Recent Developments in xia2

Strubi early 2012 update

Graeme Winter

Diamond Light Source

1 March 2012

diamond

Overview

- CCP4 2012 presentation
- Removal of frame count limitations
- Current state of multi-crystal analysis
- Conclusions & plans

diamond

Removal of frame count limitations

- Use Aimless in place of Scala
- Use Pointless in place of Reindex[1]
- Run -3daii or -3da
- No more batch limitations

[1] This is still not as efficient as it could be - but it works

Example: 20 sweeps from insulin

- Insulin test sample from I04-1
- 20 sweeps: 0.2 degree / frame x 900: 18000 images
- Radiation damage very likely
- Substantial variation between data sets
- Processing rather time consuming, use a big workstation

diamond

High resolution limit	1.66	7.43	1.66
Low resolution limit	55.18	55.18	1.70
Completeness	100.0	99.9	100.0
Multiplicity	314.3	346.7	262.8
I/sigma	81.7	228.5	7.6
Rmerge	0.106	0.071	5.227
Rmeas(I)	0.106	0.071	5.240
Rmeas(I+/-)	0.106	0.071	5.247
Rpim(I)	0.006	0.004	0.322
Rpim(I+/-)	0.008	0.005	0.450
Wilson B factor	31.457		
Total observations	2982419	44723	180512
Total unique	9488	129	687

Effects of Running Aimless

- Output effectively the same as Scala
- SD correction slightly more effective
- C++ not Fortran - no need to recompile for lots of batches
- Parallel (openMP) version in development

diamond

Effects of Running Aimless in xia2

- Since scaling performed by XSCALE very little difference in results

diamond

Current state of multi-crystal analysis

- Developed over 1 year or so
- Hiatus for most of that time
- Still very rough around the edges
- Requires use of XDS processing
- Requires R (at the moment) to generate plot

diamond

Basic principles

- Assume in processing that everything comes from the same crystal
- Try to assemble a data set from everything
- Worry about isomorphism in scaling rather than from unit cell etc.
- Look at CC's between sweeps as measure of isomorphism
- Scale CC's to ersatz distance as $d = \frac{1}{CC} - 1$[2]

[2] Which means you can *interpret* the distances.

Basic Usage

```
xia2 -3daii /here/are/my/data
```

diamond

Time passes... let's talk about options

diamond

Useful options

- -microcrystal - switch to microcrystal mode
- -failover - ignore failure of processing individual sweeps
- -spacegroup - set the spacegroup as e.g. P23
- -cell $a, b, c, \alpha, \beta, \gamma$ - needs -spacegroup

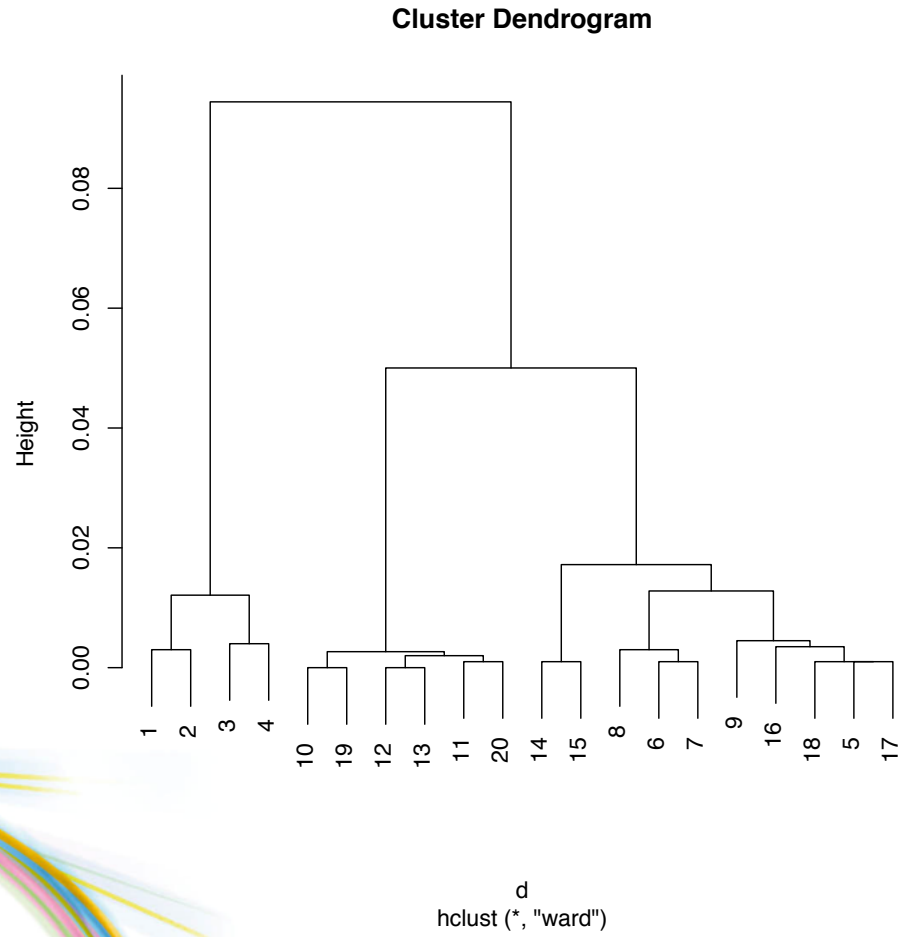
diamond

After processing

```
> cd DEFAULT/scale
> x1335 XSCALE.LP
> R --no-save < x1335.R
```

diamond

Results

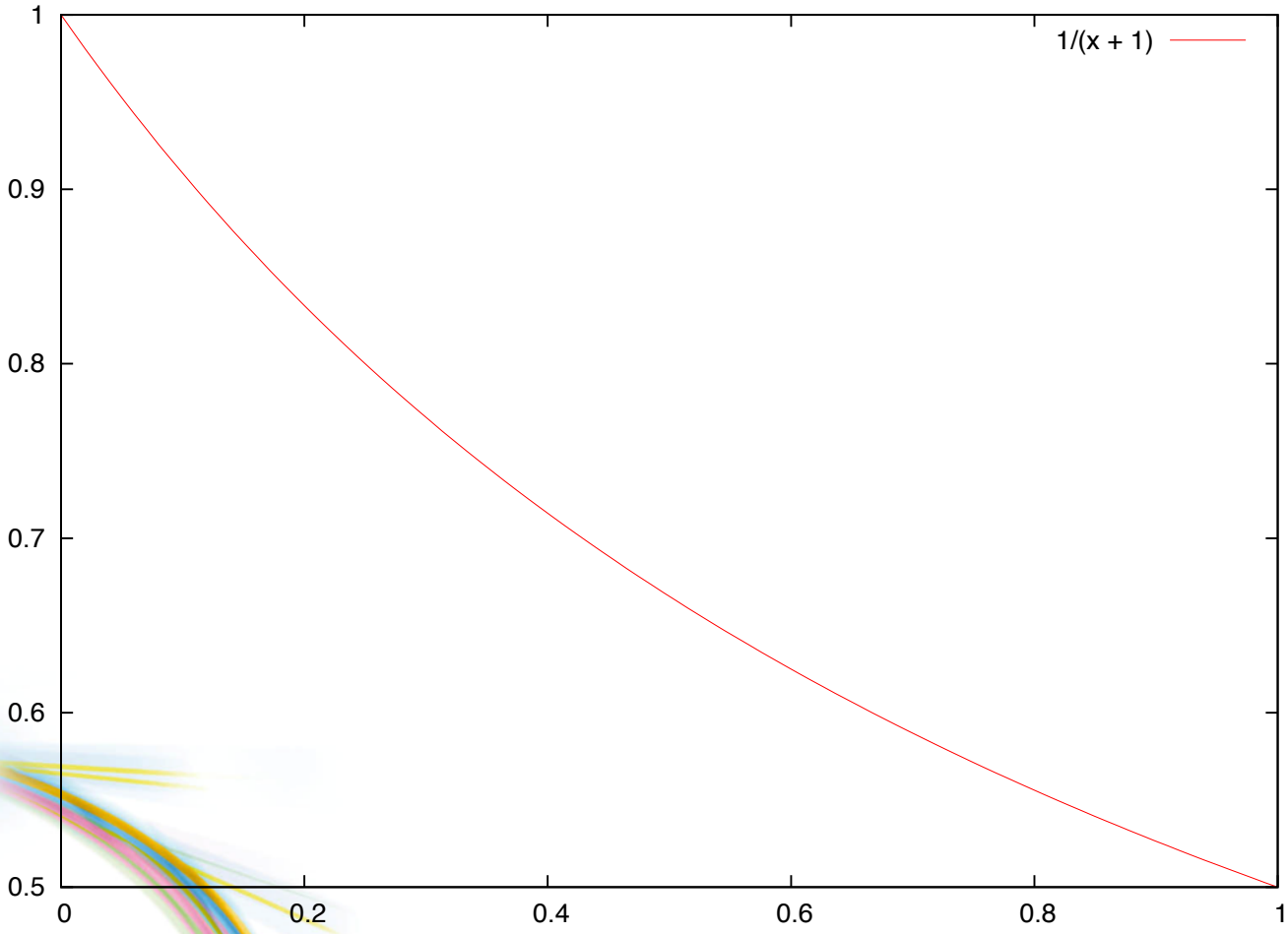

Cluster Dendrogram
Height
0.00
0.02
0.04
0.06
0.08
1
2
3
4
10
19
12
13
11
20
14
15
8
6
7
9
16
18
5
17
d
hclust (*, "ward")

Interpretation

- $d = \frac{1}{CC} - 1$
- $CC = \frac{1}{d+1}$
- $CC = 0.95 \rightarrow d < 0.05$
- Up to you to decide on your limits...

diamond

Interpretation



No.	Av. CC	I/sig	File Name
1	0.9796	2.25	NATIVE_SWEEP11.HKL
2	0.9822	3.61	NATIVE_SWEEP22.HKL
3	0.9807	7.49	NATIVE_SWEEP33.HKL
4	0.9869	9.85	NATIVE_SWEEP34.HKL
5	0.9930	10.93	NATIVE_SWEEP35.HKL
6	0.9939	2.34	NATIVE_SWEEP36.HKL
7	0.9913	3.60	NATIVE_SWEEP37.HKL
8	0.9902	7.04	NATIVE_SWEEP38.HKL
9	0.9922	9.89	NATIVE_SWEEP39.HKL
10	0.9868	13.14	NATIVE_SWEEP1.HKL
11	0.9904	1.85	NATIVE_SWEEP2.HKL
12	0.9889	1.78	NATIVE_SWEEP3.HKL
13	0.9885	1.14	NATIVE_SWEEP4.HKL
14	0.9907	1.74	NATIVE_SWEEP5.HKL
15	0.9904	2.70	NATIVE_SWEEP6.HKL
16	0.9926	4.12	NATIVE_SWEEP7.HKL
17	0.9924	2.60	NATIVE_SWEEP8.HKL
18	0.9925	2.64	NATIVE_SWEEP9.HKL

Next steps

- Book keeping - take the sweeps you wish to merge from dendogram
- Edit xinfo file
- Run again
- This needs automating, and also calculations performing to tell you how complete results will be