HARD X-RAY SCIENCE: OPPORTUNITIES AND CHALLENGES FOR COMPUTING

STEFAN VOGT

Associate Director
X-ray Science Division
Advanced Photon Source

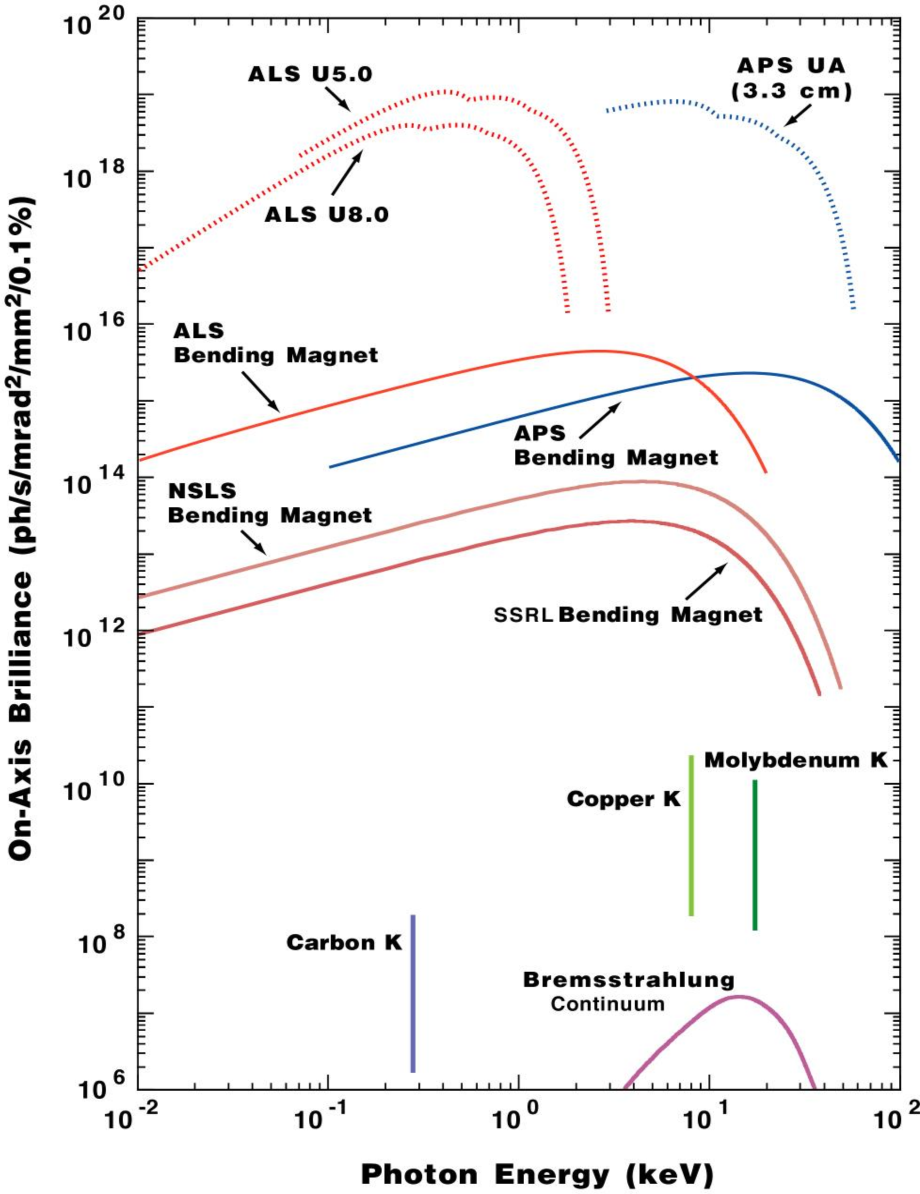
Principal Science Advisor
APS upgrade

Adj. Assoc. Professor
Feinberg School of Medicine
Northwestern University

OUTLINE:

- Overview of the Advanced Photon Source
- Science examples
  - Materials science
  - Life science
- Lensless imaging
- Data drivers
- What do we want / need ?
- A Bright Future: APS upgrade
- Summary &Outlook

SYNCHROTRON: A VERY BRIGHT X-RAY SOURCE



X-ray source:
the **Advanced Photon Source**

Premier high-energy X-ray source in U.S.

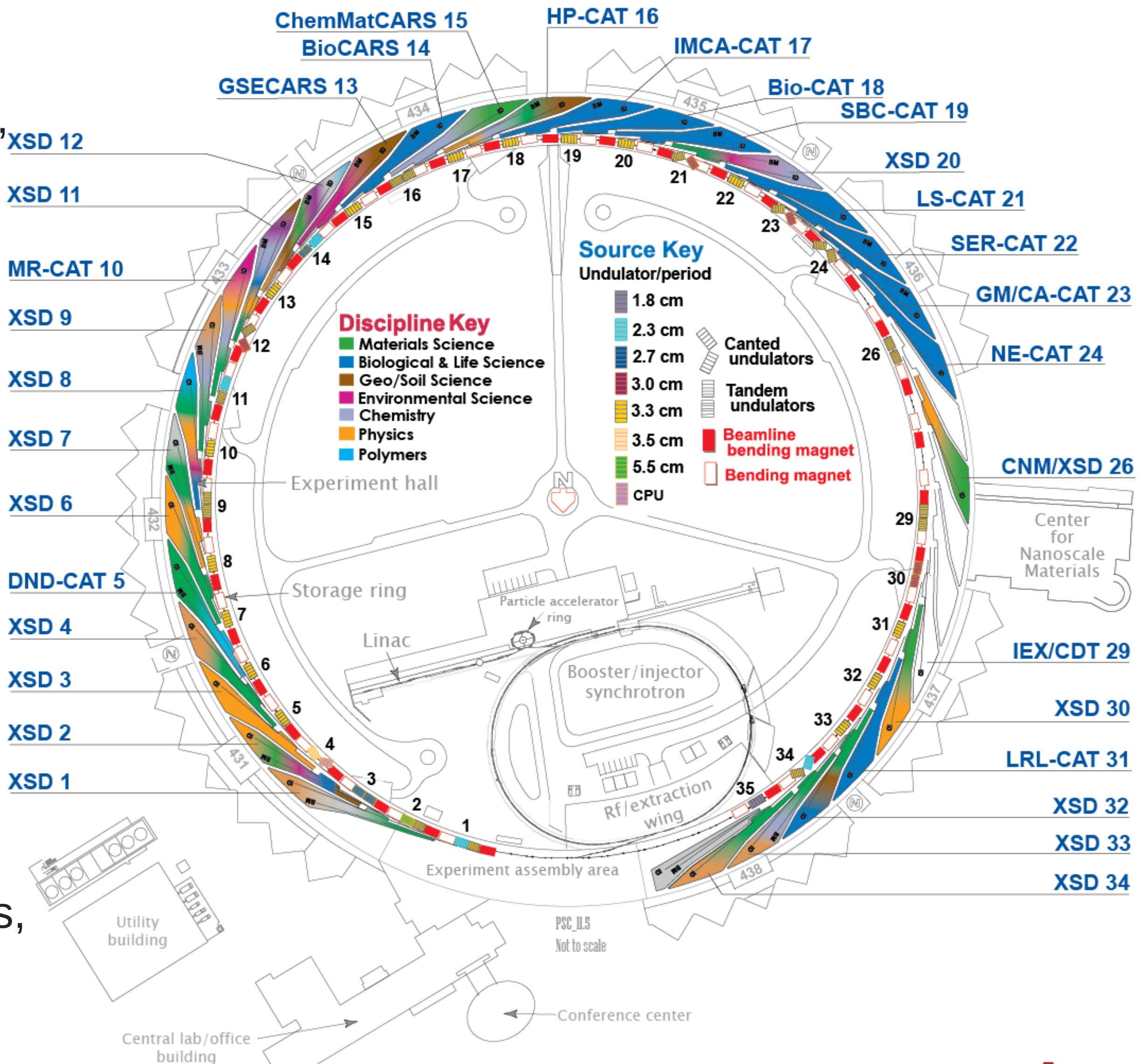
used by 5,700 researchers a year from

- All 50 states plus Puerto Rico
- 33 countries
- 150 companies
- 250 universities

Argonne NATIONAL LABORATORY

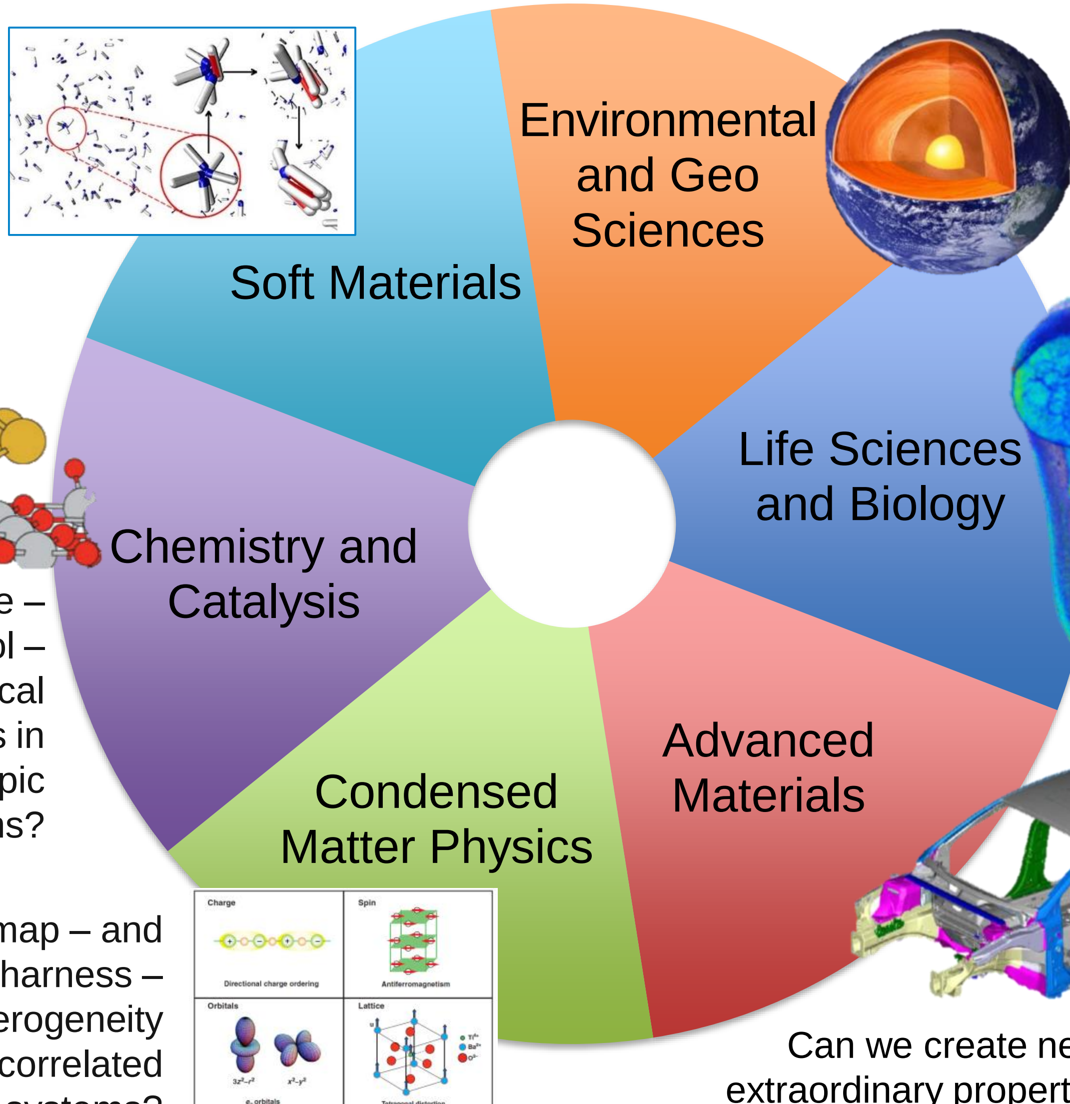
THE ADVANCED PHOTON SOURCE:

- 66 simultaneously operating beamlines, 5000 hours per year
- >5000 unique users in FY14, from every state and worldwide
- >5700 experiments in FY14
- Very diverse user groups
- 1800 peer-reviewed publications in CY14
- >1200 protein structures solved per year, many new drugs discovered
- Many industrial users (~160 companies) from pharma, energy, electronics, materials, ...
- Significant upgrade planned 100x improvement (~2023)

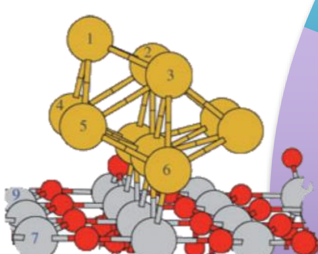


New tools are needed to answer the most pressing scientific questions

Can we determine pathways that lead to novel states and nonequilibrium assemblies?



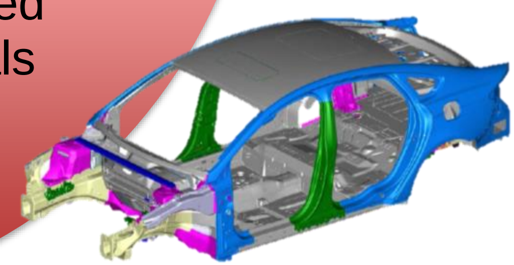
Can we understand physical and chemical processes in the most extreme environments?



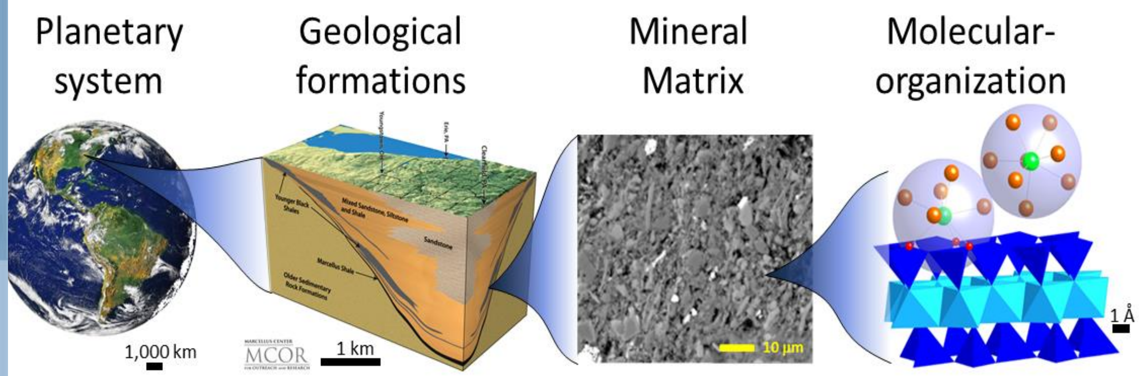
Can we observe – and control – nanoscale chemical transformations in macroscopic systems?

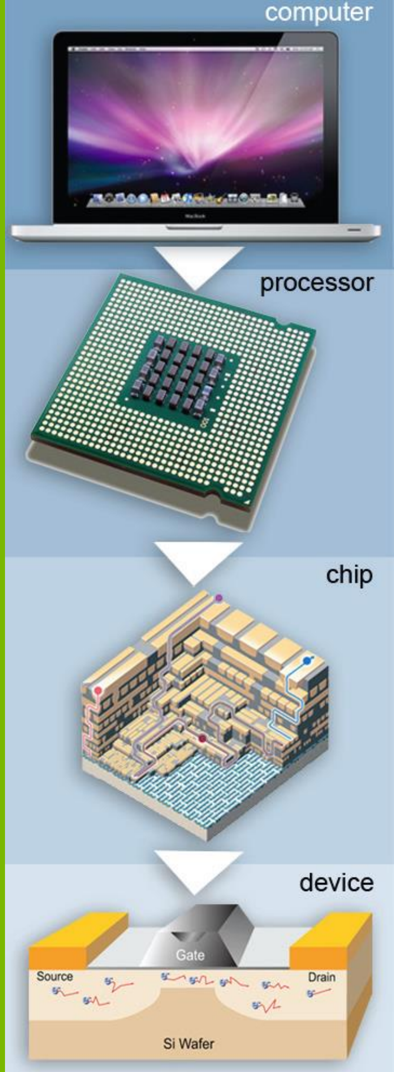
Can we unravel the secrets of biological function – across length scales?

Can we map – and ultimately harness – dynamic heterogeneity in complex correlated systems?

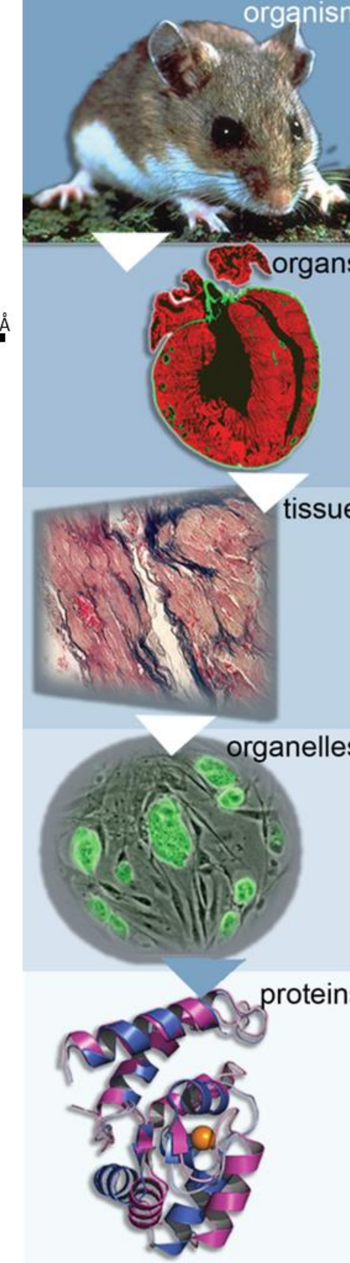


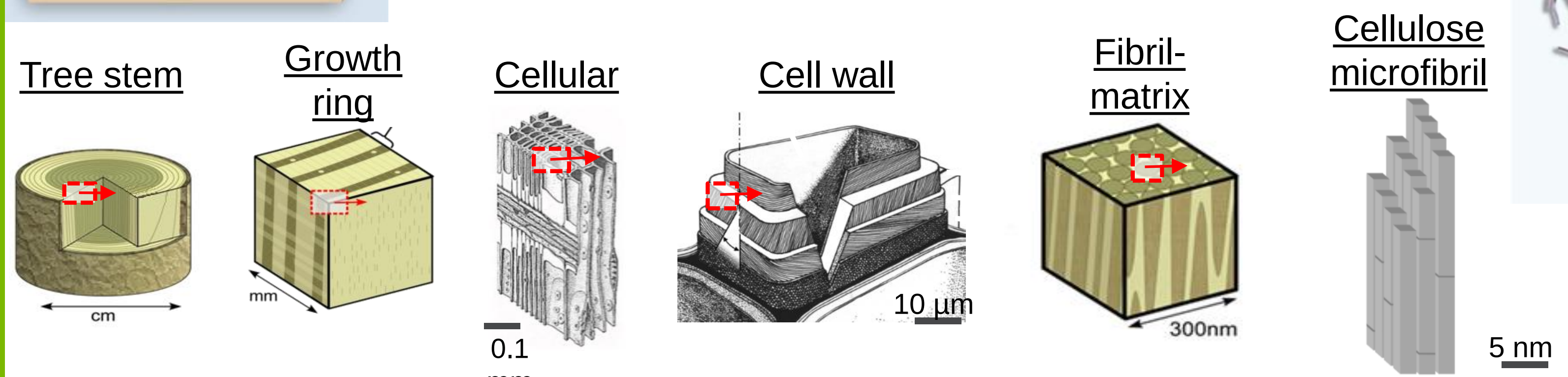
Can we create new materials with extraordinary properties – by engineering defects at the atomic scale?





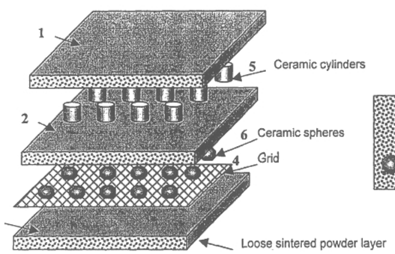
HIERARCHICAL STRUCTURE OF COMPLEX SYSTEMS
=>
NEED TO VISUALIZE STRUCTURE AND FUNCTION ON ALL RELEVANT LENGTHSCALES

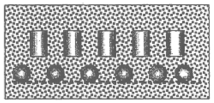




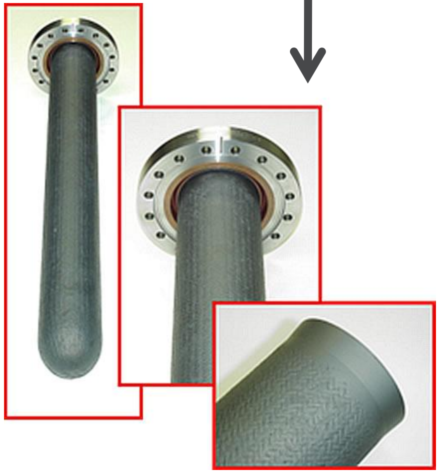
Argonne NATIONAL LABORATORY

NEED FOR 3D MATERIALS CHARACTERIZATION





- New lightweight composites
- Optimizing metal sheet forming
- High-temperature alloys

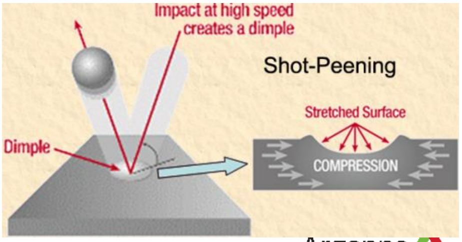


New SiC(m)/SiC(f) composites for in-core structural components in high temperature gas-cooled nuclear reactors under development

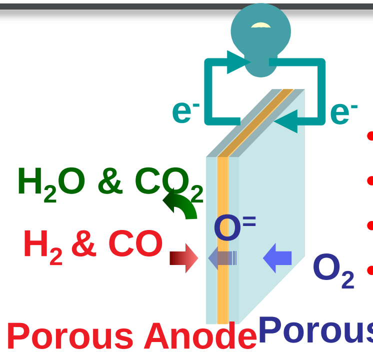
Subsurface microstructural gradients are common and used to enhance material functionality.

Advanced characterization under service conditions enables process optimization, reliability and material discovery.

Process-enhanced properties for airfo



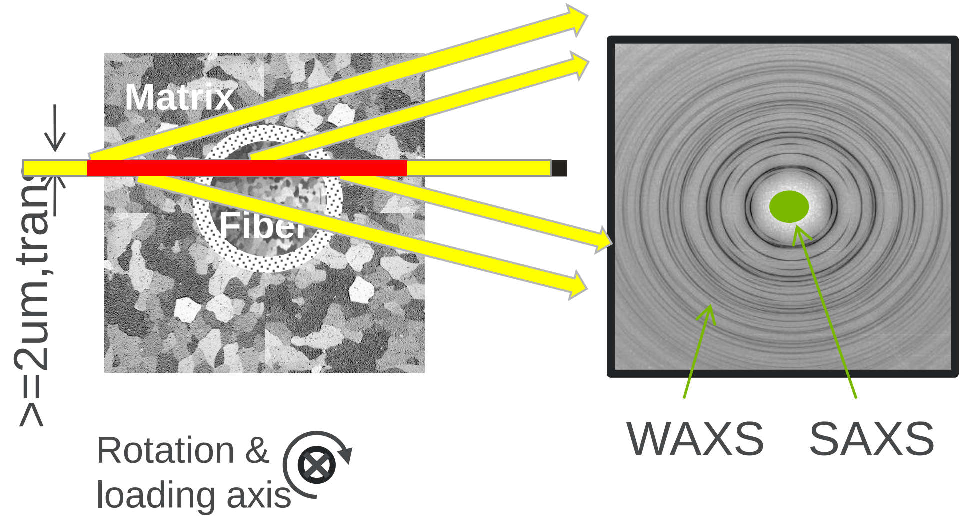
SOFC (battery)

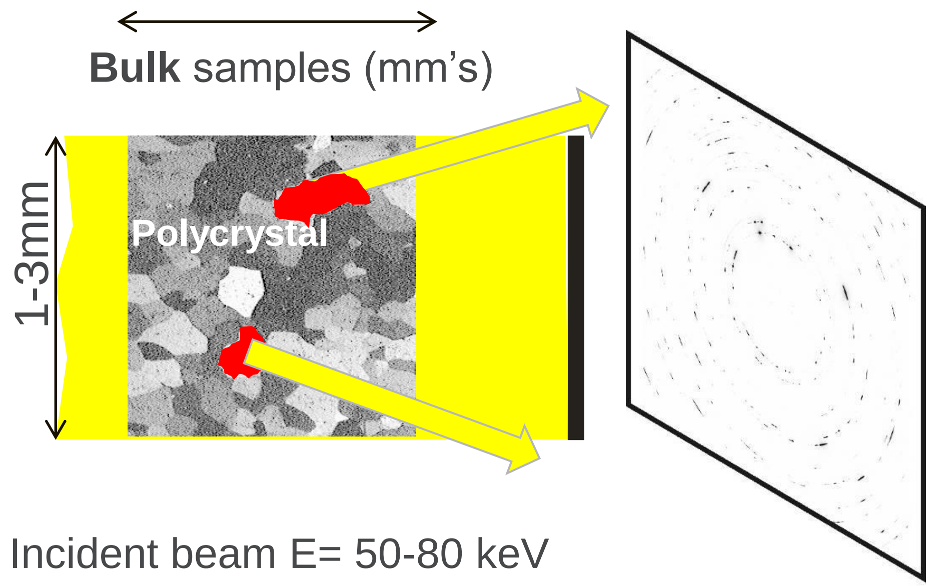


- Controlled porosity
- Thermal mismatch
- Chemical durability
- Mechanical integrity

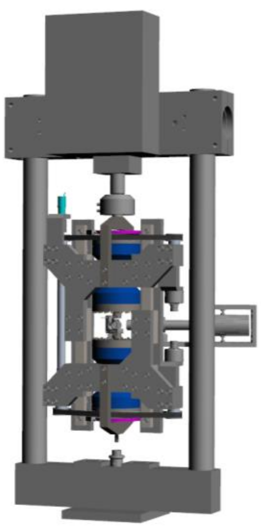
HIGH-ENERGY X-RAY STRAIN AND MICROSTRUCTURAL MAPPING

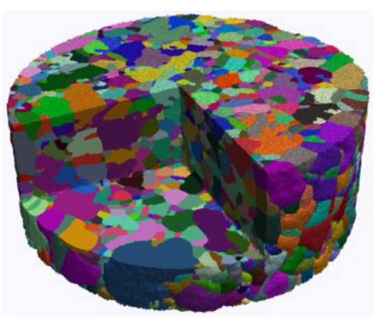
- Small-and wide-angle scattering (grain & line-averaged)
- Scattering tomography (grain-averaged)
  - Translate N times and rotate M times (NxM images) – 2D voxels
  - Reconstruct SAXS/WAXS/absorption information for each voxel
  - Vertical translation for 3D volumes
- Single-grain diffraction tomography (HEDM) + absorption tomography
  - Full-field image and rotate M times (M images)
  - Reconstruct distinct spots on detector
  - Vertical translation for 3D volumes





Incident beam E= 50-80 keV
Scattering angles <10 deg

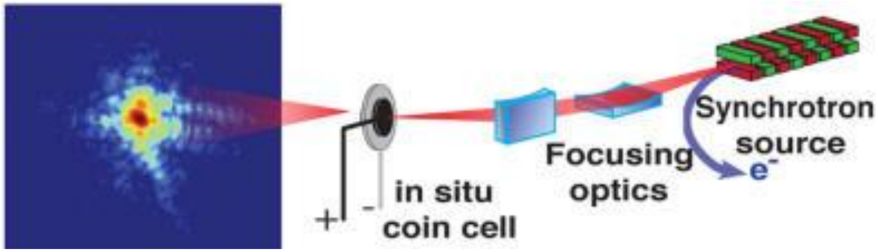


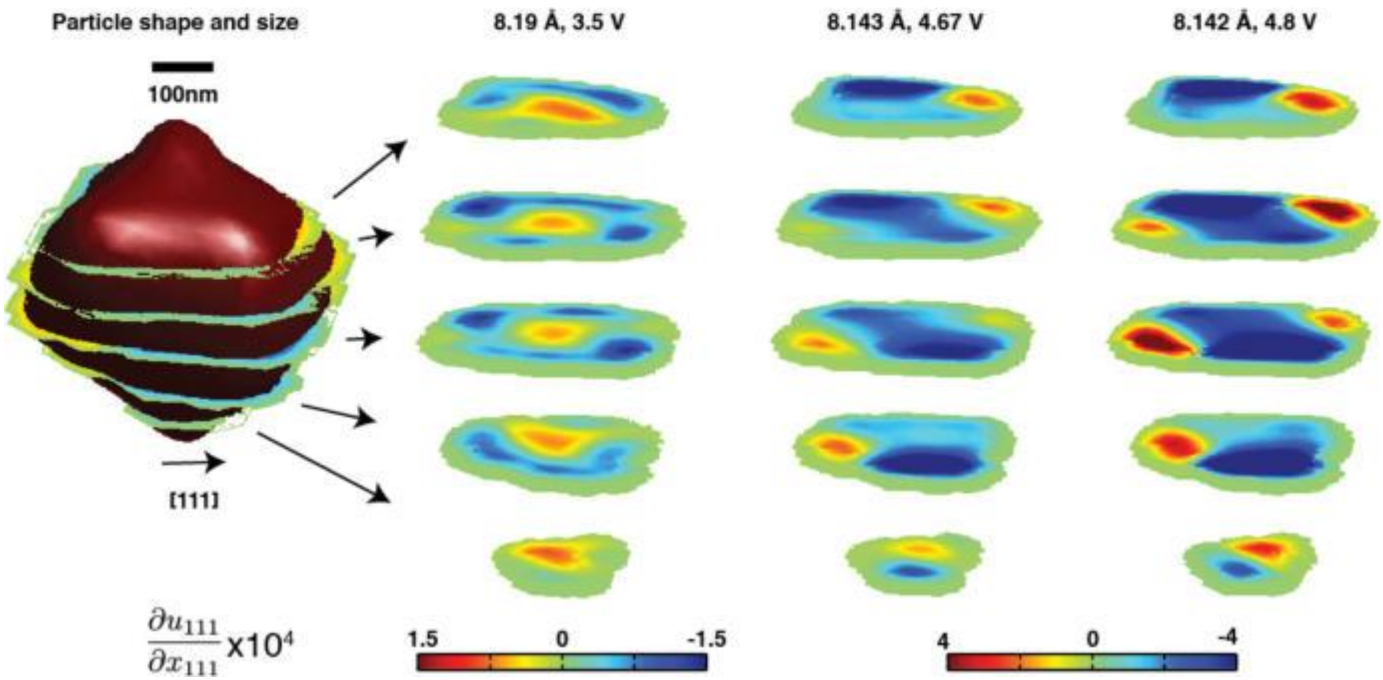


in-situ loading

nanoscale 3D coherent imaging of operating battery material

- Structural phase transitions are a fundamental aspect of the charge and discharge cycle of a lithium ion battery cathode materials.
- Capacity fade in batteries can be understood by studying strain evolution a the single particle level.
- Strain generated during, for example, the cubic-tetragonal phase transformation in LiM_2O_4 causes irreversible damage, including defect nucleation, which leads to large capacity fade.





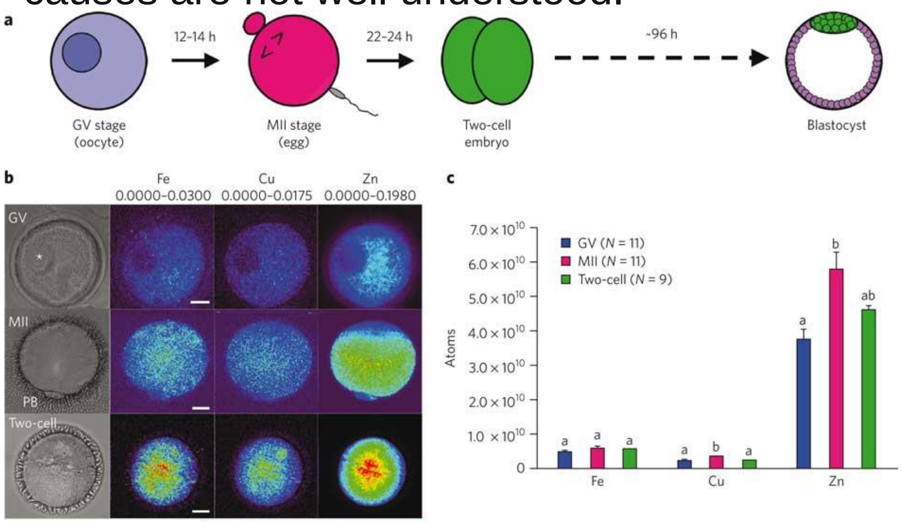
Three dimensional strain evolution in a $LiNi_{0.5}Mn_{1.5}O_4$ battery cathode particle. *As the battery is discharges, Lithium ions diffuse into the lattice causing a structural phase transition seen via strain in the crystal.*

Argonne NATIONAL LABORATORY

Ulvestad, et al. 2014. *Nano Lett.*

Understand the controlling roles of Zinc in reproductive health and embryonic development

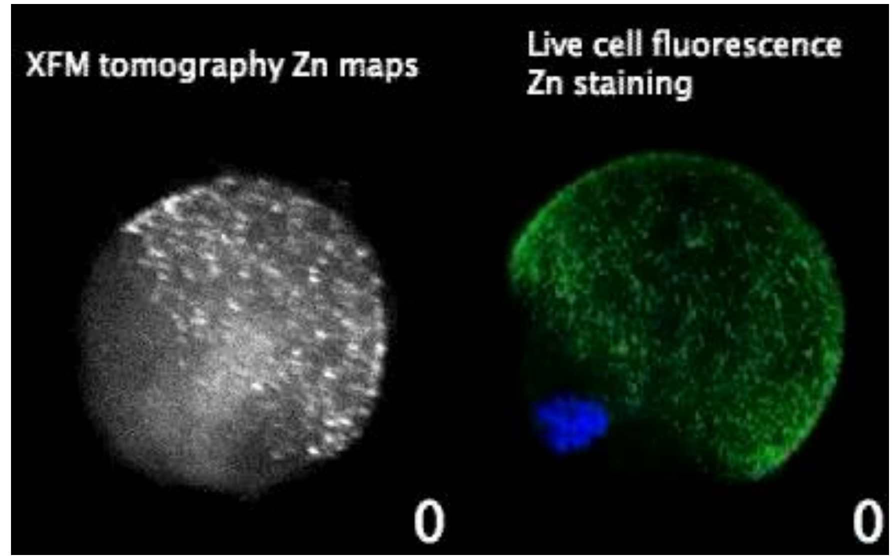
- Infertility results from diseases that impairs the body's ability to perform the basic function of reproduction. Infertility or **reduced fertility affects ~10% of the US population**, and the causes are not well understood.



- As the egg matures, it undergoes significant, dynamic changes in zinc content localization. "Zinc sparks", the concerted, release of Zinc from the egg are essential for reproduction, as is a replenishing intracellular Zinc: the molecular mechanisms are of intense interest in basic biology and medical communities.

Kim, et. al., Nat. Chem. Bio., 2010, Que, et. al., Nat. Chem., 2014, Unpublished Results, O'Halloran & Woodruff Labs

Tomography and Fluorescent Data Reveals Zinc Vesicles



Zinc is ejected during fertilization and originates from vesicles at the periphery of the egg.

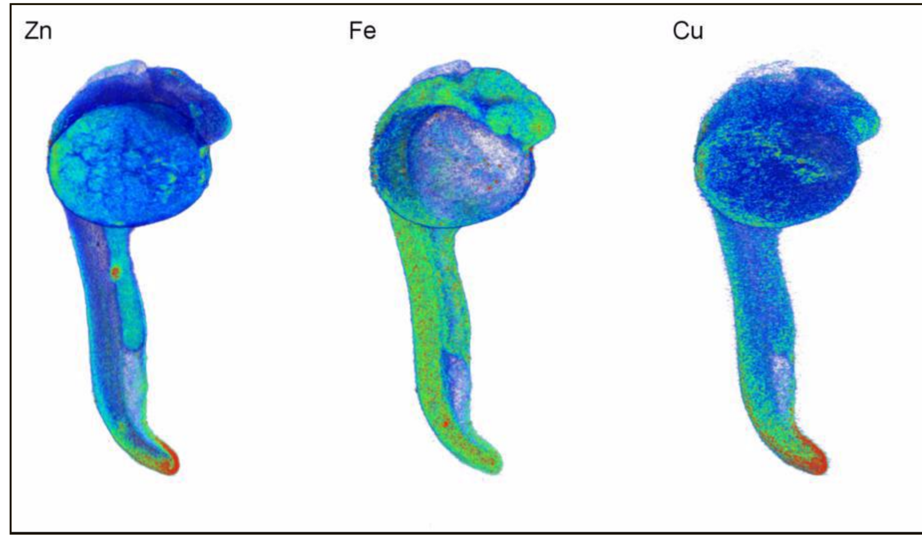
Next steps:

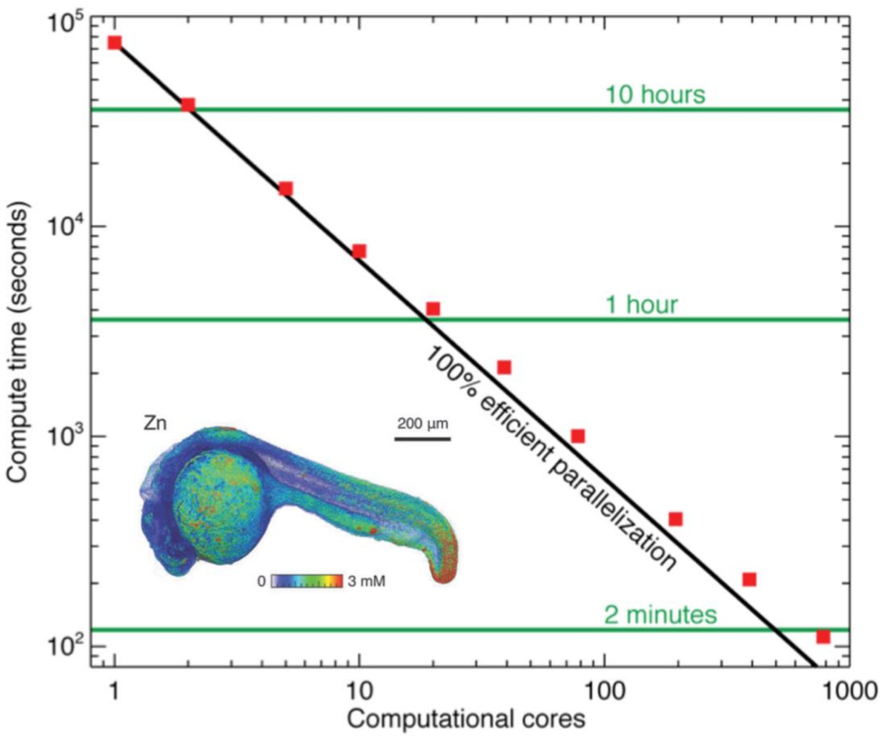
- A membrane-bound zinc transport protein localizes to the meiotic spindle: is Zinc being dynamic regulated there ?
- Sperm also undergo regulatory Zinc fluxes upon activation – how does this relate to fertility

Argonne NATIONAL LABORATORY

RECONSTRUCTION ALGORITHMS MATTER

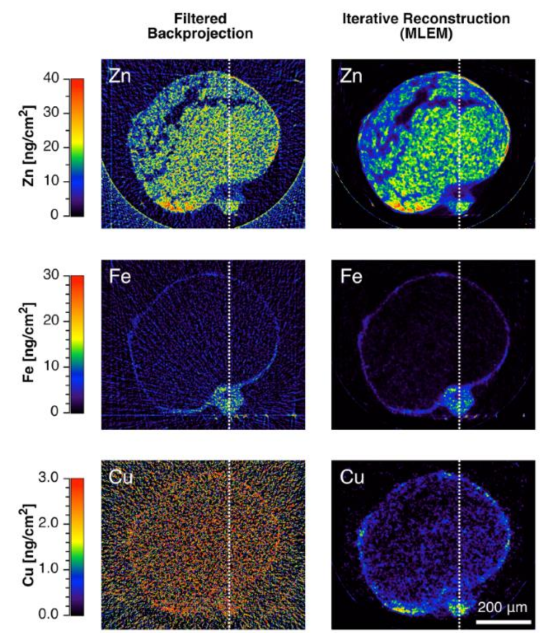
- Now: XRF tomography ~routine. Data acquisition ~routine, fairly automated.
- Reconstruction: can be difficult, need to deal with instrument errors, sample changes (desired and non-desired).
- Field of view ~800x1500um, 400x750 pixels, 60 projections, dwell:10 ms/pixel
- Here resolution limited only by available flux (scan time), and reconstruction.





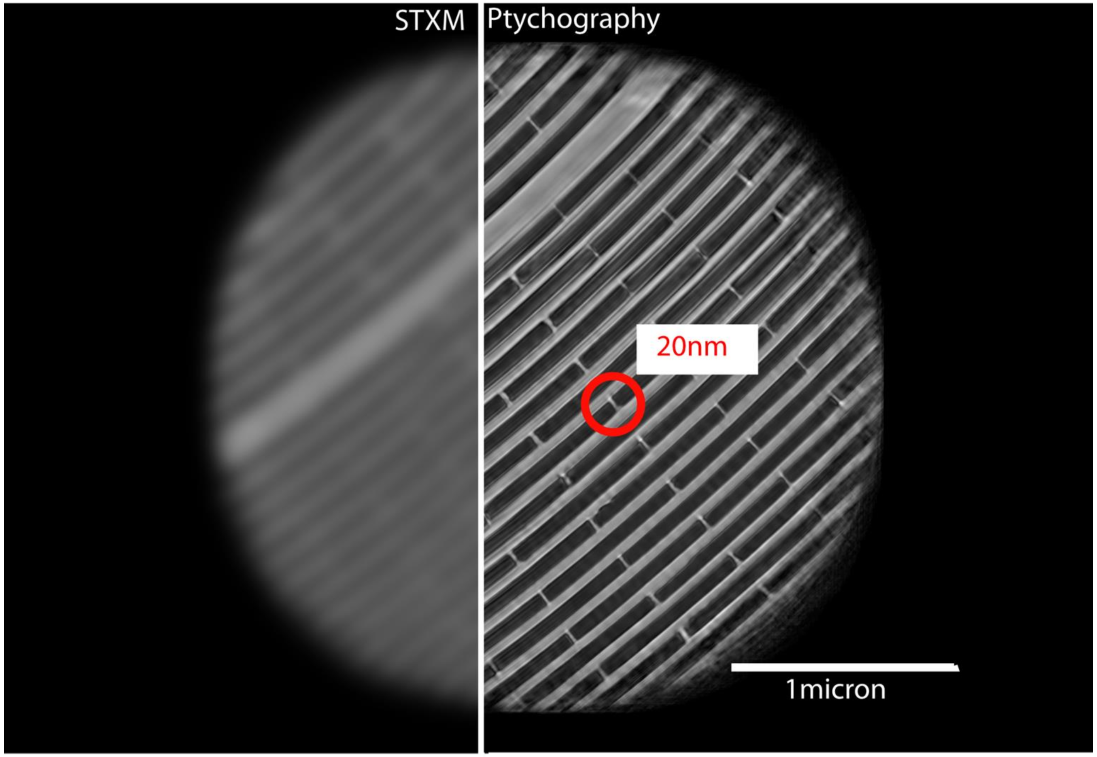
Zebra-fish: metalloprotein cofactor metal distributions correlated with characteristic anatomical features of embryonic development

D. Bourassa et al, Metallomics, 2014



D. Gursoy, T. Bicer et al *et al*

CONVENTIONAL VS LENSLESS IMAGING



- Spatial resolution primarily limited by numerical aperture of optics (somewhat also dose tolerance of sample)
- Phase contrast and amplitude contrast possible, depending on implementation

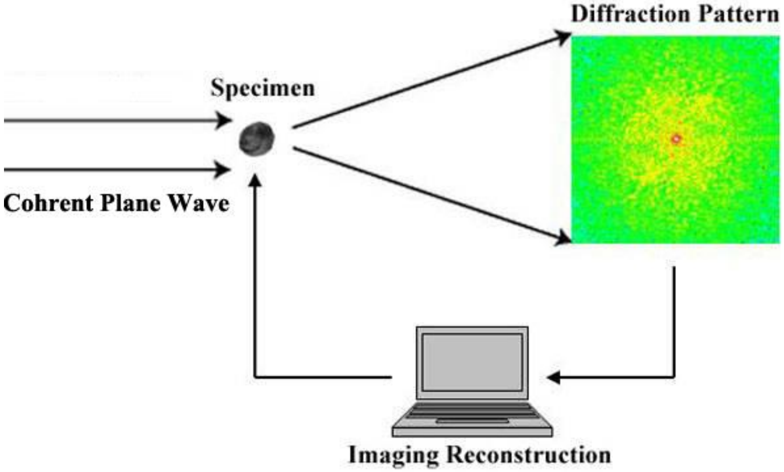
- Spatial resolution limited only by detector numerical aperture and dose tolerance of sample
- Recover both phase and amplitude

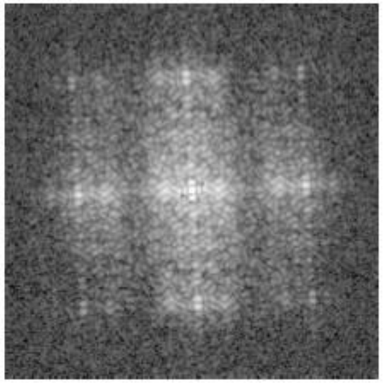
COHERENT DIFFRACTIVE IMAGING

Lensless **method**

Resolution ~ λ / angular size   limited only by wavelength and signal

- Two-step process: record coherent diffraction pattern, recover object structure numerically (iterative phase retrieval)
- Sensitive to phase as well as absorption of the specimen
- Get 3D by tomographic methods; no depth of field limit
- But: must assume some information to recover phase, e.g. known object extent or illumination profile





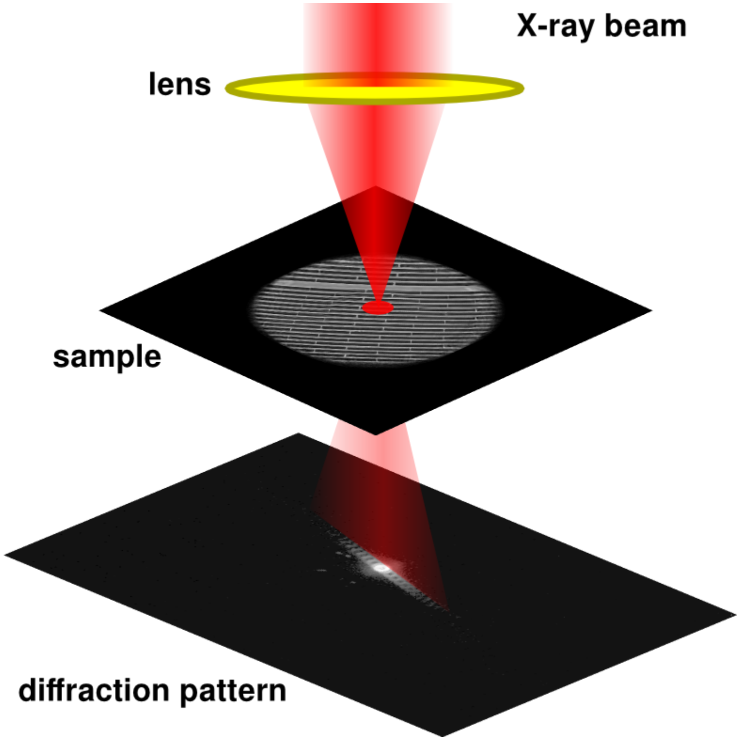
Diffraction pattern

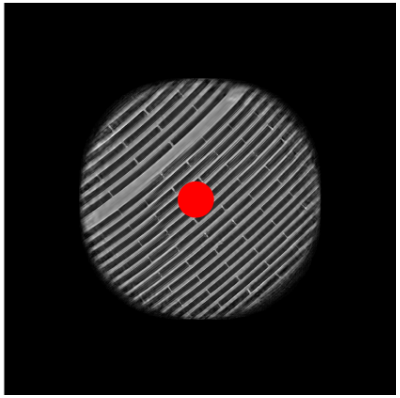
Reconstruction

J. Miao, Nature 400, 342 (1999)

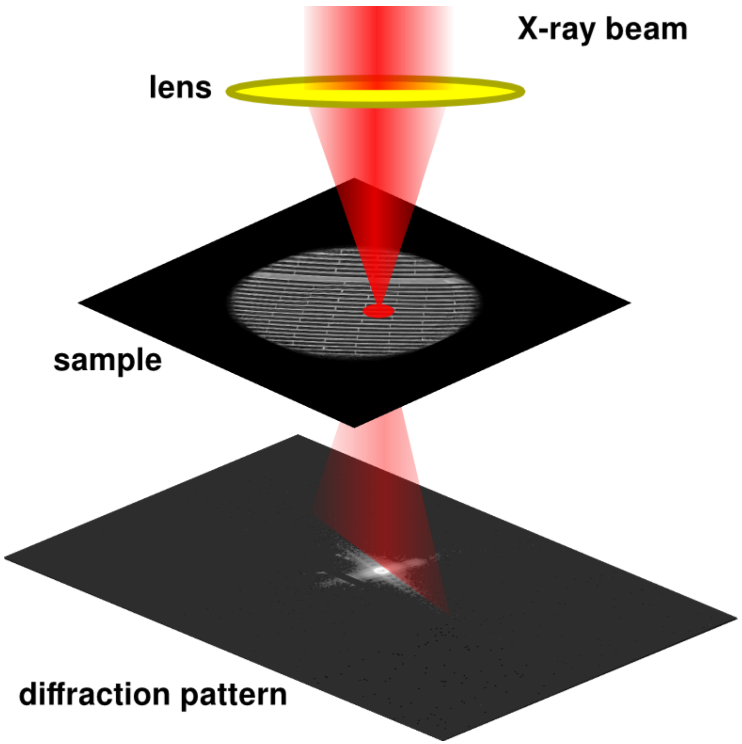
Argonne NATIONAL LABORATORY

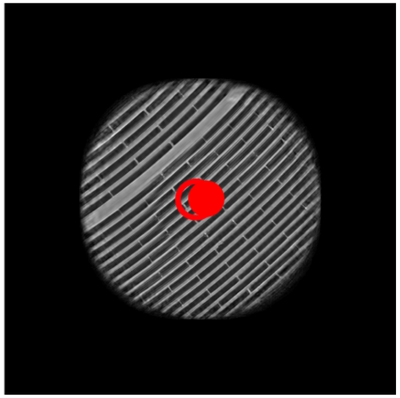
WHAT IS PTYCHOGRAPHY?



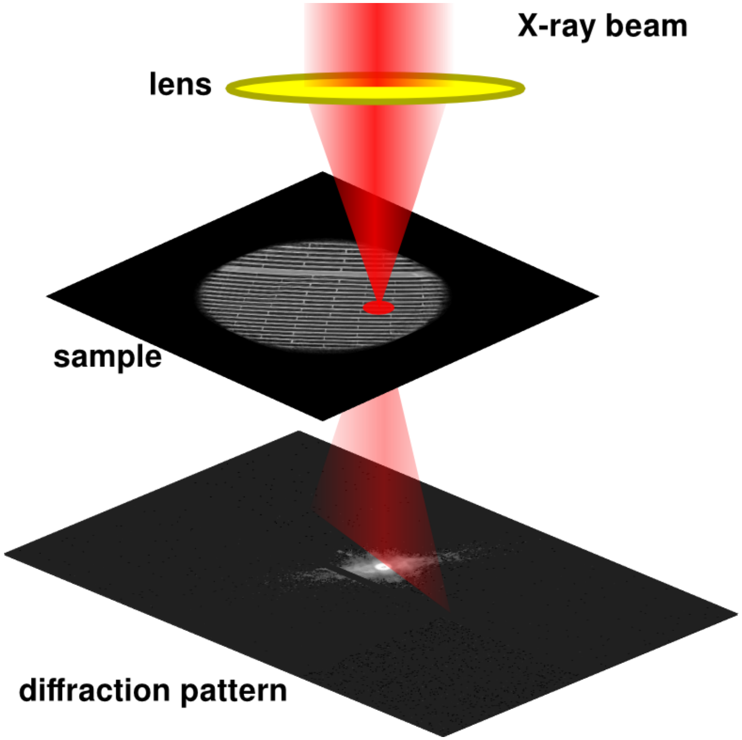


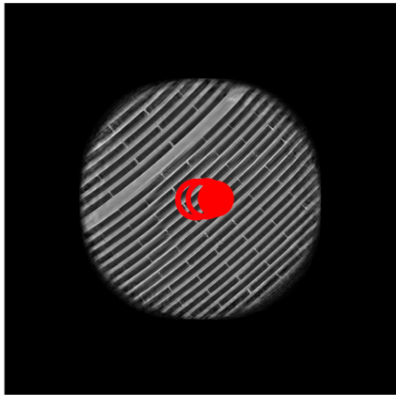
WHAT IS PTYCHOGRAPHY?



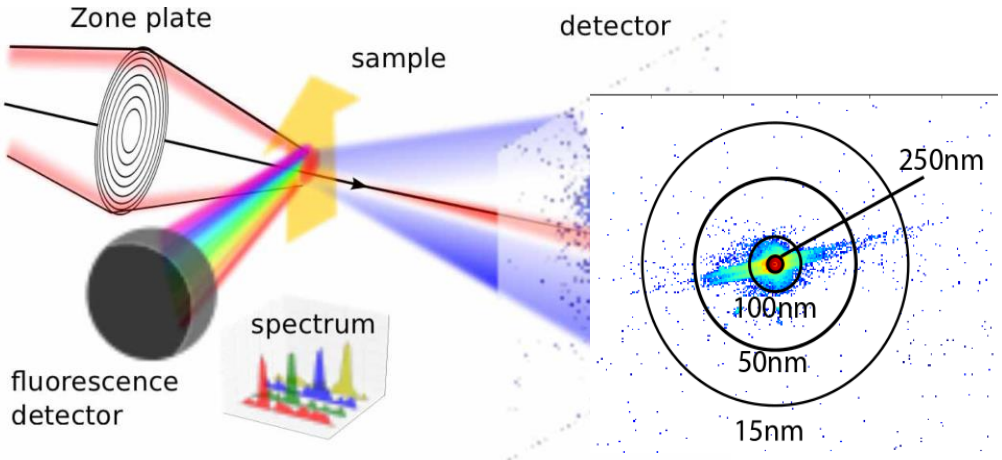


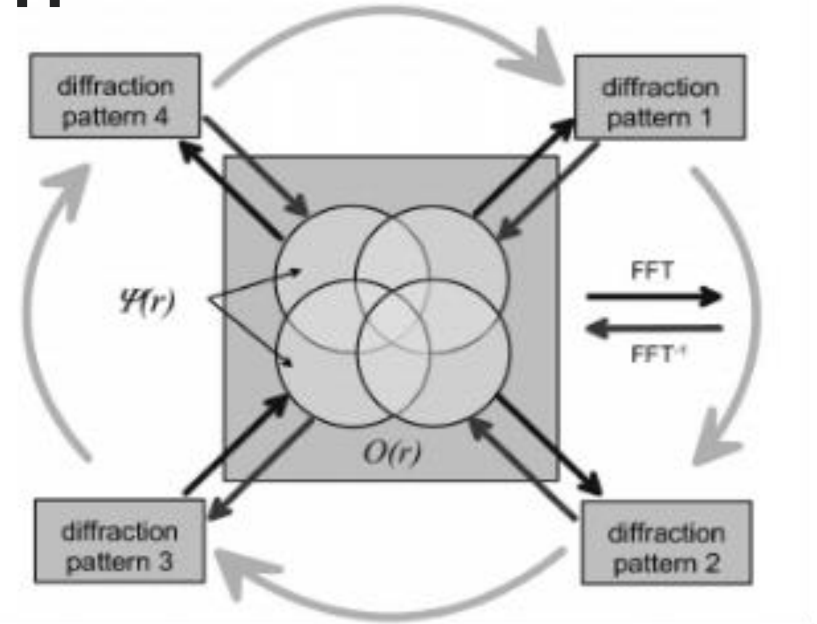
WHAT IS PTYCHOGRAPHY?





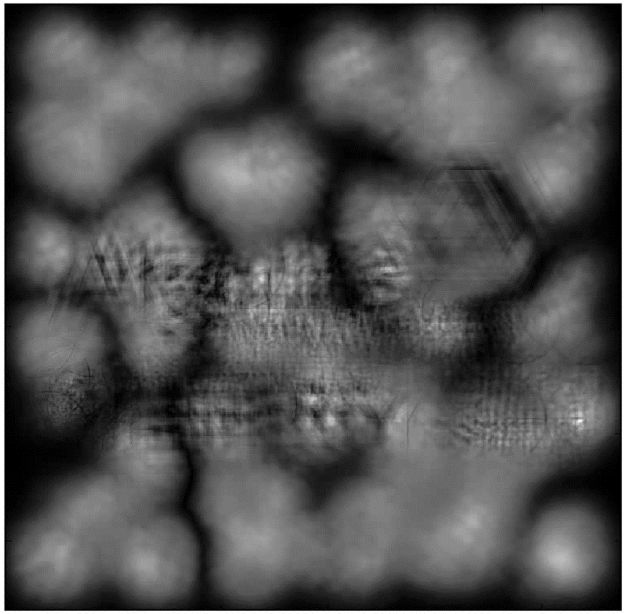
COMBINE LENSLESS IMAGING WITH SCANNING MICROSCOPY: PTYCHOGRAPHY





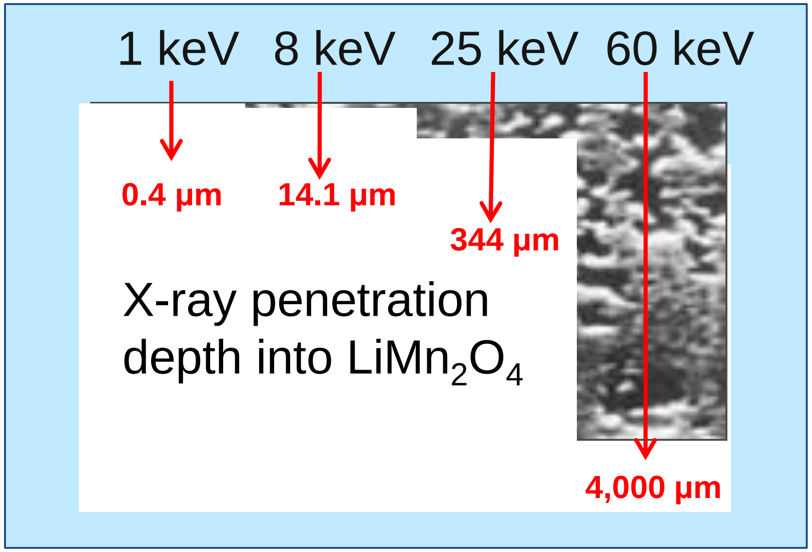
- Scanning microscopy typically only utilises red area
- Ptychography: spatial resolution only limited by detector numerical aperture and dose tolerance of sample. Also retrieve phase, absorption, as well as probe function

- R. Hegerl, W. Hoppe, *Ber. Bunsenges. Phys. Chem* 74, 1148 (1970).
- J. M. Rodenburg, H. M. L. Faulkner, *Appl. Phys. Lett.* 85, 4795 (2004).
- P. Thibault et al., *Science* 321, 379 (2008)
- D. J. Vine, *et al* Opt. Express (2012)
- J. Deng *et al*, PNAS (2015)



LENSLESS IMAGING IN COMBINATION W/ HARD X-RAYS

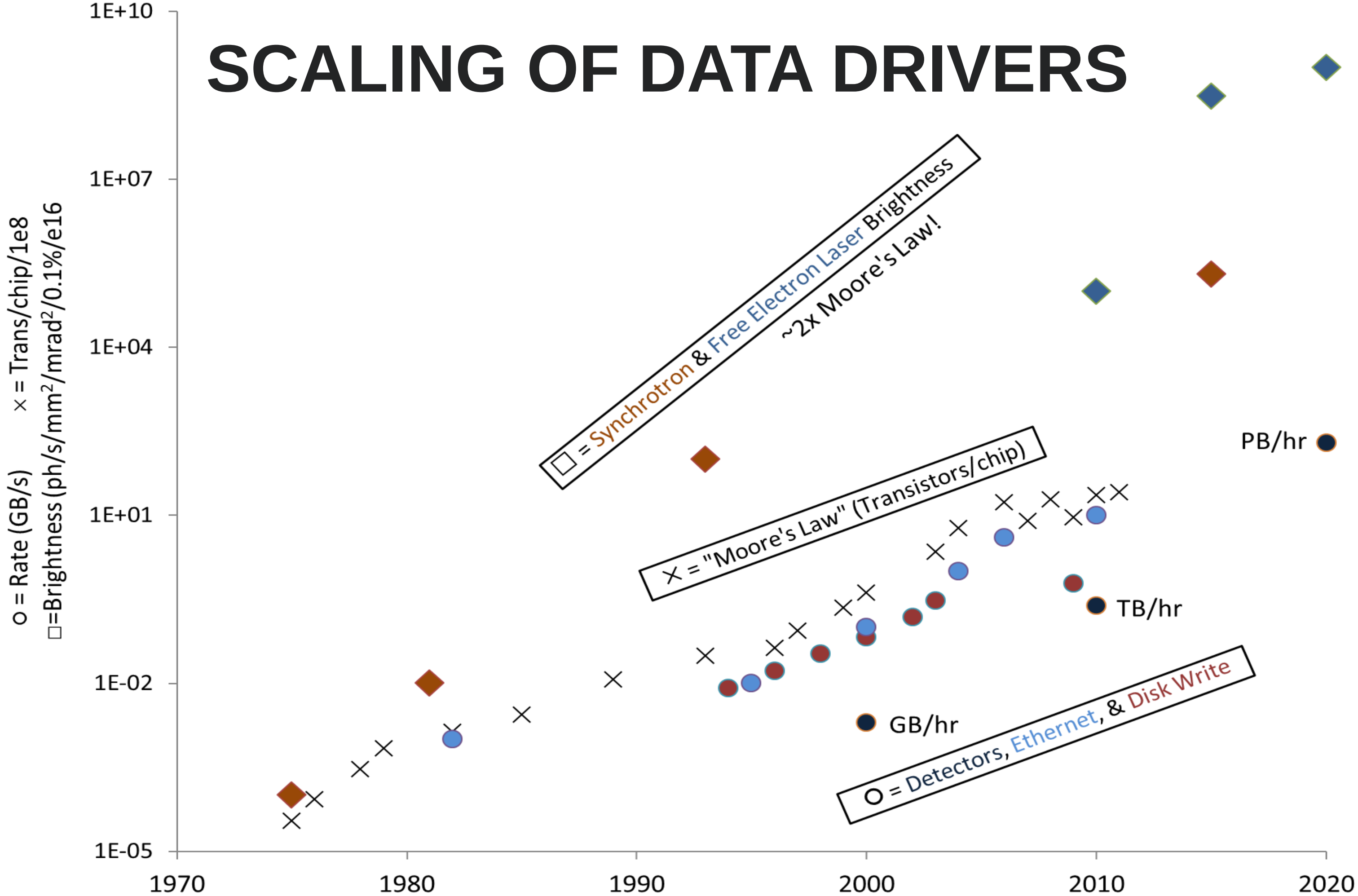
- Spatial resolution in principle only limited by the x-ray wavelength (<<1nm), and the numerical aperture (NA) of the detector
- Detectors have a ‘depth of focus’, correlated to their NA, but can reconstruct to ‘arbitrary’ focus positions.
- Hard X-rays can penetrate very deeply into materials
- Only technique I am aware of, that may have the potential to image a volume of $1x1x1mm^3$ at a 3D resolution of 10 nm.
- Data volume alone (@1 byte greyscale), = 1 petabyte



How realistic ?

- Consider IARPA/RAVEN effort (Rapid Analysis of Various Emerging Nanoelectronics)
- Five year goal: image an IC ($10x10x0.05mm^3$ @ < 10 nm 3D resolution), including reconstruction, in <25 days. => 5 PB voxels.

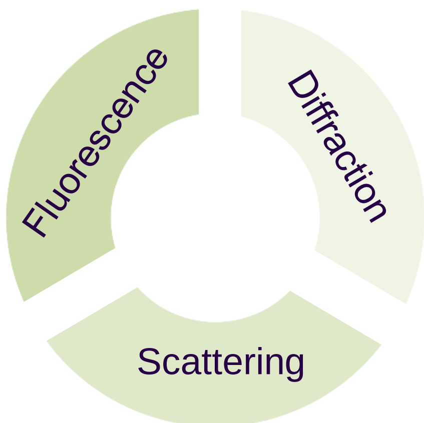
SCALING OF DATA DRIVERS

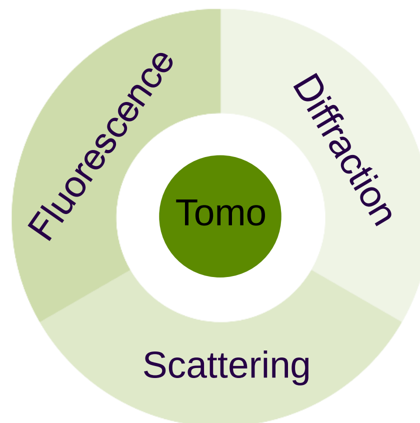


Courtesy Advanced Light Source

Example: becoming available: Jungfrau detector (PSI):
16M pixel detector running at 2kHz: 64 GB/s = 0.25PB/hr

example: Scanning X-ray Microscopy





Past

- Planar data acquisition
- 2D image quantification
- Independent analysis

Present

- Planar data acquisition
- 3D volume quantification
- Correlative analysis

Future

- Volumetric data acquisition
- 3D/4D volume quantification
- Joint analysis

Data acquisition inspired data analysis ⟷ Data analysis inspired data acquisition

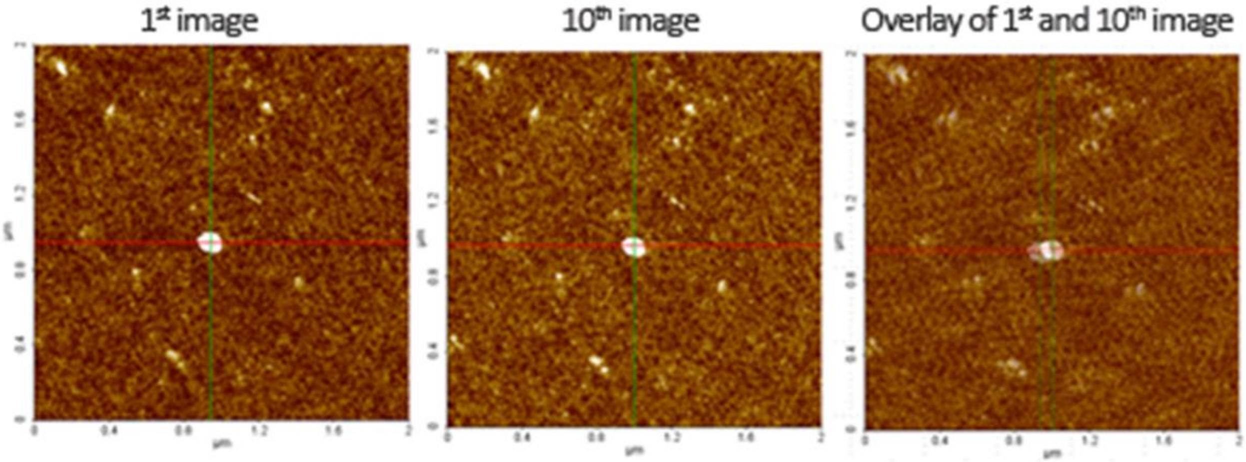
Courtesy Doga Gursoy

WHAT DO WE NEED / WANT ?

- Flexibility:
  - (nearly) every beamline instrument, and every science application is different, but can't reinvent the wheel every time
  - => a modular data analysis pipeline that an BL or experience domain scientist can assemble ?
- Ease of use: Computer literacy varies significantly, as does available institutional support (light sources are user facilities).
  - Automation !!!
  - Can beamline and domain scientists contribute code effectively ?
- Visualization: how do you show highly multimodal, multiscale data ?
- Scalability: depending on specific problem, may need to scale up to super computers, eg, to follow in real time processes.
- Accessibility, reliability, and on demand:
  - Most users probably cannot analyse their data at 'home'. Need analysis tools somewhere ('in the cloud' ?), that can access the data somewhere else, and process
  - But, Lightsources are ~24/7 facilities. Users get experimental time for 8-96 hours, typically once or twice a year. You need to be able to look at the data in order to determine next steps, ie, need (at least preliminary) analysis done within a few minutes. CanNOT (?) rely on an outside entity to keep beamlines running.
  - Need HPC resources 'on demand', but can estimate when resources will be needed.

WHAT DO WE NEED / WANT ? - 2

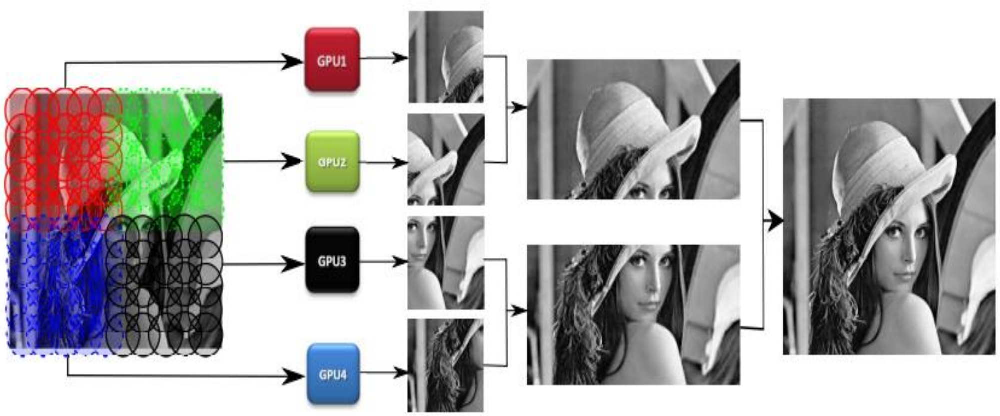
- Algorithms, algorithms, algorithms:
  - Build in error correction

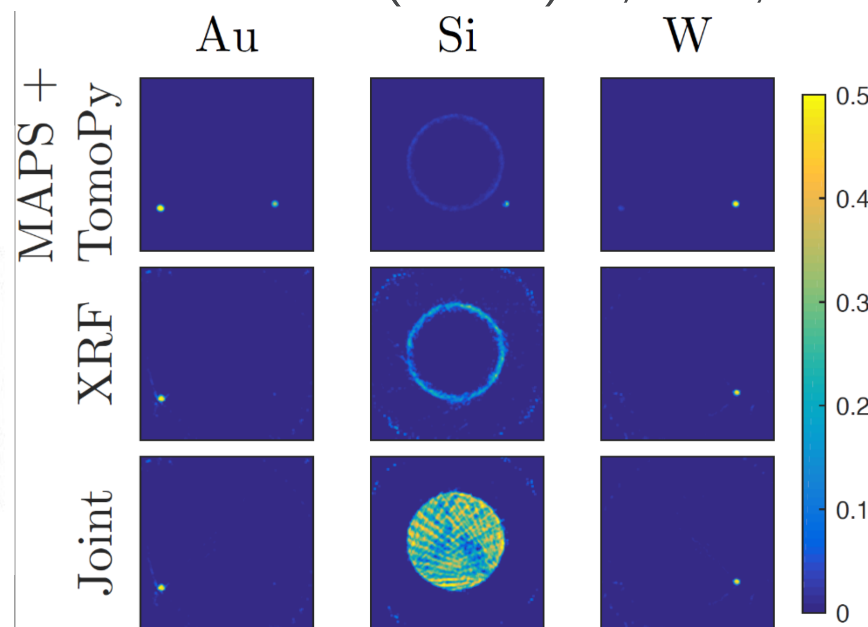


STM, etc,
X drift rate:
0.73 nm/min
Y drift rate:
0.52 nm/min

Courtesy Wendi (Zichao) Di, MCS, ANL

  - Minimize radiation damage
  - Fully exploit multimodality
  - Parallelization





  - ….

Courtesy Youssef Nashed, MCS, ANL / NU

WHAT WE WANT - SMART ALGORITHMS:

- Simulate experiment
- Train neural network on simulated experiments
- Compare acquired data to simulations to select between models
- Algorithm decides based on models, and data acquired thus far, what the next best projection to take is, to distinguish between models.

Courtesy: Mathew Cherukara (ANL-XSD)

‘DREAM’

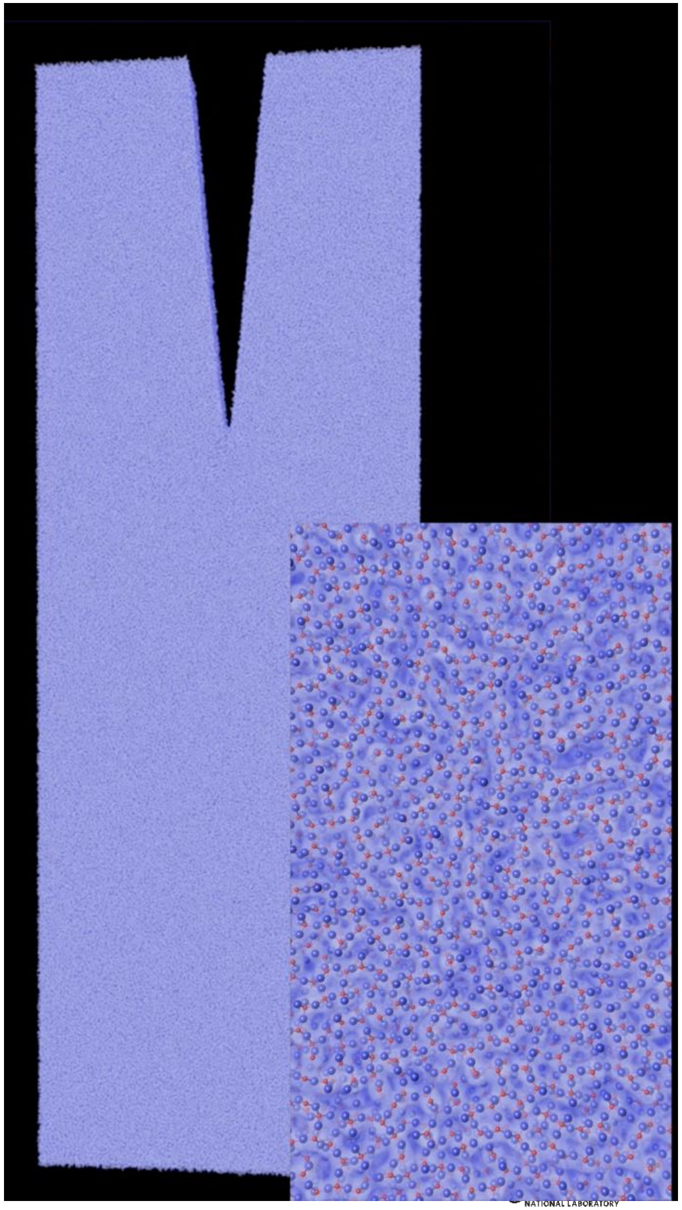
Data analysis and interpretation sufficiently developed, that the analysis drives the experiment.

Requires:

- Modelling of system
- Streaming of data into analysis pipeline
- Reduction/visualization/interpretation of data
- Decision making on how to advance experiment

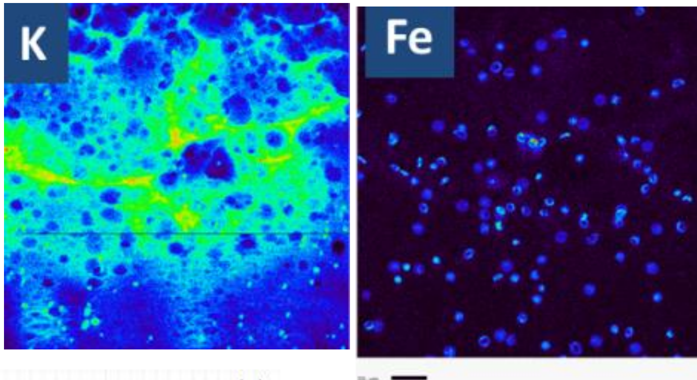
For example:

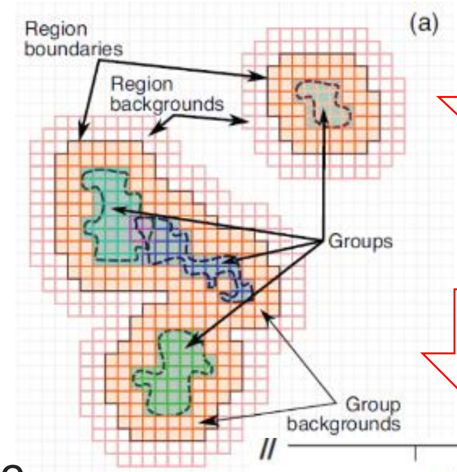
- To distinguish between models, as the experiment progresses decide which projections to take, for how long (statistics), …
- To follow dynamic processes, need to detect where change occurs, ‘zoom’ into the relevant area, and focus exposure time there.
  Problem: you do not know in advance where the change will occur.



TODAY VS TOMORROW

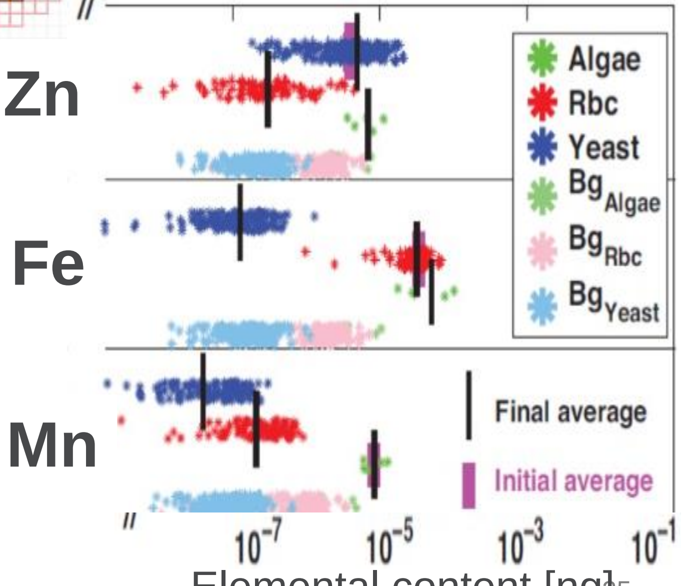
- Today
  - Manually moving, analyzing data.
  - Ad hoc tools that do not scale to the next generation of instruments
  - algorithms can be “dangerous” if not used carefully
- Tomorrow
  - Extensive toolset of scalable algorithms (e.g., machine learning, statistical)
  - Scientific knowledge integrated with analysis, visualization and simulation
  - Automatic Integration of data from multiple sources, cataloguing and transfer
  - Efficient data reduction strategies





Identification and Classification

Reduction and Visualization

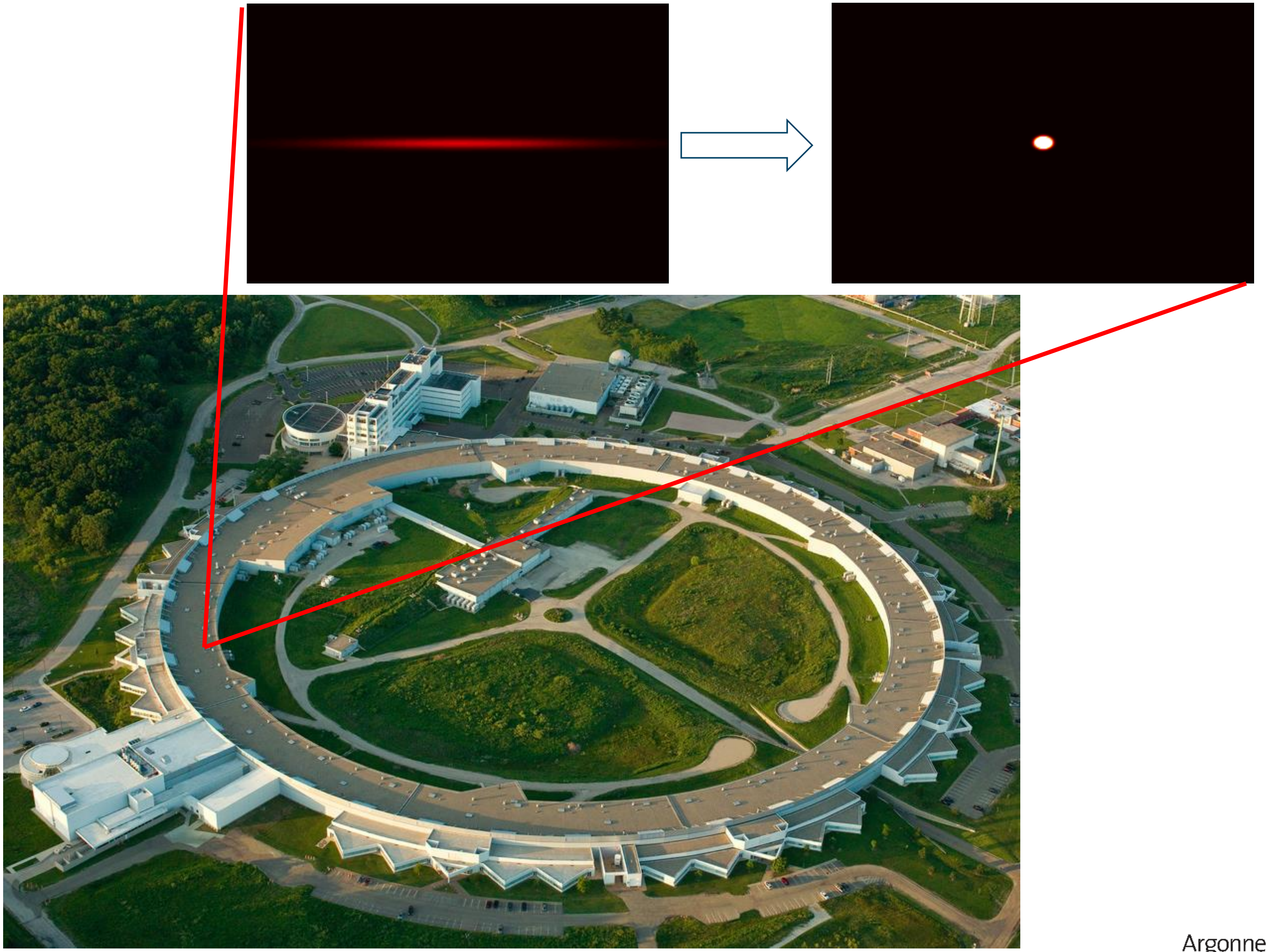


Top: X-ray fluorescence maps of different cells. Middle: software automatically identifies and classifies 3 different cell types, enabling further analysis. Comparison of the resulting average elemental content per individual cell.

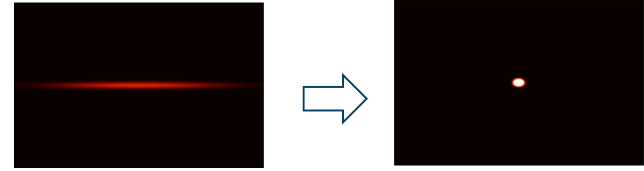
S. Wang, et al, J Synchrotron Radiation, accepted

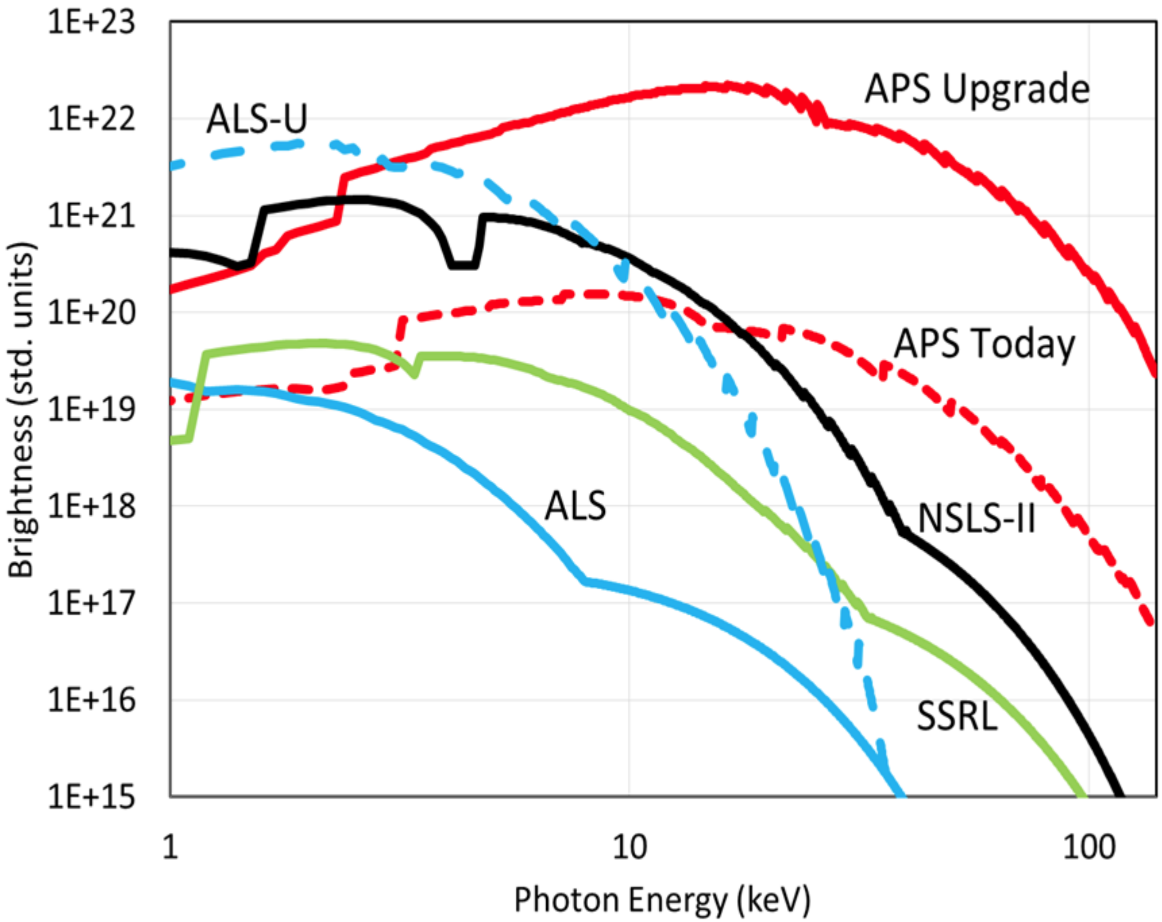
UPGRADING THE APS

AN MBA LATTICE AT APS: APPROACHING FULL LATERAL COHERENCE



APS MBA UPGRADE





Brightness vs. x-ray energy

- **Brightness increases of 100x and more compared to what we have today**
- **Micro/nanoprobes directly brightness driven**

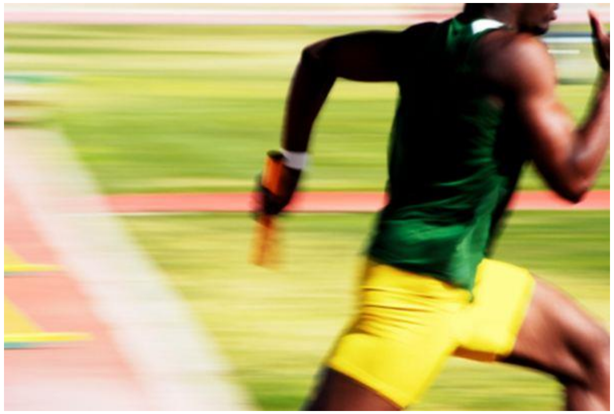
⇒ **possible to get nearly 100% of APS flux into a 0.3x0.25 um spot !!!**

⇒ **Push for highest resolution**
20 nm for full in situ & spectroscopy

⇒ **5nm and below for mapping and with CDI/Ptychography**

This upgrade will revolutionize scanning probe microscopies…

APS-U: KEY ENABLER - SPEED



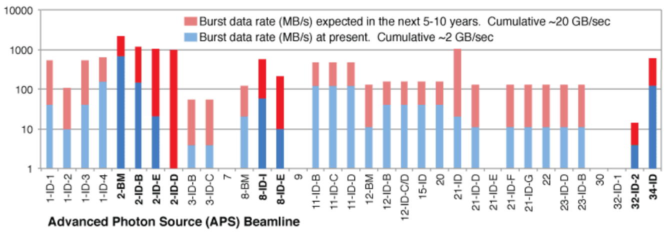
~ 50x

- Speed can be improved by factor of 100 – 10000
  - 100-200x gain in brightness due to ring
  - 10-100x gain due to additional improvements

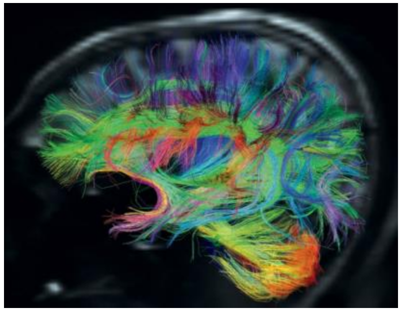
    (IDs, optics, detectors, methods such as dose fractionation, …)
- Large 2D samples at simultaneous high resolution (needle in haystack problems)
- Follow in real time *in situ* processes with Elemental mapping and Spectroscopy
- XRF Tomography on time scales of minutes (access to dynamic processes)
- 4D/5D imaging (spectroscopy, time, as a function of environment, mutations, …)

MAJOR DATA AND COMPUTATION CHALLENGES AND OPPORTUNITIES ARISE ACROSS APS. THEY EXPLODE WITH APS-U

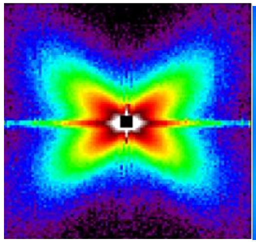
- **Huge data** from new detectors, APS-U
  - E.g., XPCS: FY14: 2MB images @ 100 Hz; FY16: 1MB images @ 2000 Hz (x 10!); Eiger: 2Mbyte @ 3000 Hz (x 3!); APS-U another 2-3 orders of magnitude.

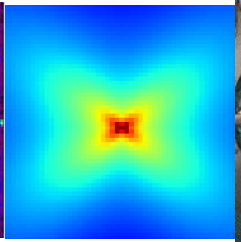


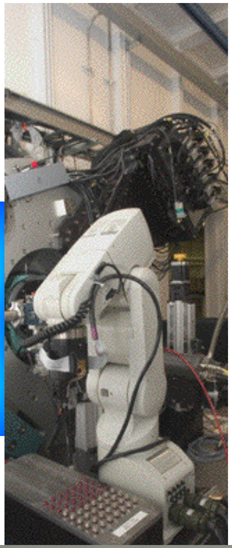
- **Complex, multi-modal data** needs advanced computation for interpretation
  - E.g., Ptychography+elemental mapping+visual images as a function of reaction conditions



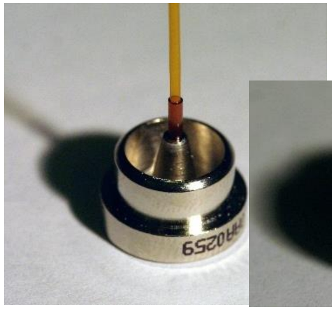
- Advanced modeling and theory enable **fitting and co-optimization** of model and experiment
  - Goal: Fit one model to all measurements







- New user demographics → **automation**
  - Scale to more and different users, many with limited or no experience with light sources



ACKNOWLEDGEMENTS

- O. Antipova, L. Li, C. Roehrig, B. Lai, E. Maxey, J. Maser, Q. Jin, S.-C. Gleber (August), and the rest of the Microscopy group, Argonne

- D. Gursoy, Computational X-ray Science Group, Argonne
- W. Di, T. Bicer, S. Wild, S. Leyffer, Y. Nashed, T. Peterka, Math / Computer Science Argonne
- J. Deng, C. Jacobsen, Argonne / Northwestern University
- T. O'Halloran *et al*, Northwestern University
- C. Fahrni *et al*, Georgia Tech

Financial support:

- Department of Energy (Basic Energy Science)
- National Institutes of Health (NIBIB, NIGMS)

Argonne NATIONAL LABORATORY


Thanks!