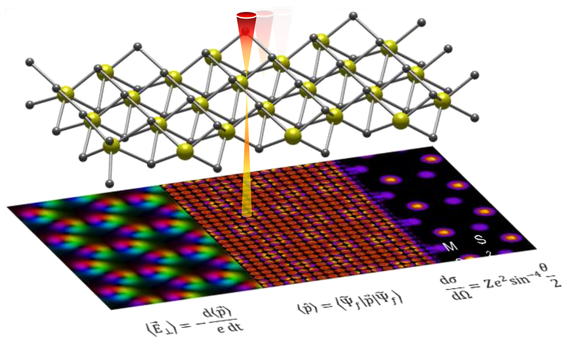
HELMHOLTZ SPITZENFORSCHUNG FÜR GROSSE HERAUSFORDERUNGEN

Ptychography 4.0 – An Information and Data Science Pilot Project Data infrastructure and applications

Alexander Clausen <a.clausen@fz-juelich.de>[1], Simeon Ehrig <s.ehrig@hzdr.de>[2], Heide Meißner <h.meissner@hzdr.de>[2], Knut Müller-Caspary <k.mueller-caspary@fz-juelich.de>[1], Dieter Weber <d.weber@fz-juelich.de>[1], Markus Wollgarten <wollgarten@helmholtz-berlin.de>[3]

[1]Forschungszentrum Jülich GmbH, Leo-Brandt-Straße, 52428 Jülich
[2]IHelmholtz-Zentrum Dresden - Rossendorf (HZDR), Bautzner Landstraße 400, 01328 Dresden
[3]Helmholtz-Zentrum Berlin für Materialien und Energie, Hahn-Meitner-Platz 1, 14109 Berlin

Ptychography – function principle



- Lensless imaging
- Resolution beyond conventional limits
- Dose-efficient imaging
- Visible light, EUV, X-rays, electrons

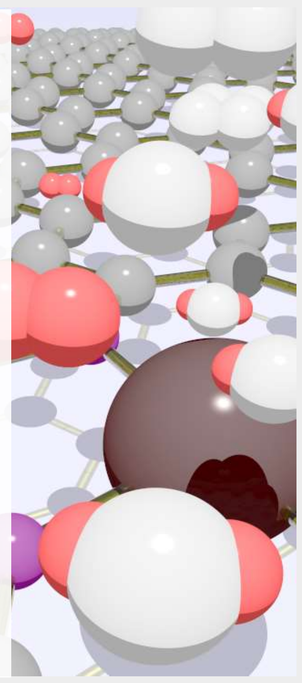
- Reconstruct object from diffraction data
- Phase is lost upon detection
- Counter by oversampling with varying illumination
- Typically scanning probe (translation)

- Large scale optimization with billions of measurements and millions of free variables

Applications of ptychography

Help solve major challenges facing society, science and the economy, including energy, earth & environment, health, matter, key technologies

- Visualize nitrogen ligands within graphene layers
  - Fix transition metals such as Ni, Co and Fe.
  - Catalyze chemical reduction of CO_2
  - Convert to hydrocarbons
  - Challenge: TEM imaging of individual atoms in dose-sensitive graphene
- Imaging bacterial infection of a host cell
  - Cell death caused by pathogen *Shigella*
  - Visualize morphological features and changes
  - Challenge: Dose sensitive, thick and complex sample, low contrast, high resolution



Challenges

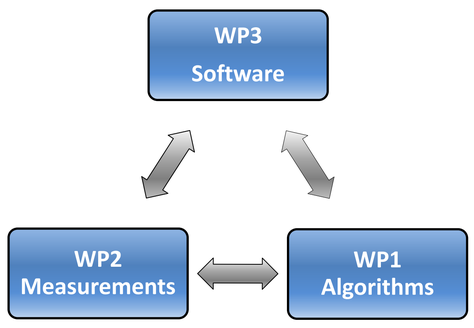
- Computationally intensive
- Variety of imaging conditions and reconstruction algorithms
- Currently experimental for many applications
- Ptychography data sets larger than RAM (GPU in particular)
- Adapt algorithms to “divide and conquer” approach
- Fast I/O and distribution of data: GB/s
- Solutions independent of computing hardware required
- Online reconstruction + visualization

Goal: Turn ptychography into a robust mainstream image processing and sample reconstruction method for interactive workflows

Cross center and cross field collaboration

The Helmholtz Information and Data Science Pilot Project brings together mathematicians developing new algorithms, experimetalists perfoming ptychographical measurements and computer scientists providing fast data management and efficient implementations of the algorithms.

Involved centers:
- HMGU
- DESY
- FZJ
- GSI (HIJ)
- HZB
- HZI
- HZDR

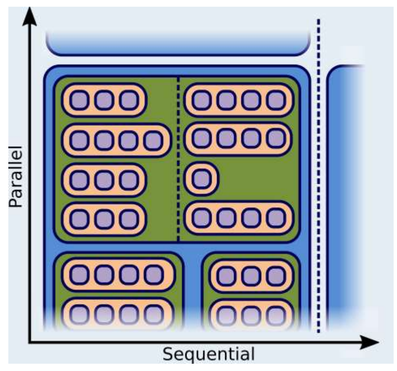


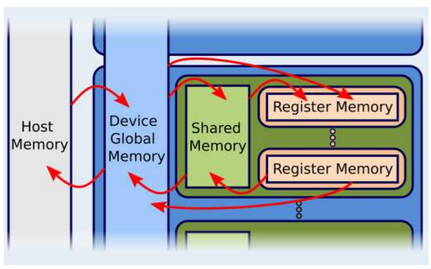
Efficient and heterogeneous computing with Alpaka

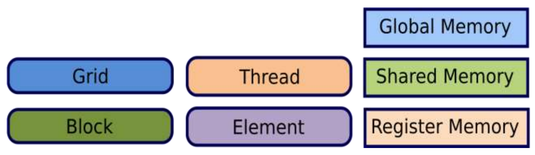
Goals:
- Modular system of reconstruction methods
- Scaling with arbitrary available hardware (GPU, CPU) using Alpaka

https://github.com/ComputationalRadiationPhysics/alpaka

Alpaka: C++ template-based abstraction library for platform independent code:
- Four nested parallel execution layers
- Each with dedicated memory layer
- Explicit copy between layers
- Mapped to existing programming models

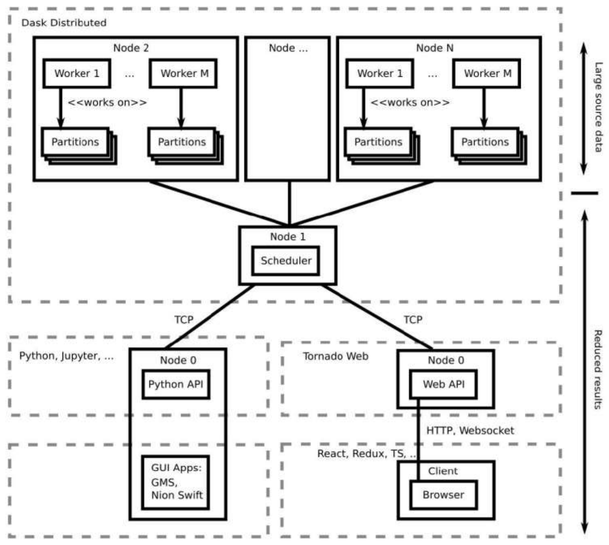






Manage computation with LiberTEM

https://libertem.github.io/LiberTEM/



- Data logistics for scientific algorithms
- Separate scientific algorithms from system and implementation details
- MapReduce-like approach:
  - Sub-divide input data
  - Distribute computation
  - Merge partial results

- Support many file formats
- Interfacing with applications
- Optimized I/O
- Python-based
- Designed for live data

JÜLICH
HZDR HELMHOLTZ ZENTRUM DRESDEN ROSSENDORF
HZB Helmholtz Zentrum Berlin