Data flows and processing at the LCLS MeV-UED facility: challenges and opportunities

AUREIS mini-workshop on data-flows for Intelligent Sensing

Alex Reid, Fuhao Ji, Yusong Liu/ LCLS MeV-UED

27 Jun 2025

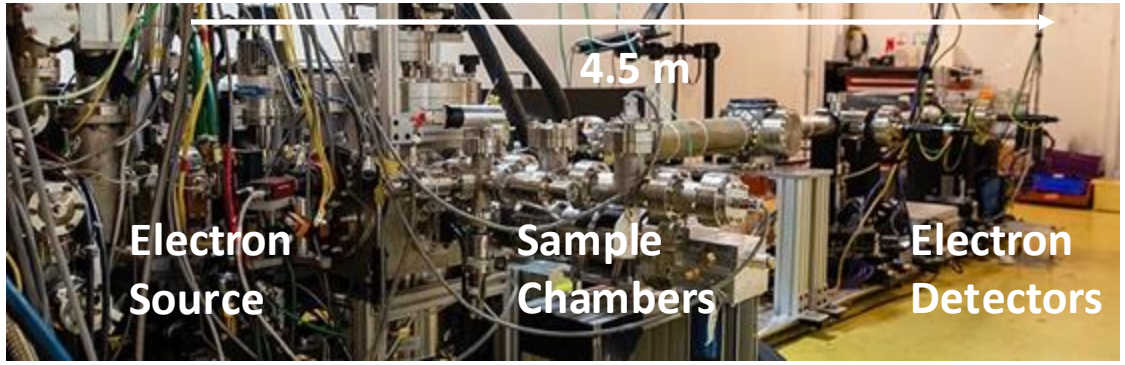
SLAC NATIONAL ACCELERATOR LABORATORY

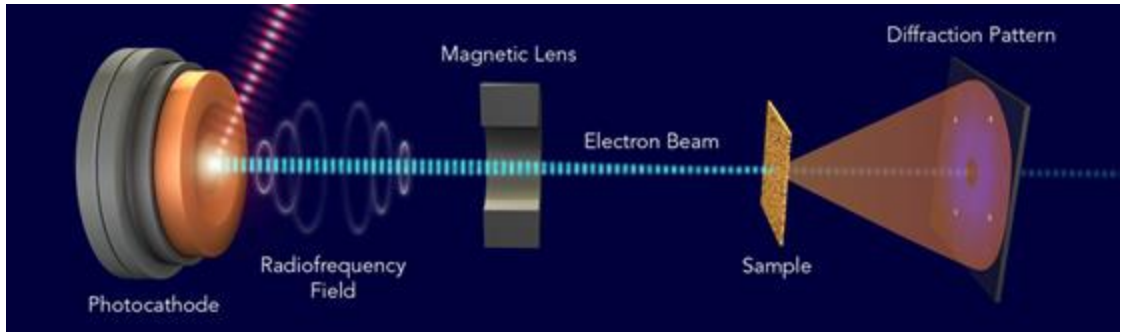
Stanford University | U.S. DEPARTMENT OF ENERGY

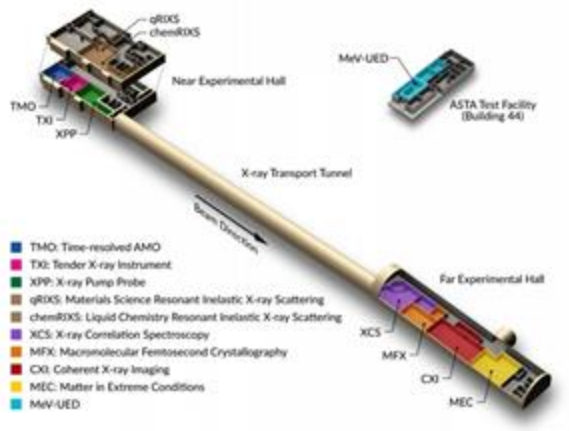
Outlines

- SLAC MeV-UED Overview
- Intelligent scientific user facility R&D at SLAC MeV-UED
- UED detectors and data flows
- Challenges and opportunities of UED data processing
- Summary

SLAC MeV-ultrafast electron diffraction facility







Key parameters

- 2-4 MeV electrons
- Pulses – single shot/360 Hz/1080 Hz
- 10^4 - 10^6 electrons / pulse
  - Measured diffraction <10%
- Time resolution ~150 fs

- MeV-UED is complementary to time-resolved x-ray science
- Program began in 2014, became an LCLS user facility in 2019
- Probe structural and electronic dynamics in solid, gas and liquid systems under optical pumps and operando excitations

- Independent operation from LCLS
- LCLS data structure and detectors
- Lower access threshold
- Relatively simple ‘standard’ experiment

S P Weathersby et al, *RSI* 86.7 (2015).
X Shen et al, *Ultramicroscopy* 184 (2018): 172-176.
F Ji et al, *Nat. Comm* 15, 4726 (2024).

LCLS MeV-UED Science Program & Data

Solid phase

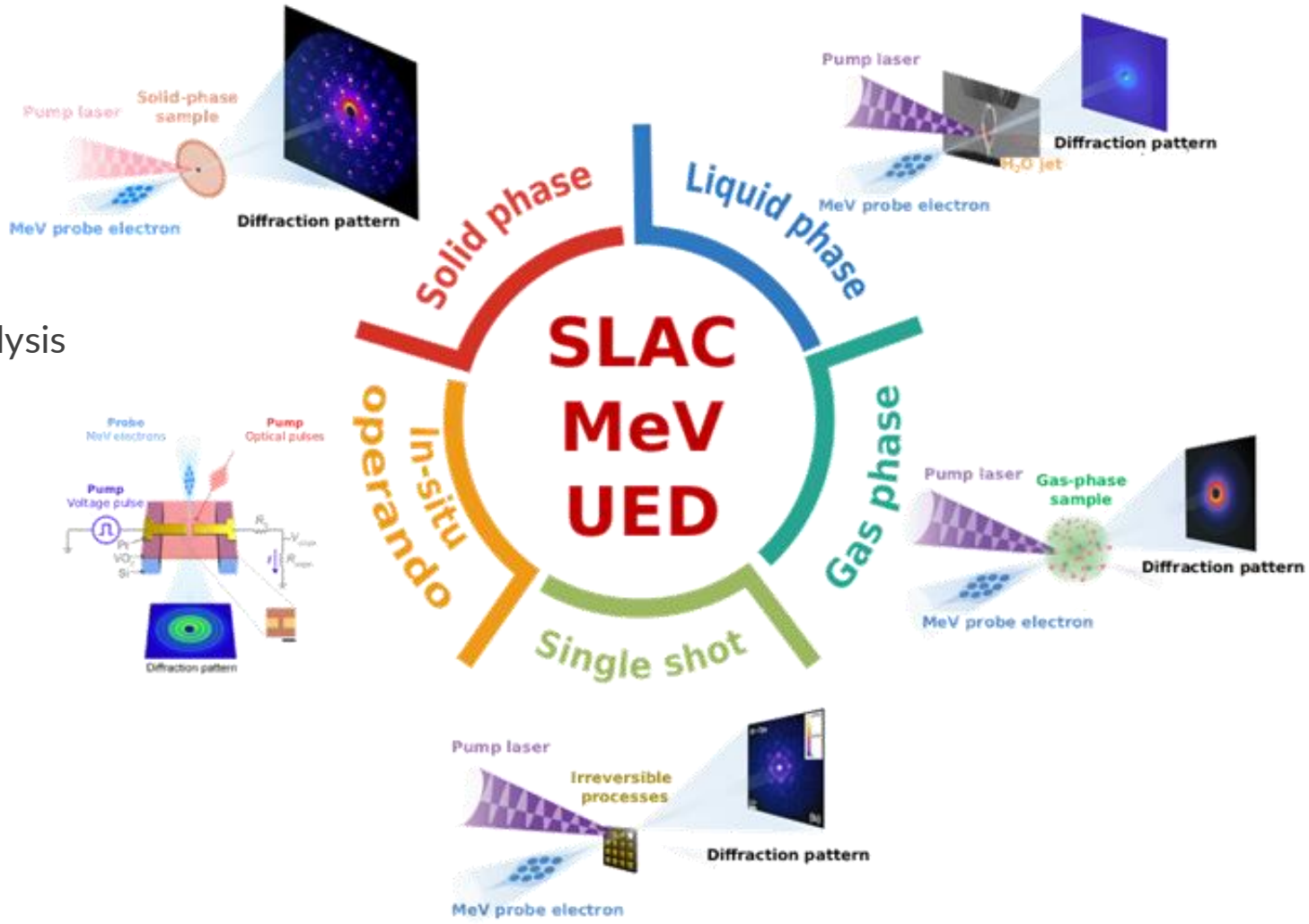
- static & dynamic crystal structure
- high dynamic range data
- user/material defined data analysis

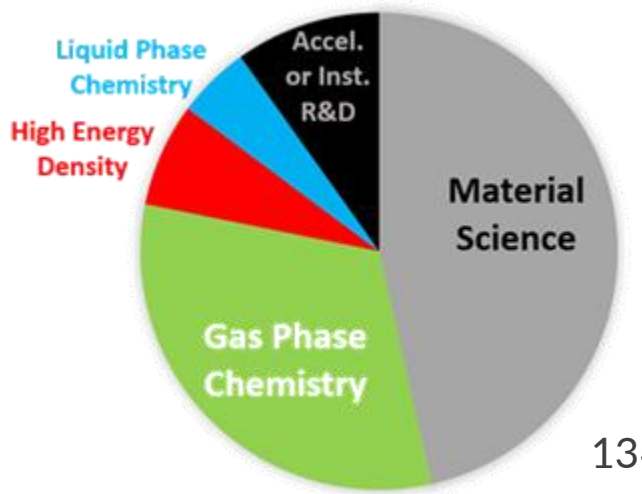
Gas-phase photochemistry

- Atomic pair distribution function analysis
- Long accumulations for SNR

Liquid-phase photo & Single shot

- Mix of short and long range orders





134 Publications

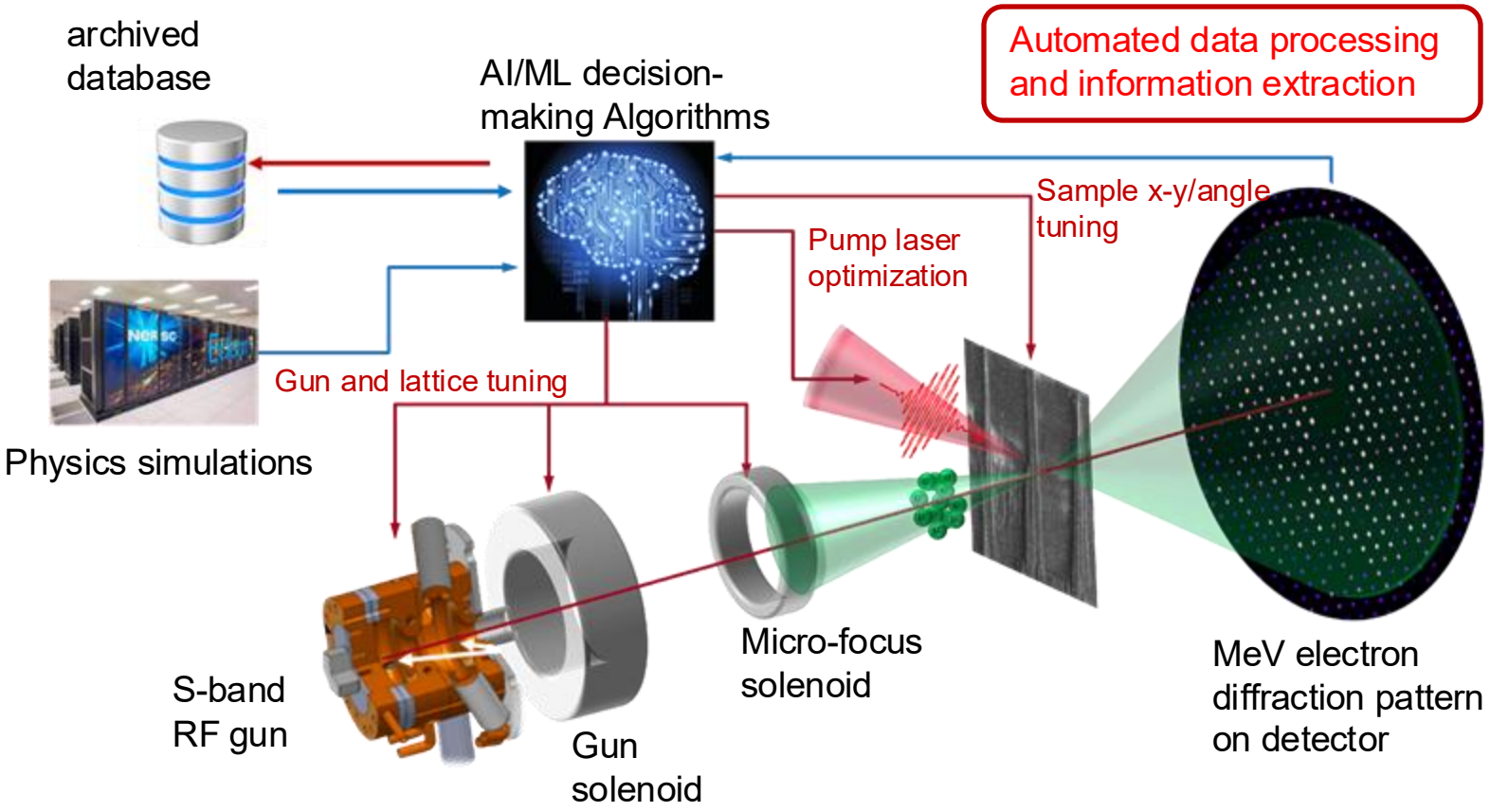
MeV-UED Data flows Opportunities and Challenges

Fuhao Ji

SLAC NATIONAL ACCELERATOR LABORATORY

Stanford University | U.S. DEPARTMENT OF ENERGY

Intelligent scientific user facility R&D at MeV-UED



- AI/ML algorithms in combination with human-in-the-loop methodologies for real-time facility tuning and autonomous experimentation
- Intelligent scientific user facility for maximizing scientific outputs of MeV-UED facilities
- Automated, high-through data processing plays a key role, currently remains a challenge at MeV-UED

Main Detectors Data Steam and Production

Andor EMCCD, Integrated Shots

From DAQ to User Friendly Data Format

LCLS-II DAQ .xtc

UED Local Storage drp-ued-cmp001/2 → **Data Mover**

Data transfer

LCLS-II DAQ

srcf_ffb (.xtc) '/sdf/data/lcls/drpsrcf/ffb/ued/.... '

Data Mover

s3df (.xtc) /sdf/data/lcls/ds/ued/uedxxxxxx/xtc

User Offline Analysis

IoC/DAQ

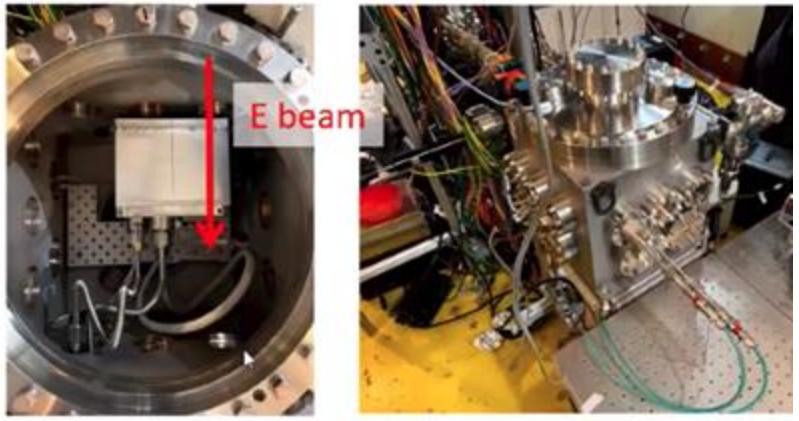
ePix10k, Single-shot

Large Data
2.35MB*360Hz = 0.83GB/Sec

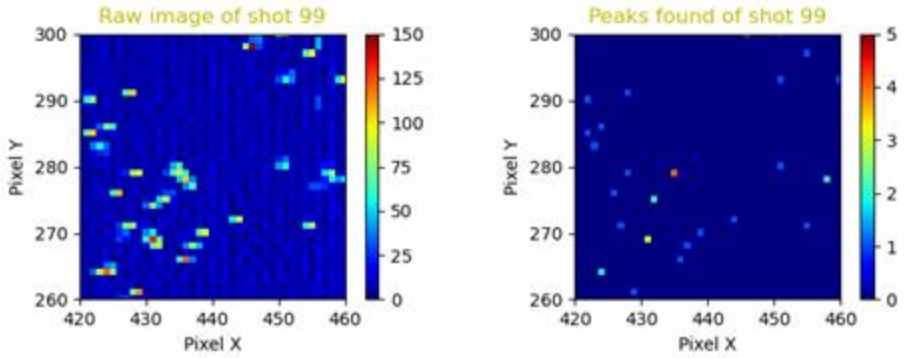
Preprocessing (.hdf5) Standard (smd, cube,..) Manual (peakFinding..)

s3df (.hdf5) /sdf/data/lcls/ds/ued/uedxxxxxx/hdf5/smalldata

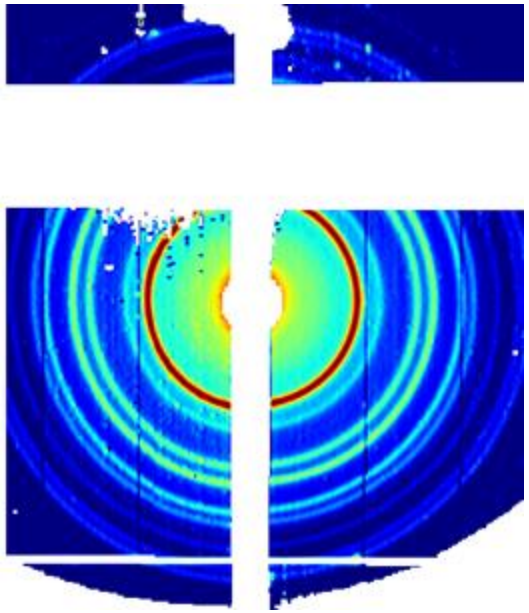
LCLS ePix10k as shot-by-shot single electron detector



Single shot electron counting



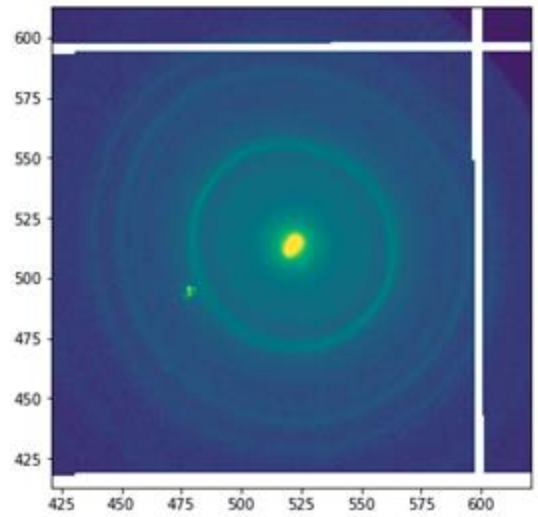
Integrated diffraction pattern

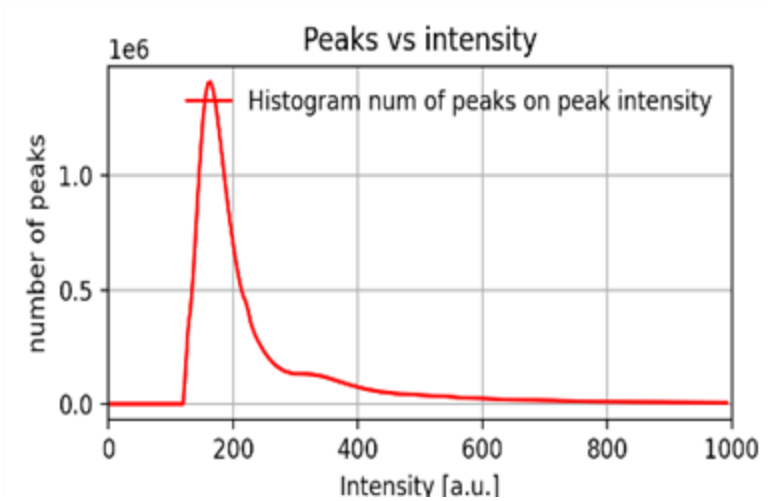


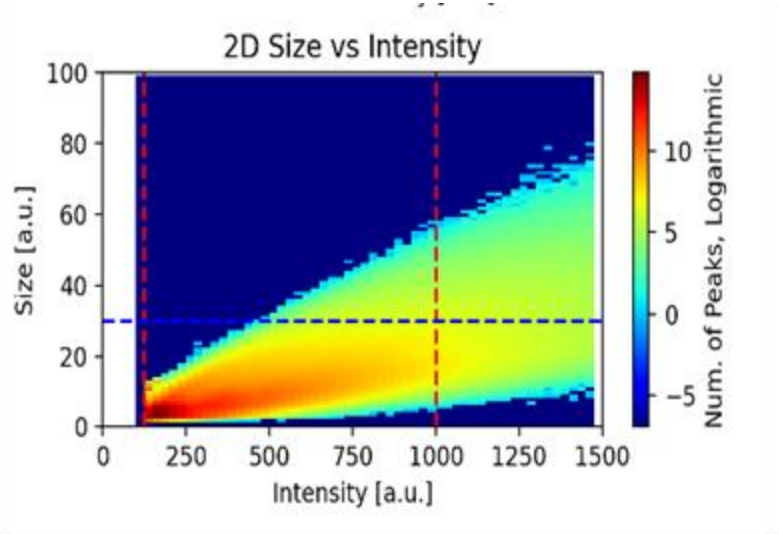
- Shot-by-shot detection up to 360 Hz repetition rate (1080 Hz for the 2nd generation detector)
- Single electron counting capability
- 3 fixed gain modes (Low, Med, High) and two auto-ranging modes
- Large raw data volume (50 - 100 TB/beamtime)
- Computationally intensive, typically > 300 cores required for online processing at 360 Hz data rate
- Time overheads in data transfer and cluster queuing results in further delays of processing time (~ 30 min)

LCLS ePix10k as shot-by-shot single electron detector

Polycrystalline Bi diffraction with main beam

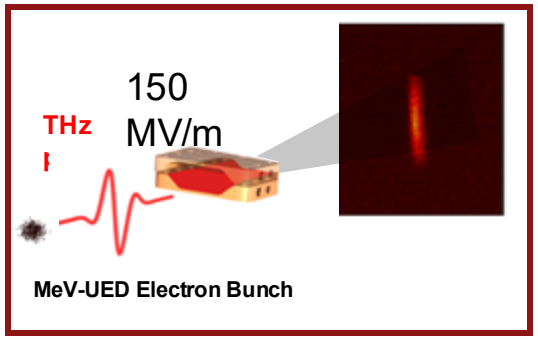






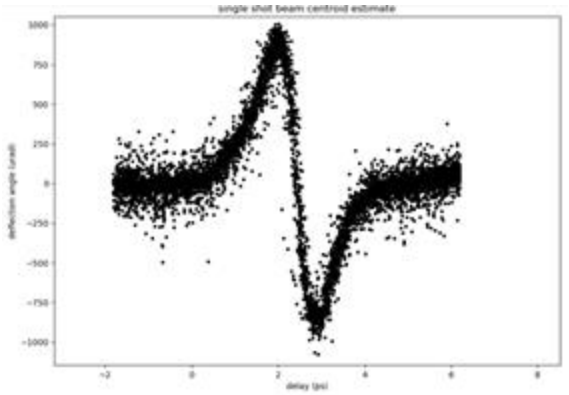
- High dynamic range data requires on-the-fly selections of most suitable detector gain settings based on experimental conditions, typically requires human inputs
- Electron counting algorithm requires human inputs for identifying suitable peak finding parameters under different experimental conditions.
- Long-tail problem: well understanding of single electron case, less understanding of multiple electron case, large data space
- Currently, real-time processing of ePix data for online experimental feedback remains a challenge

Shot-by-shot time stamping

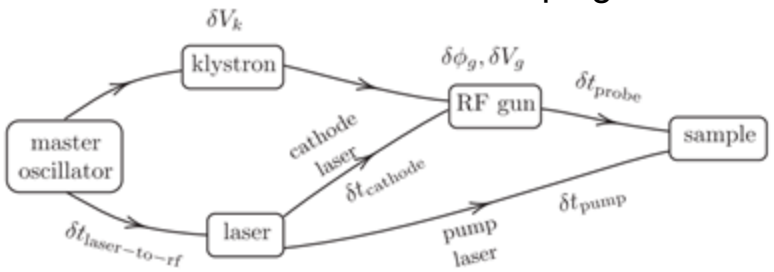


M. Othman, et al., Efficient THz Time Stamping of Ultrafast Electron Probes, 47th IRMMW-THz Conference (2022).

10K shots (30s data) takes 15 min to process



Virtual time stamping



$$dt_{\text{tot,g}}(d\phi_{i,g}, d\alpha_g) = dt_{\text{cav,g}} + dt_{\text{drift,g}}$$

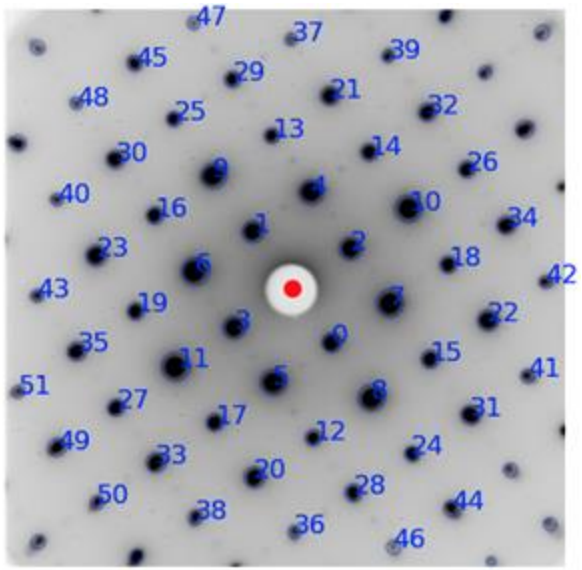
$$dt_{\text{cav,g}} = \frac{1}{\omega}(d\phi_{f,g} - d\phi_{i,g}),$$

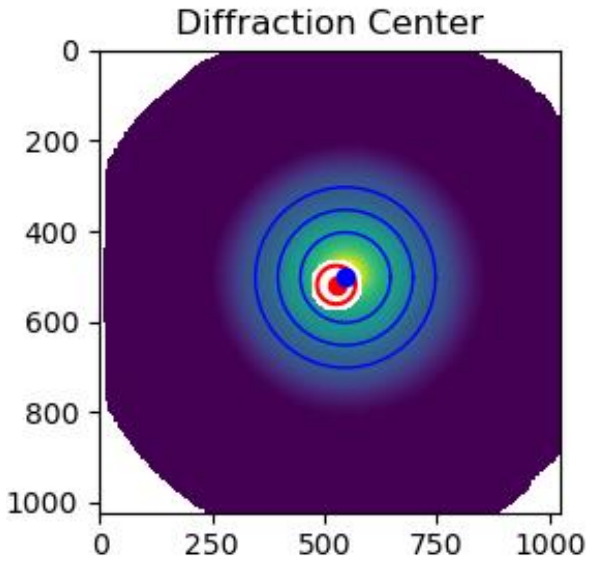
$$dt_{\text{drift,g}} = -\frac{L_g}{c\tilde{\beta}_{f,g}{}^3\tilde{\gamma}_{f,g}{}^3} d\gamma_{f,g},$$

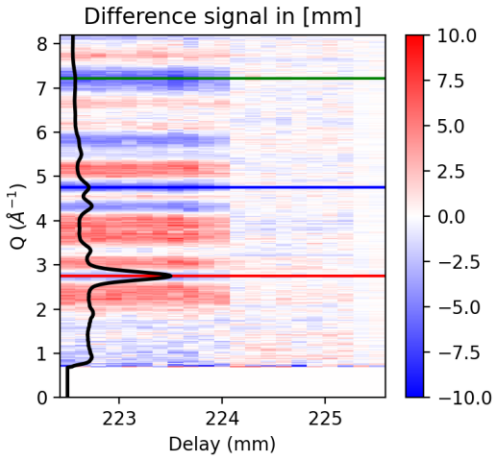
- Deployment of THz timing tool would allow shot-by-shot time-of-arrival correction and improving the overall temporal resolution to < 50 fs FWHM
- Streaked electron beam data is generated at 360-1080 Hz and put further pressure on online data processing
- The concept of virtual diagnostics, a large amount of data be compressed to a single parameter for storage or feedback
- High bandwidth electrical signals taken with 100 us windows where the data is <1 us (and vice versa)

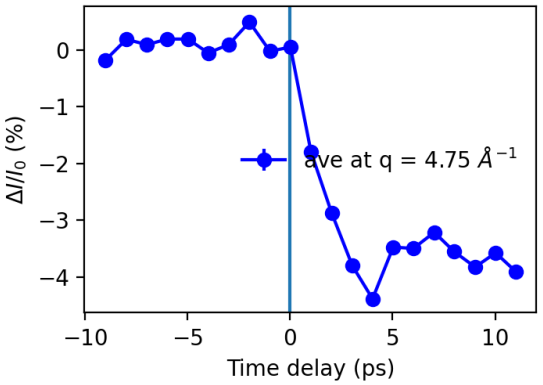
Phys. Rev. Accel. Beams 28, 024001 (2025)

Electron scattering pattern processing and information extraction









- Robust data dimension reduction: Bragg peak identification and diffraction centroid finding
- Time-zero identification and extraction of Debye-Waller time constants
- Real-time space group and orientation determination
- Extraction of weak signals including long-range order parameters and low-energy phonon excitations

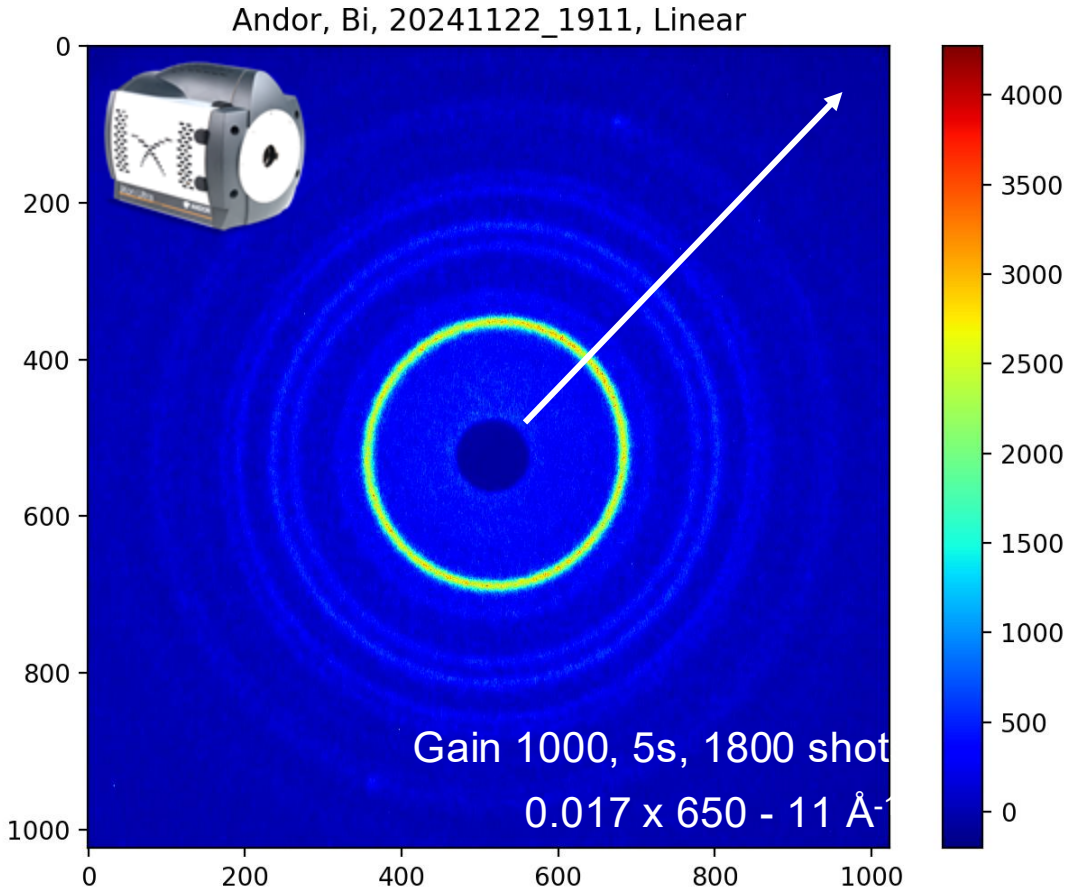
Summary

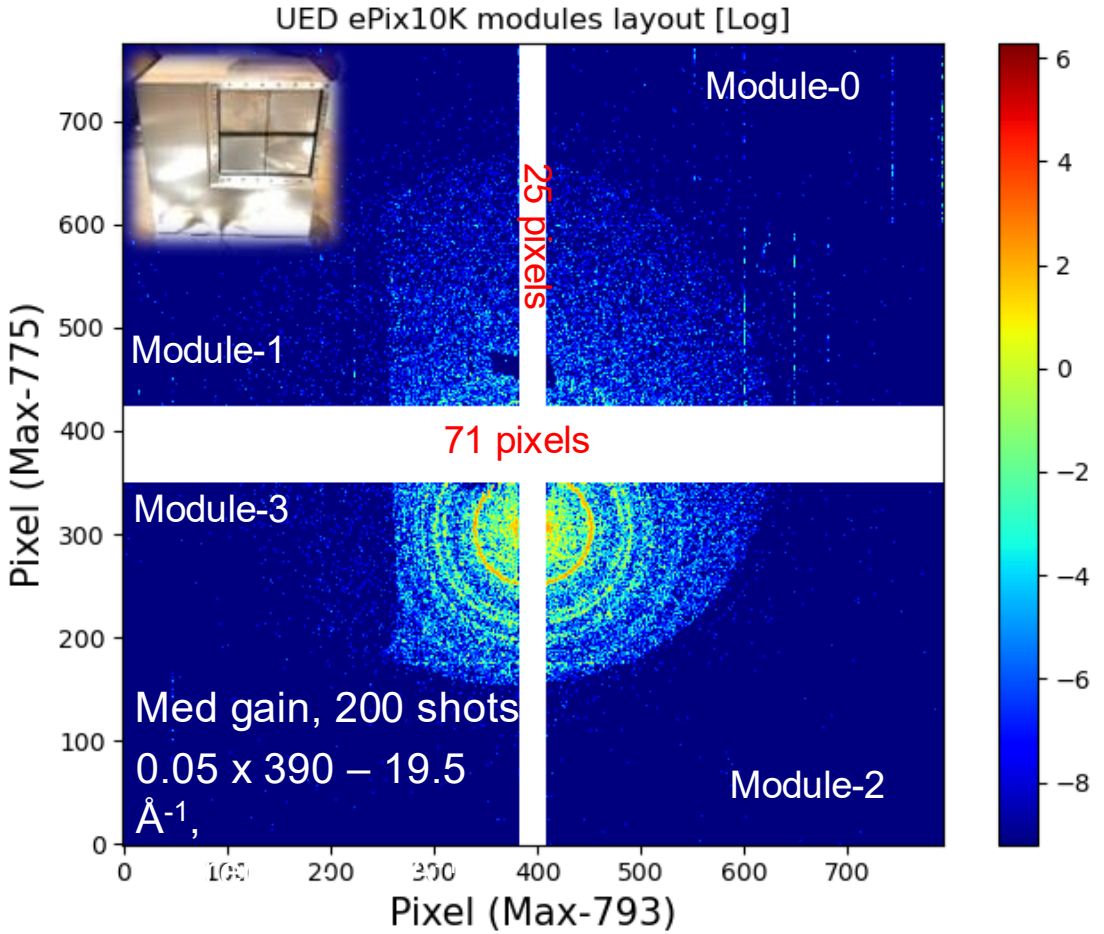
- MeV-UED is complementary to time-resolved x-ray science and holds potential of enabling sciences in key areas
- The deployment of ePix10k as single electron detector and THz timing tool not only brings tremendous scientific opportunities, but also pose significant challenges for real-time processing of data generated at high repetition-rate
- Adaptive/intelligent sensing and high throughput data processing at edge would play a key role in the intelligent UED SUF and ultimately enable large-scale, high throughput scientific discoveries at MeV-UED facilities

Back up slides

Andor vs ePix Diffraction Measurement

Andor vs ePix data





Module: 352 x 384 (V x H)
Frame: 775 x 793 (V x H)