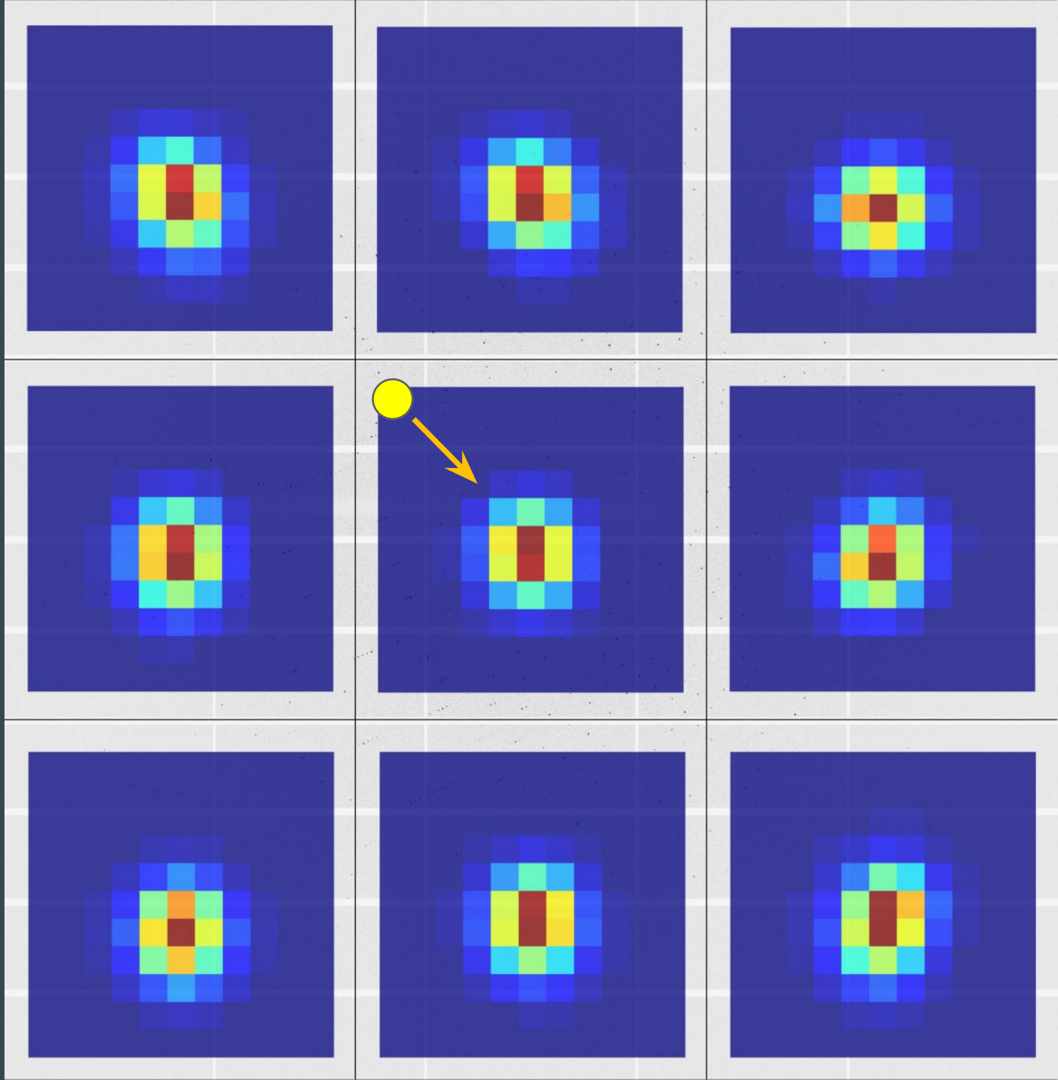
Building reference profiles

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

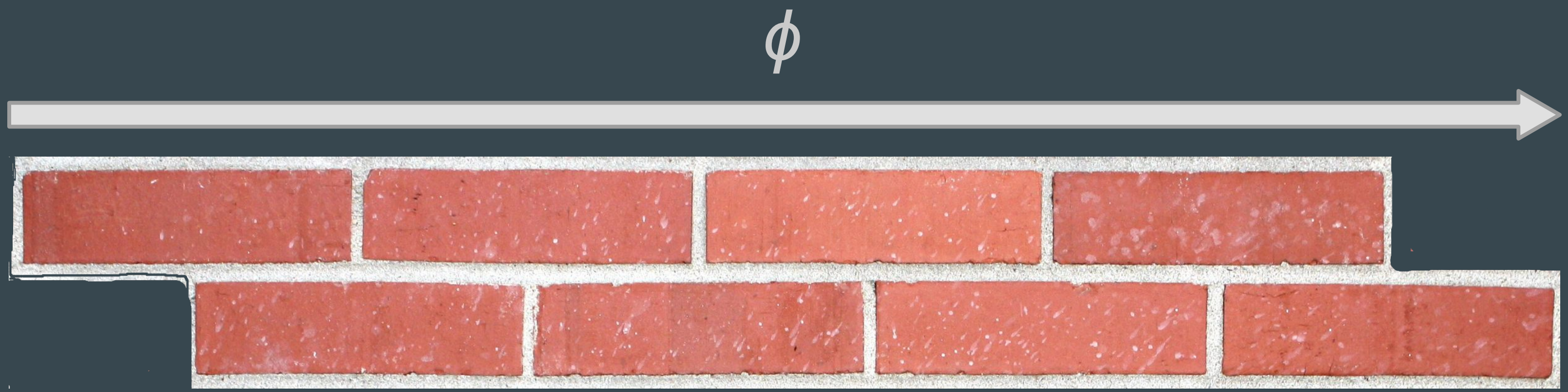


Fitting reference profiles

Each reflection is fitted against its closest reference profile



Fitting reference profiles



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

dials.report

Experiments

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

Unit cell: $\begin{pmatrix} 57.785 & 57.785 & 150.006 & 90.000 & 90.000 & 90.000 \end{pmatrix}$

U matrix:

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

B matrix:

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

A = UB:

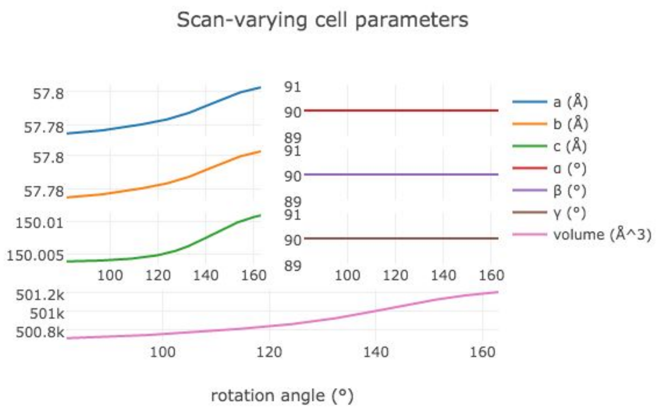
$$\begin{pmatrix} 0.0060 & -0.0045 & -0.0060 \\ 0.0154 & 0.0068 & 0.0015 \\ 0.0051 & -0.0153 & 0.0024 \end{pmatrix}$$

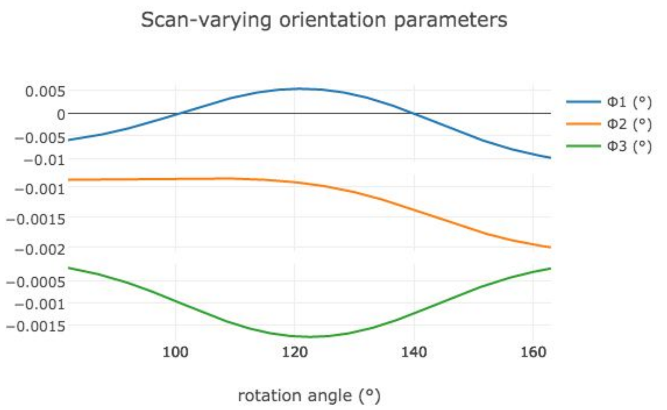
A sampled at 541 scan points

Average unit cell: $\begin{pmatrix} 57.786 & 57.786 & 150.006 & 90.000 & 90.000 & 90.000 \end{pmatrix}$

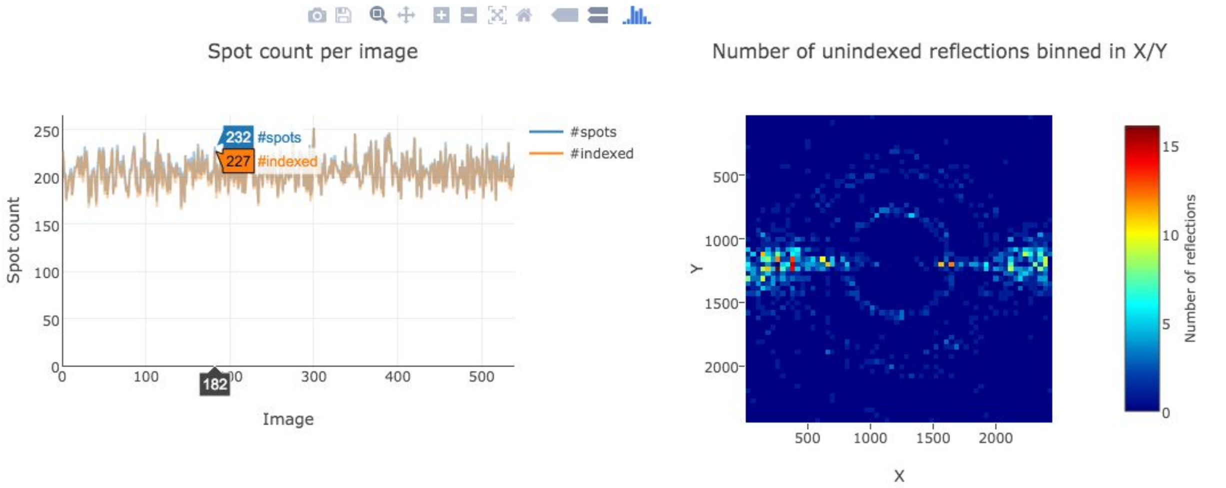
Experimental geometry

Analysis of scan-varying crystal model

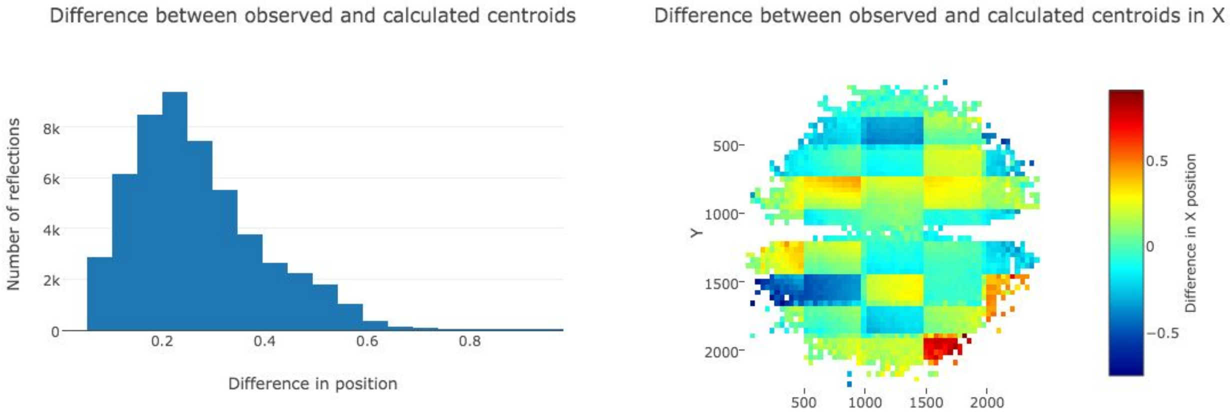




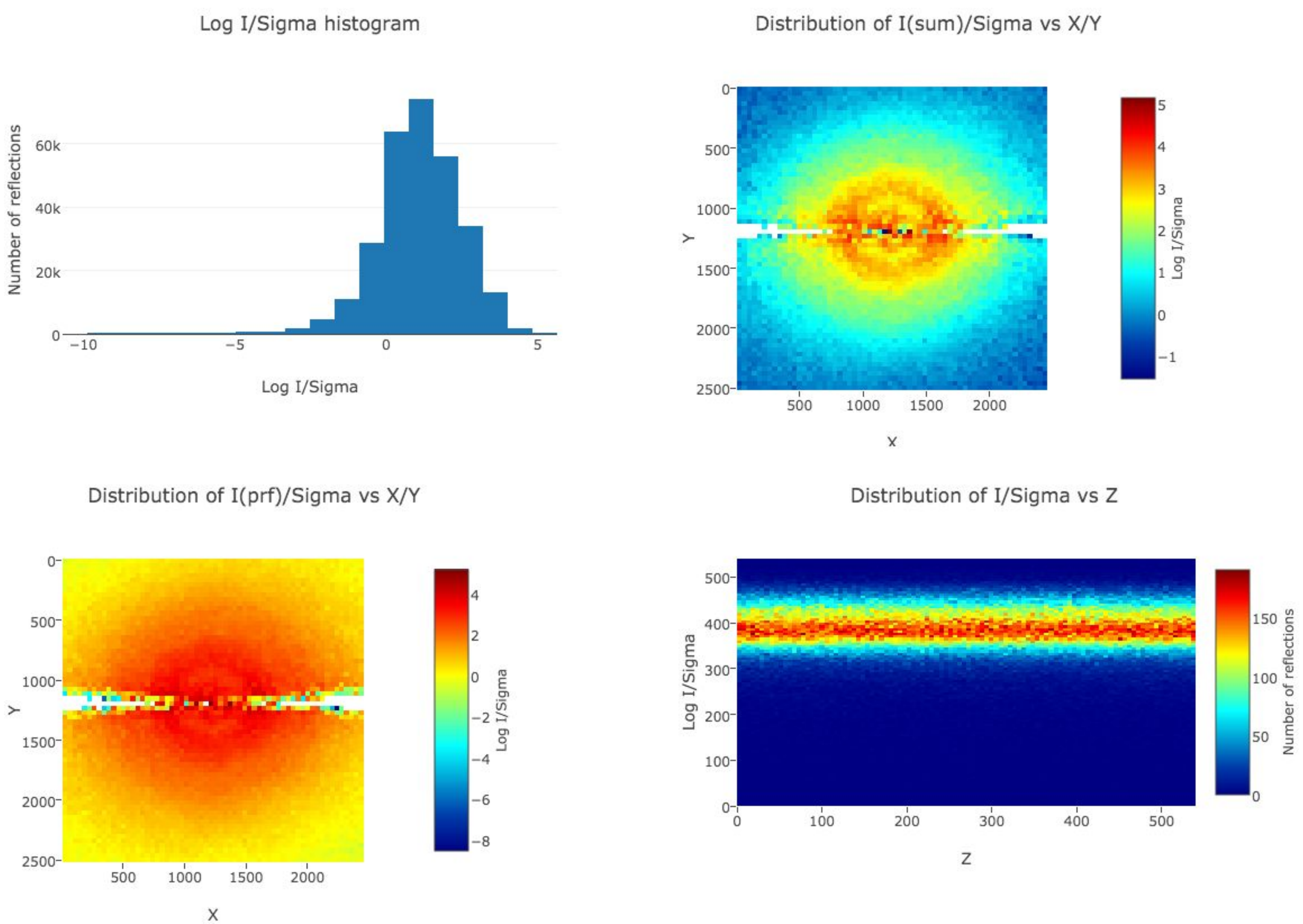
Analysis of strong reflections



Analysis of reflection centroids



Analysis of reflection intensities



DIALS

Diffraction Integration for Advanced Light Sources

http://dials.diamond.ac.uk/doc/documentation/tutorials/

DIALS

Diffraction Integration for Advanced Light Sources

Navigation

- About
- Installation
- Getting started
- Documentation
- Publications
- Links
- License

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

xia2

Navigation

Quick start guide

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. $I(+) \neq I(-)$ 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.

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 http://dials.github.io/ under BSD license
- Binary releases available for Mac and Linux
- DIALS 1.3 is included in CCP4 7.0 (including Windows)
- GUI is currently in development

Acknowledgements

DIALS East

Gwyndaf Evans, Graeme Winter, David Waterman, James Parkhurst, Richard Gildea, Luis Fuentes-Montero, Markus Gerstel, Melanie Vollmar

DIALS West

Nick Sauter, Aaron Brewster, Tara Michels-Clark, Iris Young

Lots of other people

Garib Murshudov, Andrew Leslie, Phil Evans, Harry Powell, Takanori Nakane

DIALS East- Diamond / CCP4

diamond

BioStructx

Research Complex
at Harwell

Science & Technology
Facilities Council

DIALS
Diffraction Integration for Advanced Light Sources

wellcome

BERKELEY LAB
Lawrence Berkeley National Laboratory

MRC
Laboratory of
Molecular Biology

CCP4

Thanks for listening!

...

https://dials.diamond.ac.uk