PAUL SCHERRER INSTITUT

PSI

Karol Nass : Instrument Scientist : Paul Scherrer Institut

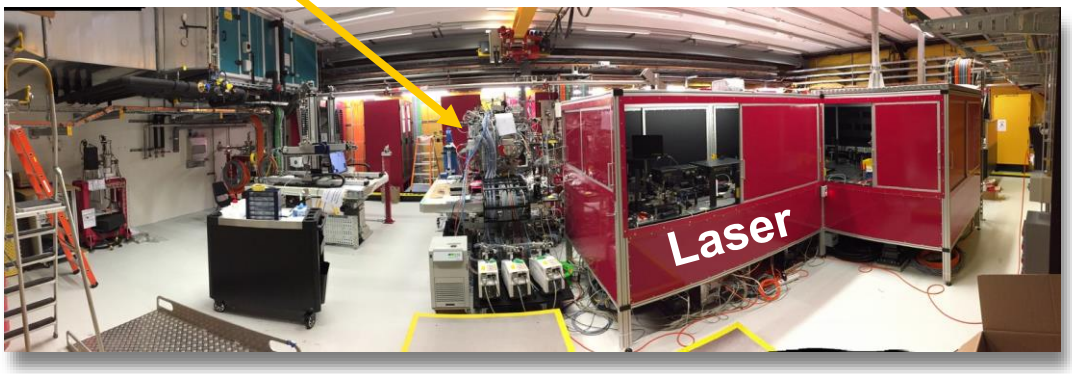
Online and offline SFX data analysis at SwissFEL

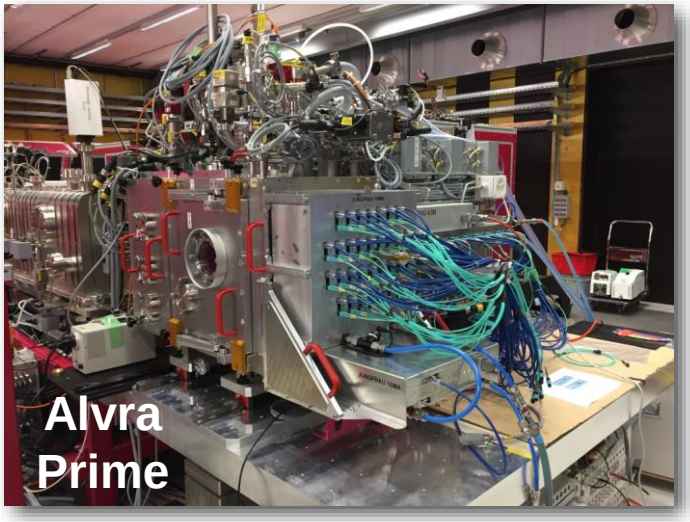
Göteborg, January 28, 2020

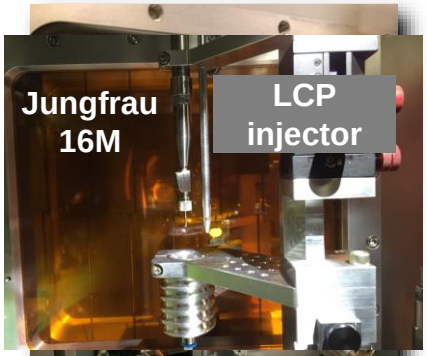
Experimental Station Alvra

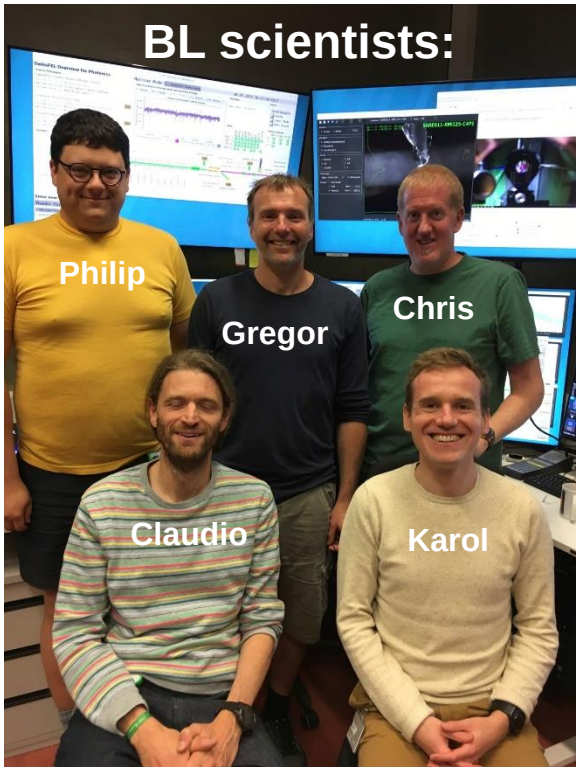
Ultrafast photochemistry and photobiology

Alvra Prime: *(Tender)* X-ray spectrometer combined with X-ray scattering







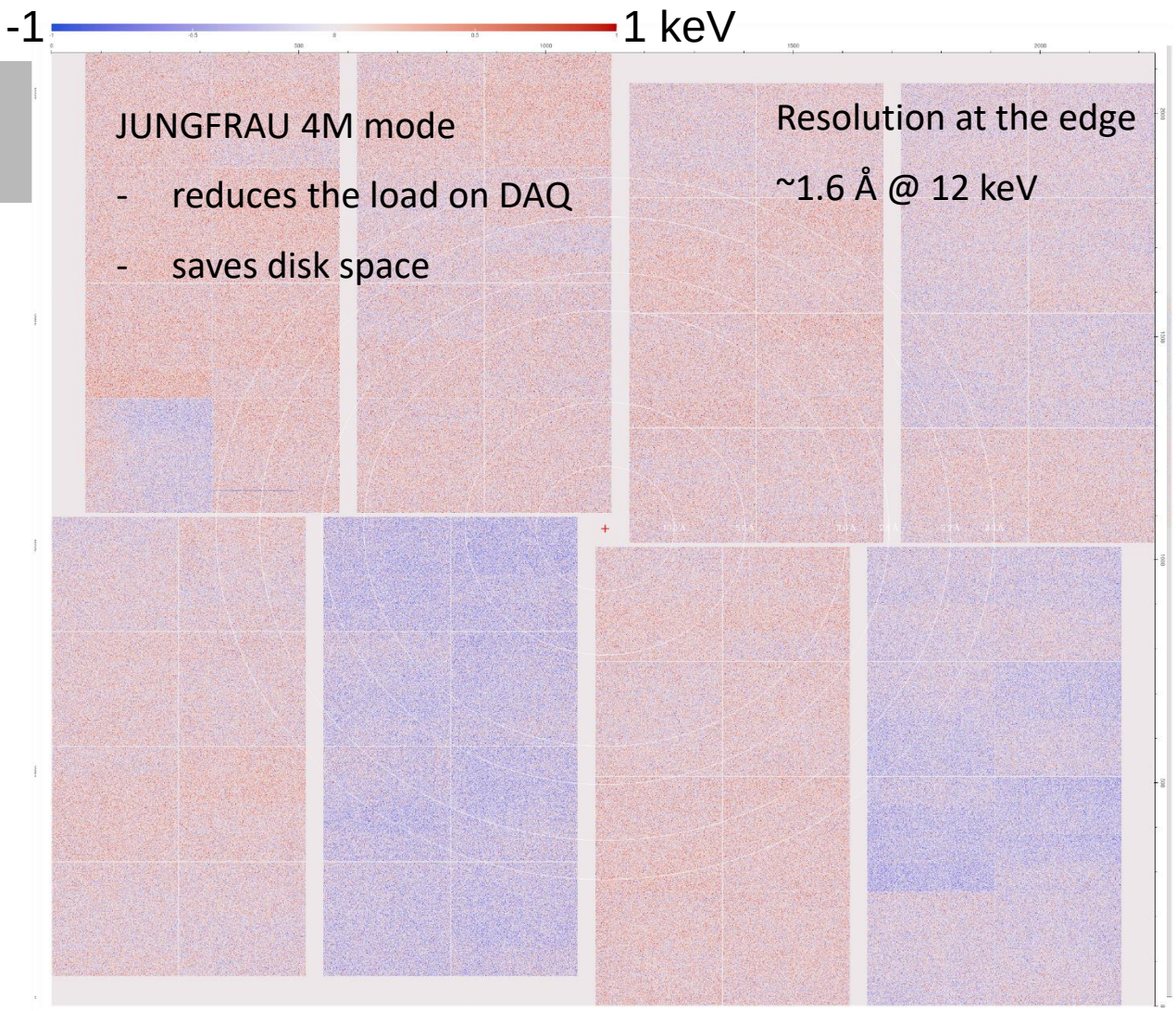


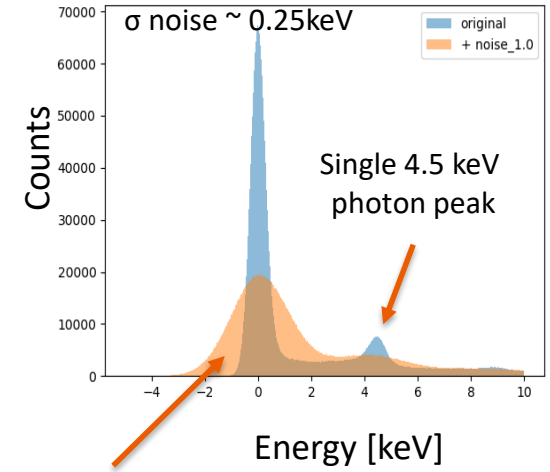
P. Jahnson, G. Knopp, C. Milne, C. Cirelli, K. Nass

IT support: Dmitry Ozerov, Ivan Usov

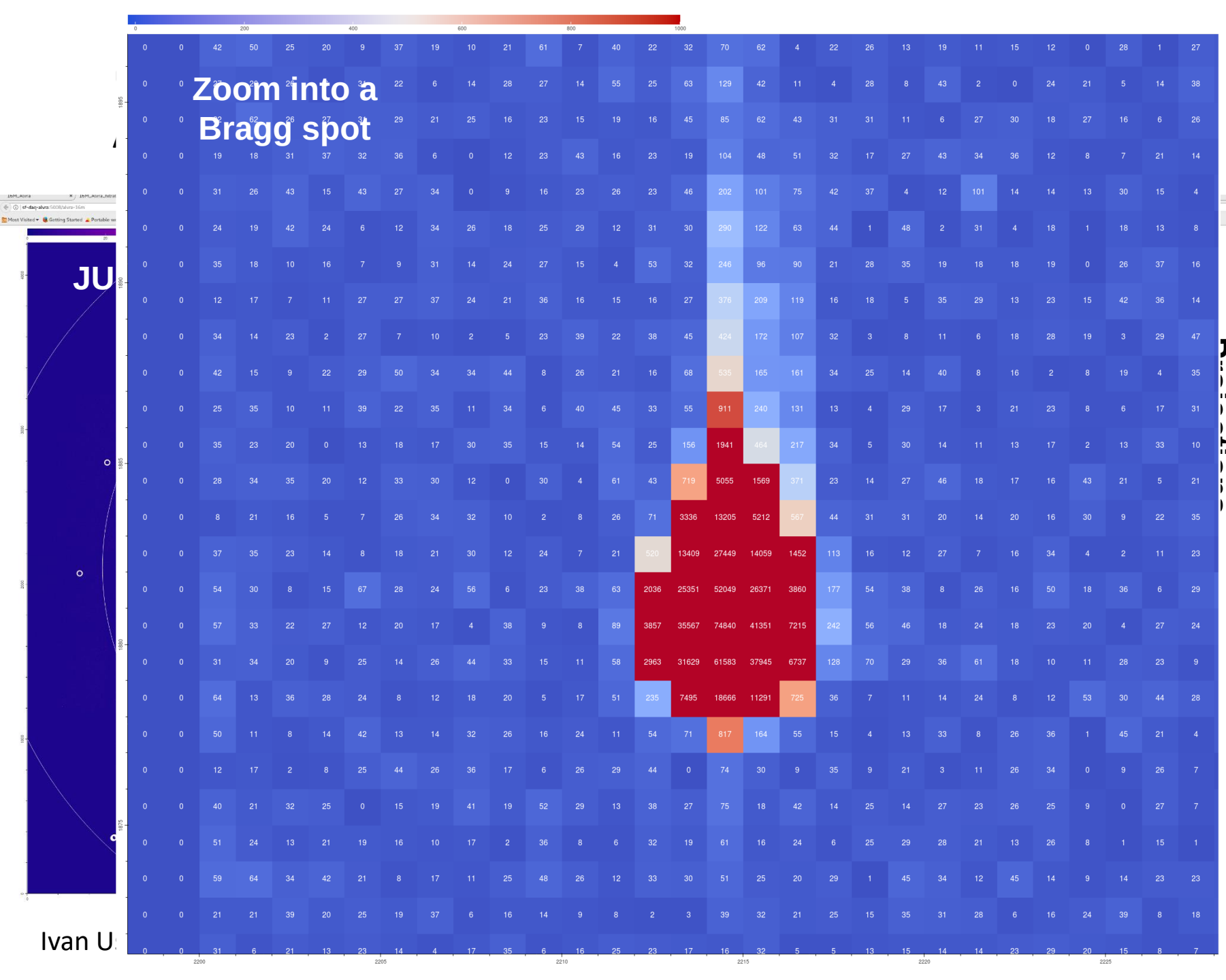
http://www.psi.ch/alvra/

JUNGFRAU 16M / 4M detector at SwissFEL



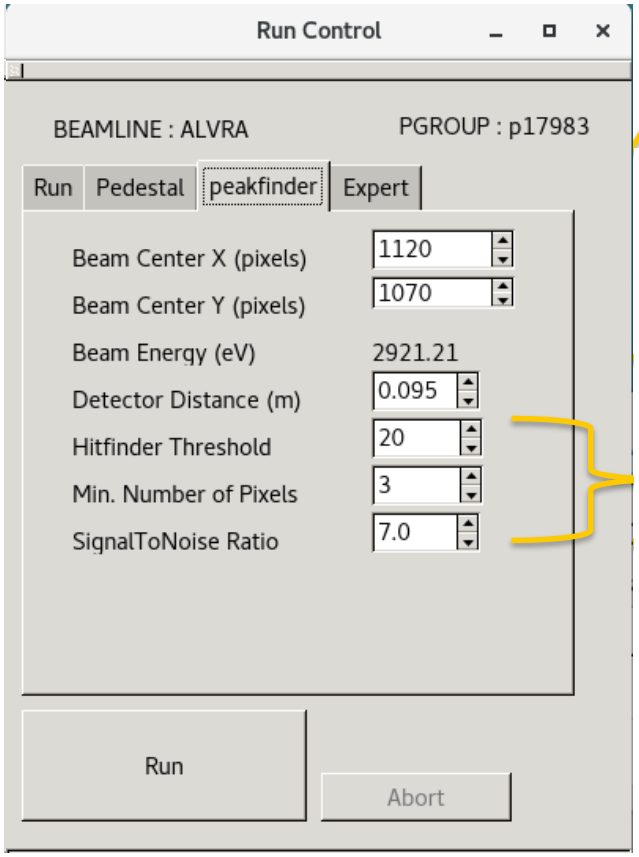


- Low noise
- 75 x 75 µm^2 pixel size
- 3 stage gain switching per pixel
- High dynamic range
- 1·10^4 x 12.4 keV photons


Zoom into a Bragg spot

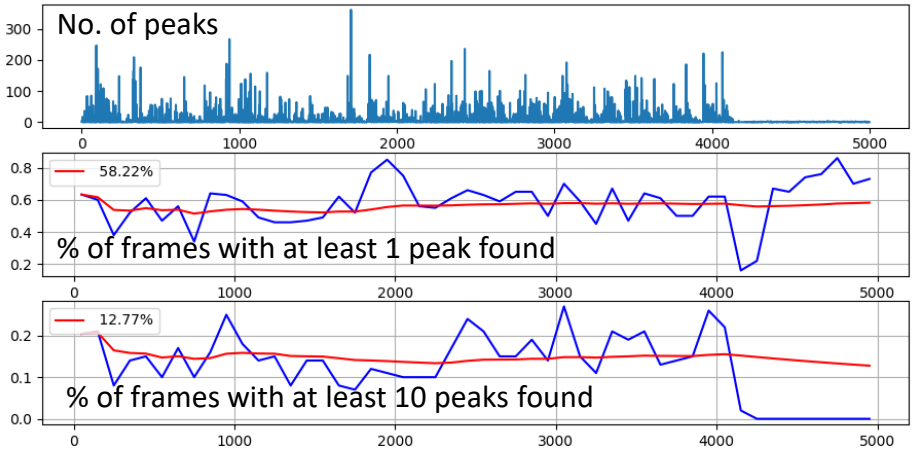
Run Control and Hit Finder at SwissFEL

Application that controls data acquisition and defines peak finder parameters for jet based SFX



1. Enter sample name and number of frames per run

Online hitrate monitor



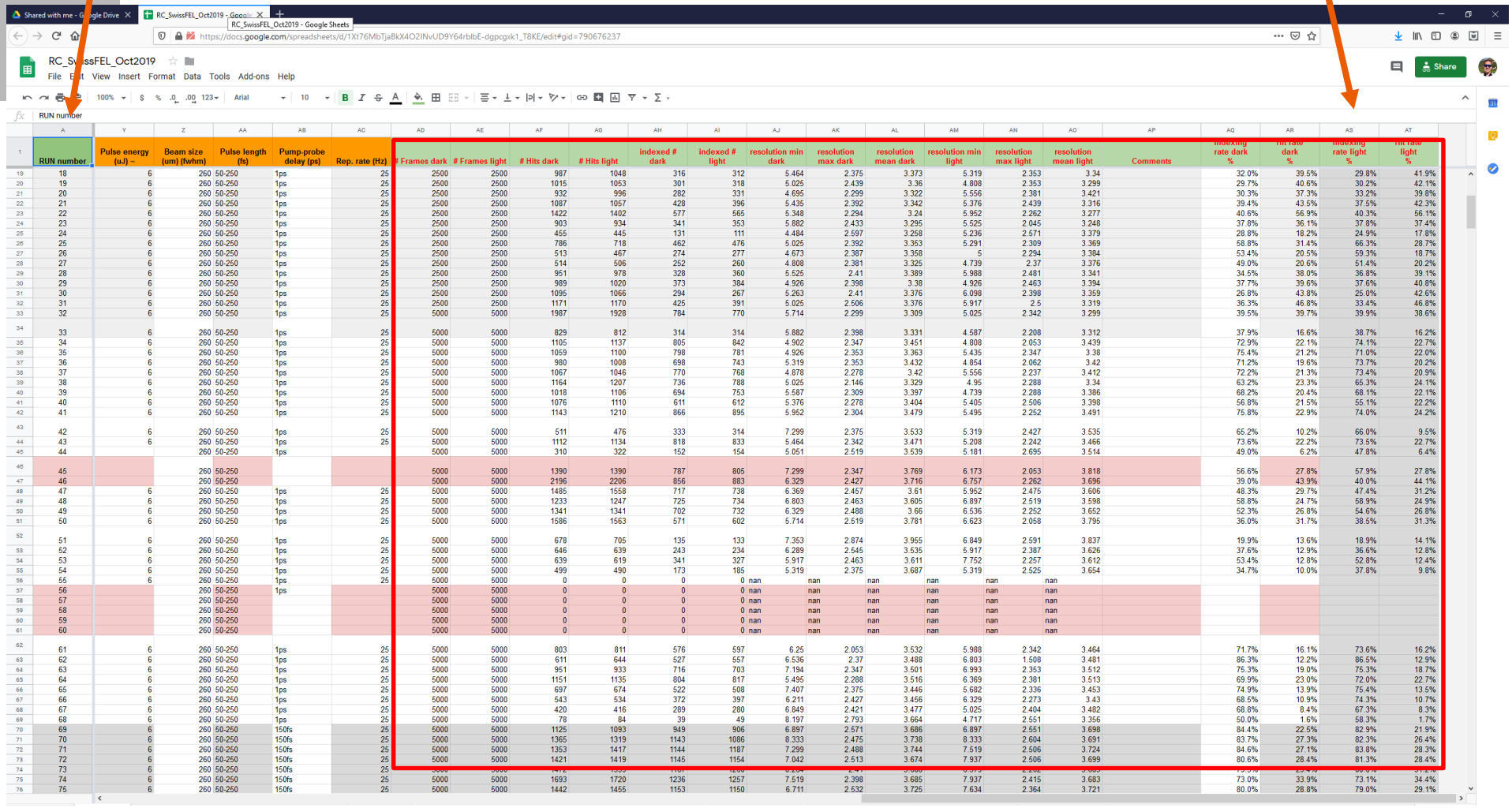
CrystFEL is integrated into the data acquisition system at SwissFEL and runs continuesly in the background. It provides live feedback and preliminary results during beam time.

Dmitry Ozerov

Automated indexing and online spreadsheet update

Reduces the load on data processing and logbook teams during the experiment!

Run number, # Frames (dark/light), # Hits (dark/light), # Indexed (dark/light), Resolution, Hit/indexing rate



RUN number	Pulse energy (uJ) ~	Beam size (um) (fwhm)	Pulse length (fs)	Pump-probe delay (ps)	Rep. rate (Hz)	# Frames dark	# Frames light	# Hits dark	# Hits light	indexed # dark	indexed # light	resolution min dark	resolution max dark	resolution mean dark	resolution min light	resolution max light	resolution mean light	Comments	indexing rate dark %	hit rate dark %	indexing rate light %	hit rate light %
18	6	260	50-250	1ps	25	2500	2500	987	1048	316	312	5.464	2.375	3.373	5.319	2.353	3.34		32.0%	39.5%	29.8%	41.9%
19	6	260	50-250	1ps	25	2500	2500	1015	1053	301	318	5.025	2.439	3.36	4.808	2.353	3.299		29.7%	40.6%	30.2%	42.1%
20	6	260	50-250	1ps	25	2500	2500	932	996	282	331	4.695	2.299	3.322	5.556	2.381	3.421		30.3%	37.3%	33.2%	39.8%
21	6	260	50-250	1ps	25	2500	2500	1087	1057	428	396	5.435	2.392	3.342	5.376	2.439	3.316		39.4%	43.5%	37.5%	42.3%
22	6	260	50-250	1ps	25	2500	2500	1422	1402	577	565	5.348	2.294	3.24	5.952	2.262	3.277		40.6%	56.9%	40.3%	56.1%
23	6	260	50-250	1ps	25	2500	2500	903	934	341	353	5.882	2.433	3.295	5.525	2.045	3.248		37.8%	36.1%	37.8%	37.4%
24	6	260	50-250	1ps	25	2500	2500	455	445	131	111	4.484	2.597	3.258	5.236	2.571	3.379		28.8%	18.2%	24.9%	17.8%
25	6	260	50-250	1ps	25	2500	2500	786	718	462	476	5.025	2.392	3.353	5.291	2.309	3.369		58.8%	31.4%	66.3%	28.7%
26	6	260	50-250	1ps	25	2500	2500	513	467	274	277	4.673	2.387	3.358	5	2.294	3.384		53.4%	20.5%	59.3%	18.7%
27	6	260	50-250	1ps	25	2500	2500	514	506	252	260	4.808	2.381	3.325	4.739	2.37	3.376		49.0%	20.6%	51.4%	20.2%
28	6	260	50-250	1ps	25	2500	2500	951	978	328	360	5.525	2.41	3.389	5.988	2.481	3.341		34.5%	38.0%	36.8%	39.1%
29	6	260	50-250	1ps	25	2500	2500	989	1020	373	384	4.926	2.398	3.38	4.926	2.463	3.394		37.7%	39.6%	37.6%	40.8%
30	6	260	50-250	1ps	25	2500	2500	1095	1066	294	267	5.263	2.41	3.376	6.098	2.398	3.359		26.8%	43.8%	25.0%	42.6%
31	6	260	50-250	1ps	25	2500	2500	1171	1170	425	391	5.025	2.506	3.376	5.917	2.5	3.319		36.3%	46.8%	33.4%	46.8%
32	6	260	50-250	1ps	25	5000	5000	1987	1928	784	770	5.714	2.299	3.309	5.025	2.342	3.299		39.5%	39.7%	39.9%	38.6%
33	6	260	50-250	1ps	25	5000	5000	829	812	314	314	5.882	2.398	3.331	4.587	2.208	3.312		37.9%	16.6%	38.7%	16.2%
34	6	260	50-250	1ps	25	5000	5000	1105	1137	805	842	4.902	2.347	3.451	4.808	2.053	3.439		72.9%	22.1%	74.1%	22.7%
35	6	260	50-250	1ps	25	5000	5000	1059	1100	798	781	4.926	2.353	3.363	5.435	2.347	3.38		75.4%	21.2%	71.0%	22.0%
36	6	260	50-250	1ps	25	5000	5000	980	1008	698	743	5.319	2.353	3.432	4.854	2.062	3.42		71.2%	19.6%	73.7%	20.2%
37	6	260	50-250	1ps	25	5000	5000	1067	1046	770	768	4.878	2.278	3.42	5.556	2.237	3.412		72.2%	21.3%	73.4%	20.9%
38	6	260	50-250	1ps	25	5000	5000	1164	1207	736	788	5.025	2.146	3.329	4.95	2.288	3.34		63.2%	23.3%	65.3%	24.1%
39	6	260	50-250	1ps	25	5000	5000	1018	1106	694	753	5.587	2.309	3.397	4.739	2.288	3.386		68.2%	20.4%	68.1%	22.1%
40	6	260	50-250	1ps	25	5000	5000	1076	1110	611	612	5.376	2.278	3.404	5.405	2.506	3.398		56.8%	21.5%	55.1%	22.2%
41	6	260	50-250	1ps	25	5000	5000	1143	1210	866	895	5.952	2.304	3.479	5.495	2.252	3.491		75.8%	22.9%	74.0%	24.2%
42	6	260	50-250	1ps	25	5000	5000	511	476	333	314	7.299	2.375	3.533	5.319	2.427	3.535		65.2%	10.2%	66.0%	9.5%
43	6	260	50-250	1ps	25	5000	5000	1112	1134	818	833	5.464	2.342	3.471	5.208	2.242	3.466		73.6%	22.2%	73.5%	22.7%
44		260	50-250	1ps		5000	5000	310	322	152	154	5.051	2.519	3.539	5.181	2.695	3.514		49.0%	6.2%	47.8%	6.4%
45		260	50-250			5000	5000	1390	1390	787	805	7.299	2.347	3.769	6.173	2.053	3.818		56.6%	27.8%	57.9%	27.8%
46		260	50-250			5000	5000	2196	2206	856	883	6.329	2.427	3.716	6.757	2.262	3.696		39.0%	43.9%	40.0%	44.1%
47	6	260	50-250	1ps	25	5000	5000	1485	1558	717	738	6.369	2.457	3.61	5.952	2.475	3.606		48.3%	29.7%	47.4%	31.2%
48	6	260	50-250	1ps	25	5000	5000	1233	1247	725	734	6.803	2.463	3.605	6.897	2.519	3.598		58.8%	24.7%	58.9%	24.9%
49	6	260	50-250	1ps	25	5000	5000	1341	1341	702	732	6.329	2.488	3.66	6.536	2.252	3.652		52.3%	26.8%	54.6%	26.8%
50	6	260	50-250	1ps	25	5000	5000	1586	1563	571	602	5.714	2.519	3.781	6.623	2.058	3.795		36.0%	31.7%	38.5%	31.3%
51	6	260	50-250	1ps	25	5000	5000	678	705	135	133	7.353	2.874	3.955	6.849	2.591	3.837		19.9%	13.6%	18.9%	14.1%
52	6	260	50-250	1ps	25	5000	5000	646	639	243	234	6.289	2.545	3.535	5.917	2.387	3.626		37.6%	12.9%	36.6%	12.8%
53	6	260	50-250	1ps	25	5000	5000	639	619	341	327	5.917	2.463	3.611	7.752	2.257	3.612		53.4%	12.8%	52.8%	12.4%
54	6	260	50-250	1ps	25	5000	5000	499	490	173	185	5.319	2.375	3.687	5.319	2.525	3.654		34.7%	10.0%	37.8%	9.8%
55	6	260	50-250	1ps	25	5000	5000	0	0	0	0	nan	nan	nan	nan	nan	nan					
56		260	50-250	1ps		5000	5000	0	0	0	0	nan	nan	nan	nan	nan	nan					
57		260	50-250			5000	5000	0	0	0	0	nan	nan	nan	nan	nan	nan					
58		260	50-250			5000	5000	0	0	0	0	nan	nan	nan	nan	nan	nan					
59		260	50-250			5000	5000	0	0	0	0	nan	nan	nan	nan	nan	nan					
60		260	50-250			5000	5000	0	0	0	0	nan	nan	nan	nan	nan	nan					
61	6	260	50-250	1ps	25	5000	5000	803	811	576	597	6.25	2.053	3.532	5.988	2.342	3.464		71.7%	16.1%	73.6%	16.2%
62	6	260	50-250	1ps	25	5000	5000	611	644	527	557	6.536	2.37	3.488	6.803	1.508	3.481		86.3%	12.2%	86.5%	12.9%
63	6	260	50-250	1ps	25	5000	5000	951	933	716	703	7.194	2.347	3.501	6.993	2.353	3.512		75.3%	19.0%	75.3%	18.7%
64	6	260	50-250	1ps	25	5000	5000	1151	1135	804	817	5.495	2.288	3.516	6.369	2.381	3.513		69.9%	23.0%	72.0%	22.7%
65	6	260	50-250	1ps	25	5000	5000	697	674	522	508	7.407	2.375	3.446	5.682	2.336	3.453		74.9%	13.9%	75.4%	13.5%
66	6	260	50-250	1ps	25	5000	5000	543	534	372	397	6.211	2.427	3.456	6.329	2.273	3.43		68.5%	10.9%	74.3%	10.7%
67	6	260	50-250	1ps	25	5000	5000	420	416	289	280	6.849	2.421	3.477	5.025	2.404	3.482		68.8%	8.4%	67.3%	8.3%
68	6	260	50-250	1ps	25	5000	5000	78	84	39	49	8.197	2.793	3.664	4.717	2.551	3.356		50.0%	1.6%	58.3%	1.7%
69	6	260	50-250	150fs	25	5000	5000	1125	1093	949	906	6.897	2.571	3.686	6.897	2.551	3.698		84.4%	22.5%	82.9%	21.9%
70	6	260	50-250	150fs	25	5000	5000	1365	1319	1143	1086	8.333	2.475	3.738	8.333	2.604	3.691		83.7%	27.3%	82.3%	26.4%
71	6	260	50-250	150fs	25	5000	5000	1353	1417	1144	1187	7.299	2.488	3.744	7.519	2.506	3.724		84.6%	27.1%	83.8%	28.3%
72	6	260	50-250	150fs	25	5000	5000	1421	1419	1145	1154	7.042	2.513	3.674	7.937	2.506	3.699		80.6%	28.4%	81.3%	28.4%
73	6	260	50-250	150fs	25	5000	5000	1472	1505	1161	1200	6.804	2.41	3.668	6.993	2.262	3.665		79.0%	29.4%	80.0%	31.2%
74	6	260	50-250	150fs	25	5000	5000	1693	1720	1236	1257	7.519	2.398	3.685	7.937	2.415	3.683		73.0%	33.9%	73.1%	34.4%
75	6	260	50-250	150fs	25	5000	5000	1442	1455	1153	1150	6.711	2.532	3.725	7.634	2.364	3.721		80.0%	28.8%	79.0%	29.1%

Depending on the number of frames and hit rate, online CrystFEL results appear ~2 – 10 min after run ends.

Computing resources at SwissFEL/PSI

- **online resources:**
  - Used during beamtime for data conversion, hit finding, auto indexing and auto merging with CrystFEL, no user access
  - 16 nodes (2 used as login), each with 36 cores, 256GB memory
  - Total online resources: 14 computing nodes, 504 cores
- **offline resources (Ra cluster):**
  - Used for data analysis, remote access (NoMachine, ssh)

• 16 nodes	24 cores	256 GB memory
• 16 nodes	32 cores	256 GB memory
• 14 nodes	36 cores	256 GB memory
• 2 login nodes	36 cores	512 GB memory
• 1 gpu node		

  - Total offline resources: 46 computing nodes, 1400 cores

Where is the data on Ra and how to access it?

How to configure NoMachine to access the PSI offline cluster (Ra):
https://www.psi.ch/en/photon-science-data-services/remote-interactive-access

```
[nass_k@ra-1-002 p17984]$ pwd
/das/work/p17/p17984
[nass_k@ra-1-002 p17984]$ tree -d -C RC_1ps
RC_1ps
├── data
│   ├── dark
│   ├── light
│   └── pixel_mask
├── index
│   ├── dark
│   └── light
└── merge
    ├── dark
    │   ├── final
    │   ├── final_previous
    │   ├── merged_10000
    │   └── merged_15000
    └── light
        ├── final
        ├── final_previous
        ├── merged_10000
        └── merged_15000
```

Work directory

Frames with hits found online are arranged in directories named as entered in “Run Control”

Sorted into light/dark

Indexing results from auto processing

Merging results

- Every 5k indexed images
- Final (all indexed images in RC_1ps)
- Sorted into dark/light

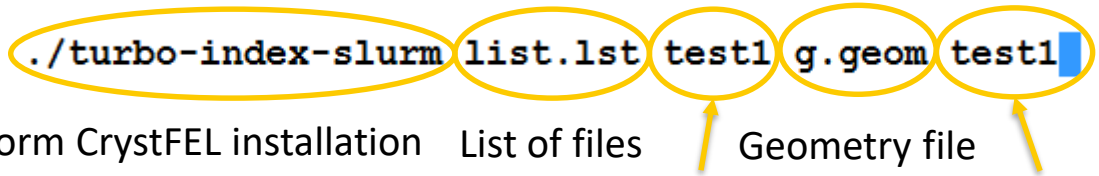
How to re-process the data on Ra?

Access to computing nodes on Ra is controlled by Slurm, a modern work load manager for Linux clusters.

More info how to use the batch system:

https://www.psi.ch/en/photon-science-data-services/offline-computing-facility-for-sls-and-swissfel-data-analysis

Typical command used to submit indexing jobs to the cluster:

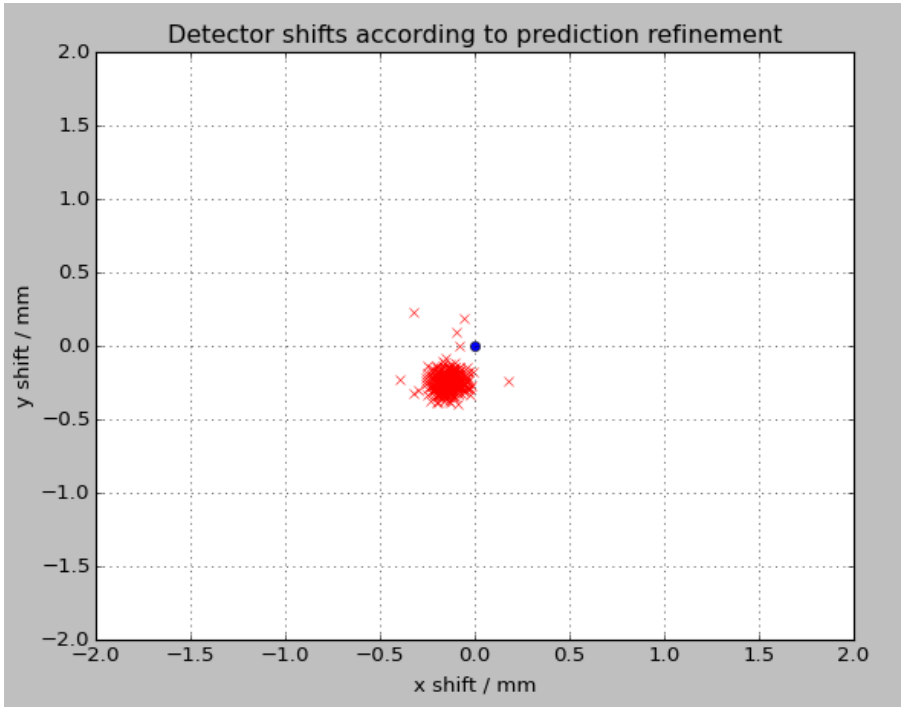



```
SLURMFILE="$PWD/$STREAMDIR/${NAME}.sh"

echo "#!/bin/sh" > $SLURMFILE
echo >> $SLURMFILE

echo "#SBATCH --partition=hour" >> $SLURMFILE  # Set your partition here
echo "#SBATCH --time=00:30:00" >> $SLURMFILE
echo "#SBATCH --nodes=1" >> $SLURMFILE
echo "#SBATCH --nice=100" >> $SLURMFILE # Set priority very low to allow other jobs through
echo >> $SLURMFILE

echo "#SBATCH --workdir   $PWD/$STREAMDIR" >> $SLURMFILE
echo "#SBATCH --job-name  $NAME" >> $SLURMFILE
echo "#SBATCH --output    $NAME-%N-%j.out" >> $SLURMFILE
echo "#SBATCH --error     $NAME-%N-%j.err" >> $SLURMFILE
echo "##SBATCH --mail-type END" >> $SLURMFILE
echo "##SBATCH --mail-user $MAIL" >> $SLURMFILE
echo >> $SLURMFILE

echo "source /etc/scripts/mx_fel.sh" >> $SLURMFILE  # Set up environment here (again) if necessary
echo >> $SLURMFILE

echo "nproc=\`grep proce /proc/cpuinfo | wc -l\`" >> $SLURMFILE

command="indexamajig -i $FILE -o $PWD/$STREAMDIR/$STREAM --indexing=xgandalf-latt-cell,asdf-latt-cell \
--geometry=$PWD/$GEOM --pdb=$PWD/rc-new2.cell --peaks=peakfinder8 --integration=rings-grad \
--tolerance=2.0,2.0,2.0,2,2,2 --threshold=50 --min-gradient=1000 --min-snr=5 --int-radius=3,4,6 \
-j \${nproc} --no-multi --no-retry --check-peaks --max-res=3000"

echo $command >> $SLURMFILE

sbatch $SLURMFILE
```

Geometry calibration

JUNGFRAU modules are fixed to the base plate. SwissFEL provides very well optimized geometry.

What possibly needs to be re-calibrated?



1. **Beam center**

```
python detector-shift test1.stream
```

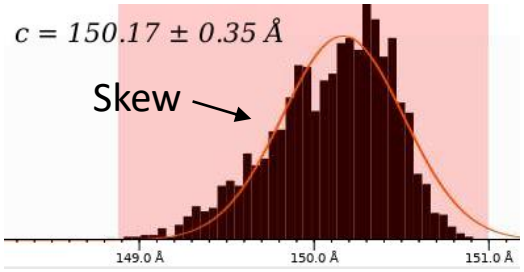
2. **Detector distance**

```
./submit_crystfel_optimise_detdist_CMC-kn.sh 1.lst g.geom
```

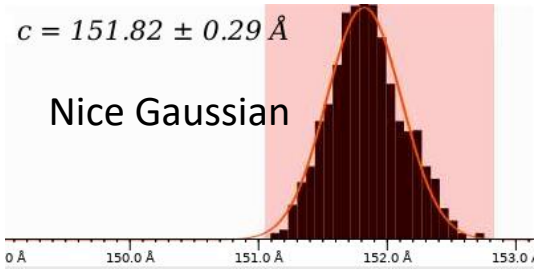
Detector distance calibration, why?

It has an effect on the unit cell parameters, width of their distribution and accuracy of the data!
Wrong detector distance introduces unwanted systematic errors.

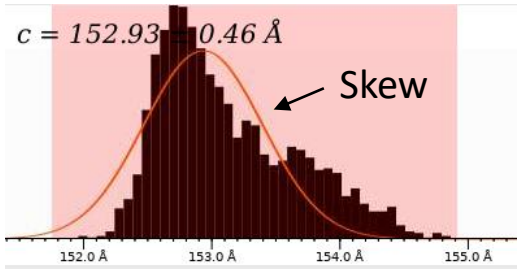
Too short by 100 µm

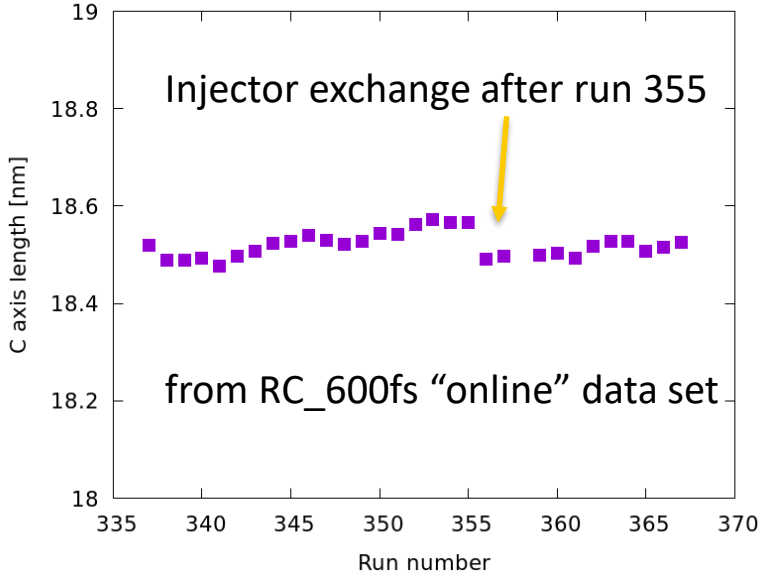


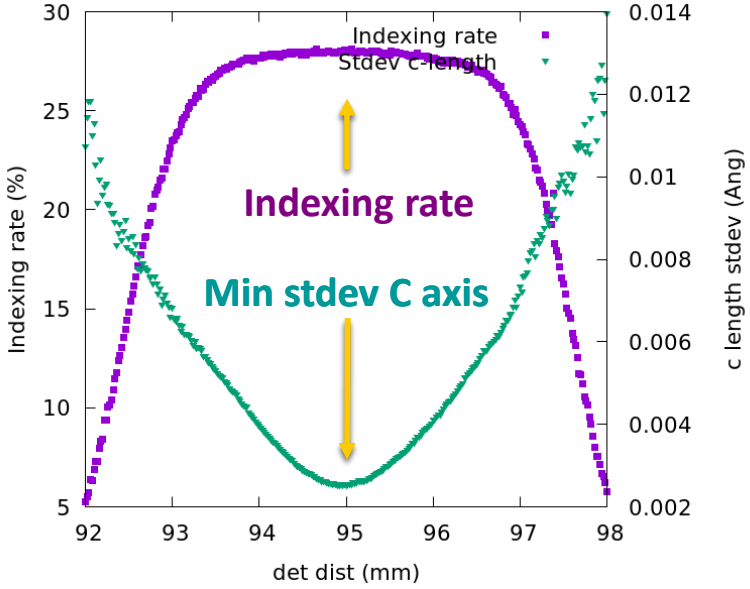
Correct detector distance



Too long by 100 µm







PAUL SCHERRER INSTITUT

Thank you for your attention!