A Light for Science

ESRF

European Synchrotron Radiation Facility

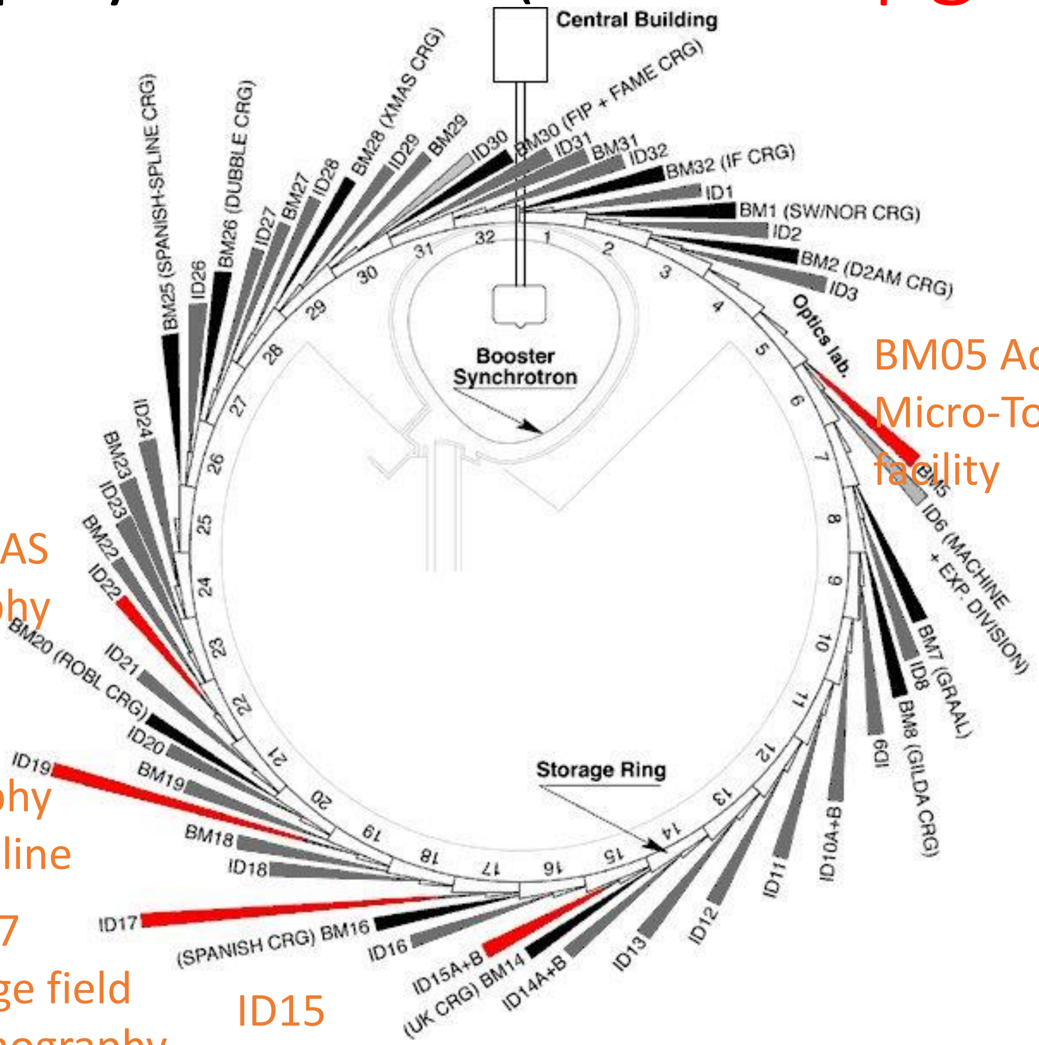
Microtomography at the ESRF:

Acquisition – Reconstruction – Storage and Analysis

Outlines:

- Data acquisition
- Tomographic reconstruction
- Post-reconstruction processes
- Data storage
- Data analysis

Tomography at ESRF (under upgrade):



BM05 Additional Micro-Tomography facility

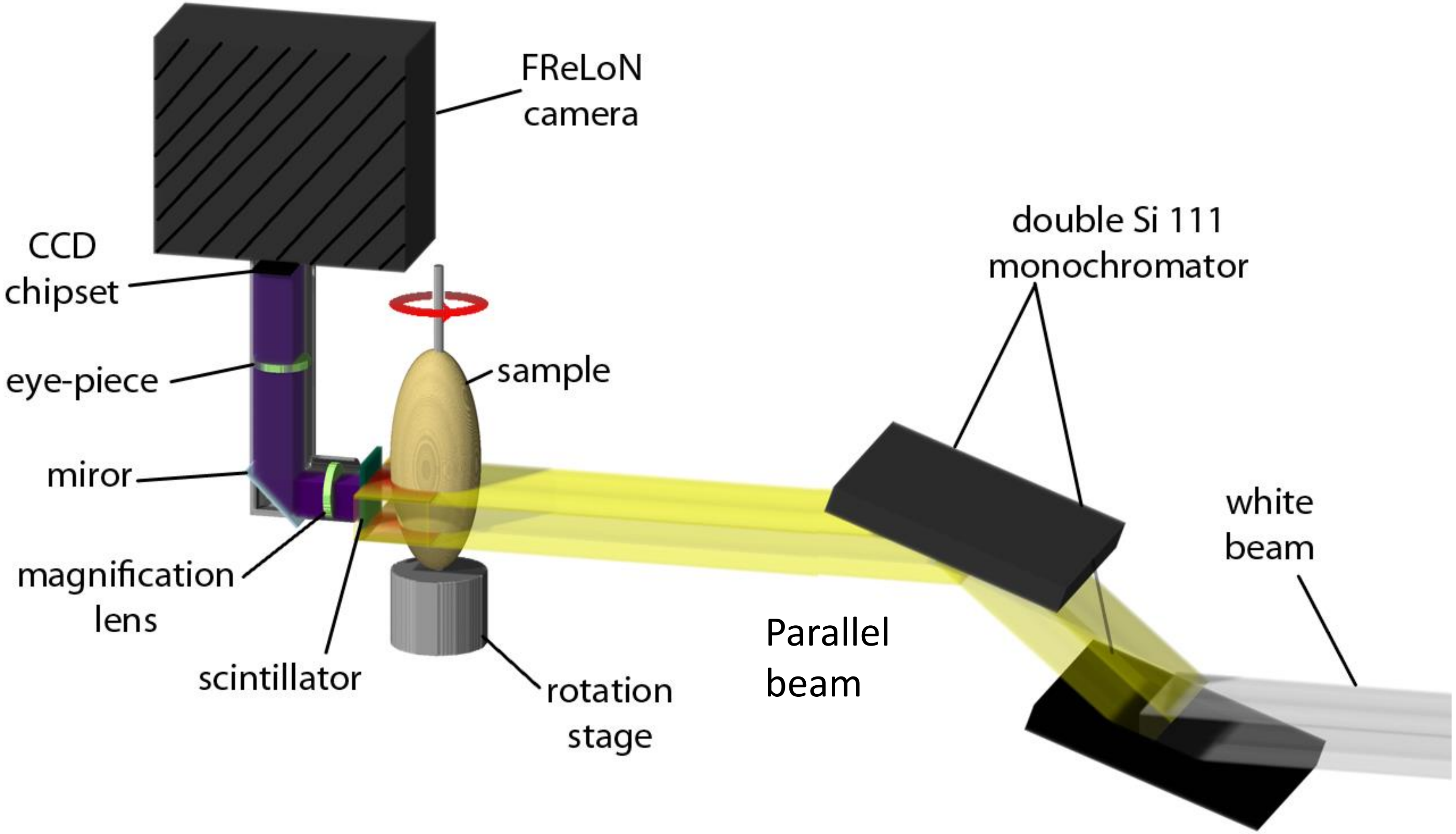
ID22 Micro-XRF/XRD/XAS combined to tomography

ID19
Micro-tomography dedicated beamline

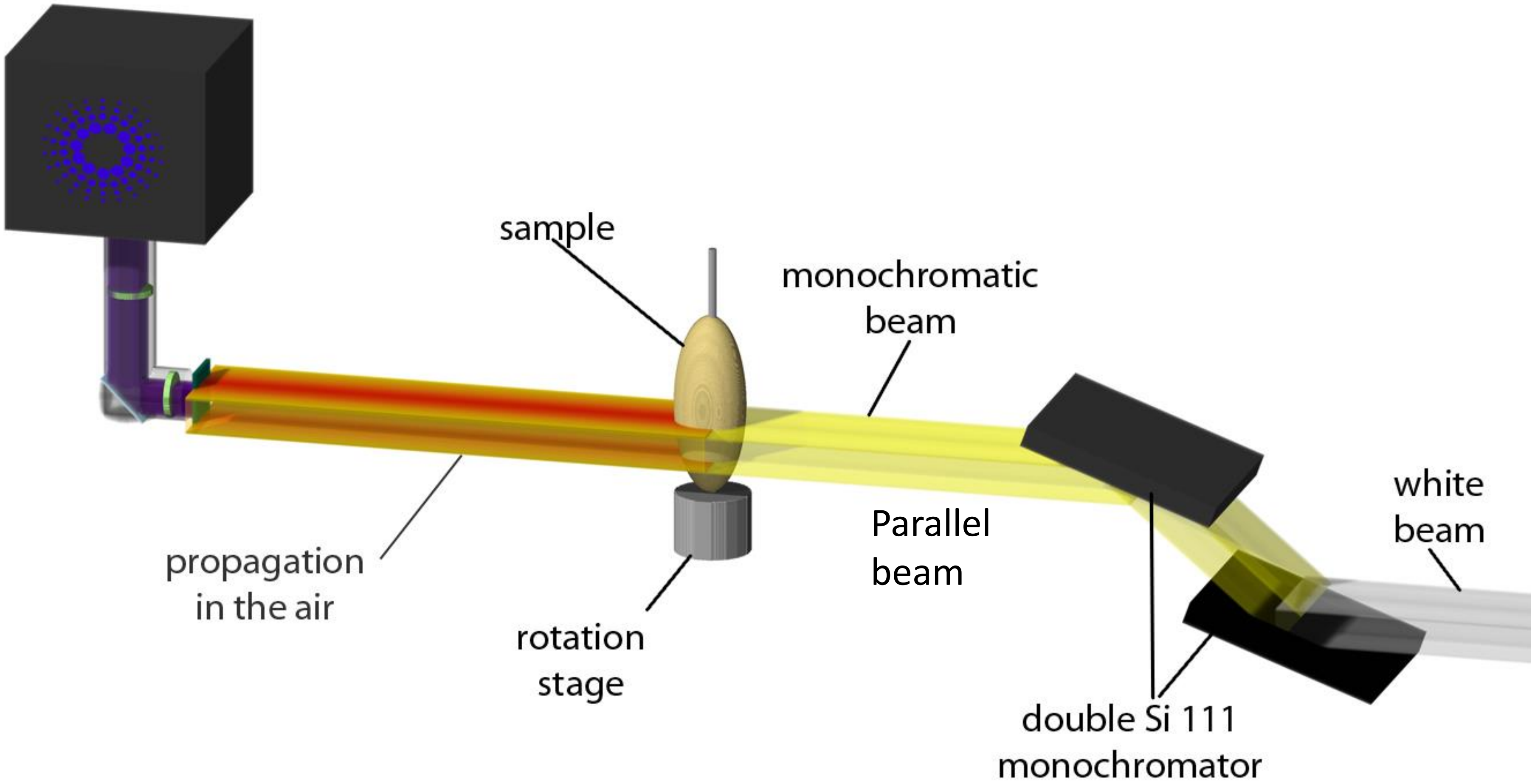
ID17
Large field tomography

ID15
High energy Tomography

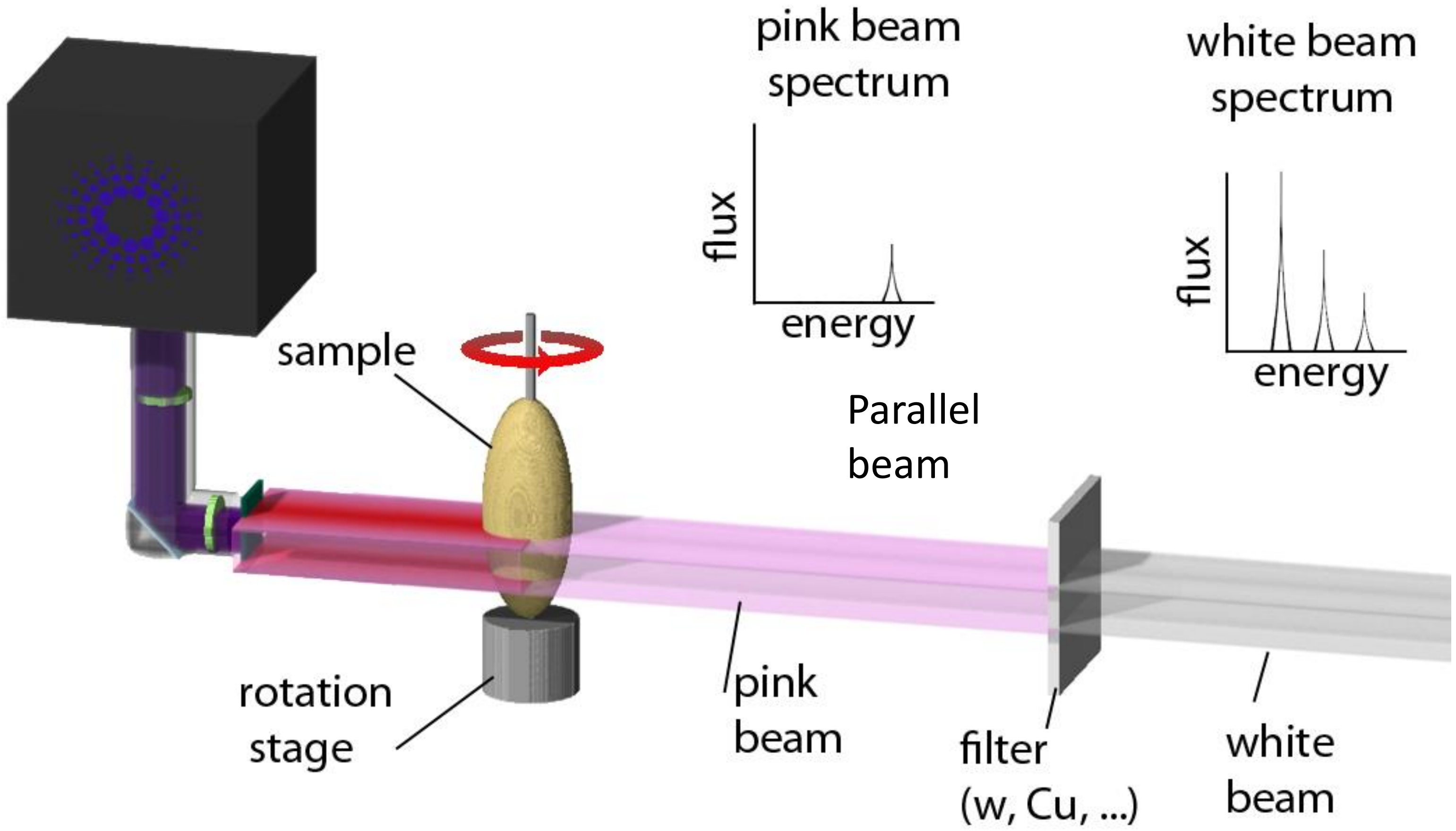
Tomography @ ESRF
absorption based CT



Tomography @ ESRF
propagation phase contrast CT



Tomography @ ESRF filtered white beam CT



Detectors

- FReLoN 2k camera
- (Fast Read out Low Noise)
- 2000 x 2000 CCD chip (4 quadrants)
- 14 bits
- Full Frame Mode (FFM): 5 images/sec
- Frame Transfer Mode (FTM): 20 images/sec
- Quantum efficiency:

33% in green spectrum (LuAg, LSO, YAG)

6.5% in red spectrum (europium)

Detectors

- PCO dimax
- 2000 x 2000 cmos chip (4 quadrants)
- 11 bits
- Full Frame Mode (FFM): 5 000 images/sec
- Frame Transfer Mode (FTM): 75 000 images/sec
- Quantum efficiency:

75% in the whole spectrum

Detectors

- PCO 4000
- 4000 x 4000 ccd chip (4 quadrants)
- 12 bits
- Full Frame Mode (FFM):
  2.8 images/sec
- Quantum efficiency:

70% in the whole spectrum

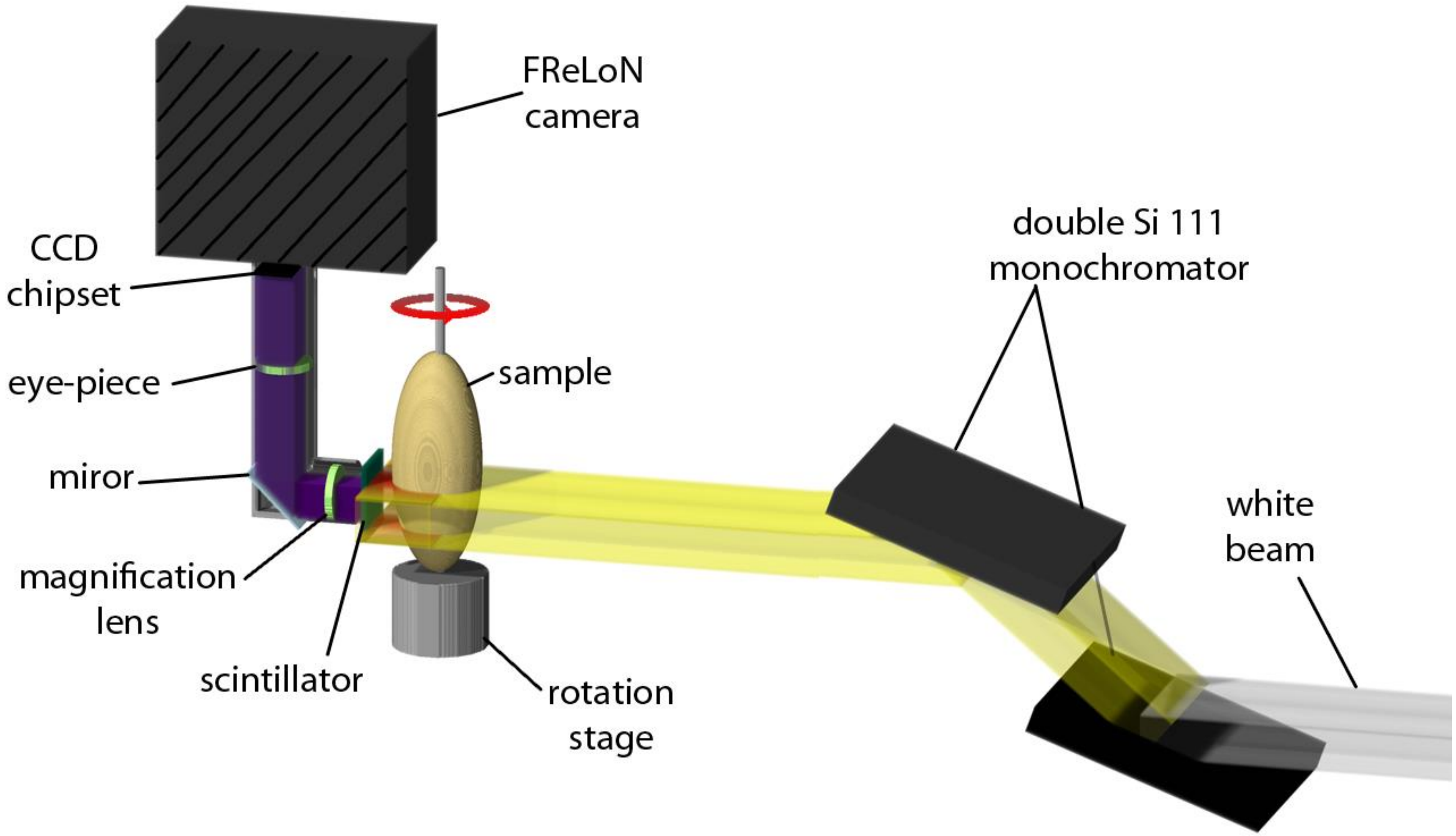
Acquisition time at maximum speed

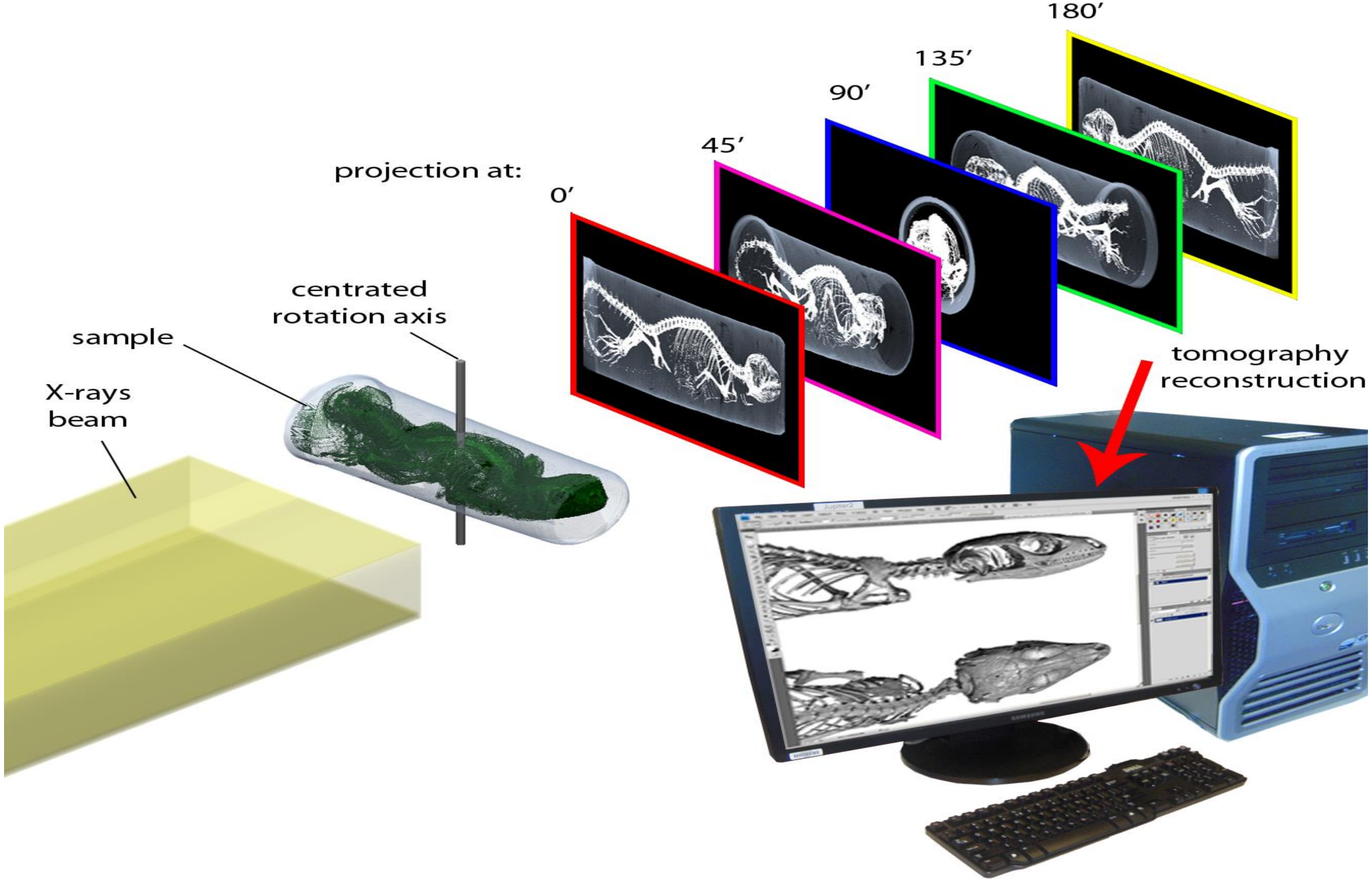
	FReLoN 2k (14 µm)	FReLoN e2v (12 µm)	PCO dimax (11µm)	PCO edge (6 µm)	PCO 4000
Mode	FTM	Rapid FFM	FFM	FFM	FTM
Image size	1024x2048	2048x2048	2016x2016	2160x2560	1024x2048
Exposure T	0.05 sec	0.34 sec	0.001 sec	0.05 sec	0.35 sec
1500 images	75 sec	510 sec (8.5 min)	1.5 sec	75 sec	535 sec (8.9 min)
Size of data	5.13 GB	7.42 GB	11.35 GB	13.13 GB	4.39 GB

Field of View (FOV)

- A set of lenses and eye-pieces to magnify or de-magnify the incoming beam to fit the chip of the detector:
- 50 µm optic (ID17-BM5): ~ 90 mm horizontal field of view
- 30 µm optic (ID17/19-BM5): 61 mm
- 13 µm : 26.62 mm
- 6 µm: 12.29 mm
- 3.5 µm: 7.17 mm
- 2.8 µm: 5.73 mm
- 1.4 µm: 2.87 mm
- 0.7 µm: 1.43 mm
- 0.16 µm: 0.32 mm

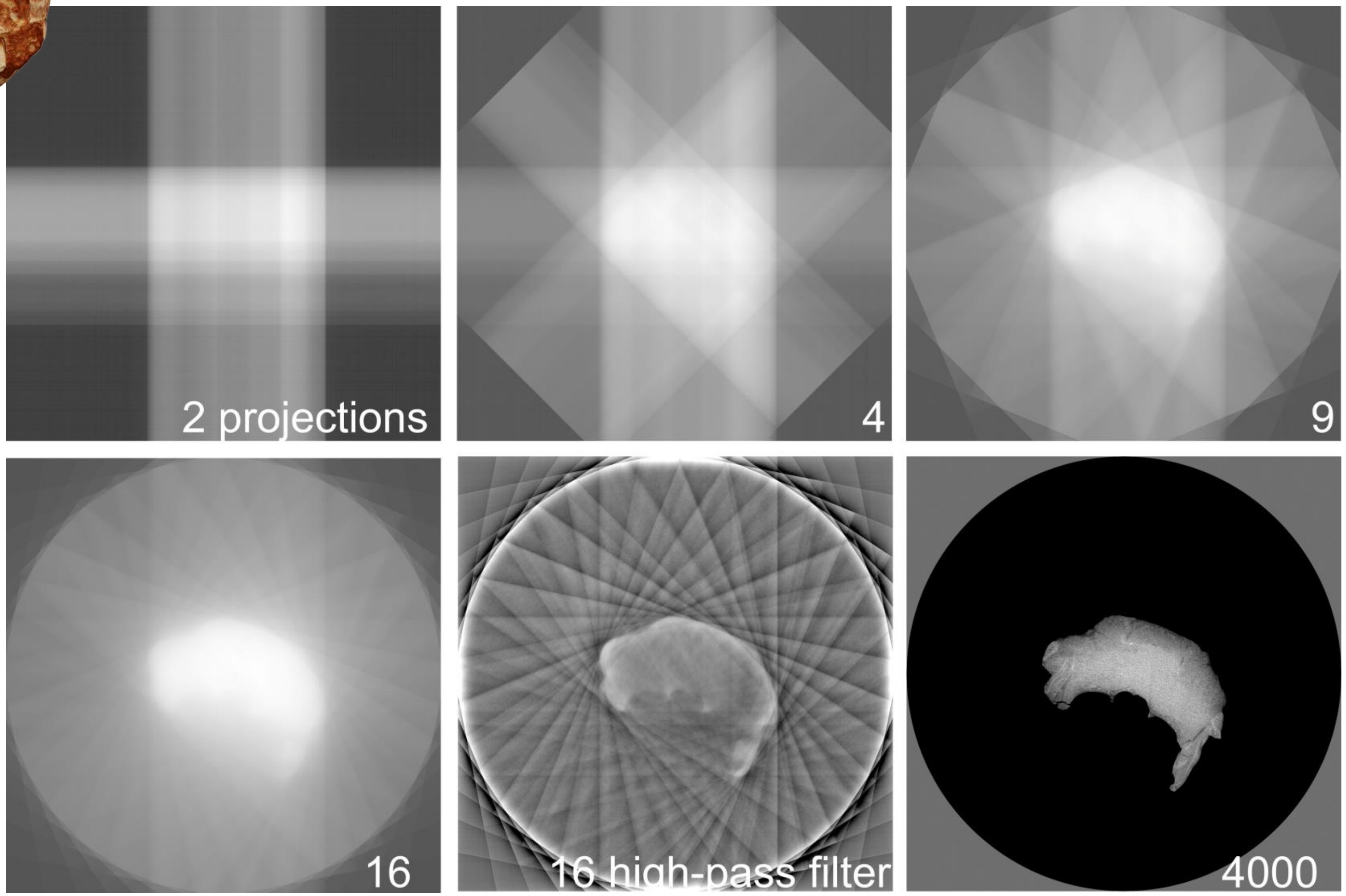
1. Example of absorption setup



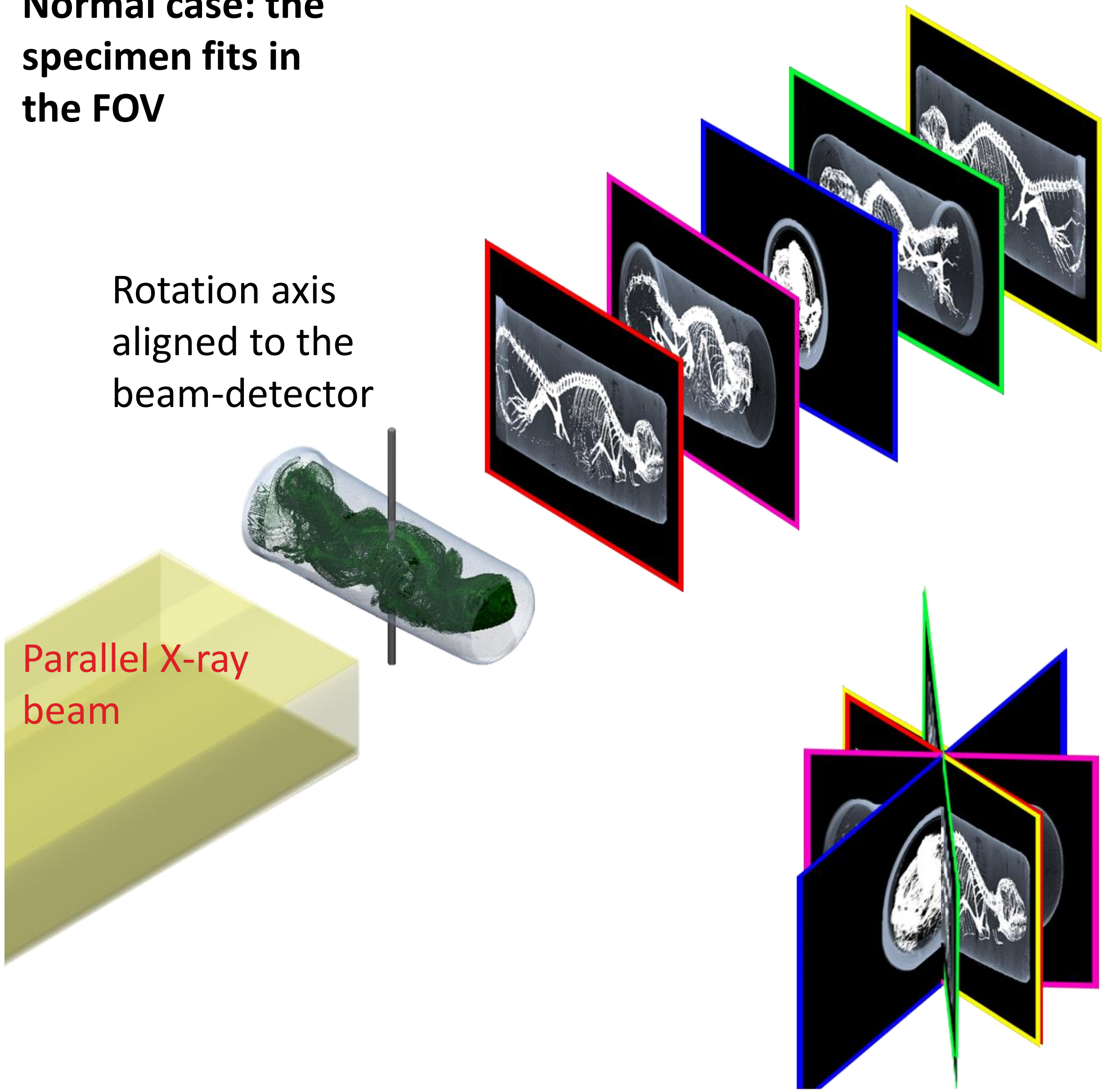

180’
135’
90’
45’
projection at:
0’
centrated
rotation axis
sample
X-rays
beam
tomography
reconstruction

Effect of the number of projections

The number of projection required equals the number of pixel defining the reconstruction circle: π * CCD width/2



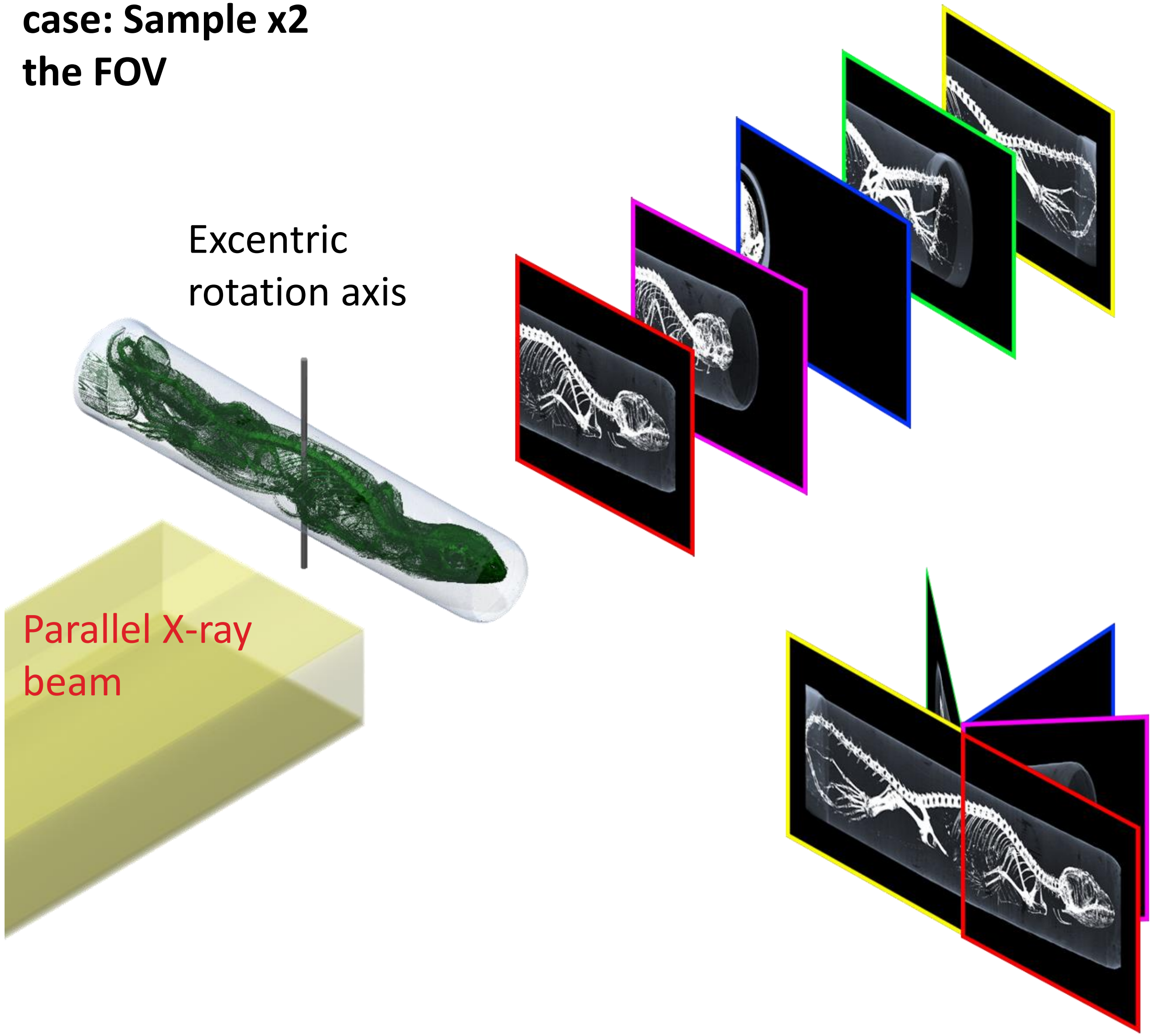
Normal case: the specimen fits in the FOV



A series a radiograph of the scanned object over 180 or 360 degrees

References images:
-A flat field image
-A dark current (noise) image

Half-acquisition case: Sample x2 the FOV



Comparison in term of data size

	Centered rotation axis (normal)	Half-acquisition
Field Of View (FOV)	2000 x 2000 px	2000 x 2000 px
Number of projections	1999 over 360°	5000 over 360°
Slide dimensions	2000 x 2000 px	3800 x 3800 px (e.g. for x1.8)
Weight of the scan (in 14 bits)	~14 GB	~37 GB
Weight of reconstruction (in 32 bits)	~32 GB	~108 GB
16 bits conversion	~16 GB	~54 GB

Problem of modern Synchrotron tomography

- Several Go/scan
- Generally 1-2 TB (1 TB = 1024 GB) of scan data generated
- Up to tens of TB in a few days
- Reconstruction process add several TB of data.
- Example, radiographs 1024 x 2048, HA= 3048 x 3048, 5 scans

	Scan data	32 bits EDF	16 Bits TIF	Center scans	Concatenation	Total	Unsharp 3D	Ring	Crop - 20%	Total
1 scan	20 GB	54 GB	26.9 GB	26.9 GB	26.9 GB		26.9 GB	26.9 GB	24.11 GB	
5 scans	200 GB	538 GB	135 GB	135 GB	67 GB	**~700 GB**	67 GB	67 GB	53.8 GB	**~900 GB**

Why and how high-throughput computing?

Basic problem: process user data “fast”

- bigger and faster detectors
- data rate increases
- time available for analysis gets shorter

Brute-force solution: buy more and faster computers

- Moore’s law: transistor density doubles every 18 months
- limits: mono-atomic layers
- higher clock speed $\Rightarrow$ exponential temperature rise
- may need more support staff, new software licenses, electrical power, cooling, rooms...

$\Rightarrow$ total cost of ownership can become unacceptable

Thus: brute force no longer sufficient

$\Rightarrow$ look at precise needs of “fast data processing”

Two scenarios are typically found:

– many (possibly small) sets of independent data

$\Rightarrow$ important is overall throughput,
not execution time for each job

– (possibly few) sequence-dependent data sets with much data and / or complex processing

$\Rightarrow$ important is elapsed time / job,
as job n must finish before $n + 1$ can start

Improve overall throughput:

– split task in several (many) independent jobs
– distribute jobs to several (many) different processors
– select “most appropriate” processor for each job
– scales well with number of processors
– no change to program code needed

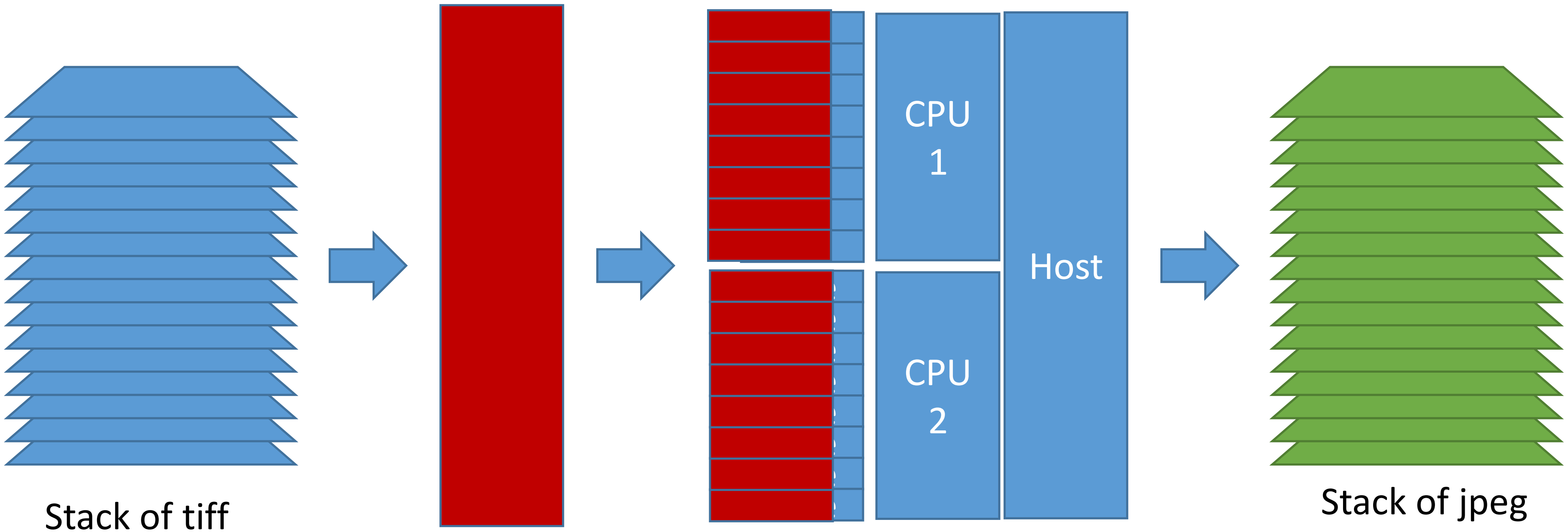
$\Rightarrow$ simple parallel processing

Reduce elapsed time for each job:

- distribute each job over several processors
- (re-)structure program into independent loops
- optimize data access for each processor
- needs change (possibly restructuring/rewrite) of program code

$\Rightarrow$ task for parallel programming

Example of simple parallel processing:
convert of stack of tif in jpeg



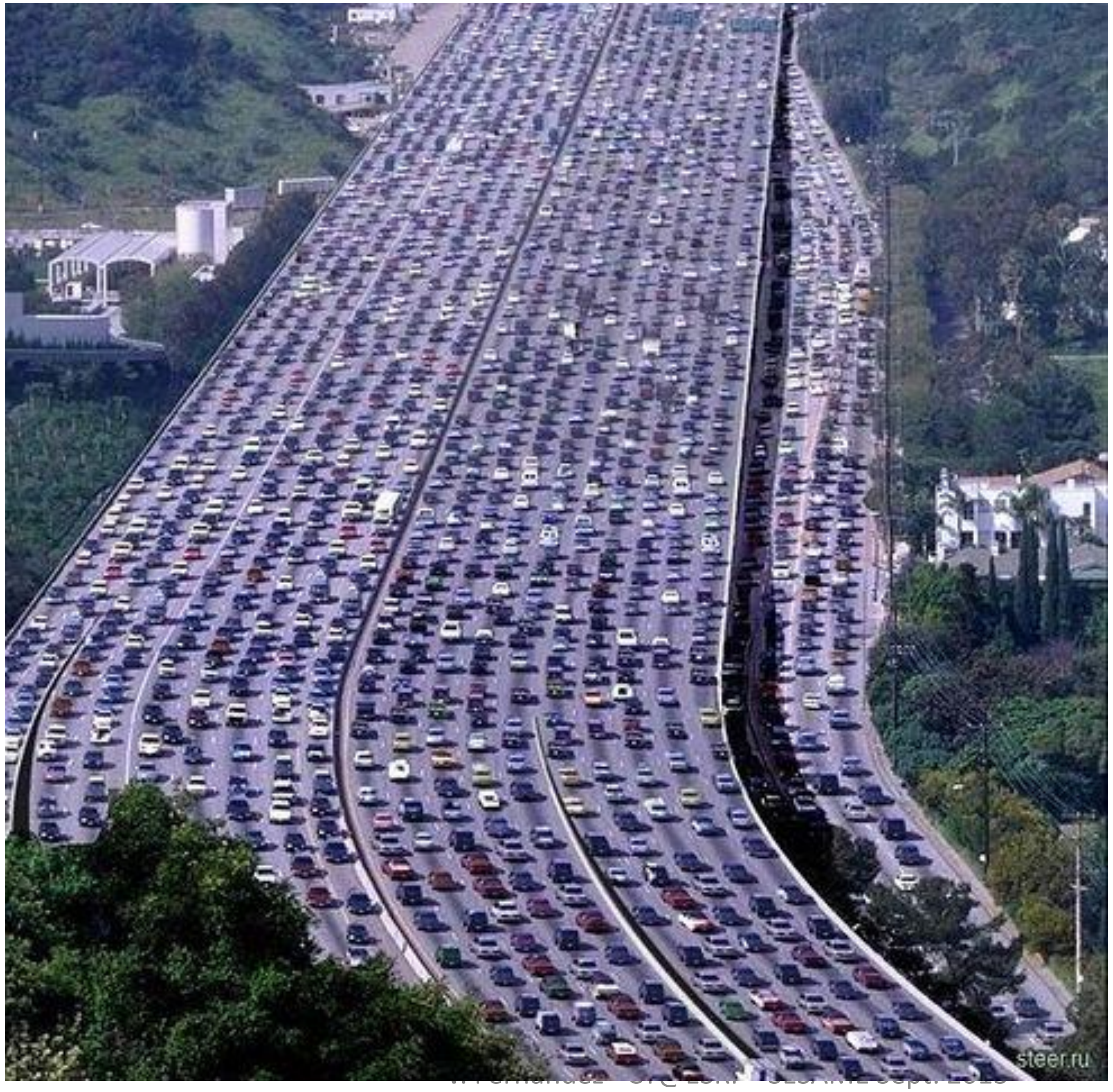
Stack of tiff

Single job

Parallelization of the job on several cores

Stack of jpeg

Even parallelization has its limits



Very expensive to allow for “worst case”

Better: use existing resources more efficiently

- select “most appropriate” processor for each job
- balance processing load between processors
- fill under-used periods (nights, weekends, ...)

⇒task for resource management system

Required features:

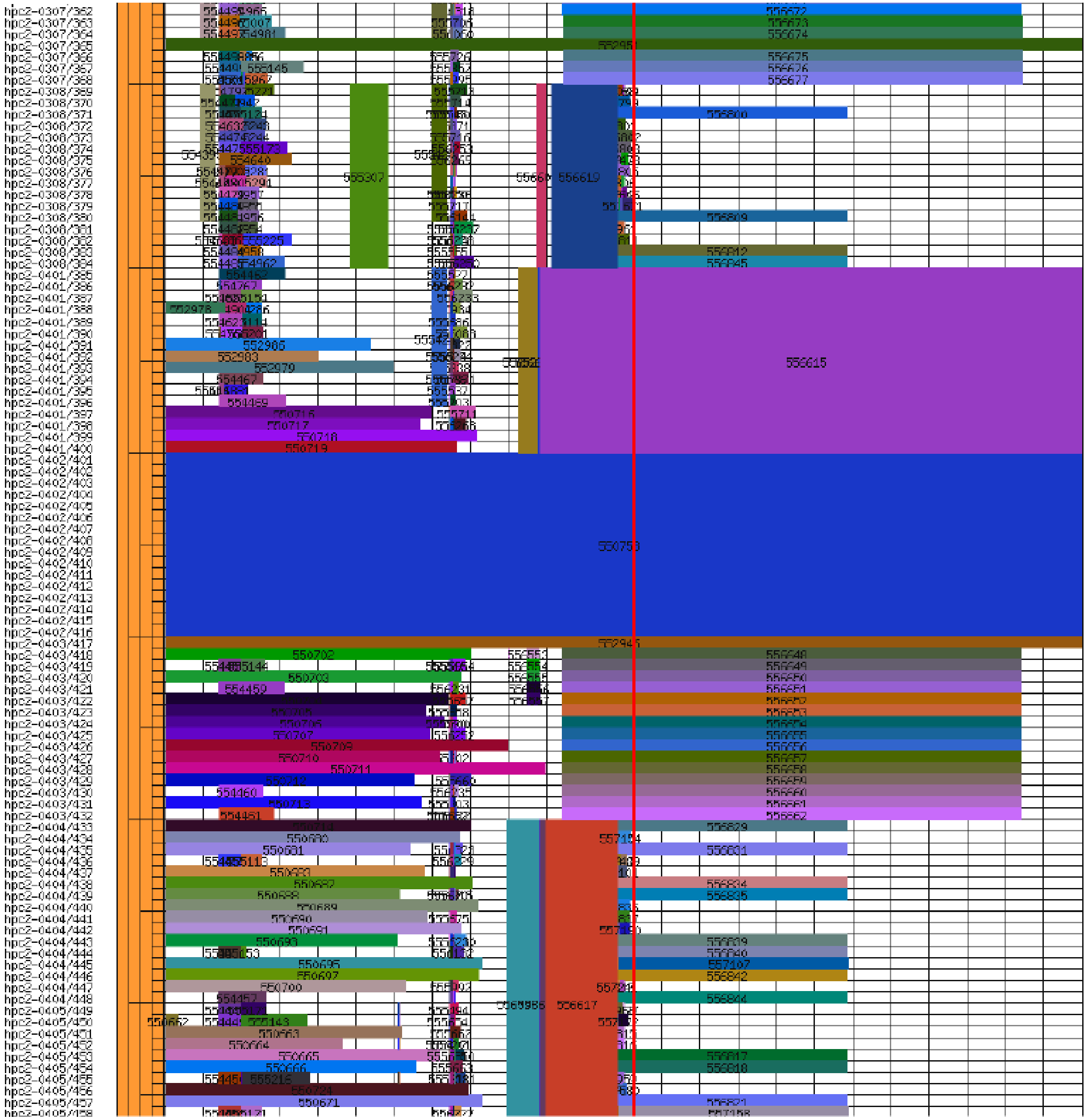
- resource distribution
- resource monitoring
- job queuing mechanism
- scheduling policy
- priority scheme

ESRF uses **OAR** (ENSIMAG Grenoble: oar.imag.fr)

Resource Manager OAR: Basic Use

- OAR features:
  - interactive or batch mode
  - controls processor and memory placement of jobs (cpuset)
  - request resources (cores, memory, time...)
  - specify properties (manufacturer, speed, network access...)
  - can define installation-specific properties and rules

- manage OAR with 3 basic commands:
  - request resources (submit job) (oarsub)
  - inquire status of requests (oarstat)
  - if necessary, cancel requests (oardel)
- after submitting jobs, user can log off and go home
- OAR starts jobs whenever suitable, delivers results



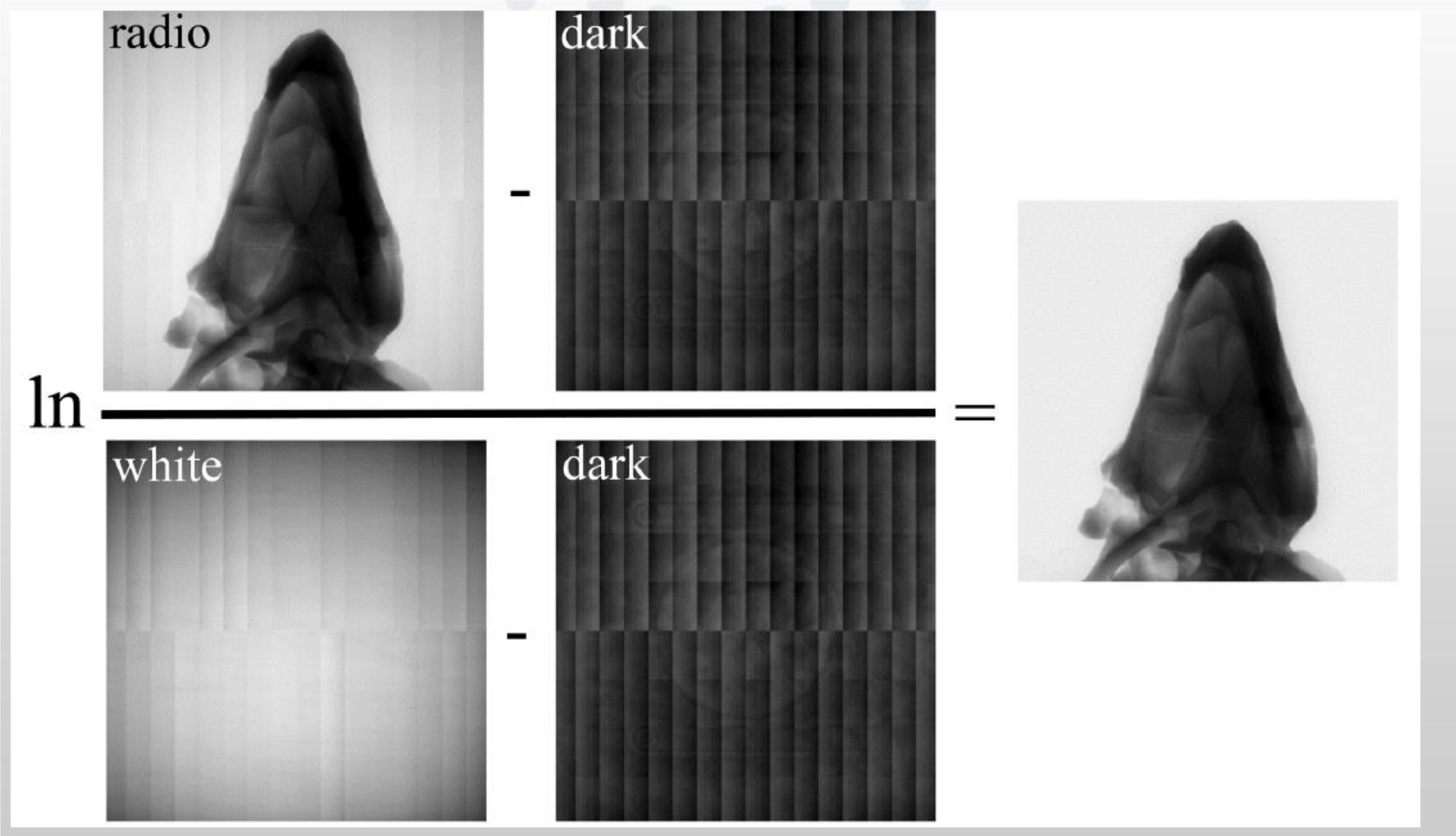
Reconstruction process

- PyHST2: a code specifically developed for the need of microtomography in 3rd generation synchrotrons (~10To data/experiment)

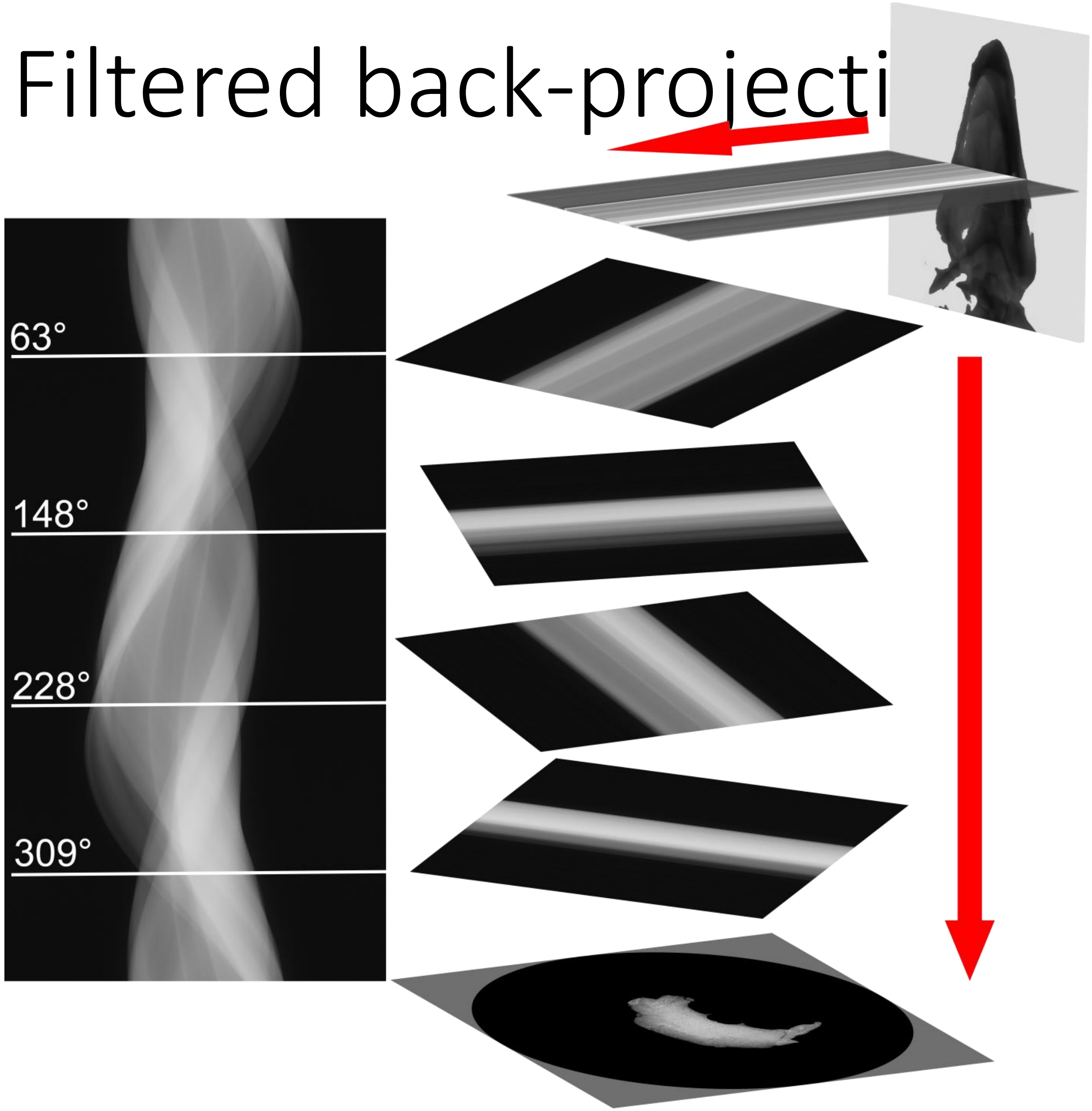
- Handle half-acquisition reconstruction

- Open access

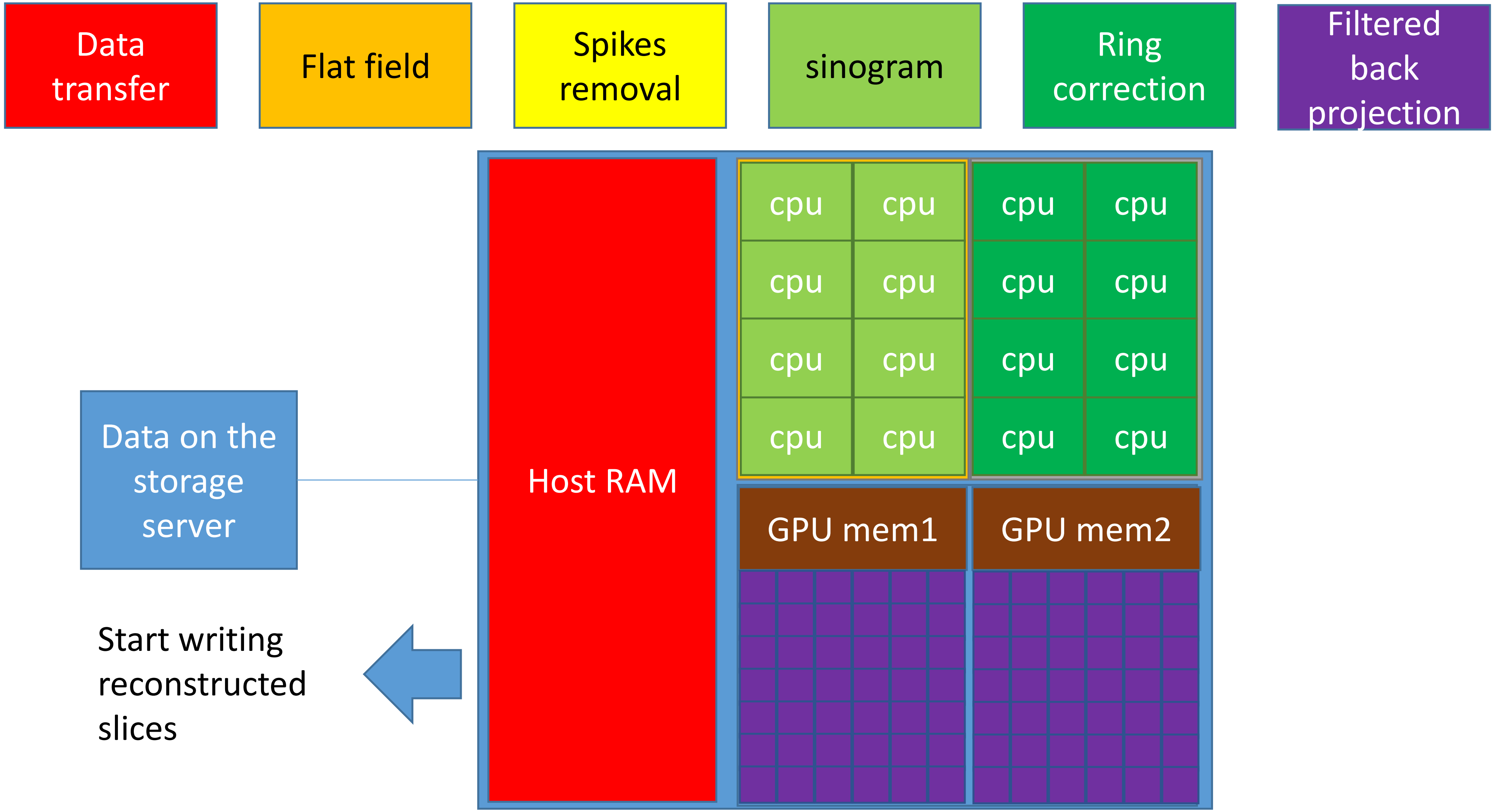
PyHST2: Flat field Correction



PyHST2: Filtered back-projecti



PyHST2: Processing pipeline



Post processing

Why a master_program?

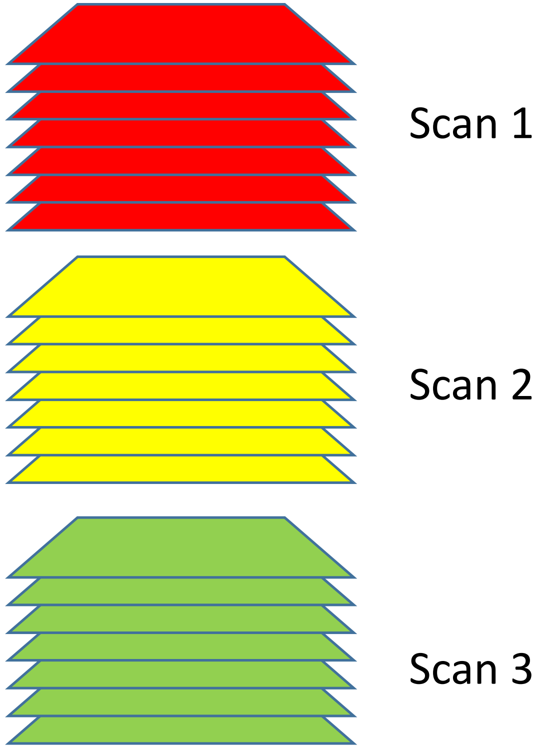
A scan series must present some overlapping.
From experience, we use this overlap to increase the signal to noise ratio
⇒Each slice of the object is scanned twice from 2 different parts of the detector.
⇒Space is not a problem for acquisition and does not change the final size of the volume.

Post processing – 2) align and crop scan

- As the center of rotation is calculated independently for each scans, there might be some variation in the reconstructed image, especially for half acquisition as the final reconstructed volume depends on the offset

Post processing – 2) align and crop scan

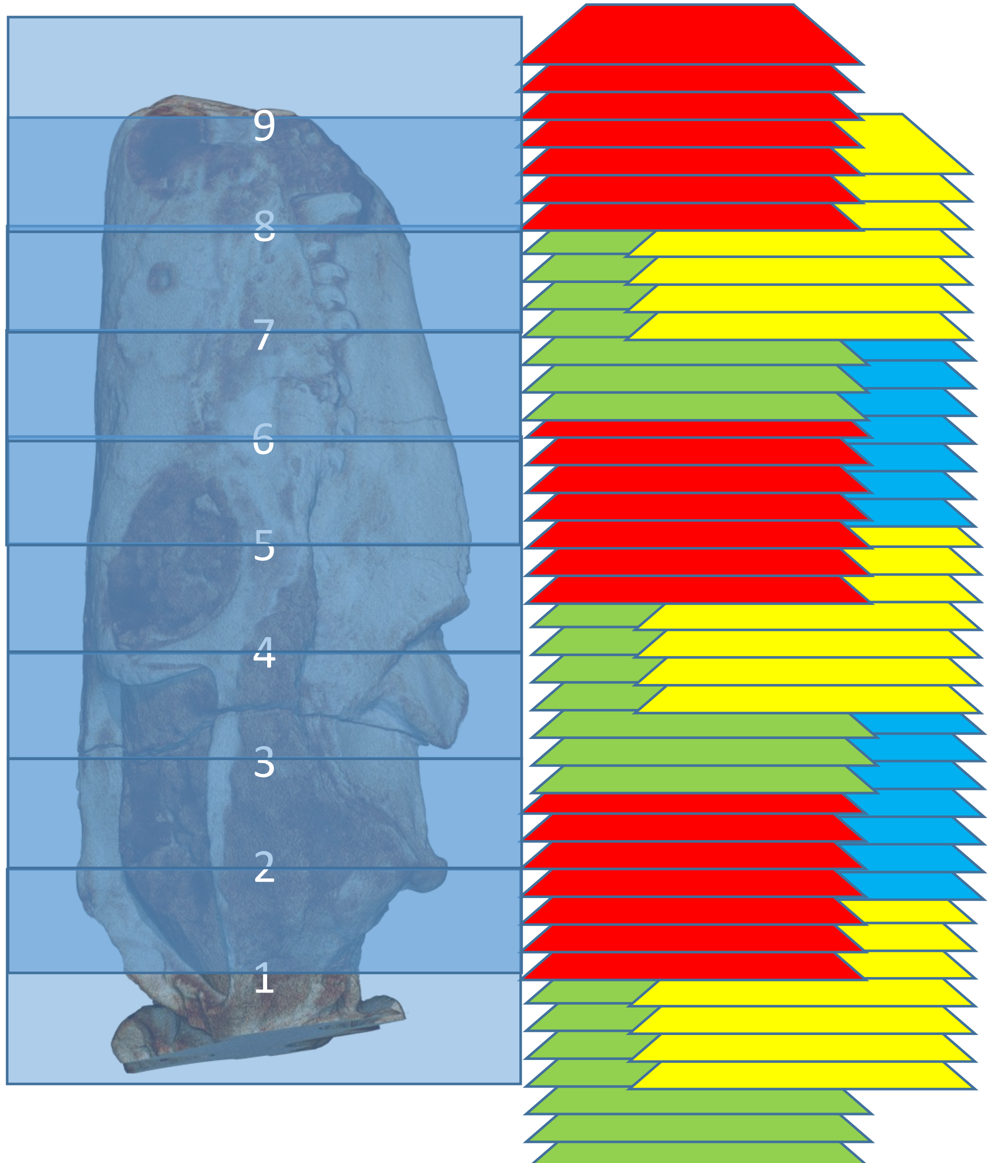
- All sub-scans are grouped together



- The program checks a few slices per scan. Comparison of dimensions of all slices
- If they are all the same, then nothing happens
- If they are different, it takes the smallest ones and crops all scans to this value

- Simple parallelization: 1 job per core
- Each job process 100 slices
- Limit to 50 jobs as a complete volume could results of several tens of scans => then it increase the number of slices per job

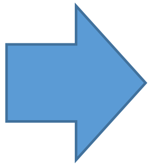
Post processing – 3) concatenate



Generating a single scan representing the object.
3 methods:

1) Remove redundant slices
2) Average redundant slices
3) Average taking into account the shape of the X-ray beam

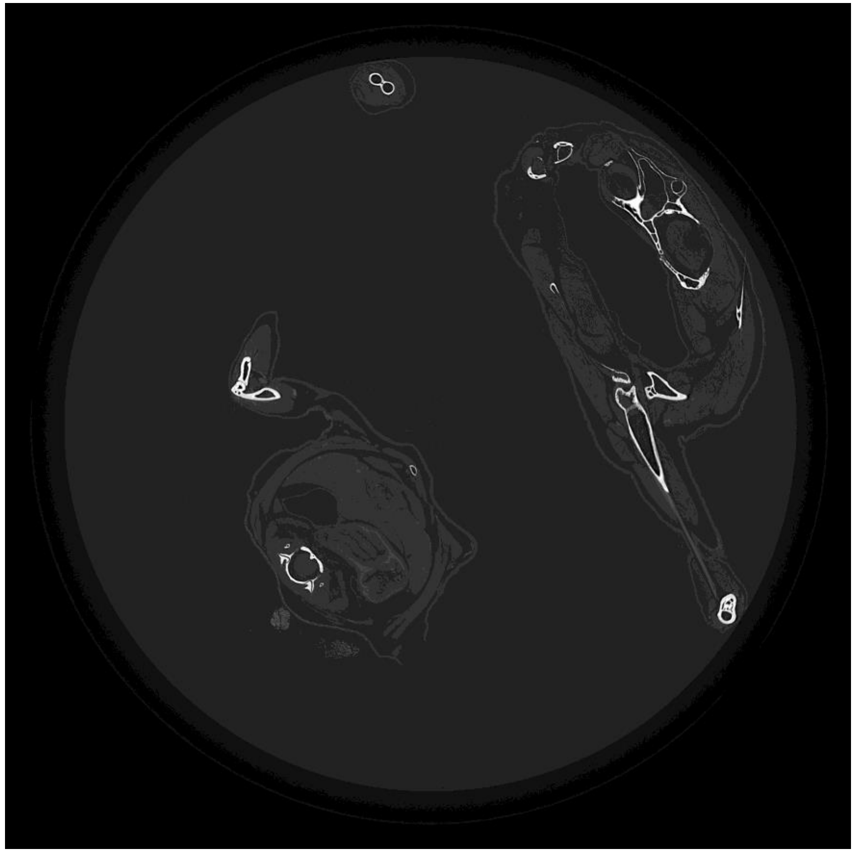
This program does not go through OAR except with interactive session. Parallelization through Matlab



After this step, the volume might be ready

Post processing – optional steps unsharp mask

Imaging technique to increase the acutance of the image
 - it de-blurs the image but can significantly increase the noise



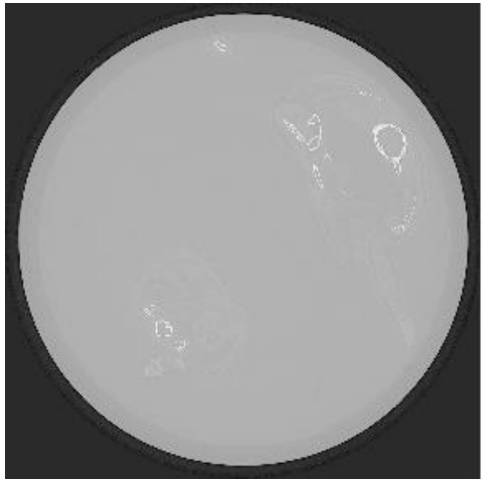
Post processing – optional steps ring correction

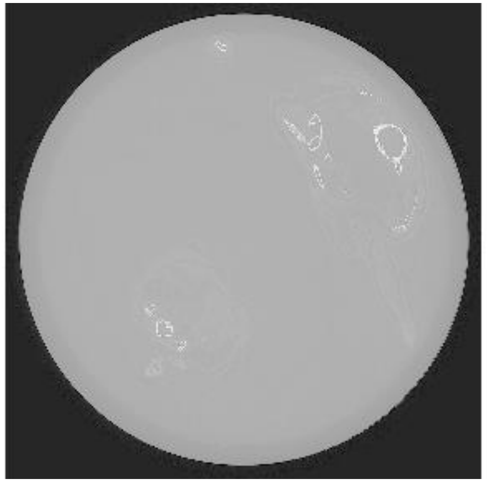
Not all rings are not corrected via flat field

The slice is unrolled

1) Median filter to remove low frequencies
2) Floating mask to identify what is a ring and what is a structure

=> Creates a mask on the ring that is removed from the image





Post processing – optional steps ring correction

1) unroll in polar coordinates

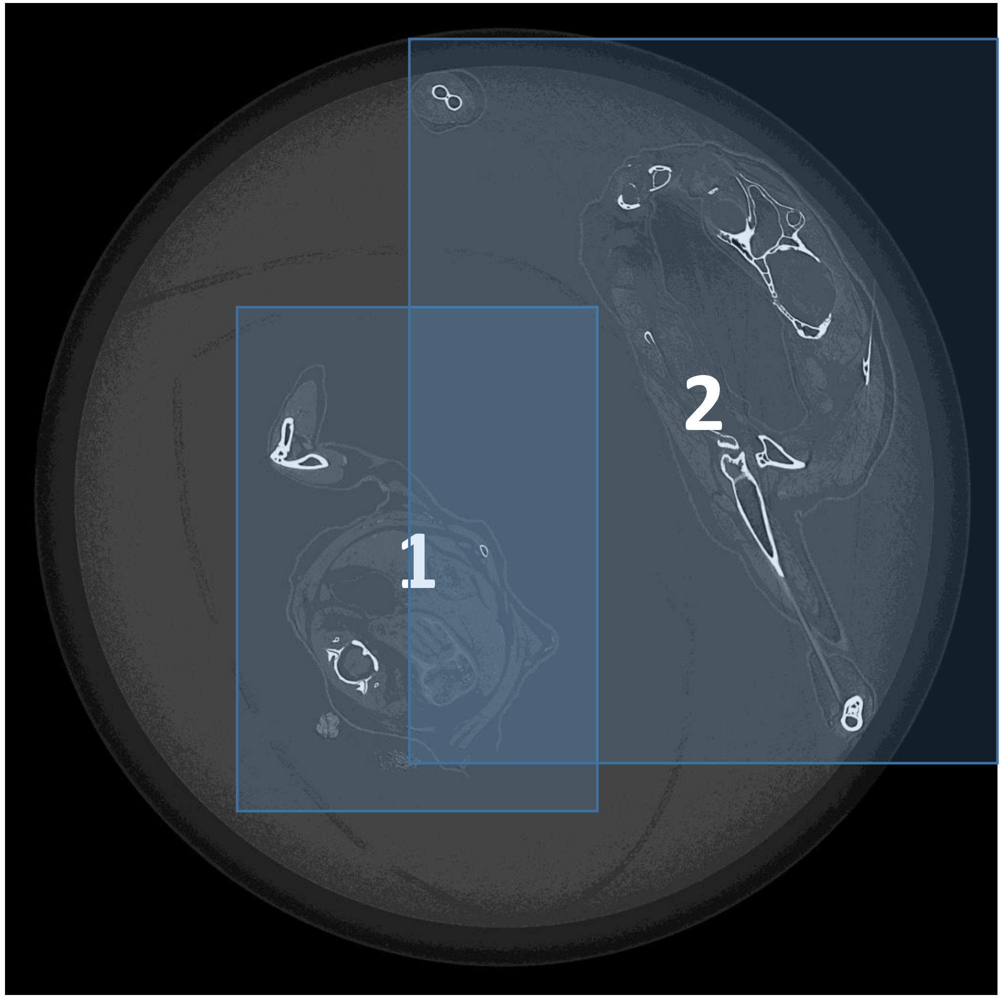
2) Directional median filter

3) Directional blurring with a floating mask to isolate high frequencies

4) From the high and low frequencies the final mask is applied to the original image to remove the rings

Post processing – even more optional steps
3D crop of the volume

- The program reads 20 % of the slices
- These slices are reduced in size to accelerate the calculation
- Maximum projection on each direction to determine a bounding box and crop the region of interest



<u>Example:</u>

1) Separate two specimens scanned at the same time

<u>Problem</u>, if the specimens are not well positioned from the beginning, they cannot be separated

Post processing – even more optional steps
3D binning

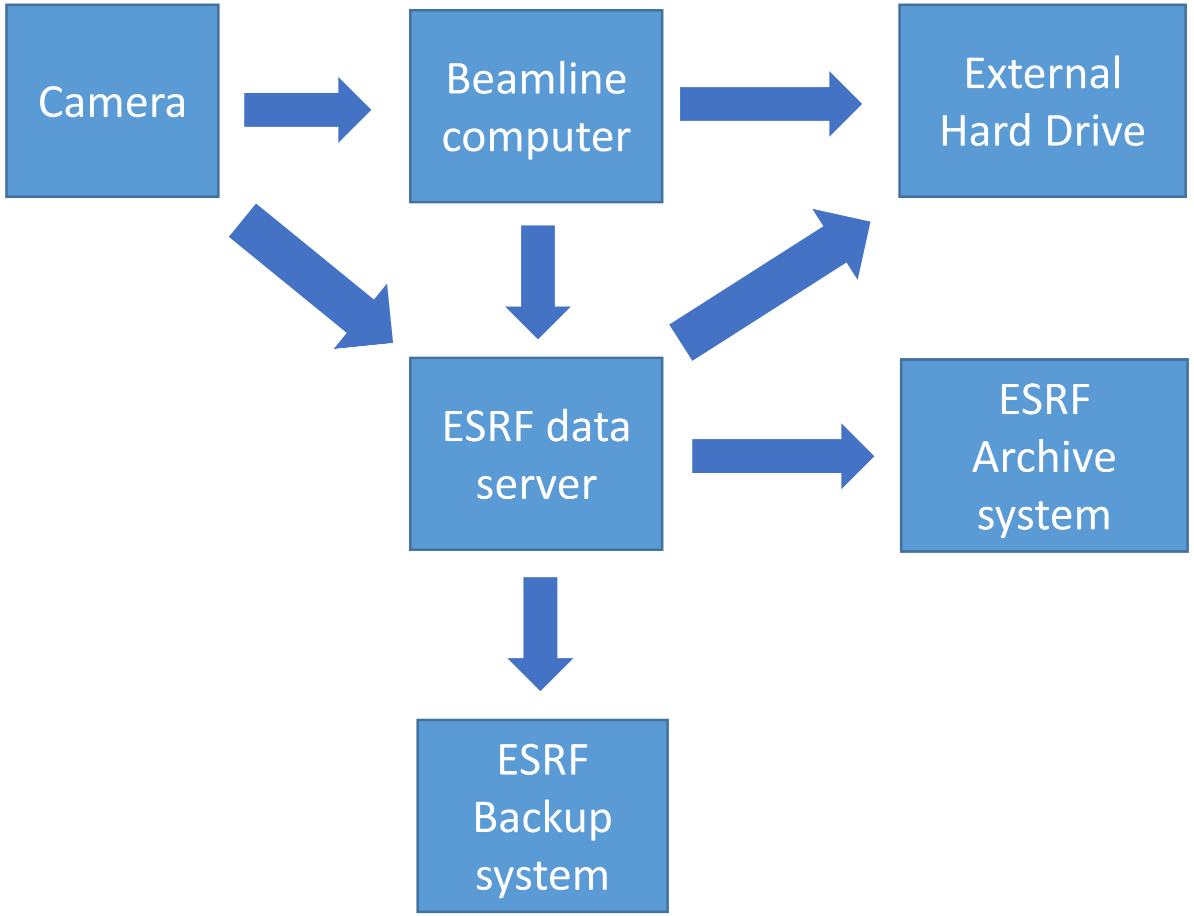
•Generating a binned volume at the end increase the overall dynamics
•It also allows the user to perform a pre-analysis on a relatively small volume and transfer it to the full resolution data set for final rendering.

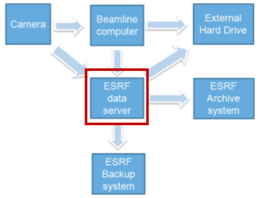
•Example on a series of 10 scans of 1000 slices 3800x3800:
- final stack of TIF: ~135 GB
- Binning 2x2: ~16.8 GB
- Binning 4x4: ~2.1 GB

Image analysis

5 Servers on Windows server system
5 Wacom touch screens for manual segmentation
4 licences for Vgstudio Max 2.2

Where do the data go?





ESRF data server

Data server - currently

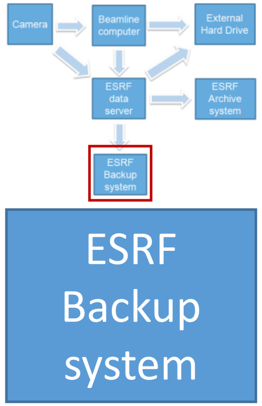
NAS server: Array of 2 Tb – **240 disks**

500 TB

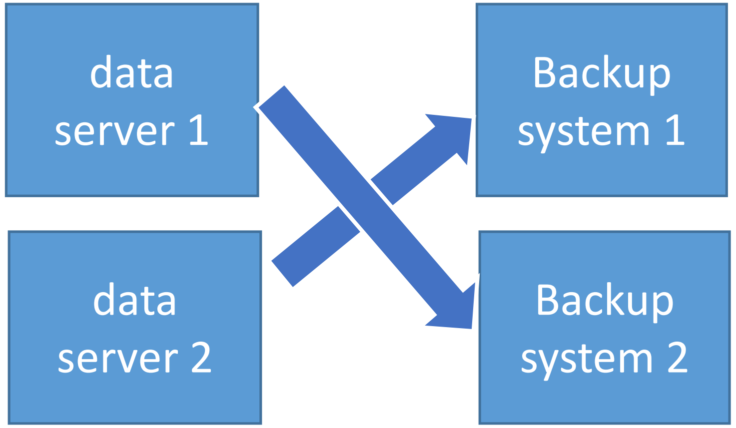
80 % used for X-ray imaging purpposes

50 % of which is used by ID19

Currently:
About **1.8 PB** (1 800 TB) available on the data server
-**1.3 PB** for local usage
-**0.5 PB** for visitors



Backup system

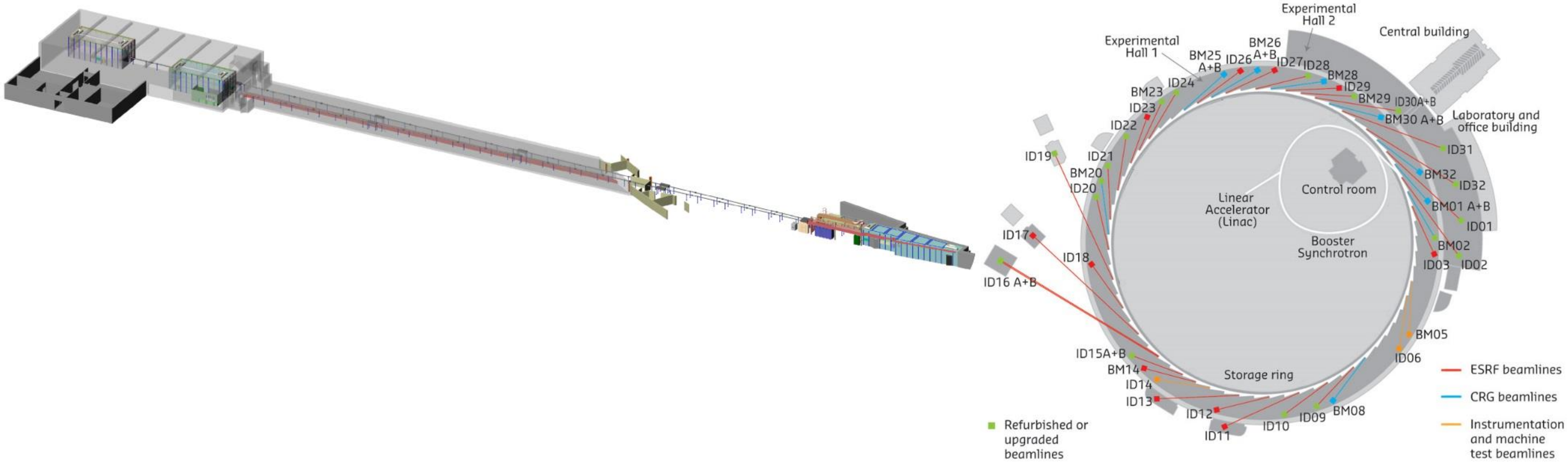


Backup: 150 GB/ hour:
-Daily backup checking modification and new files.
-Full backup from beginning every 3 months – takes 1 week. The limit to 1 week influence the number or writing slots and size of file system.
-Data are present 3 or 4 times to avoid problems.

Backup cost as much as storage and necessitates a full-time job to maintain it

Magnified holotomography at the nano-imaging beamline

ID16A – NI



Nano Imaging beamline

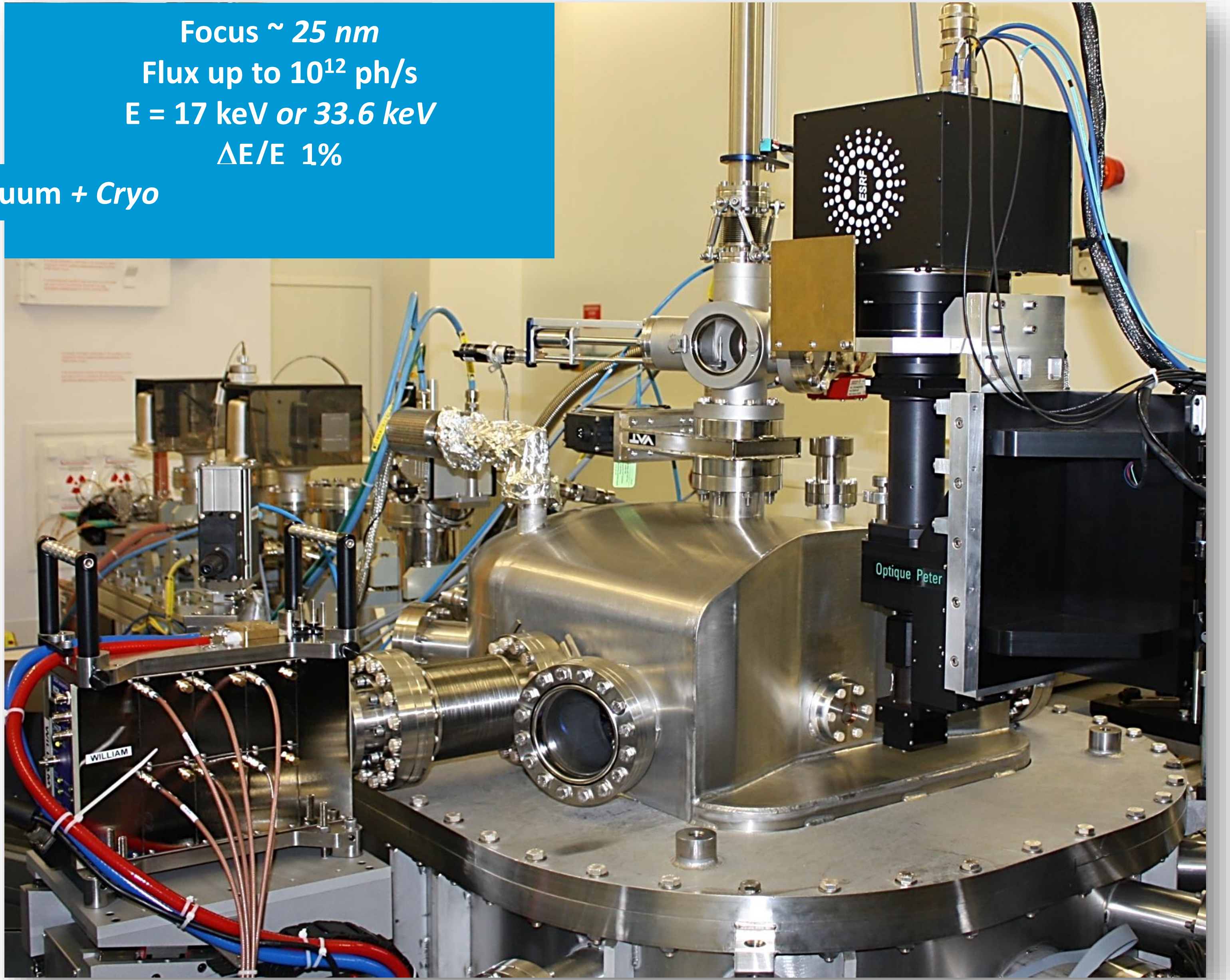
In-vacuum + *Cryo*

Focus ~ *25 nm*

Flux up to 10^{12} ph/s

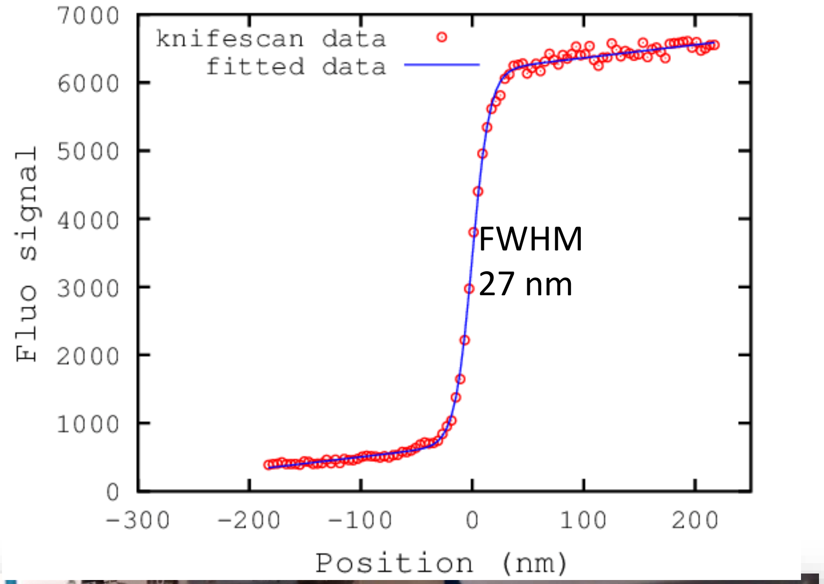
E = 17 keV *or 33.6 keV*

$\Delta E/E$ 1%

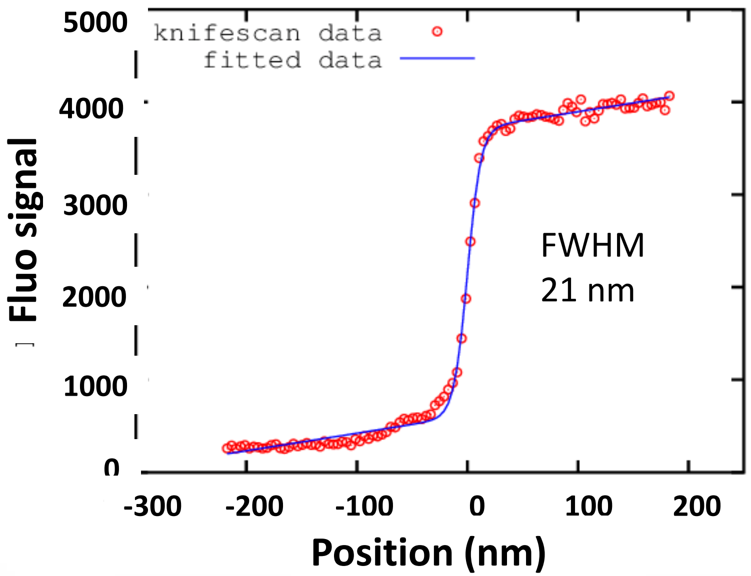


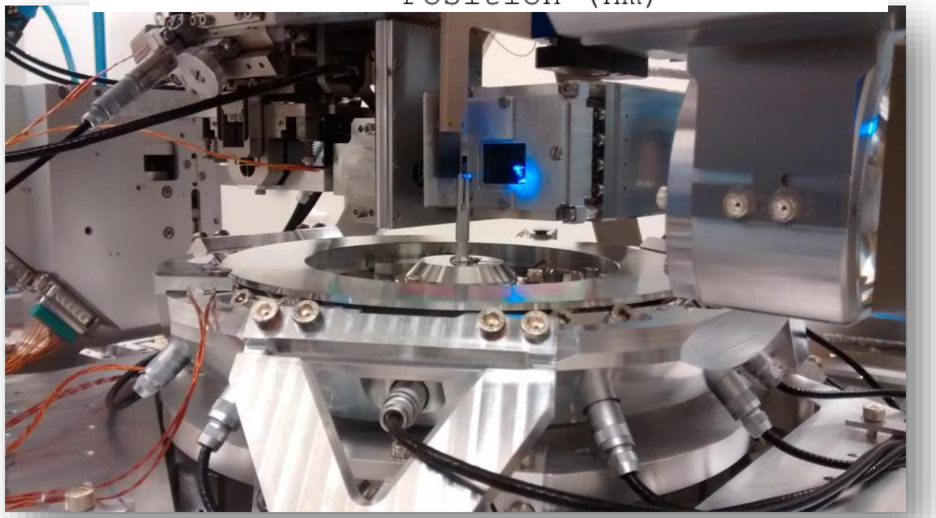
hard X-ray nano-focused beam

Horizontal

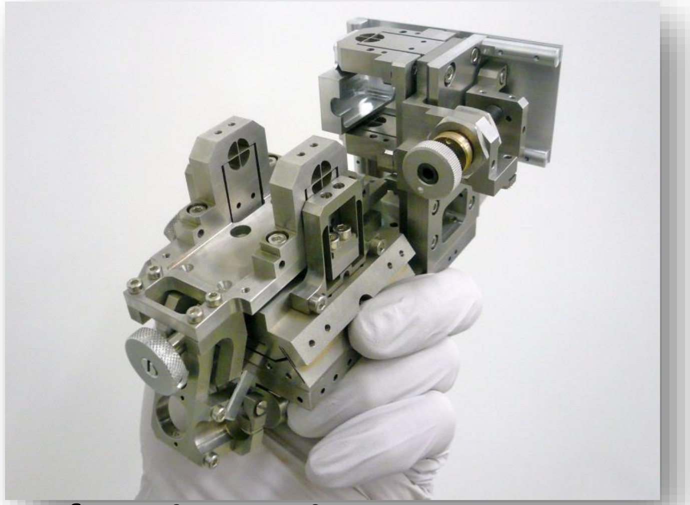


Vertical



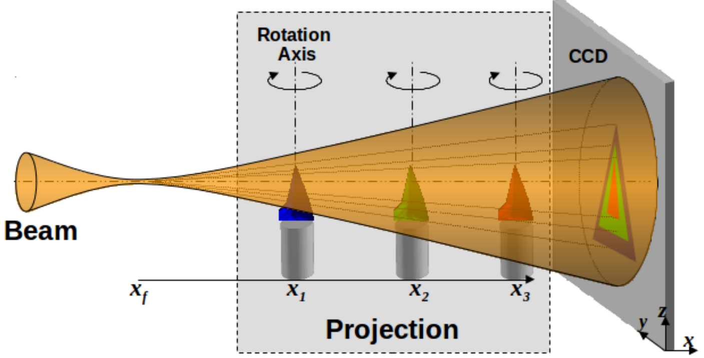


Nano-positioning stage



X-ray focusing optics

Magnified HoloTomography



Record 4 tomographic sets of in-line holograms	
4xN	angles

↓

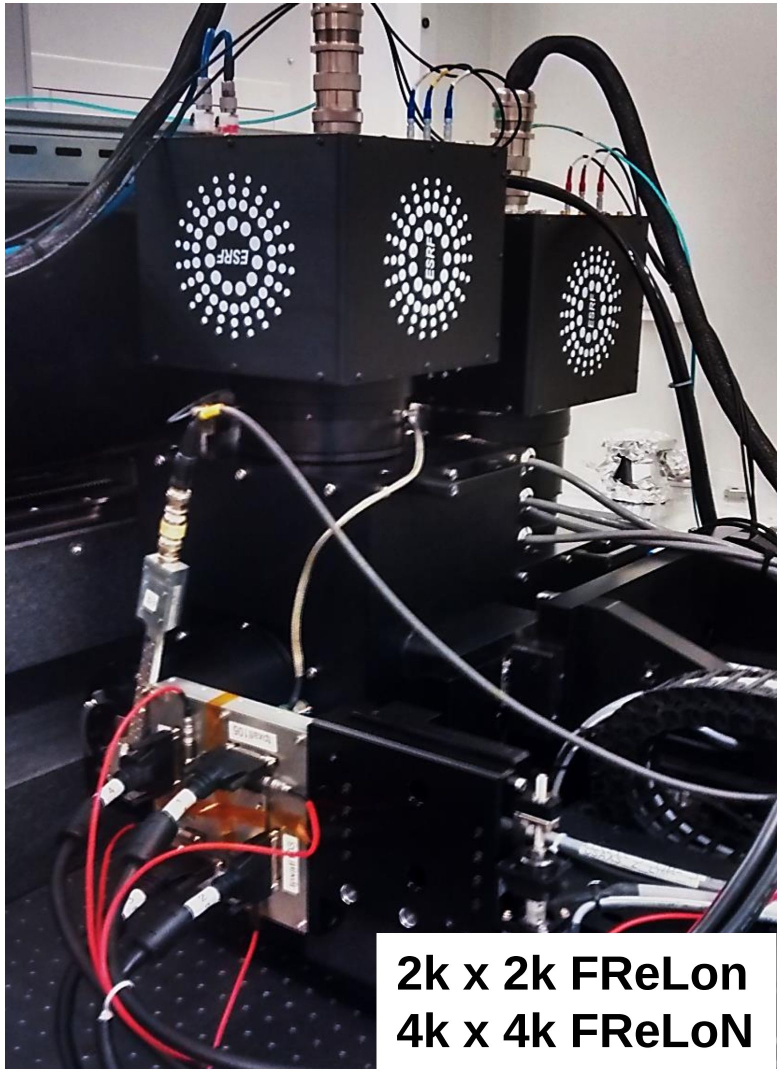
Phase retrieval	
N	angles

↓

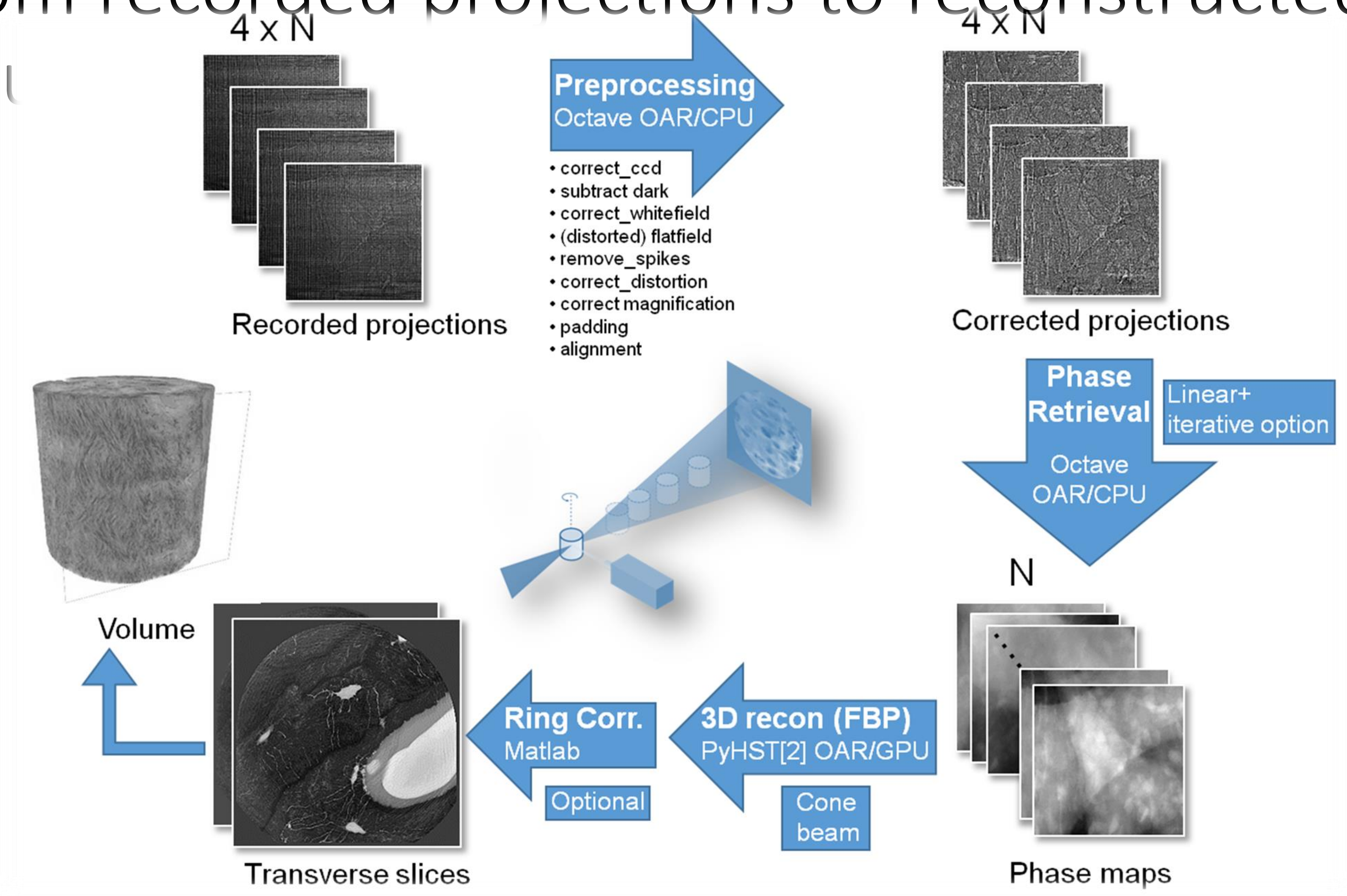
Tomographic reconstruction	
32 GB	volume

Two detectors: 2k x 2k and 4k x 4k

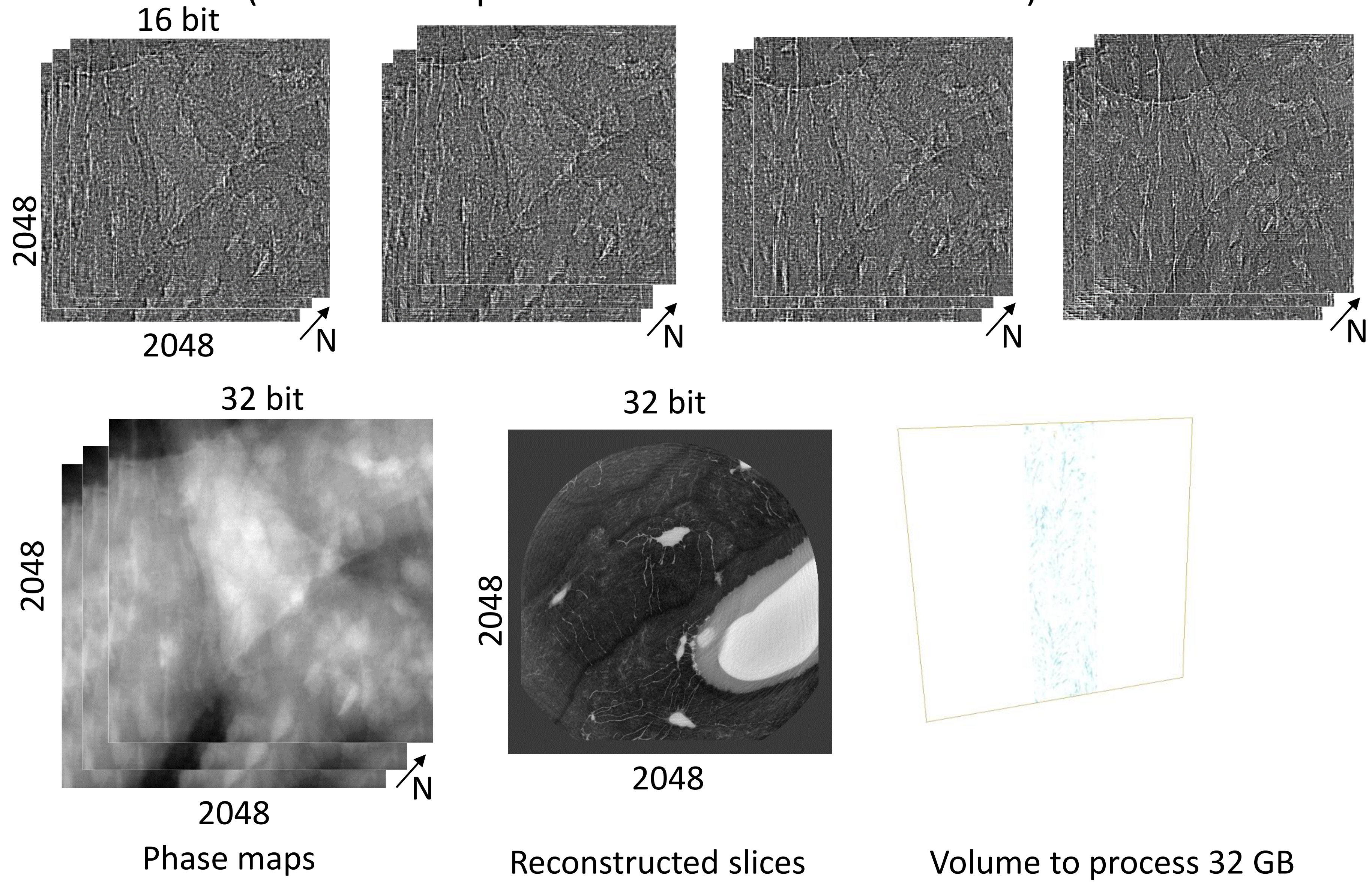
Aim to use both (record 2 distances simultaneously)



From recorded projections to reconstructed volu

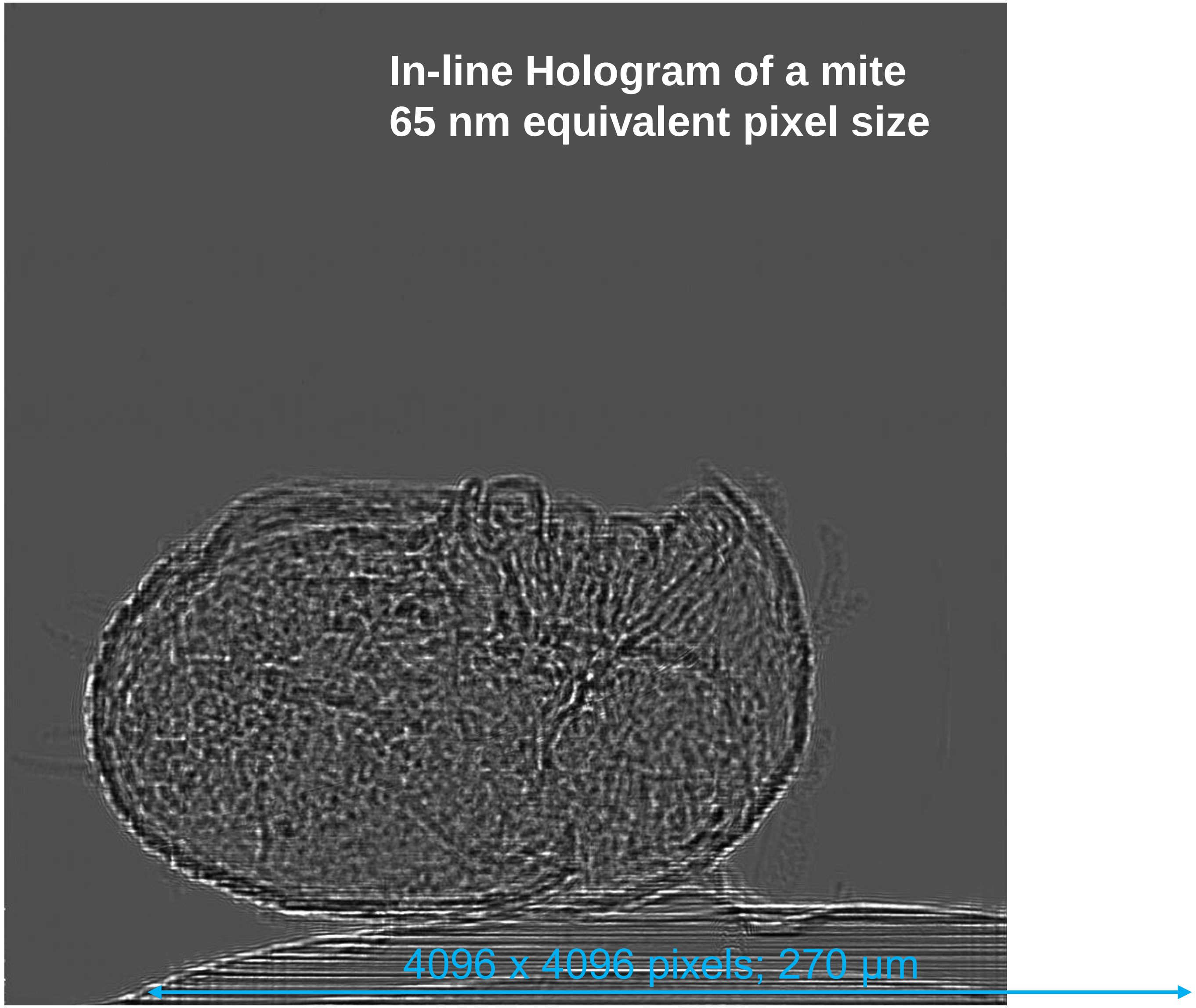


One data set (raw data plus reconstructed data) ~ 100 GB



15 GB x 4 + 30 GB + 32 GB

4k x 4k detector ~ 600 GB Per data set



Raw holograms
~ 60 GB x 4

Phase maps
~ 120 GB

Reconstructed
volume
~ 256 GB

Not ready yet

cluster	Aug-Oct Exectime*cores [%]	Aug-Oct Nr of Jobs [%]	Oct-Dec Exectime*cores [%]	Oct-Dec Nr of Jobs [%]
ASD – 40 nodes (768 cores)	17.57	49.07	16.08	17.60
ID16A – 2 nodes (40 cores)	0.29	0.07	1.25	0.19
TOMO – 5 nodes (46 cores)	1.92	0.64	1.85	0.89
NICE - 75 nodes (1208 cores)	80.16	41.23	80.77	66.73

Statistics on queue usage changed during the two periods, Aug-Oct and Oct-Dec: the behavior of the users is changing, depending on beam mode, scheduled experiments, and other projects.

OAR scheduled a lot more besteffort jobs in Oct-Dec than in Aug-Oct.

Statistics about user groups

In the next tables the percentage computation time and number of submitted jobs per user group are listed:

UNIX group	Aug-Oct Exectime *cores [%]	Aug-Oct Nr of Jobs [%]	Aug-Oct Nr of cores per job	Oct-Dec Exectime *cores [%]	Oct-Dec Nr of Jobs [%]	Oct-Dec Nr of cores per job
id11	3.56	1.07	3	2.2	0.91	3
Id16a	2.81	0.98	1	10.70	7.42	1
id16b	2.32	5.23	1	5.41	2.21	1
Id17	23.75	3.02	15	9.35	1.45	7
Id19	12.97	32.68	1	19.64	57.75	1
Inel	12.70	0.02	26	0.24	0.01	13
Jsbg	2.99	1.14	1	3.32	2.43	1
Machine	12.51	45.91	1	9.05	10.07	1
Theory	7.26	0.07	17	22.13	1.11	3
Users	8.17	0.51	15	0.95	0.09	6

The top 10 users of OAR are different in the time periods Aug-Oct and Oct-Dec.
Depending on experiments, projects or time, the OAR cluster is used by different user groups at

Dynamic XRD Computed Micro Tomography

- Metals and metal oxides anchored to porous support materials are widely used as heterogeneous catalysts in a number of important industrial chemical processes.

- Efficiency of a catalytic reactor depends on the behavior and efficiency of the millimeter-sized pelletized catalyst body

- Crucial to the catalyst design is an understanding of the factors which influence the distribution and nature of the active phase during preparation.

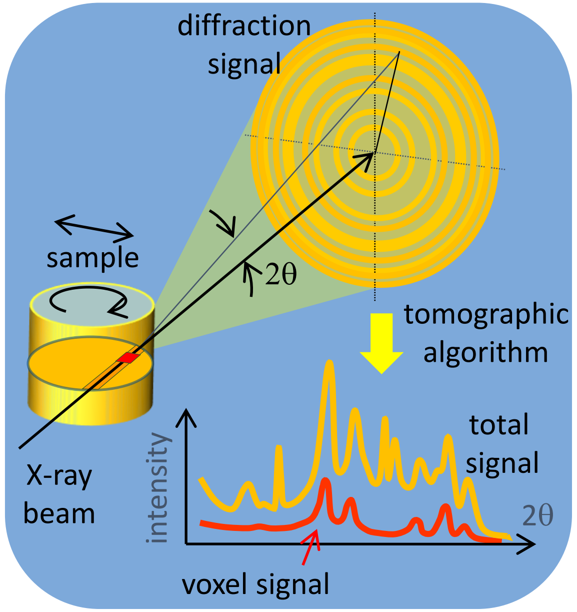
- The traditional in-situ X-ray diffraction measurement is made at one position in the sample which is often not representative of the whole

- Desirable to mimic the real conditions as closely as possible - i.e. real samples (not just powder models)

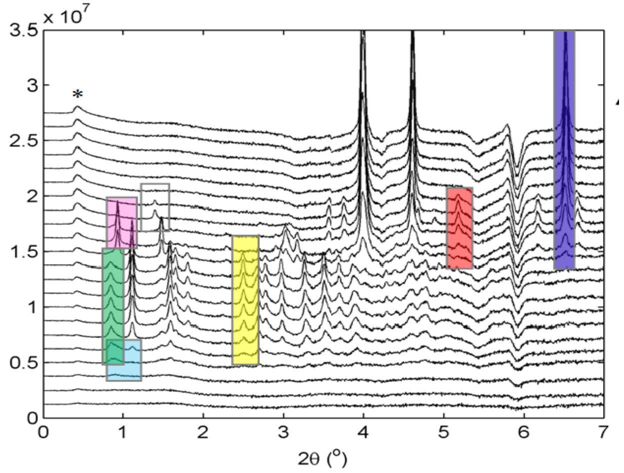
- And to look at the chemistry in different parts of the sample

- Combine XRD with Computed micro Tomography methods and fast acquisition to map chemistry in time and space

Real industrial catalyst body



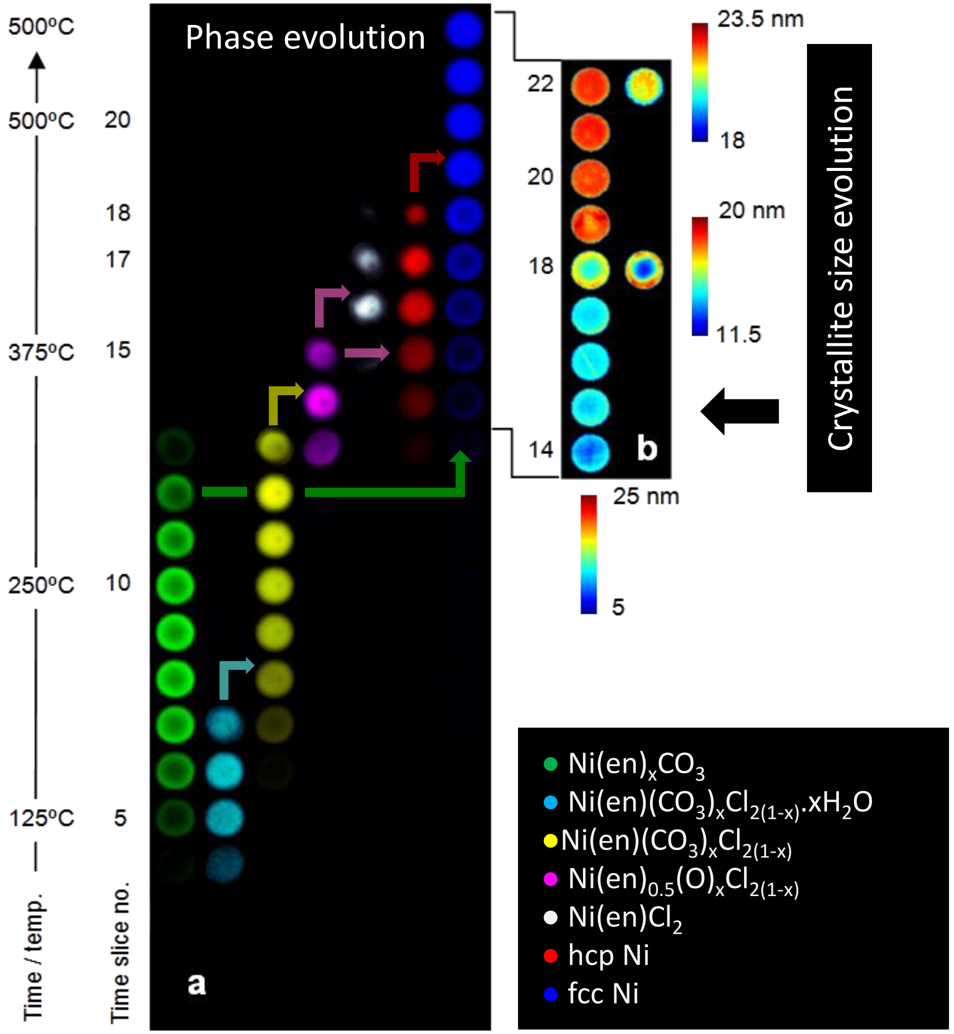
“Traditional” *in-situ* result



- Two routes to the formation of metallic fcc Ni active phase from two different decompositions of the precursor (green and cyan) with different spatial distribution (egg-shell and egg white) and different nano particle size

- important implications for the preparation and subsequent catalytic activity of industrial catalyst bodies

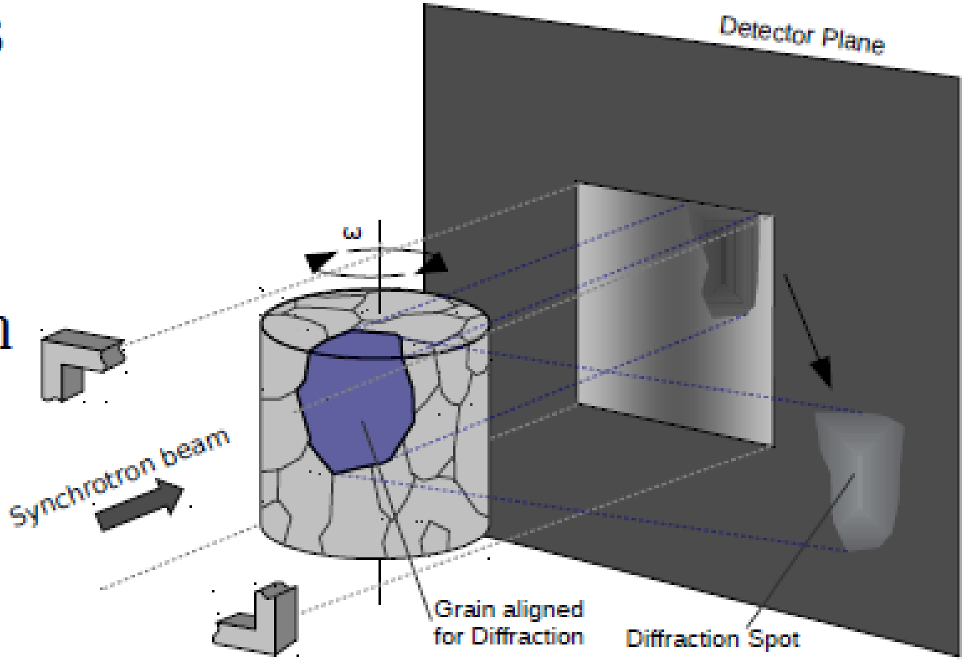
XRD-CT *in-situ* result



S. D. M. Jacques et al., Angew. Chem. doi:10.1002/anie.201104604

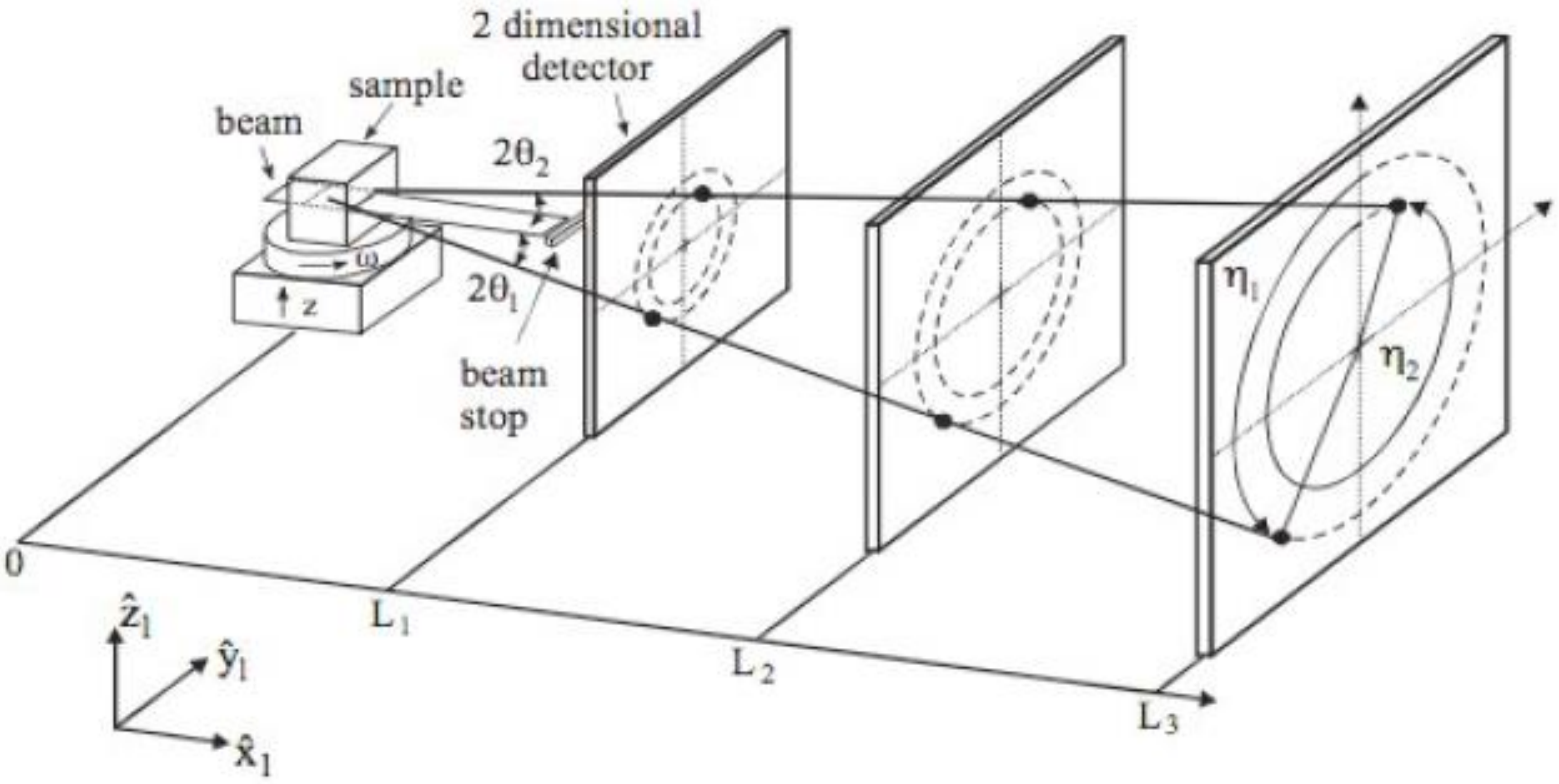
About Diffraction Contrast Tomography

- Diffraction Contrast Tomography is a non-destructive characterization technique of 3D grain microstructure:
  - Assumes undeformed materials
  - Uses 2D monochromatic beam
  - Simultaneous acquisition of transmitted and diffracted beam
  - Acquisition time: 0.1-10h
  - Performs a continuous rotation over 360°, with steps of 0.05-0.1° (7200-3600 images)

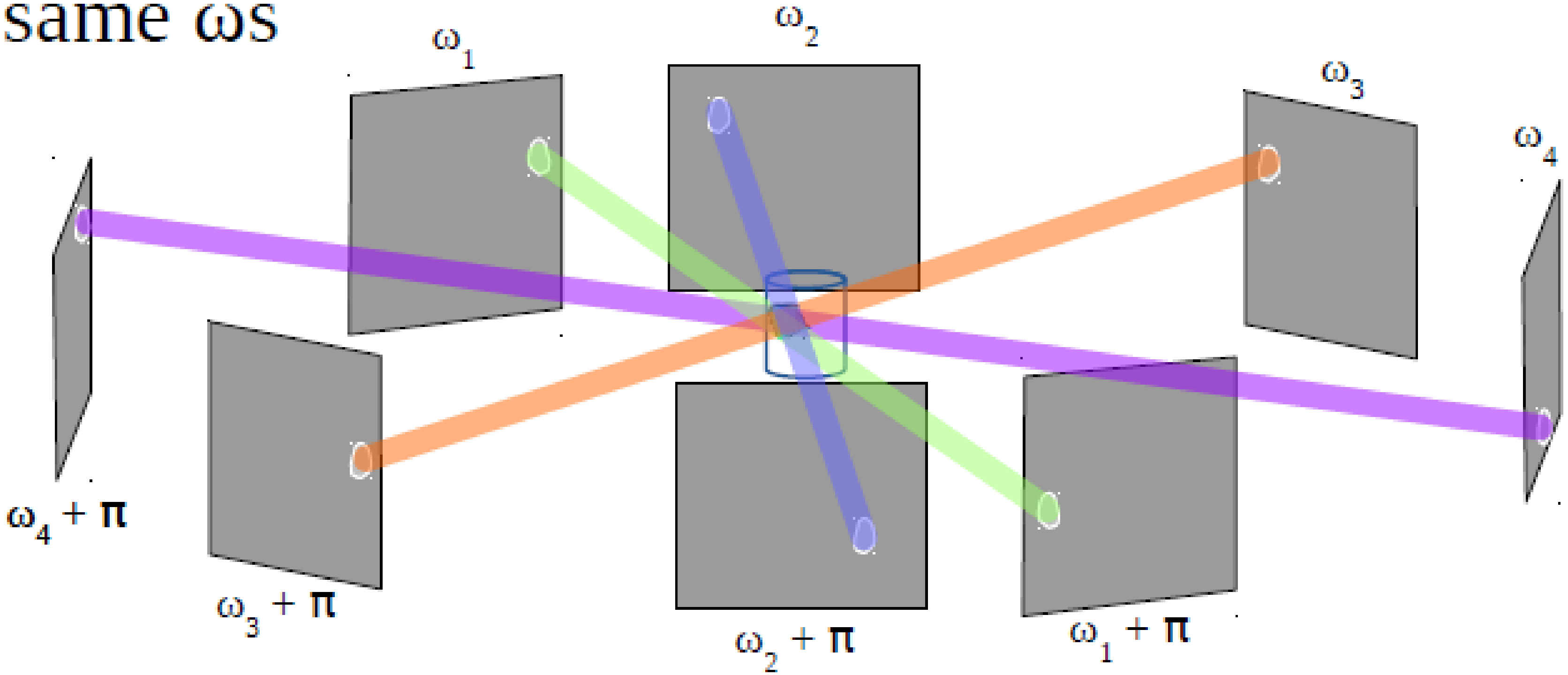


W. Ludwig et al., Rev. Sci. Instr. 2009

- Diffraction Contrast Tomography is traditionally working in a near-field regime
  - What does that mean, and what is "far-field", instead?

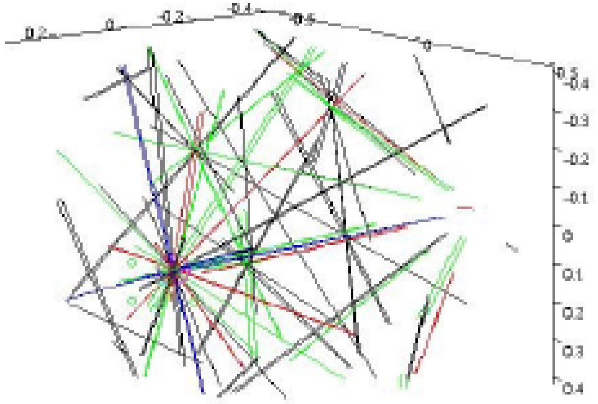


- The whole volume of undeformed grains will project to the same ωs

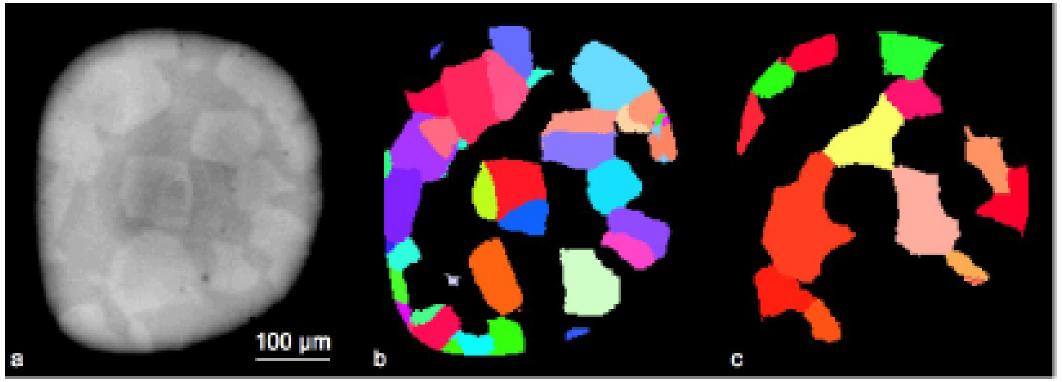


- Plane normals will reflect at ω and ω + π (Friedel pairs)
- The line connecting the centroid of the blobs of a Friedel pair, will pass through the grain centroid (in real-space)

- Using the Friedel pairs it is possible to index the grains:
  - In orientation-space (Rodrigues' space) each reflection is represented by a geodesic
  - Geodesics cross in orientation-space (same grain)



- Diffraction spots serve as projections:
  - Oblique angle reconstruction, using traditional tomography algorithms (e.g. SIRT)
- This also works for multi-phase materials!



a) Phase Contrast Tomography
b) DCT – Austenite
c) DCT – Ferrite

P. Reischig et al., J. of Applied Crystal 2013