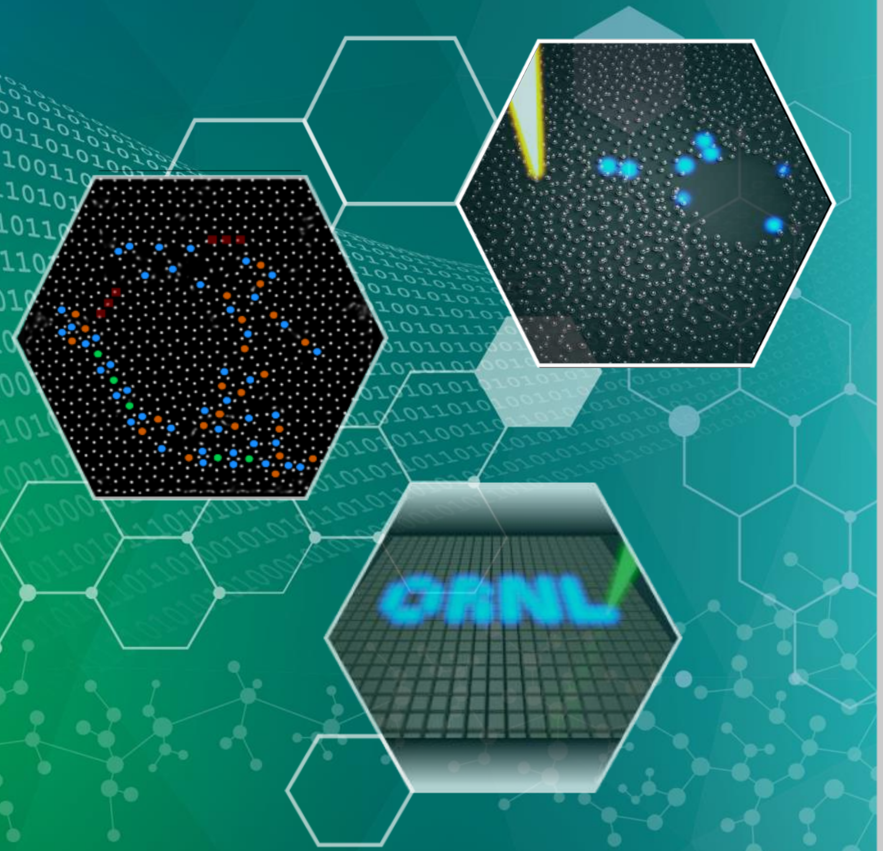
Lecture XX

Deep kernel learning in EELS and 4D-STEM

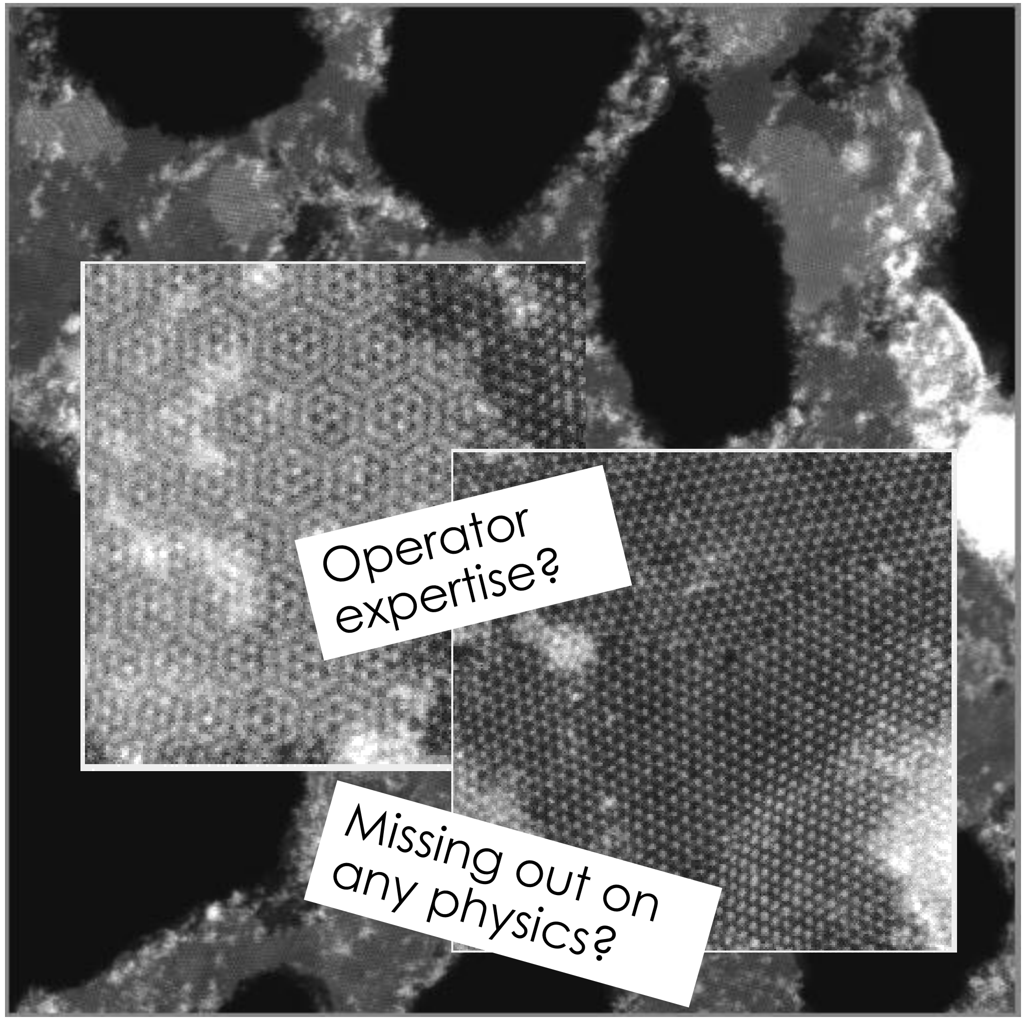


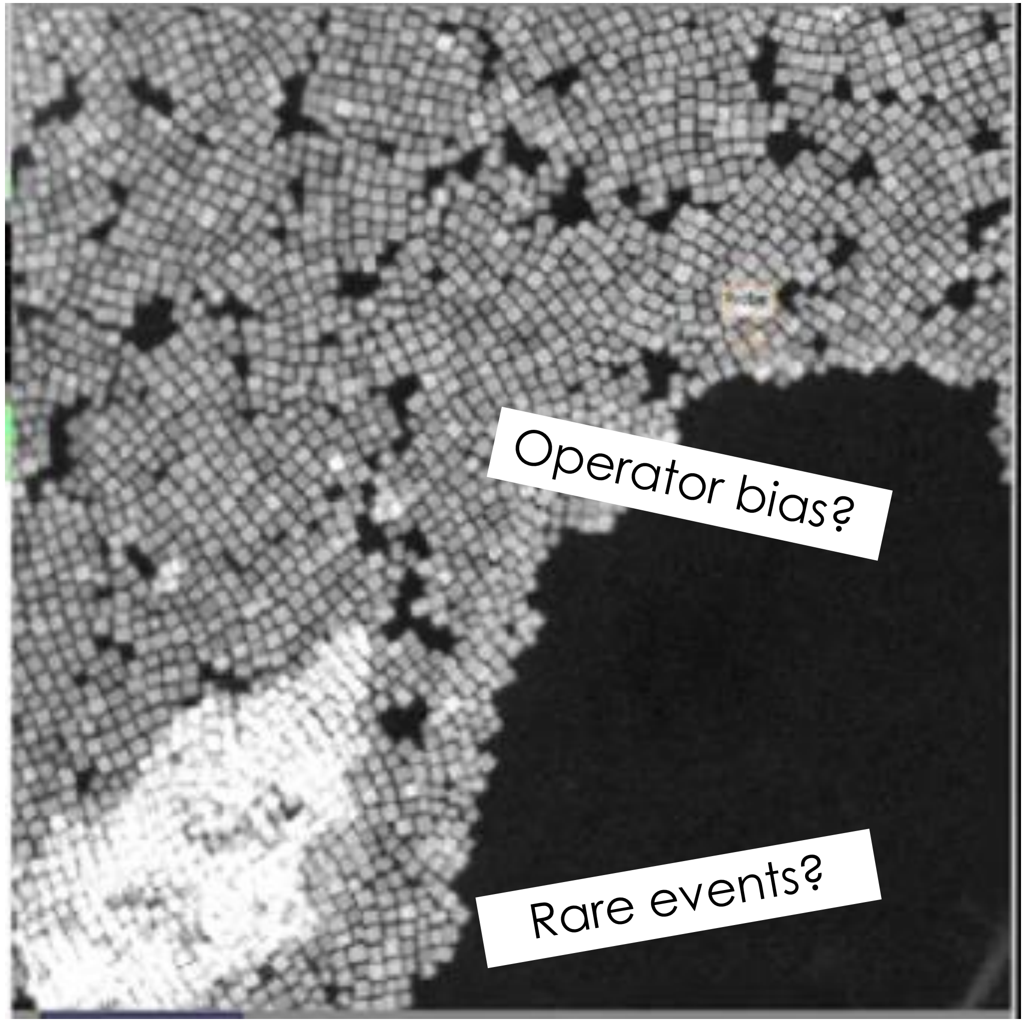
Kevin M. Roccapriore, Maxim Ziatdinov, Sergei Kalinin

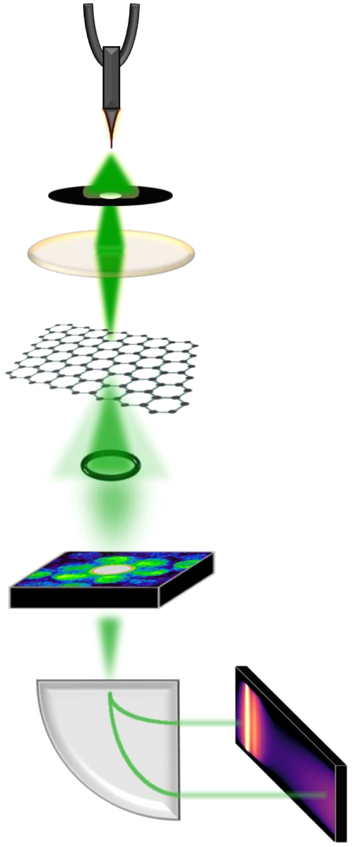
August 11, 2023

Center for Nanophase Materials Sciences
Oak Ridge National Laboratory

How do we understand structure property relationships

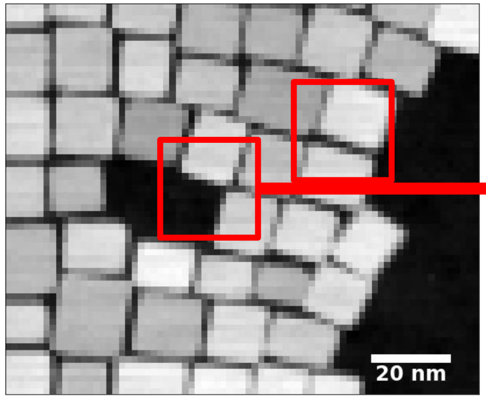


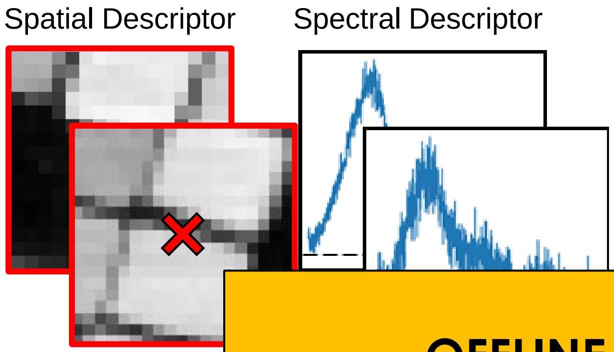


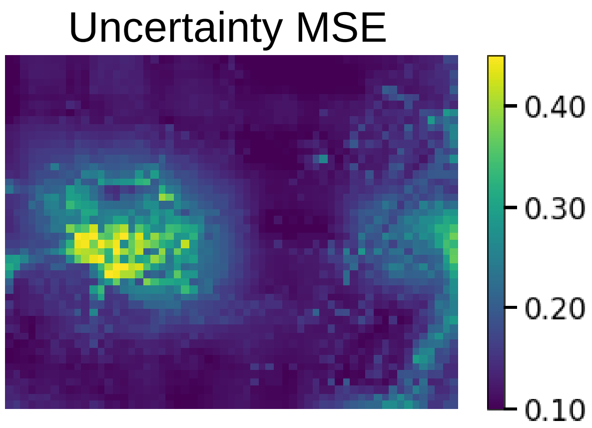


Q: Where should (would) you acquire analytical measurements? EELS spectrum image / 4D STEM, point spectra, where? **You can't say all points!**

Recall:
Structure-property correlations:
Autoencoder neural network





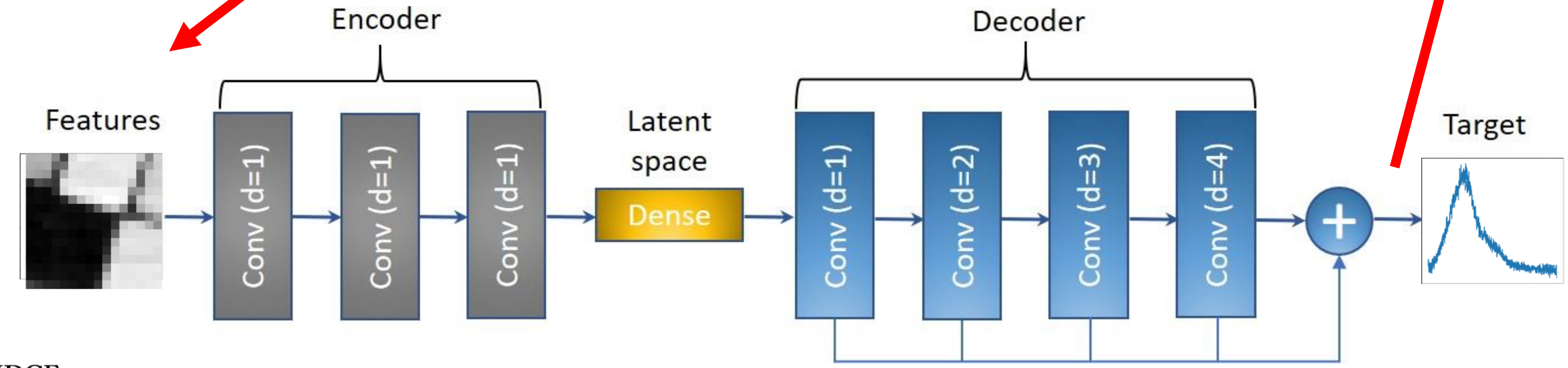


OFFLINE analysis!

Training can take some time.

Training

Evaluation



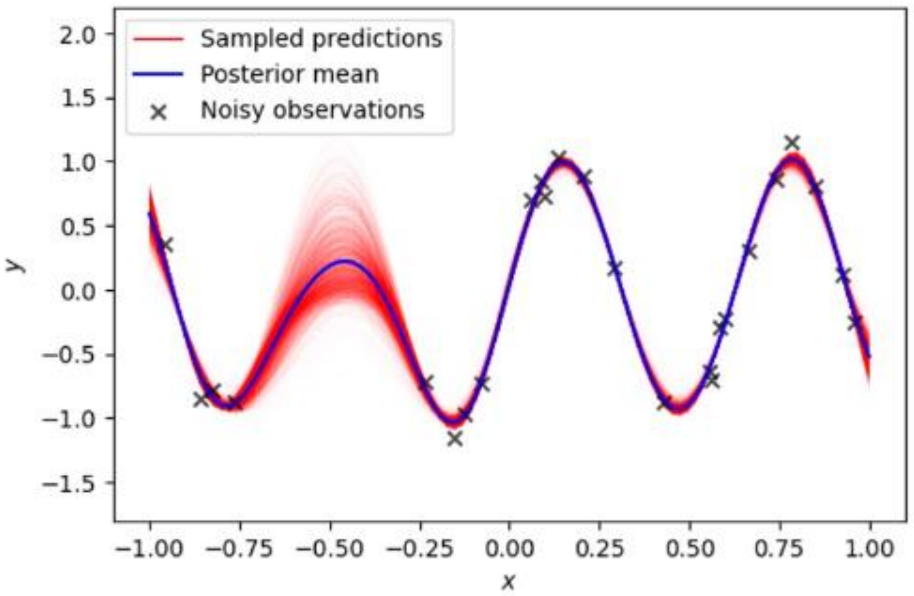
Automated Experiment

How to perform sparse measurements (intelligently) in real time?

- Initial attempts ➔
  1. use gaussian process (**GP**) to "interpolate" the measurement space,
  2. then decide next measurements with Bayesian Optimization (**BO**)

- Gaussian Process → requires a kernel for reconstruction.
-
  - ➢ What kernel do we use?
  - ➢ Kernel length?

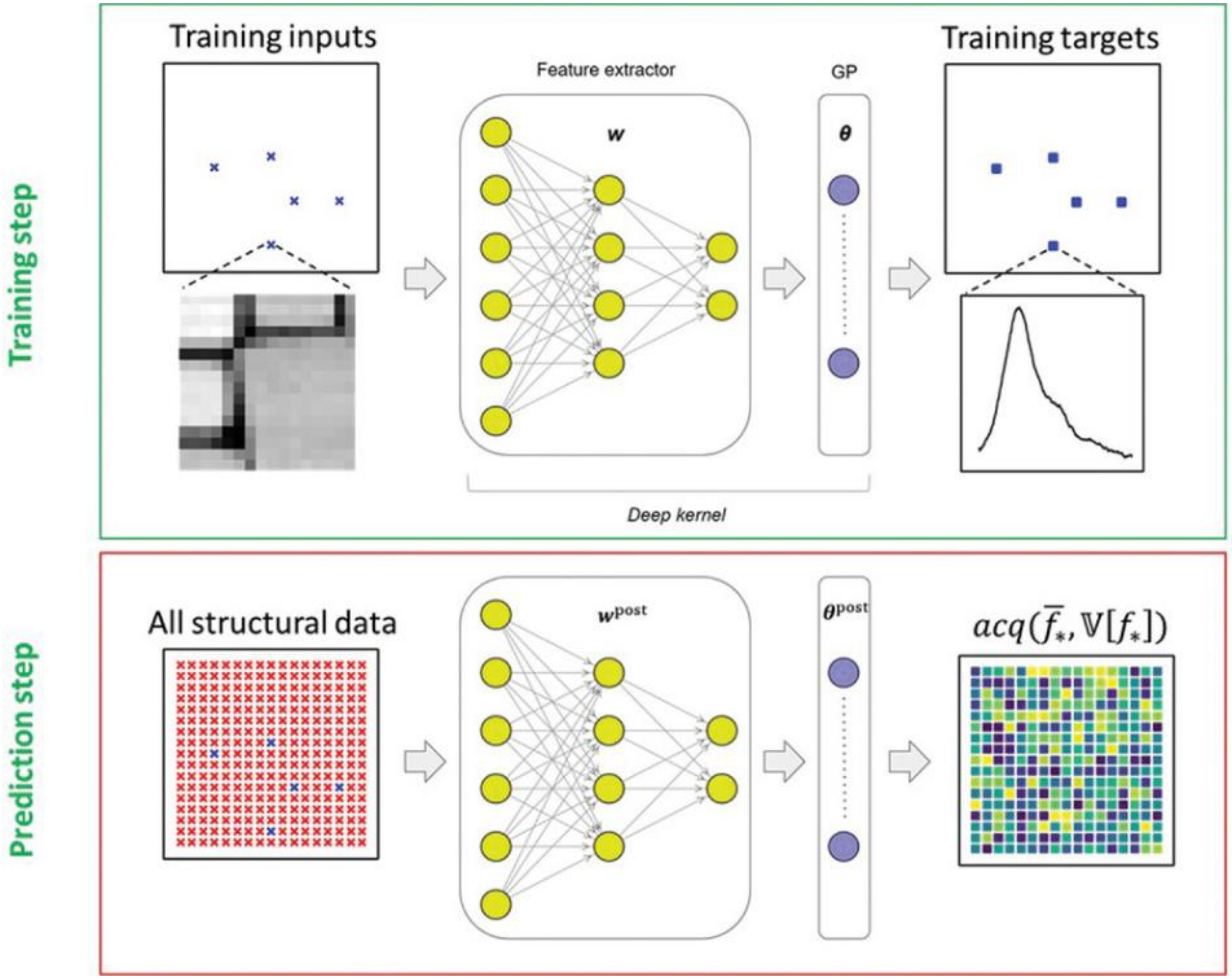
  - Periodic? Atoms… ok, lot of tuning… anything else?
  - Radial basis function (RBF) – maybe..
  - Custom kernel? But how to develop it?



If we could intelligently select a kernel… quickly!

See **lecture 14** for GP and BO details if you missed it

Deep kernel learning (DKL)



Tuning knobs

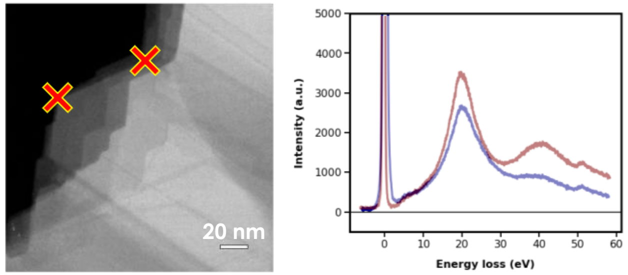
- "Scalarizer"
  - **Physics enters here**

- Acquisition function
  - How to navigate given the predicted space
  - Expected improvement
  - Upper confidence bound
  - Uncertainty-based
  - Thompson sampling

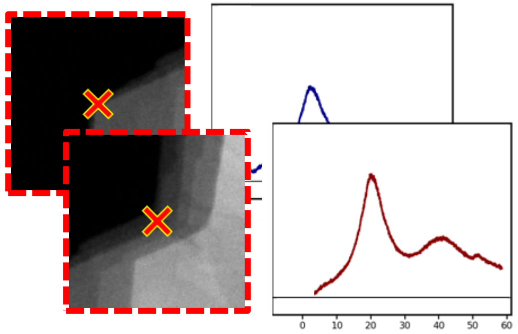
Roccapriore, K. M., Kalinin, S. V., Ziatdinov, M., **Physics Discovery in Nanoplasmonic Systems via Autonomous Experiments in Scanning Transmission Electron Microscopy**. *Adv. Sci.* **2022**, 9, 2203422.

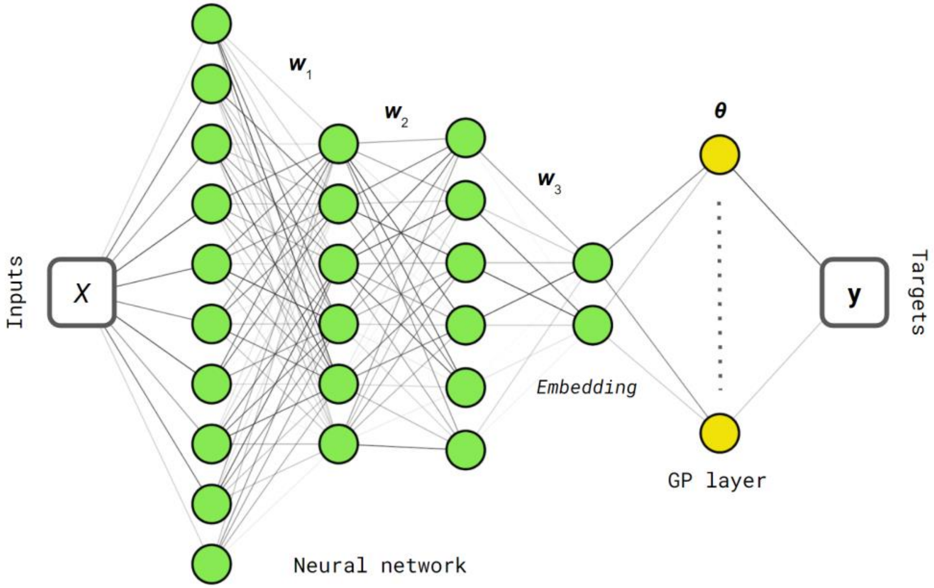
Physics-based feature engineering:
Deep kernel learning – Bayesian optimization

Specify physics criteria



Active learning





Acquire **structural data**

Measure a **spectrum**

Train DKL model with new data

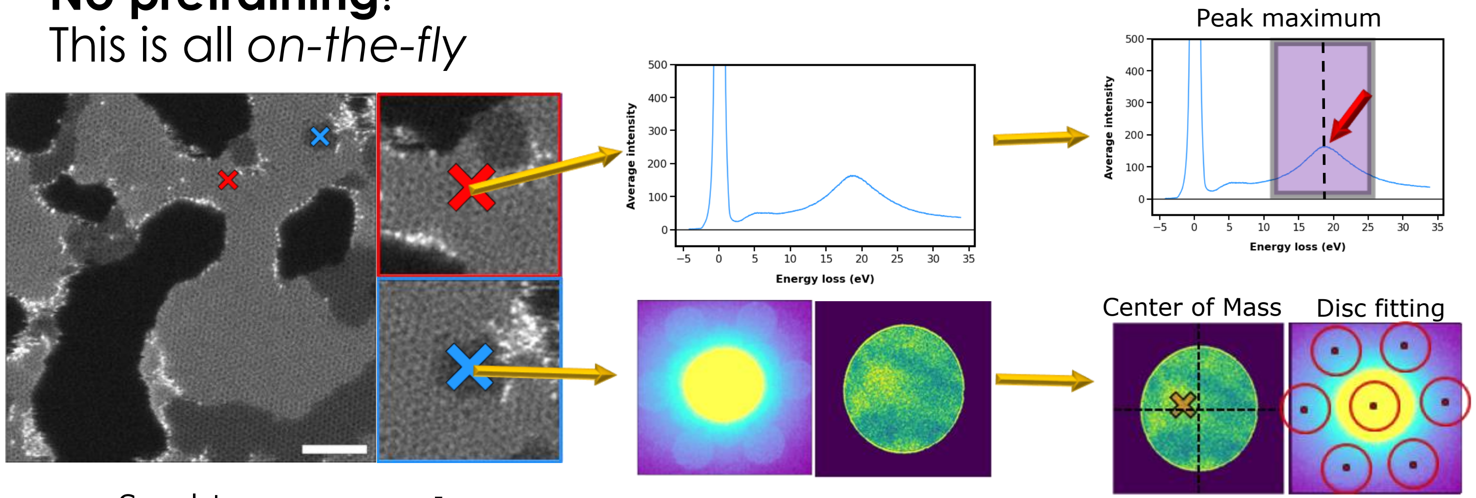
Decide next position (optimize physics criteria)

Allows navigation of the system to search for physics

OAK RIDGE National Laboratory

Actively **learning structure-property relationships: Deep kernel learning (DKL)**

No pretraining!
This is all *on-the-fly*

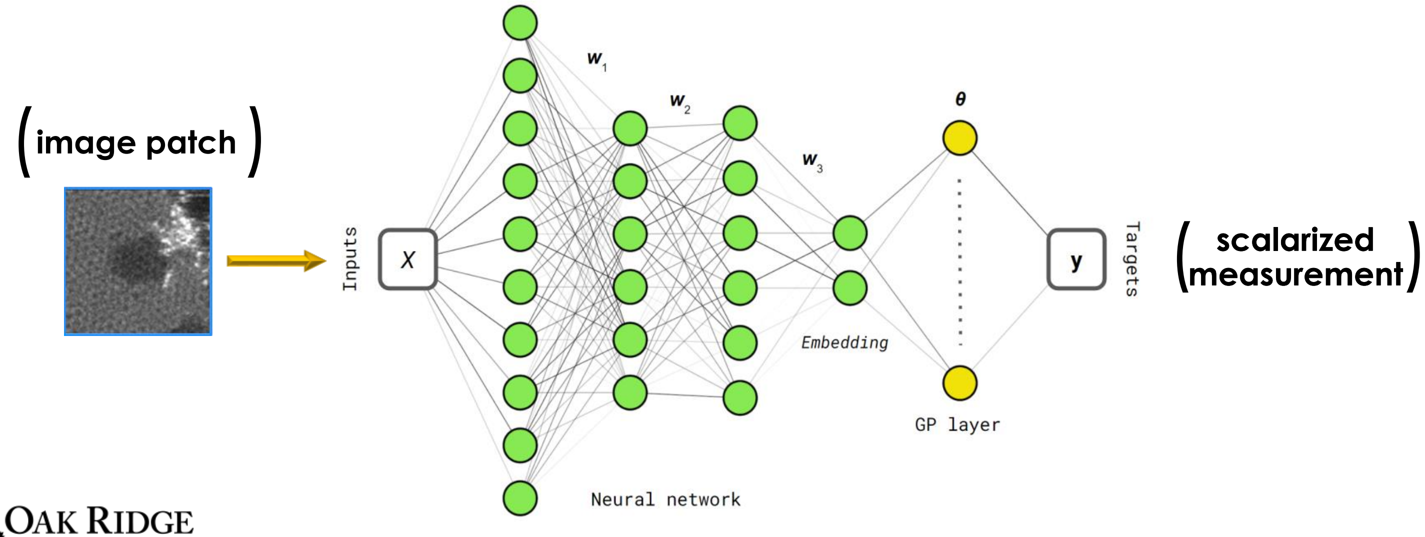


Complete **structural** image

Image patch

Measurement (1D or 2D)

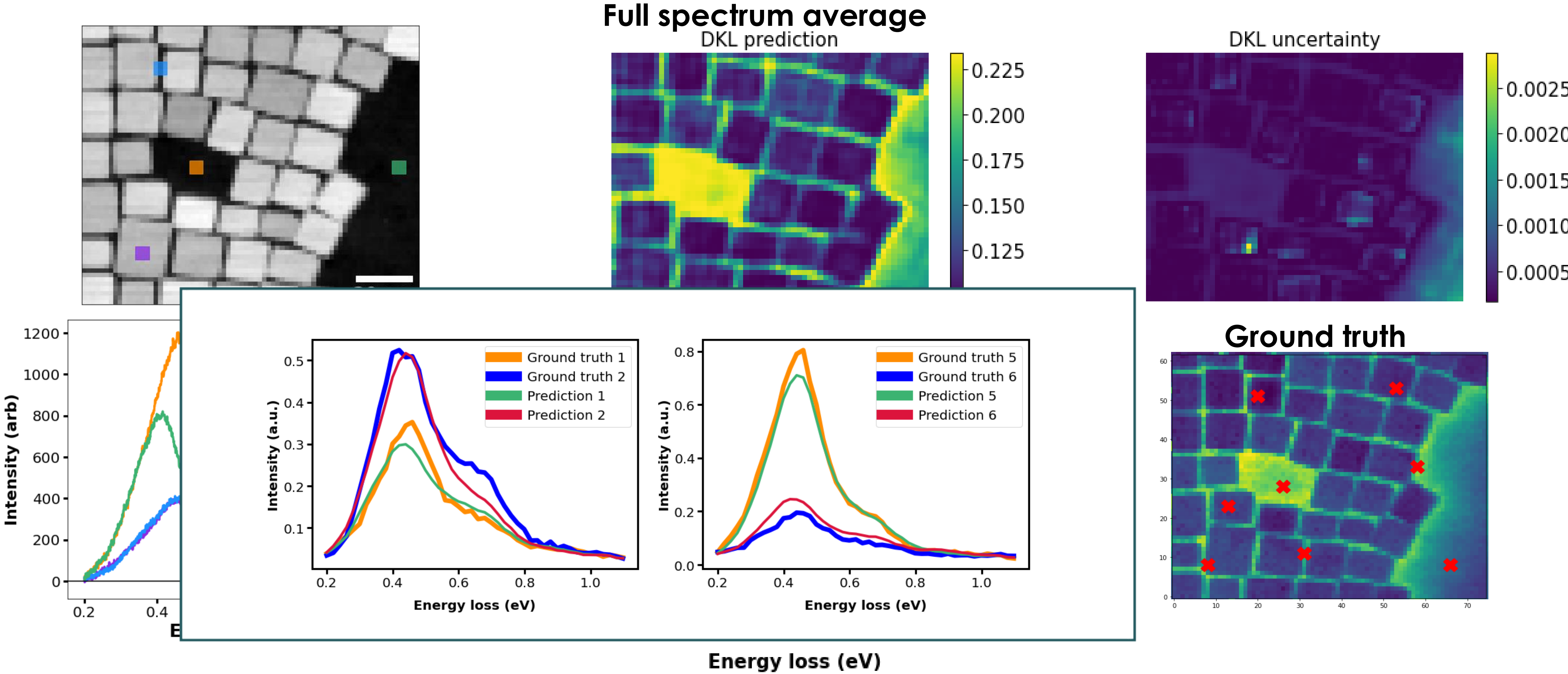
Scalarize (**physics enters here**)



1. A complete **structural image**
2. Image **patches**
3. Perform a **measurement** @ (x1,y1)
4. "**Scalarize**" (**physics** incorporation)
5. Learn **correlation** between image signal and measurement signal (training step)
6. **Next measurement** @ (x2,y2) based on learned correlation
7. Repeat

Oak Ridge National Laboratory

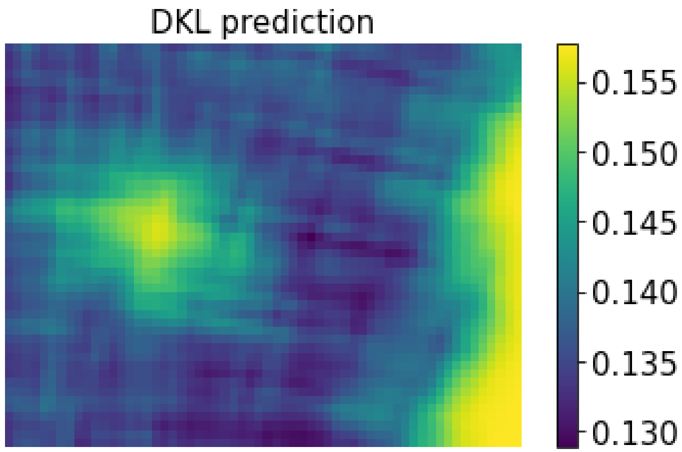
DKL with FULL pre-acquired data
(no active learning or BO)

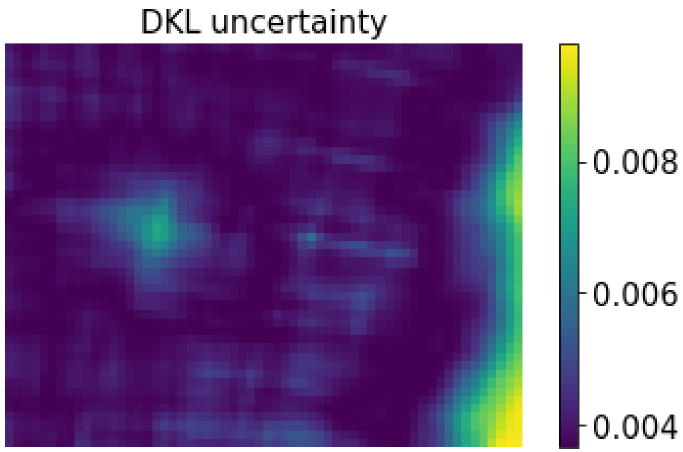


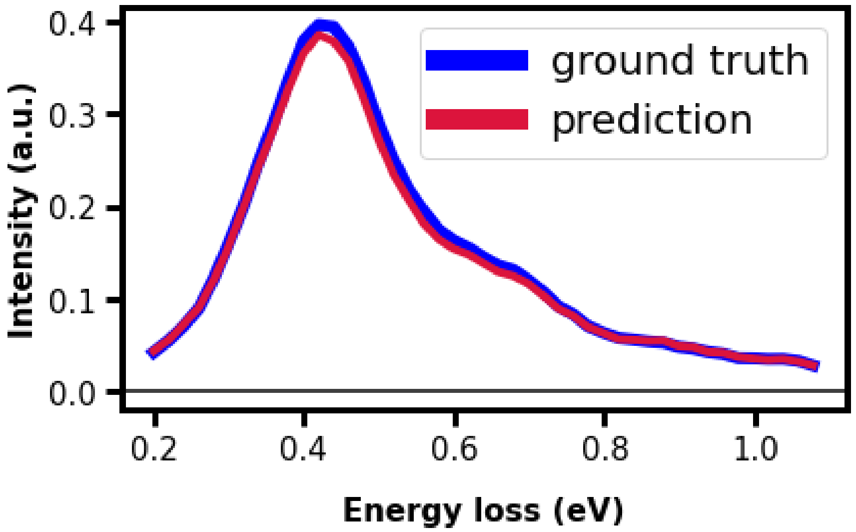
- This is actually "**vector**" DKL which **reconstructs entire 1D spectrum (or 2D image)**
- Is not so amenable for automated experiment, however

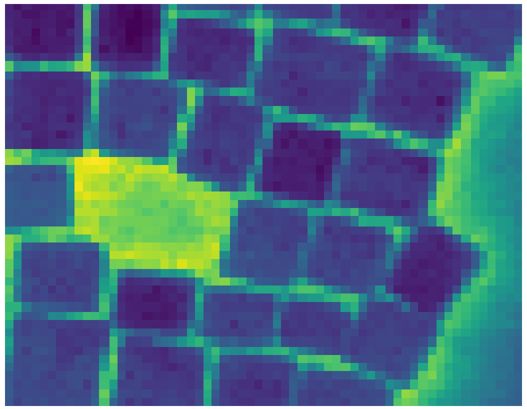
DKL with partial pre-acquired data
(no active learning or BO)

1% random sampling









Active learning with DKL still on pre-acquired data using a scalarizer

Scalarizers

- Peak max in spectral range
- Ratio of peaks (e.g., L2,3 ratio, surface plasmon / bulk plasmon)
- Number of peaks
- Peak width
- …..

- Center of mass (magnitude, angle…)
- Charge density
- Strain (diffraction disc separation in fourier space)
- …….

Acquisition functions

- Expected improvement
- Upper confidence bound
- Uncertainty-based
- Thompson sampling

➢ **Handles the next measurement choice**, based on the currently learned structure-property relationship (i.e., prediction)

➢ "exploration" vs "exploitation"

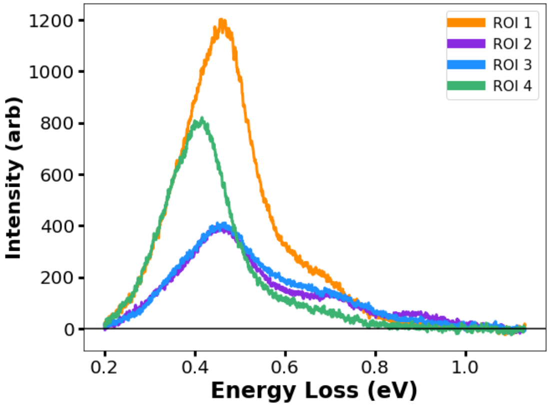
Active learning with DKL still on pre-acquired data using a scalarizer

Measure & scalarize

Train

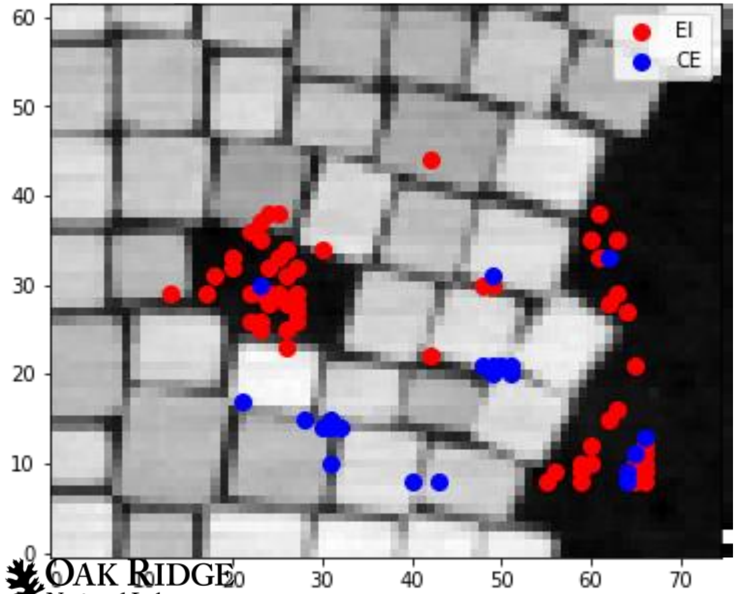
New measurement **based on A.F.**. Repeat

```python
def scalarize_acq(obj: np.ndarray) -> np.ndarray:
    """
    Scalarize acquisition function (n_samples x vector_size)
    by averaging over a selected energy band.
    """
    band = [0, -1]
    obj = obj[:, band[0]:band[1]].mean(-1)
    return obj
```



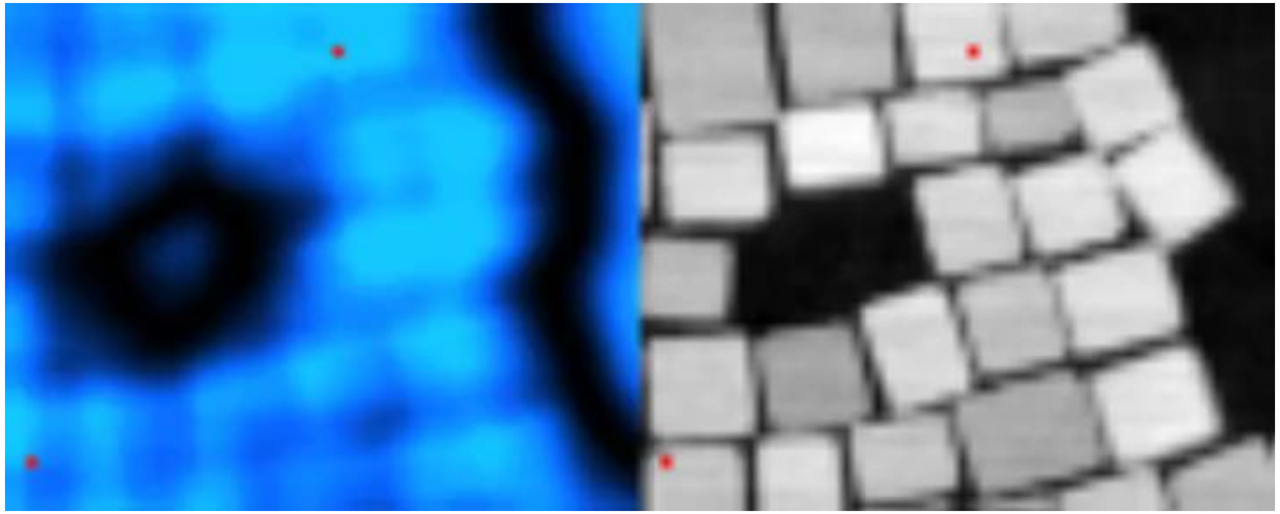
Next measurement point: $argmax()$

- Averaging whole (background subtracted) spectrum ➔ **favoring the strong dipole plasmon**



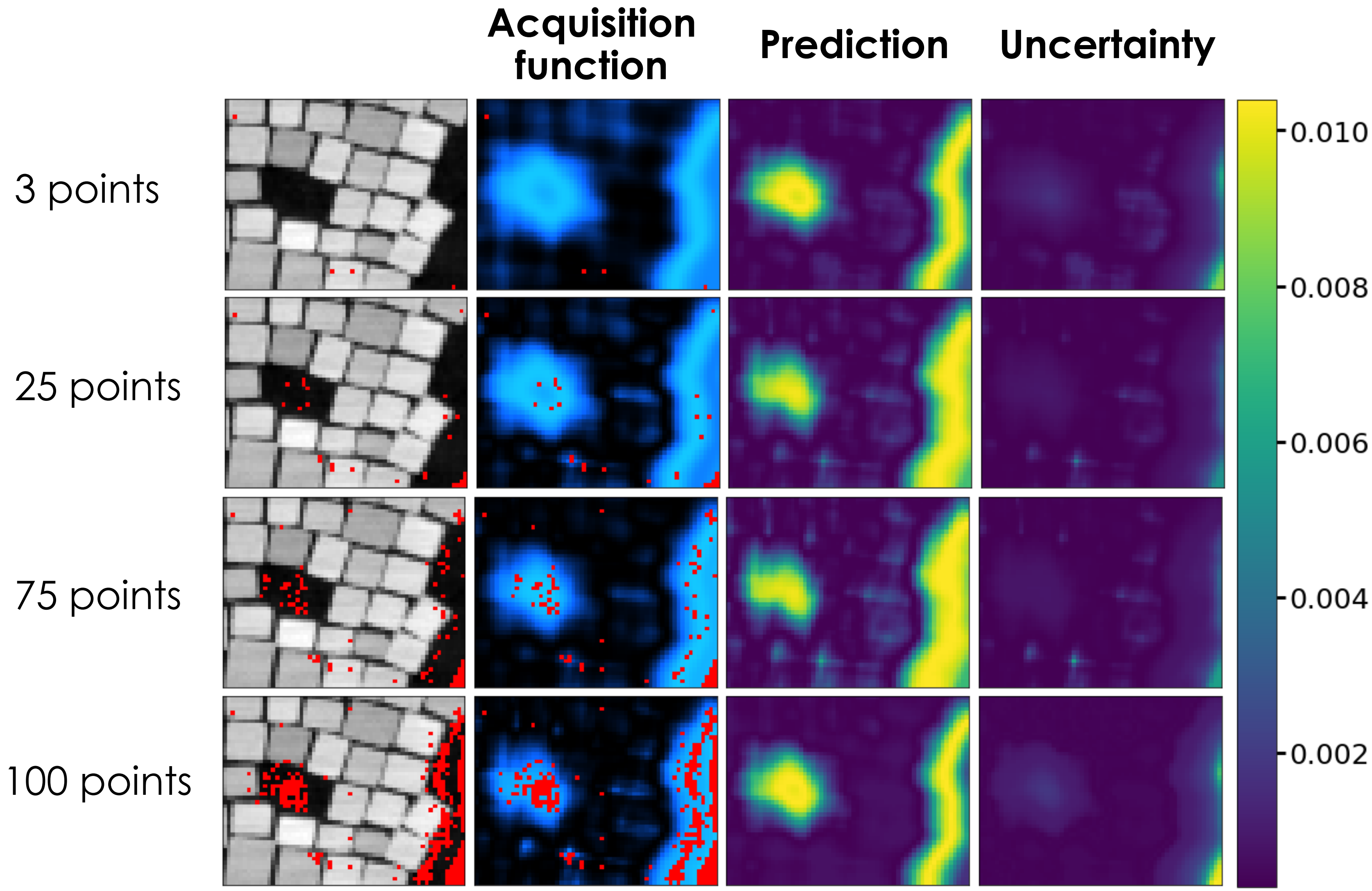
Choice of **acquisition function** affects pathway

Acquisition function



Evolution and performance in time

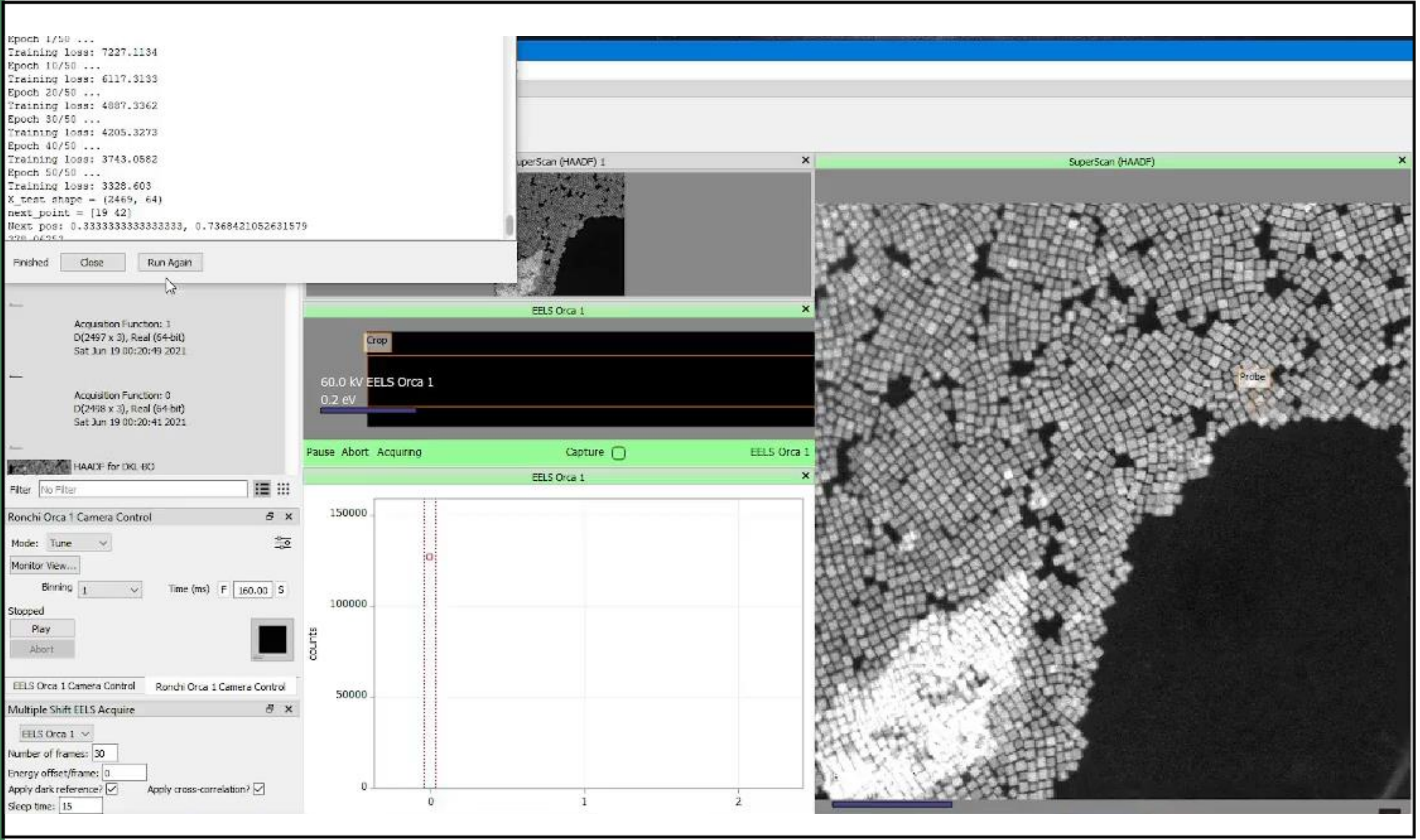
Have a look at predictions



- A good approximate structure-property relationship is **learned very quickly**
- In practice, we do not need to train at every new measurement
- Alternatively, can halt training after a few measurements (or some criterion met), then follow the acquisition function map
- Remember, **this is a prediction of the scalarized value** not the full spectrum

A real automated experiment

Can ML run experiment as a scientist?

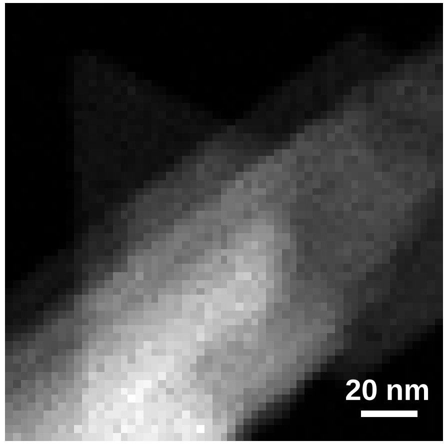


With **access to microscope controls** (Python API), this is deployed on an instrument

- **Continuously learning** structure-property relationships with each new measurement
- **Dose** reduction!
- In some sense, this is sparse sampling, but each subsequent point depends on the previous

Beam is blanked during training steps

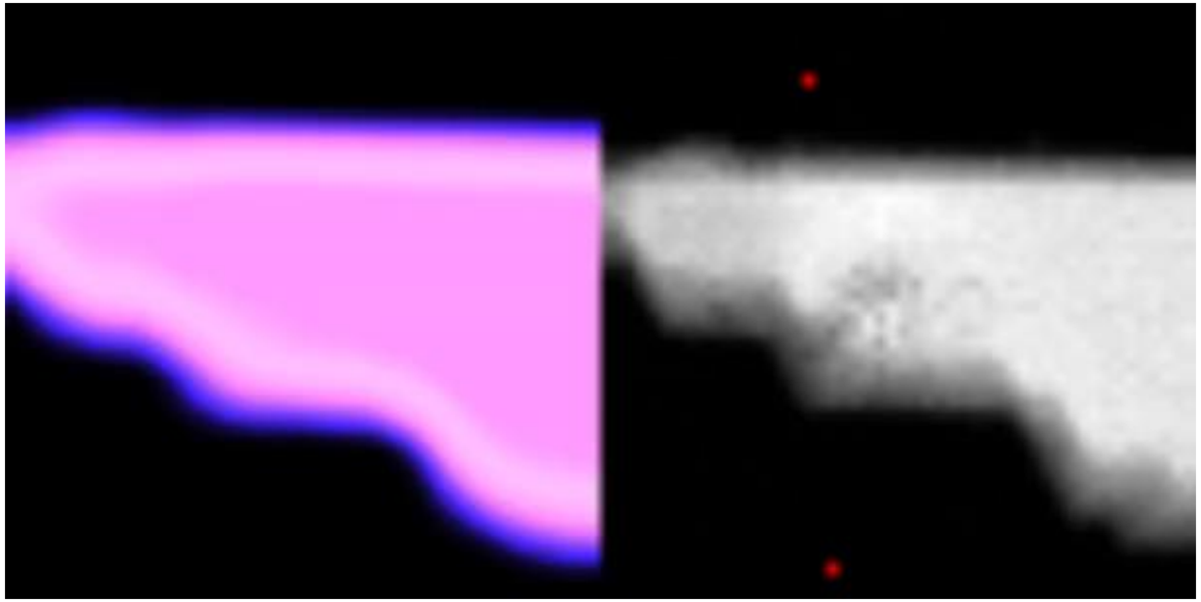
Physics-based feature engineering: $MnPS_3$



- **Discovering physics in a "new" material $MnPS_3$**

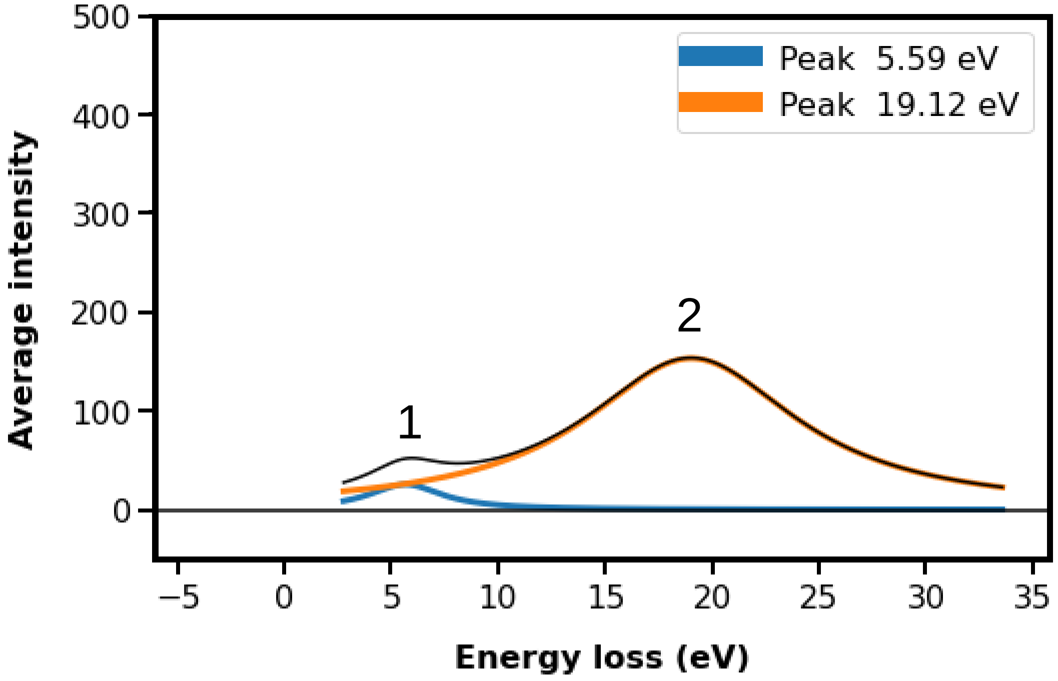
"Acquisition function"

HAADF-STEM
+ points visited



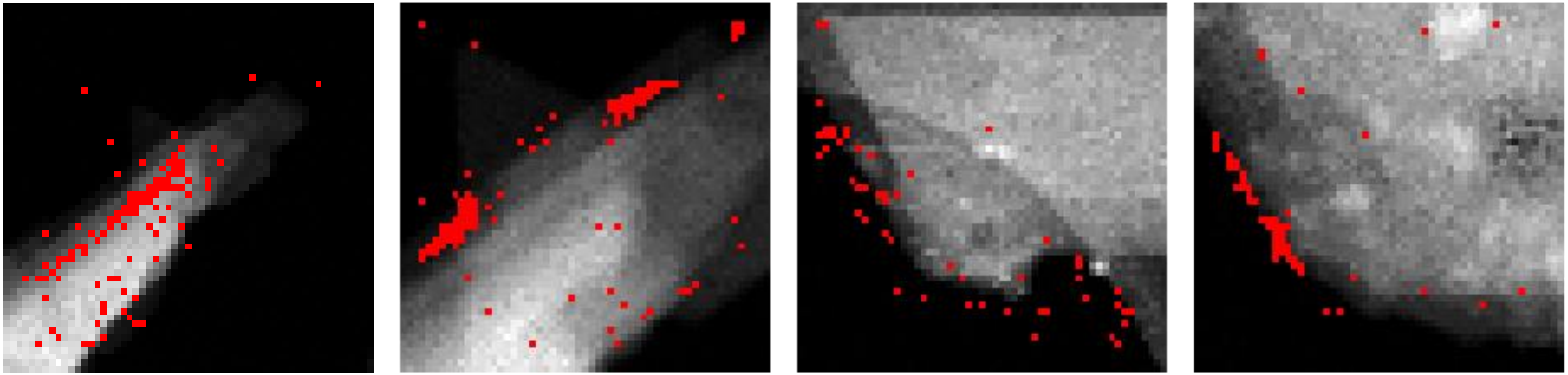
Physics search criteria:

$$\boldsymbol{Ratio = Peak\,1\ /\ peak\,2}$$



Reproducible

- Very similar behavior when searching for the same criteria elsewhere!
- Success!



Discovery pathway depends on the reward structure (scalarizer that defines signature of physics we want to discover)!

Change the scalarizer… change the physics

- (Same region) **Simple physics search:** peak max in selected region

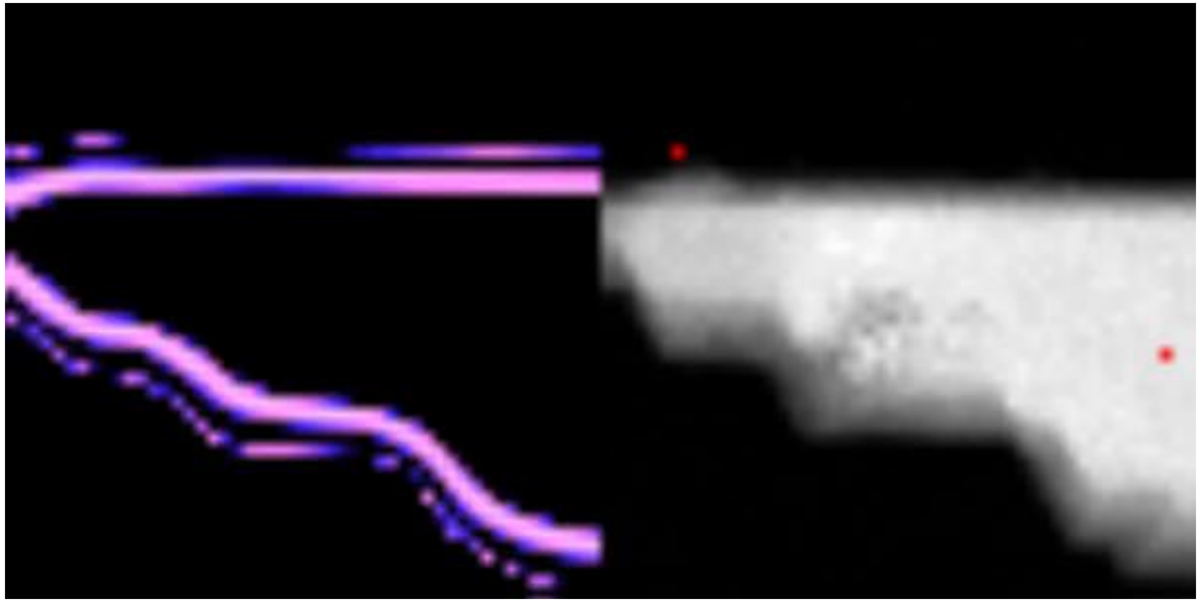
Physics search criteria:

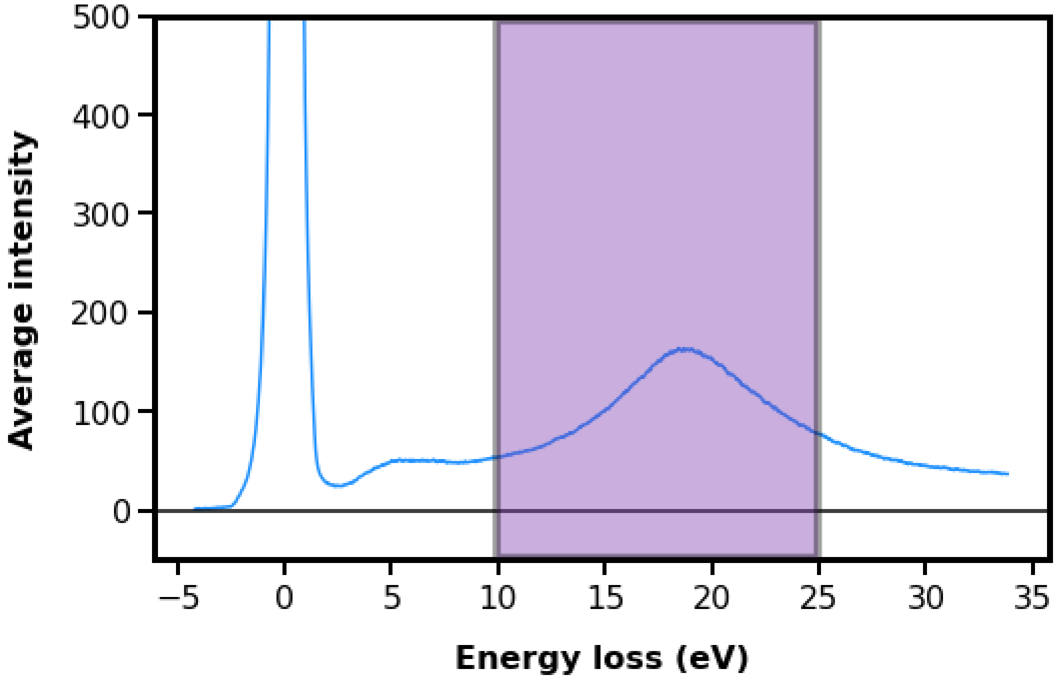
$$\boldsymbol{Maximize(f)}$$

(Specific peak intensity)

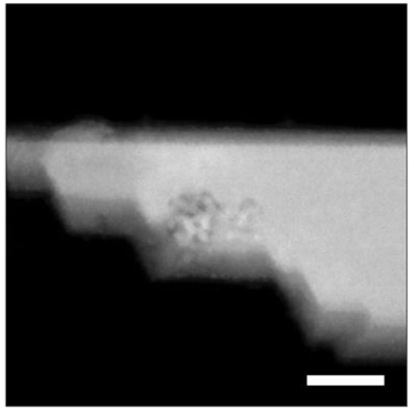
"Acquisition function"

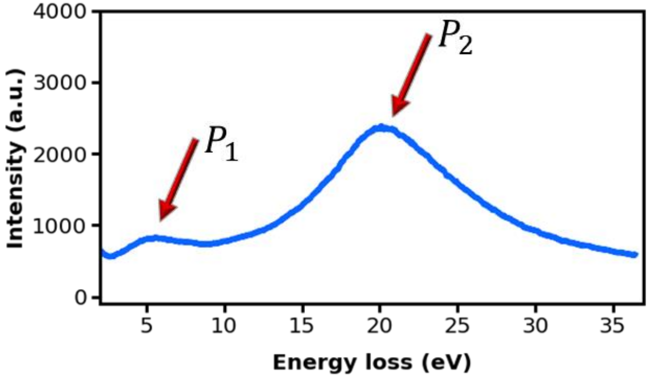
HAADF-STEM
+ points visited



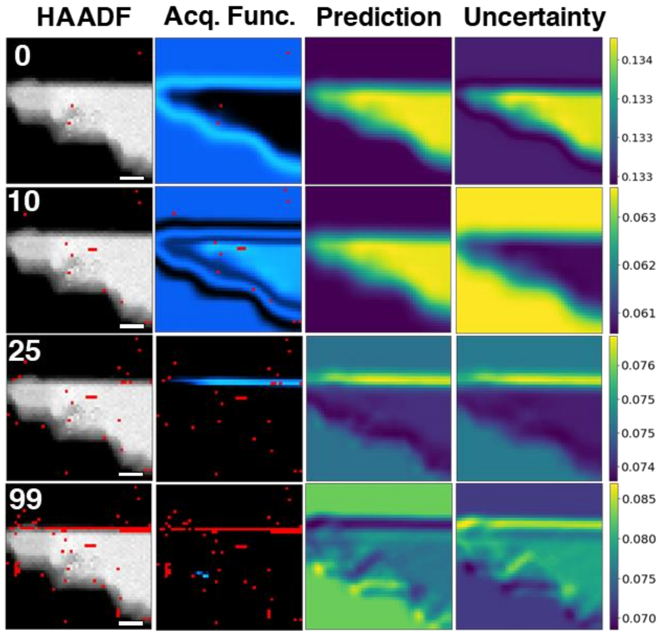


Summarize these automated experiments..

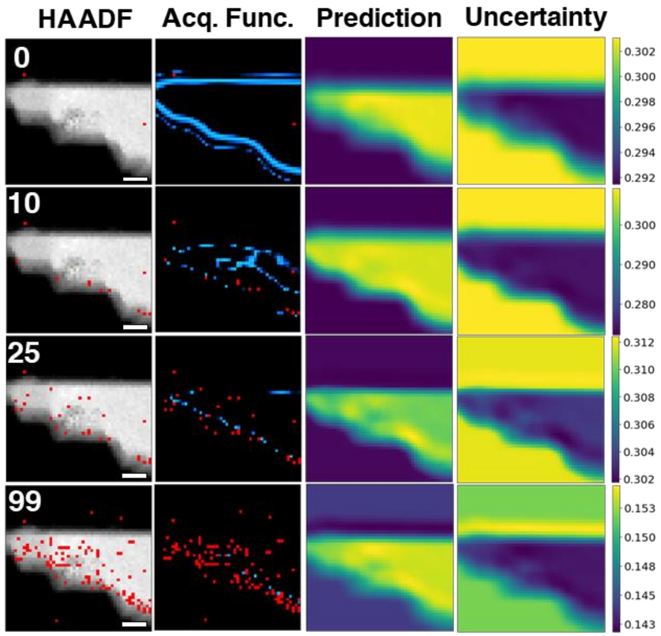




$\boxed{\max(P_1/P_2)}$

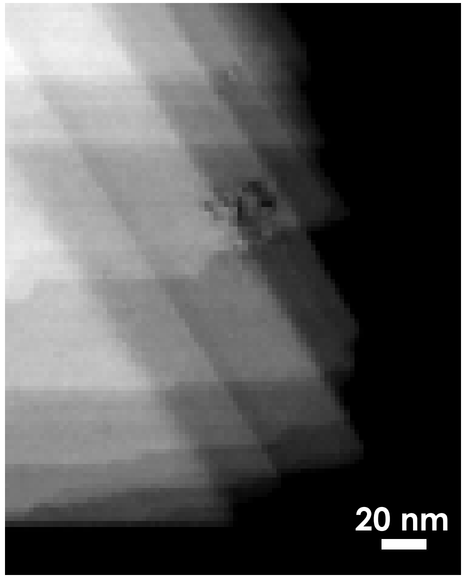


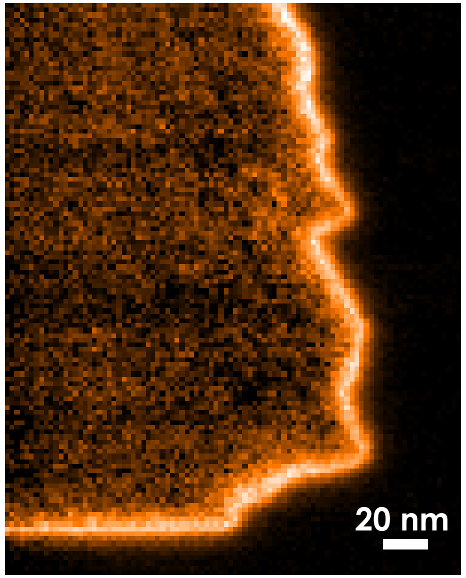
$\boxed{\max(P_2)}$

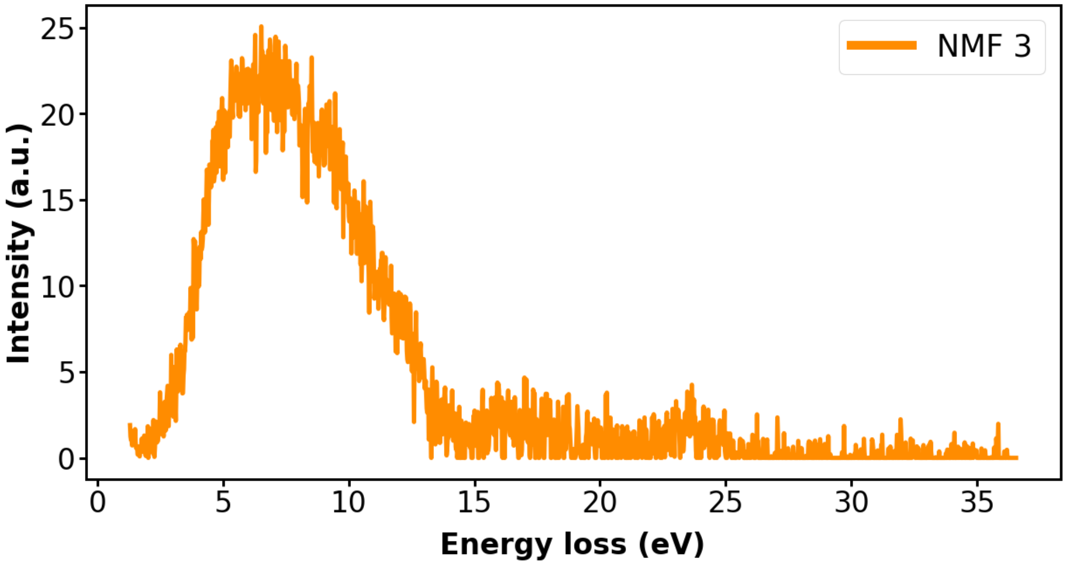


Confirmation

- Full hyperspectral EELS set collected and decomposed into few number of components (NMF)
- Indeed, there is edge plasmon activity near the energies used in the DKL







OAK RIDGE National Laboratory

4D STEM

Center of mass

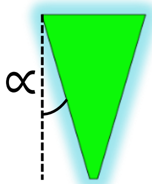
- Electric field
- Potential
- charge density

NBED disc fitting

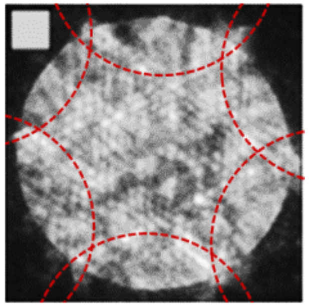
- Strain
- Grains

Require different probe conditions

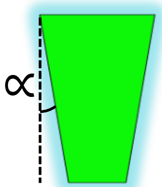
Overlapping diffraction discs



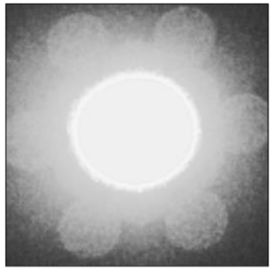
$\propto \approx 30\,mrad$



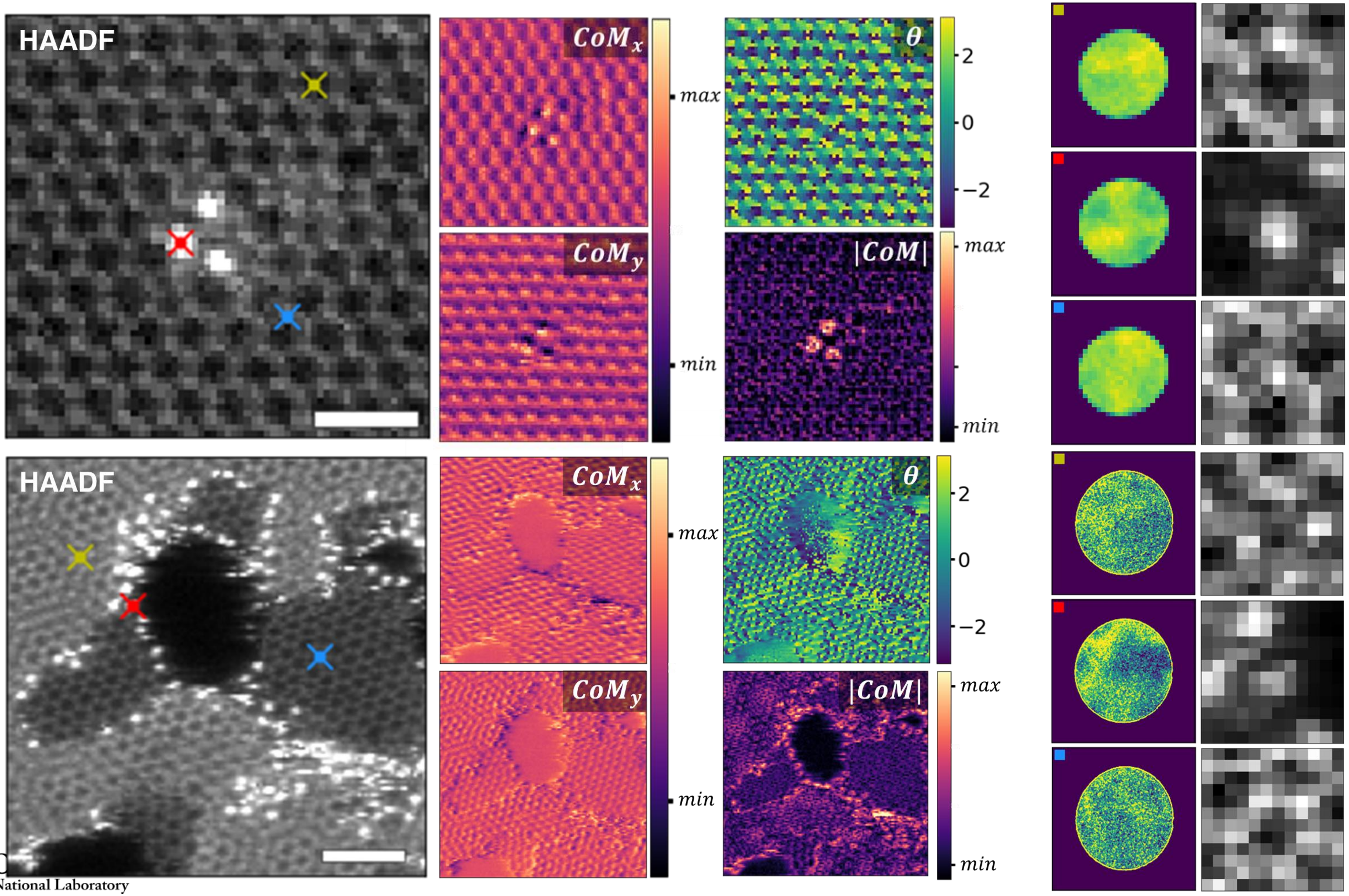
Non-overlapping diffraction discs



$\approx 5\,mrad$



DKL: 4D STEM – center of mass (DPC)



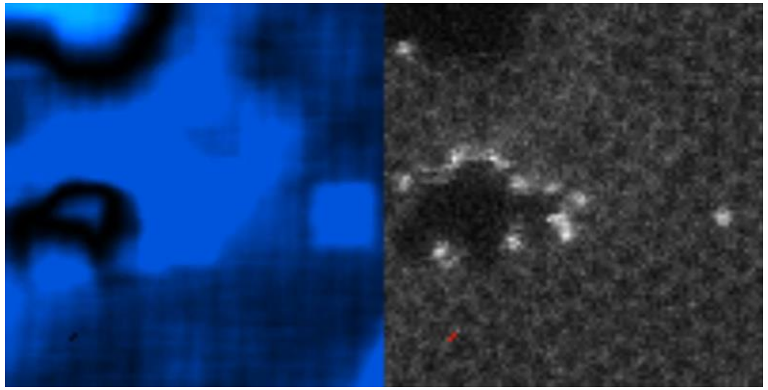
Do not normally care about the local image information

But again, let's make use of it!

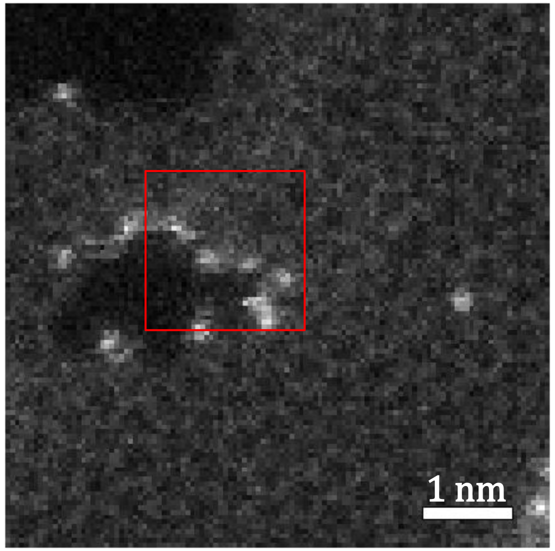
DKL: 4D STEM – center of mass

Example experiment:
- search for **maximum sample electric field**
- calculate |center of mass| of diffraction pattern

Acquisition function

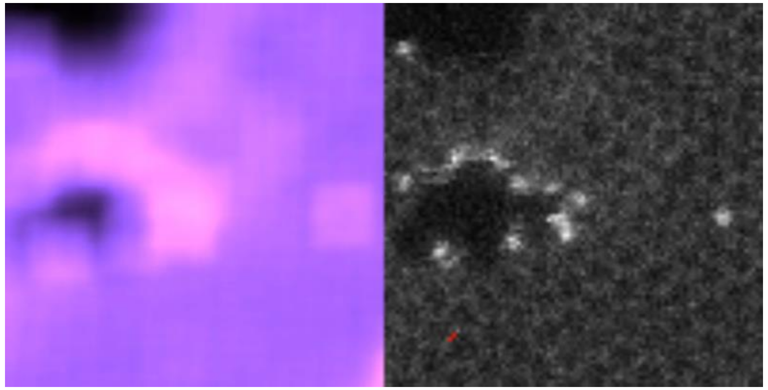


Full structural image



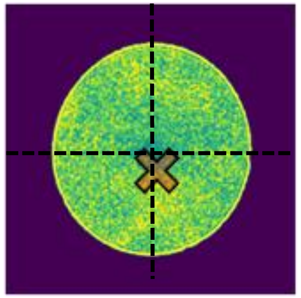
Measured quantity

Prediction map

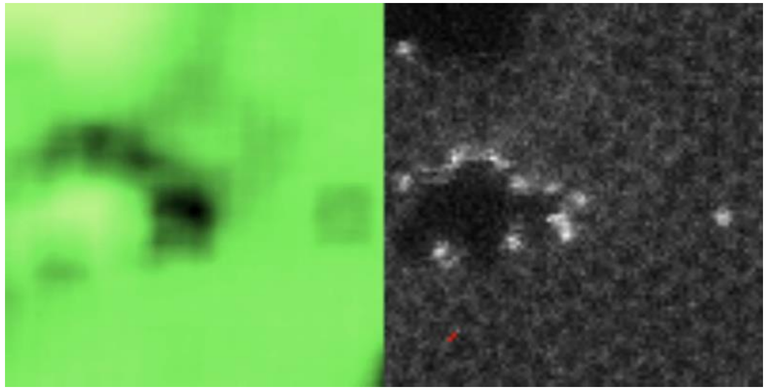


Optimize physics
(center of mass, strain, etc.)

|CoM|

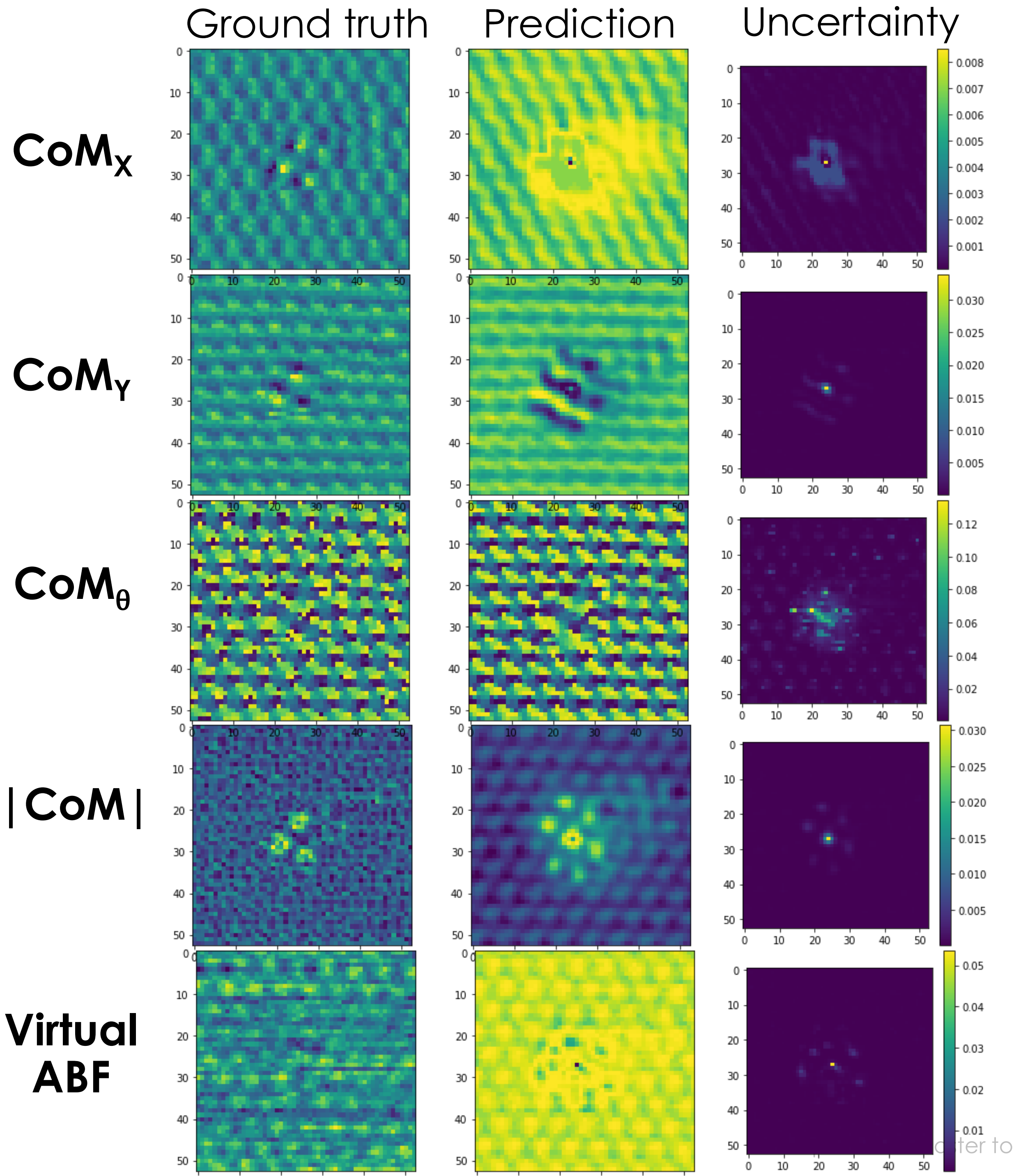


Uncertainty map



DKL: 4D STEM (pre-acquired data)

- Choosing **different scalarizers**, does one perform "better"?
- Some interesting (real) artifacts appear in some predictions

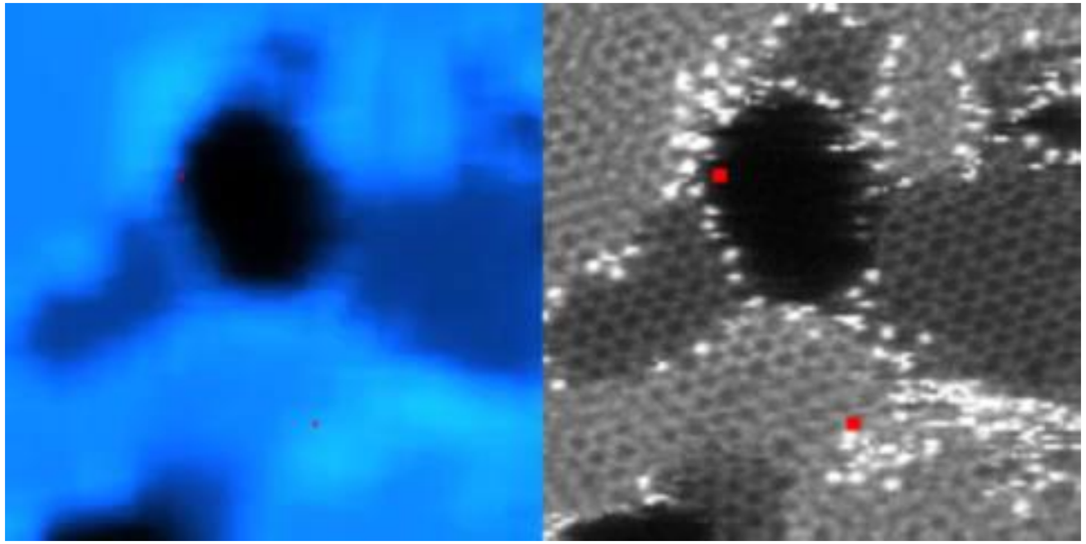


Oak Ridge National Laboratory

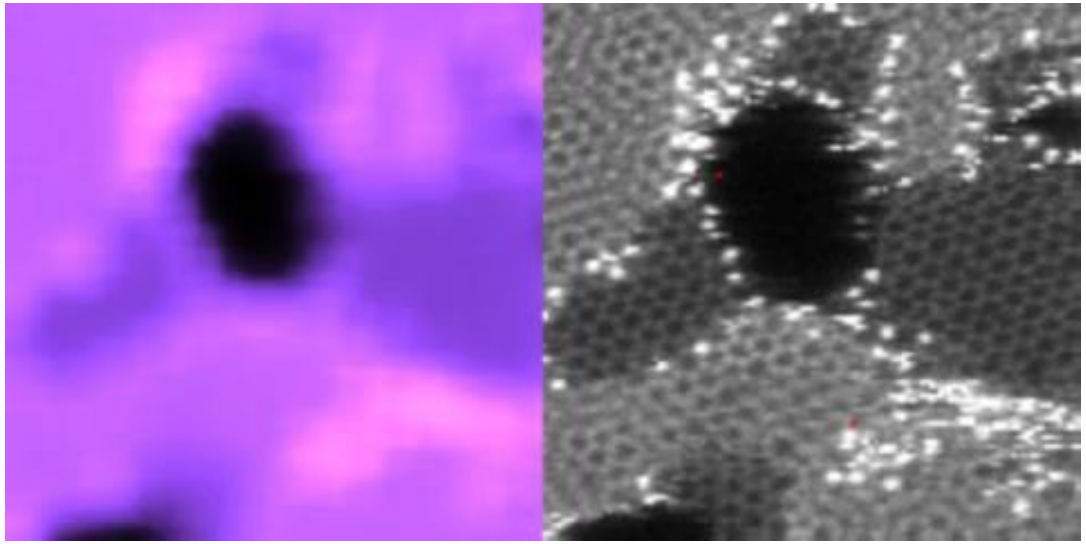
DKL: 4D STEM – active learning (pre-acquired data)

Sample: TBG (twisted graphene)
Scalarizer: DPC |CoM|

Acquisition function



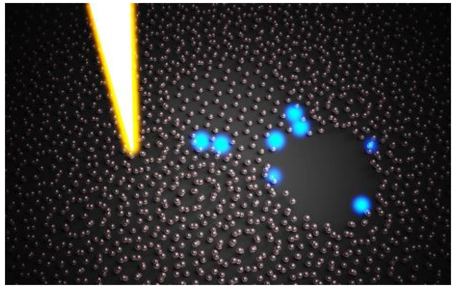
Prediction



Next measurement position is based on $\boldsymbol{argmax}(Acq.Func.)$

- Different **acquisition functions** can be used:
  - Expected Improvement (**EI**) (usually what was used)
  - Upper Confidence Bound (**UCB**), etc.

- Usually based on some combination of **prediction** and **uncertainty**.

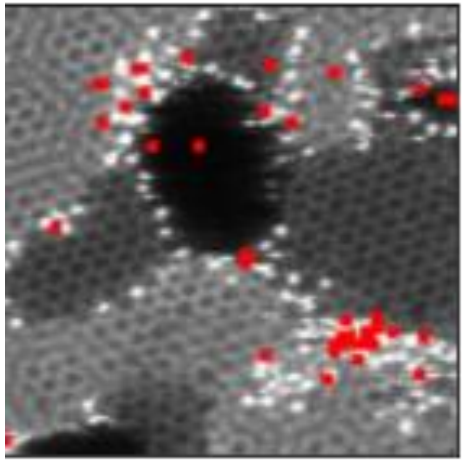


DKL: 4D STEM active learning (pre-acquired data)

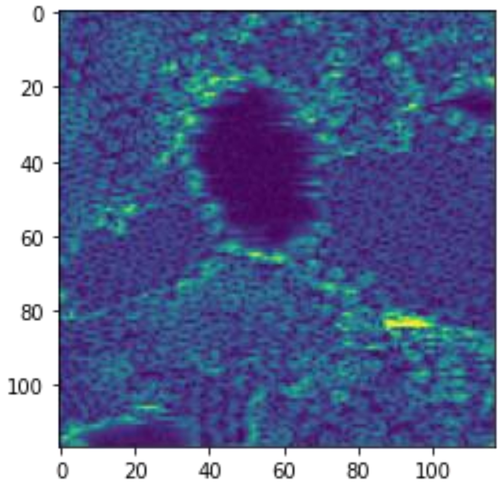
Scalarizer: |CoM|

After 25 measurements

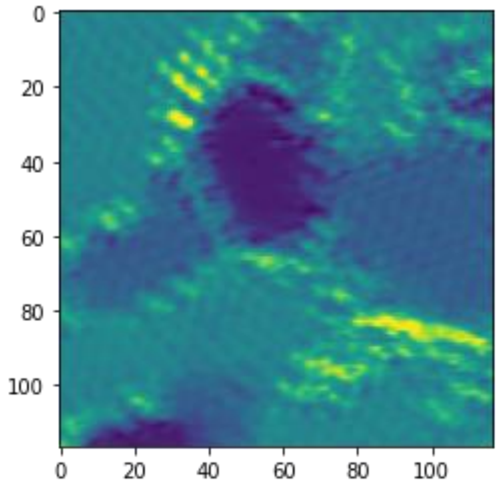
ADF with measurement points



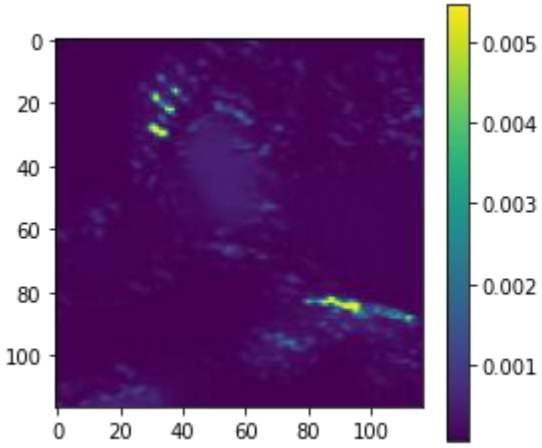
Ground truth



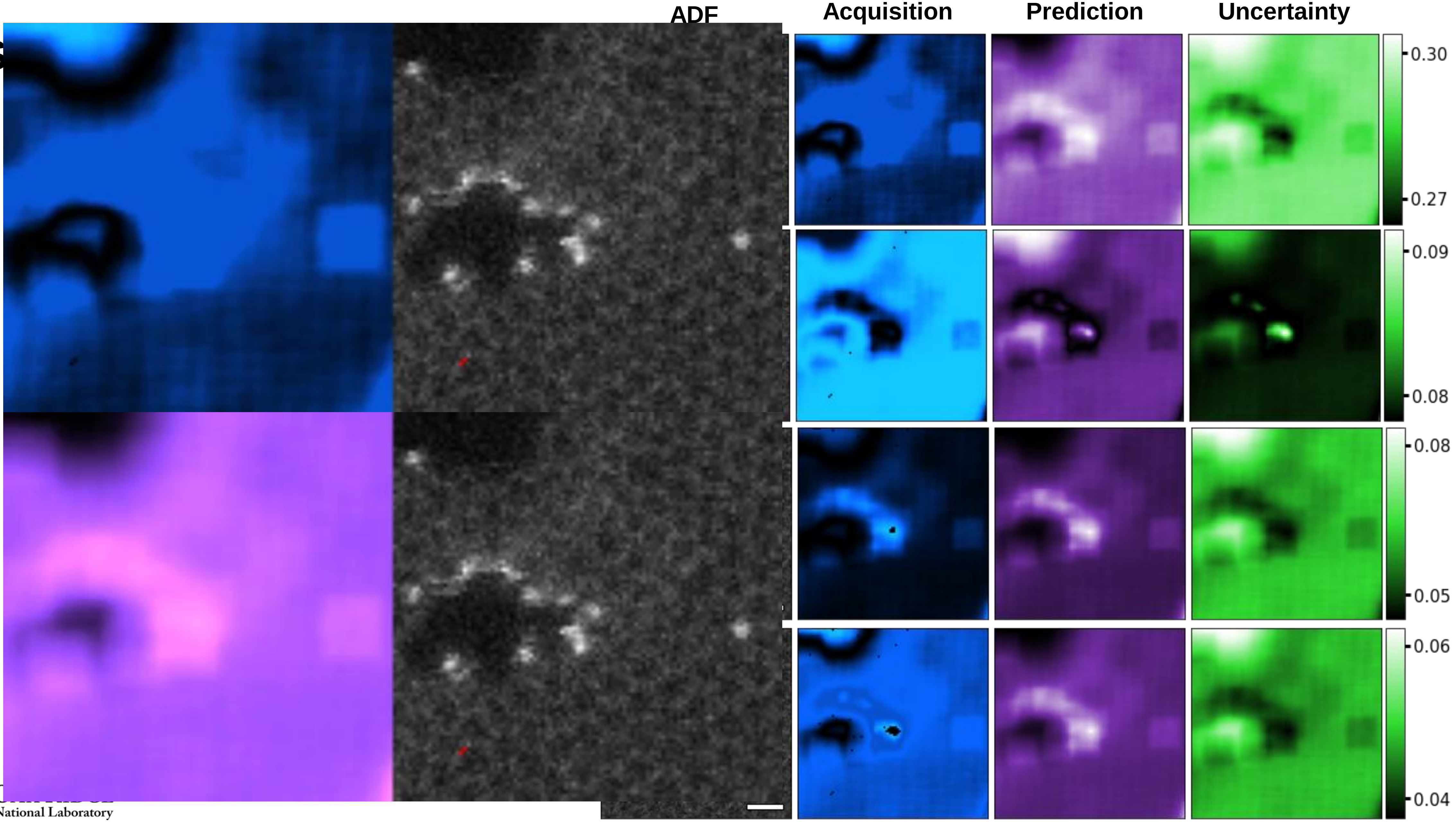
Prediction



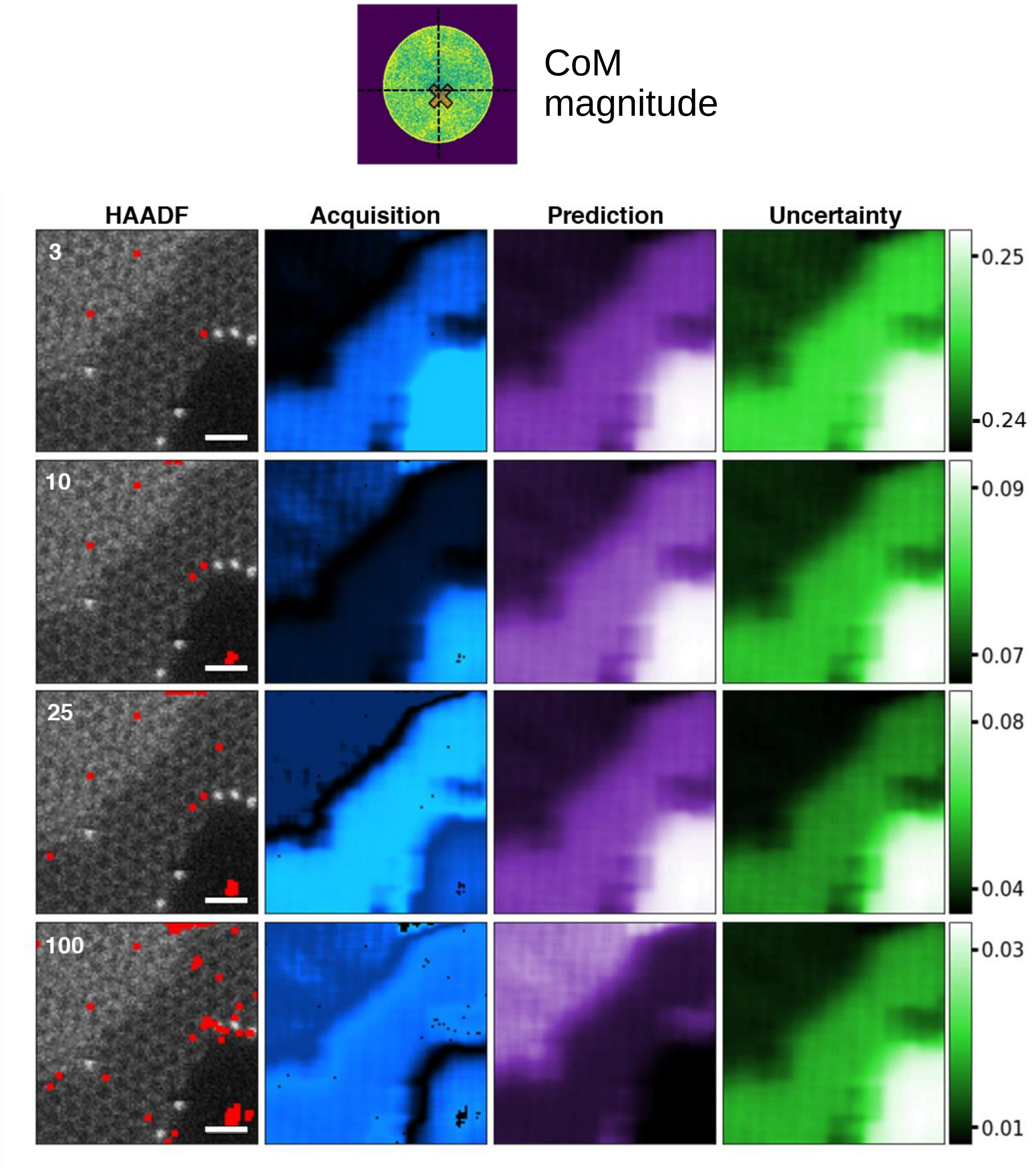
Uncertainty

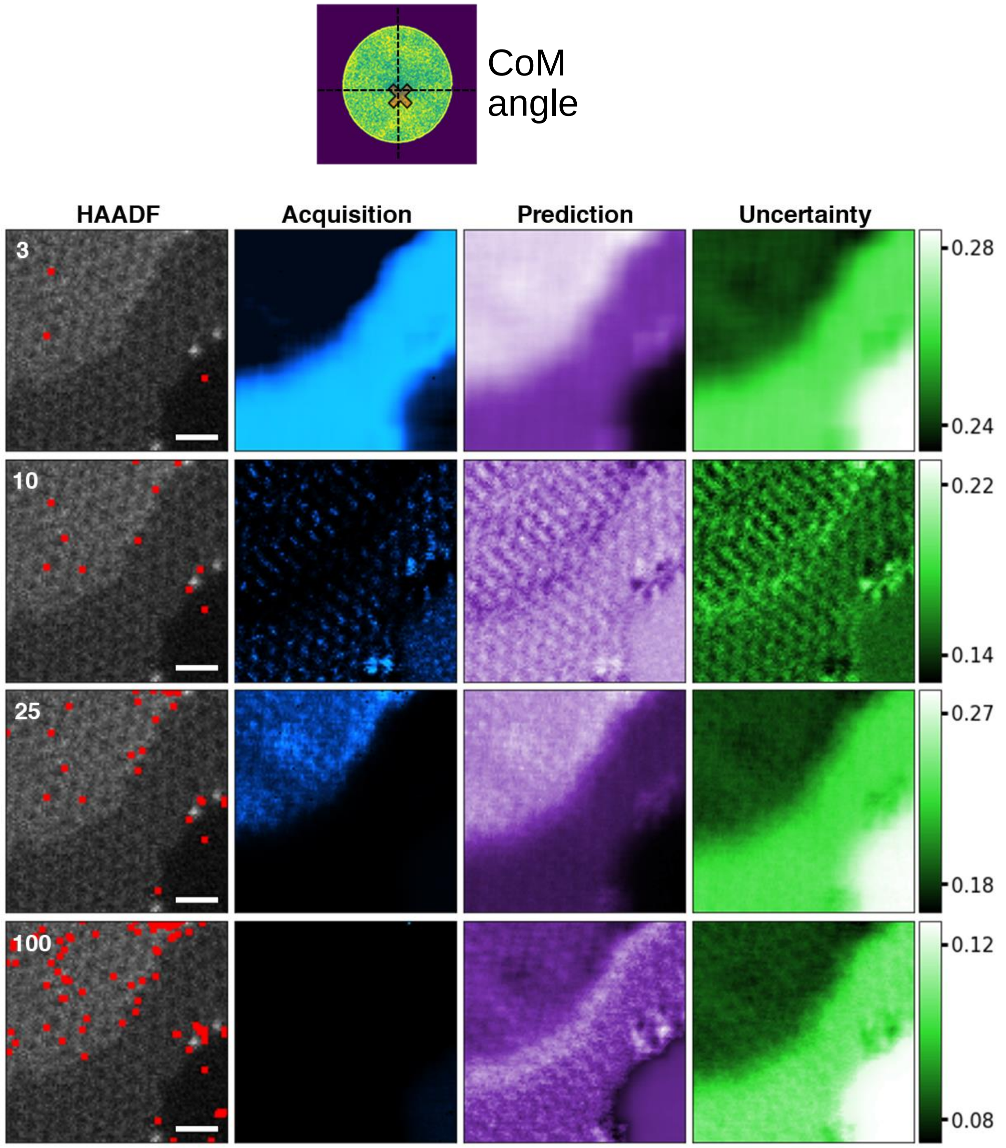


Active learning: real automated experiment



Effect of physics choice (scalarizer) on acquisition





4D STEM

Center of mass

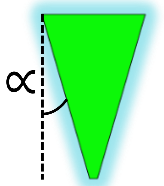
- Electric field
- Potential
- charge density

NBED disc fitting

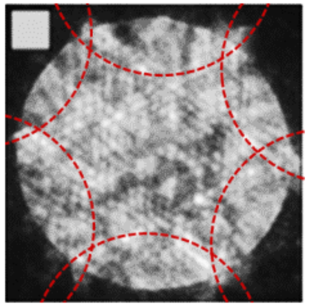
- Strain
- Grains

Require different probe conditions

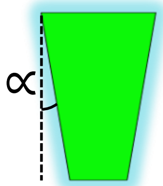
Overlapping diffraction discs



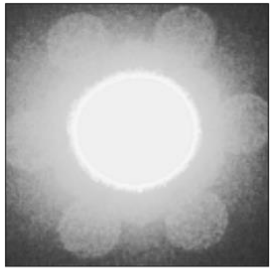
$\propto \approx 30\, mrad$



Non-overlapping diffraction discs

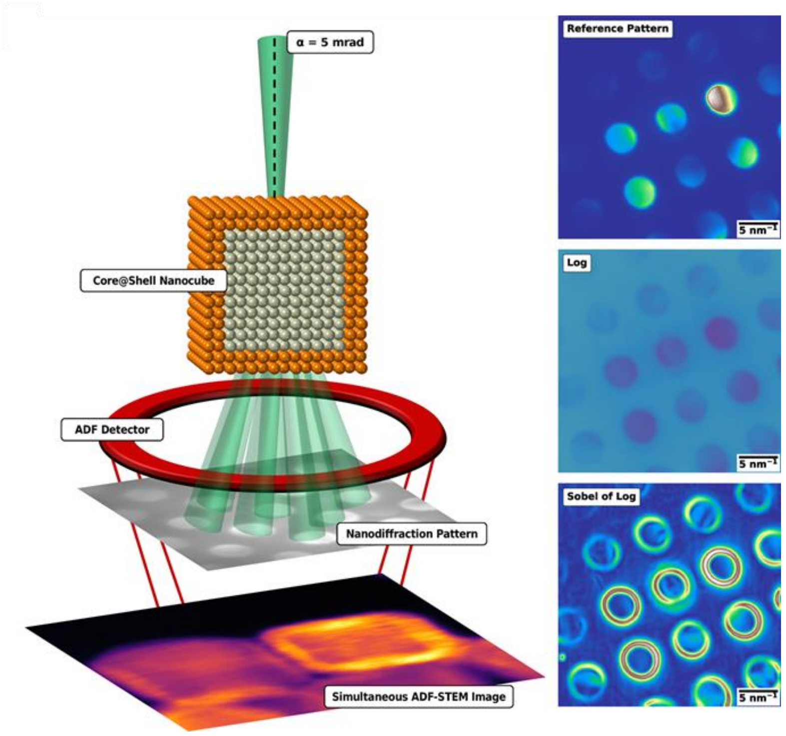


$\approx 5\, mrad$

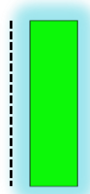


Measuring strain in the STEM

Using local diffraction



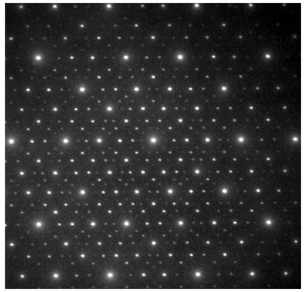
Mukherjee D. et. al., *ACS Catal.* **2020**, 10, 10, 5529−5541



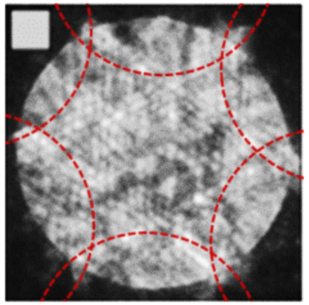
Parallel illumination

$\propto$ $\approx 5\, mrad$ $\propto \approx 30\, mrad$

Convergent beam



(Conventional TEM conditions)



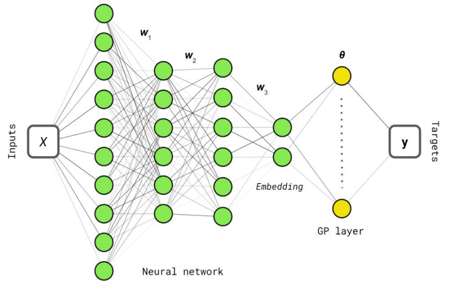
Strain, differential phase contrast (DPC), electric field, atomic potential maps

Requirements:

- ➢ Pixelated detector
- ➢ Diffraction ***disc*** fitting

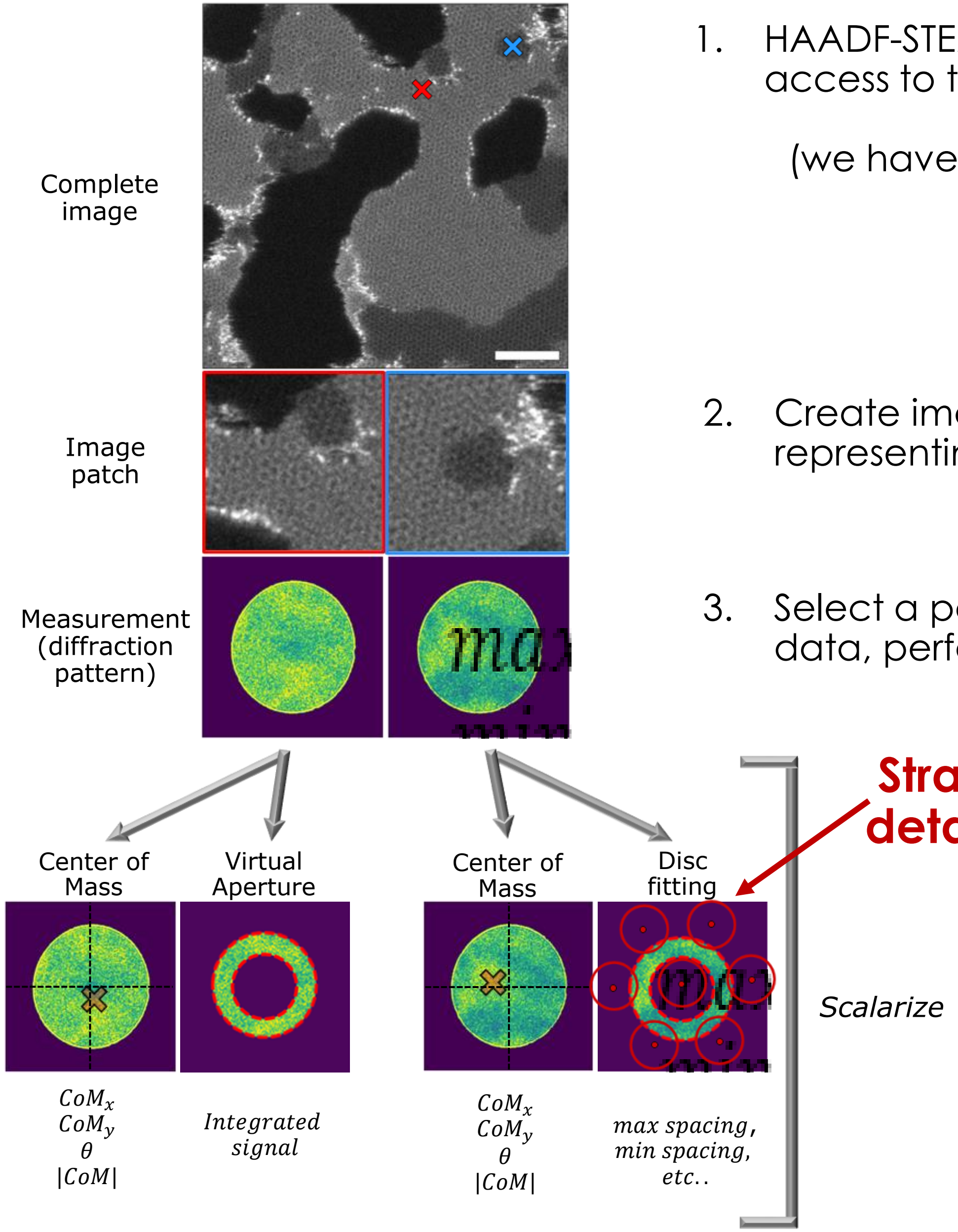
Strain is calculated **after** experiment (offline analysis)

Neural Network Architecture



1. HAADF-STEM is acquired first, we have **complete** access to this

   (we have not yet collected diffraction data yet)



2. Create image patches at each pixel, representing local geometry/structure

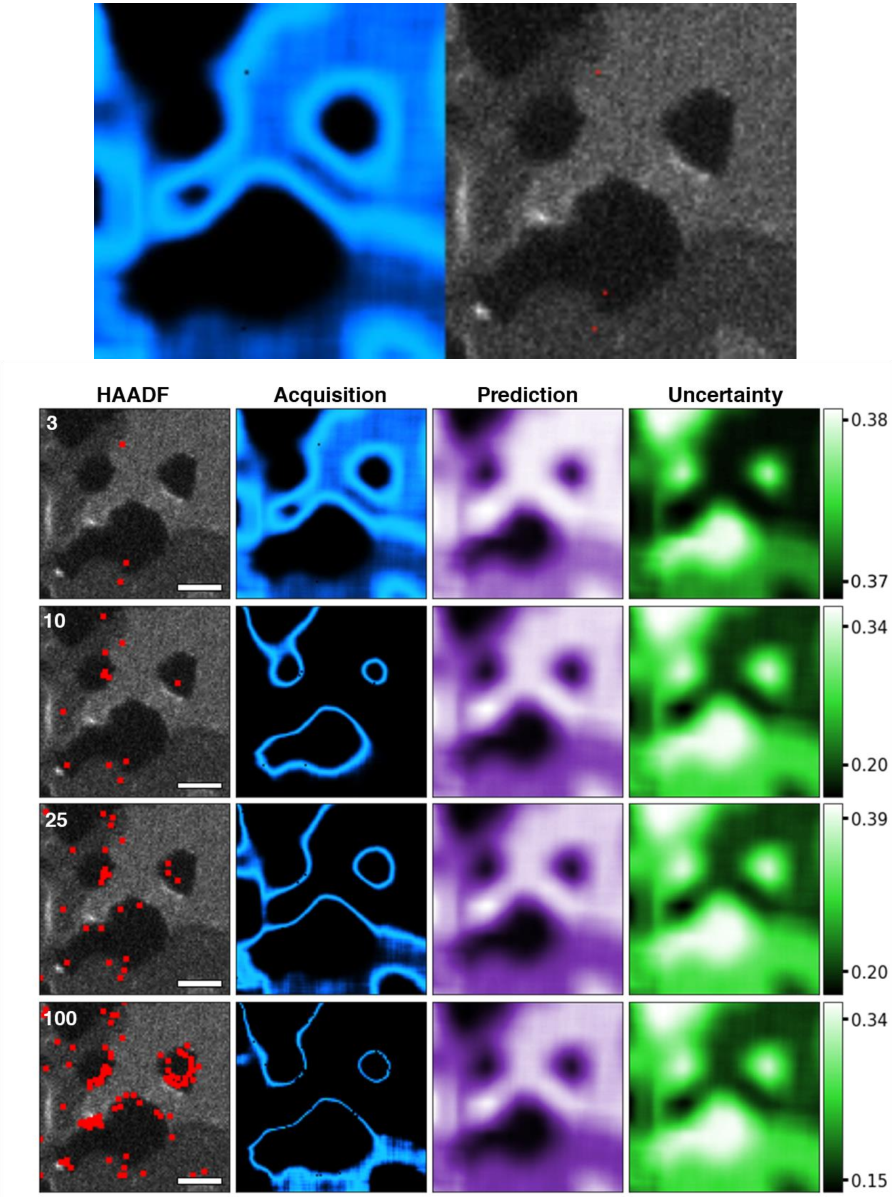
3. Select a position in space to acquire diffraction data, perform computation (scalarize).

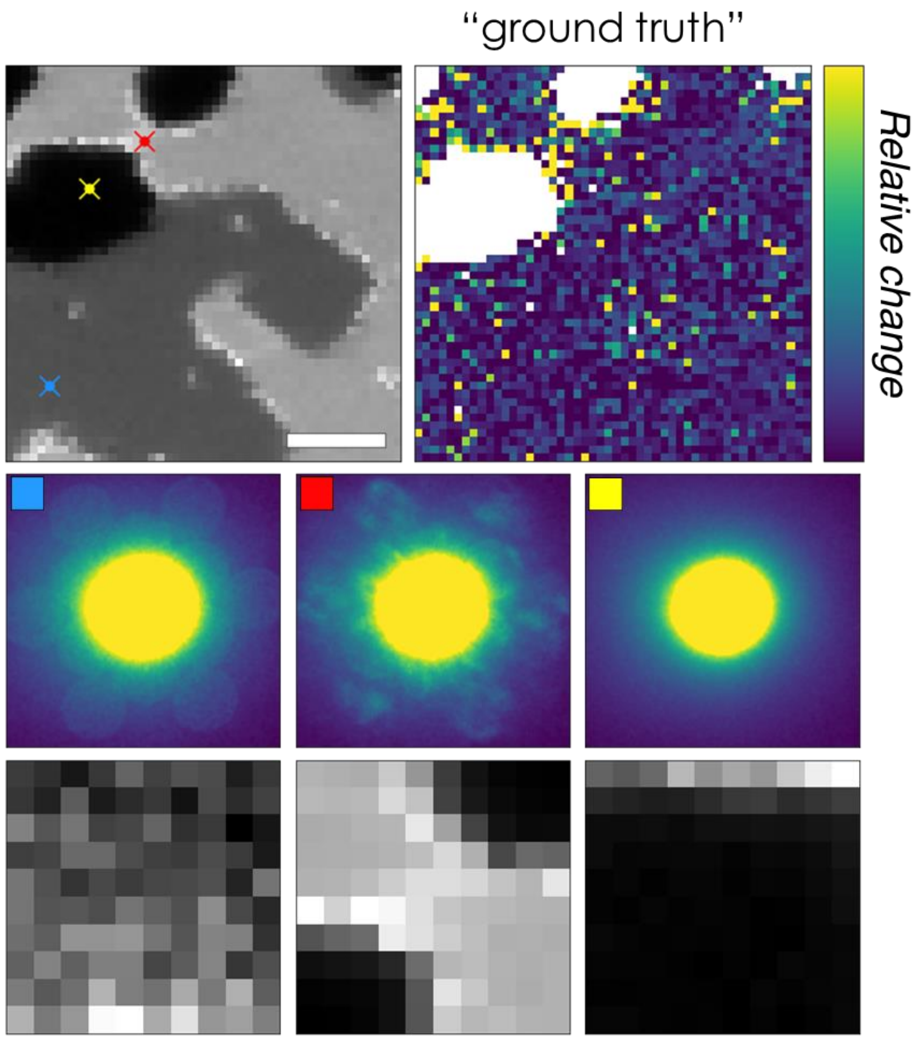
Strain details

4. Train / update network with **image patch** and **scalarized quantity PAIR**. Repeat 3&4 (automated)

This is controlled by <u>physics-based</u> "acquisition function"

Automated Experiments in 4D STEM: strain





OAK RIDGE National Laboratory

Colab!

- Change to GPU runtime if not already
- from skimage.draw import **disk** (not circle)
- spectralavg   = **128**
- dklgp = aoi.models.dklGPR(data_dim, embedim=2, precision=''=**single**'')