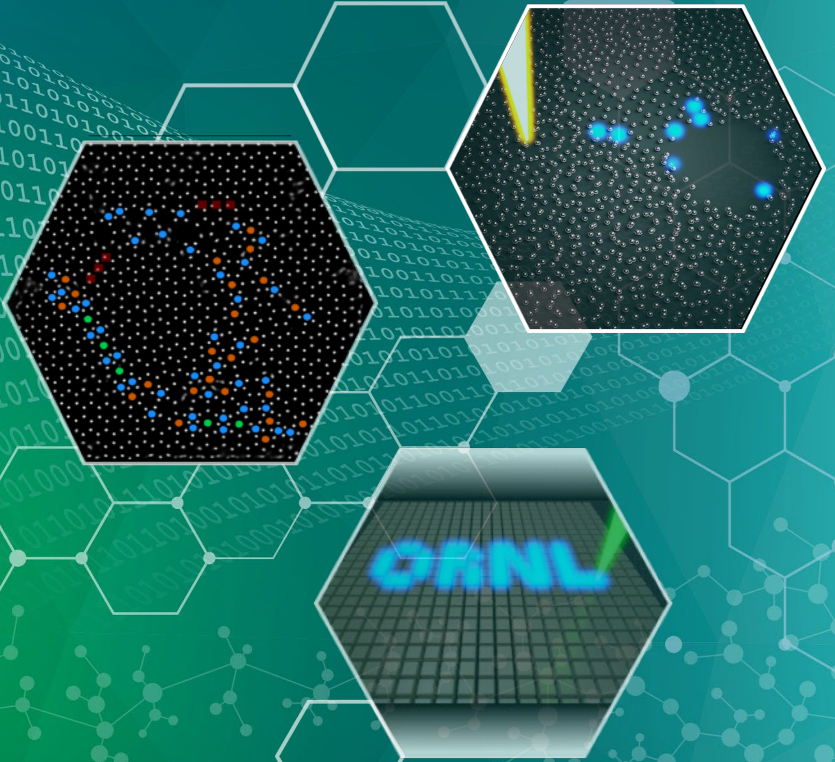
# Lecture 06

## Automated EELS and 4D STEM: Deep kernel learning

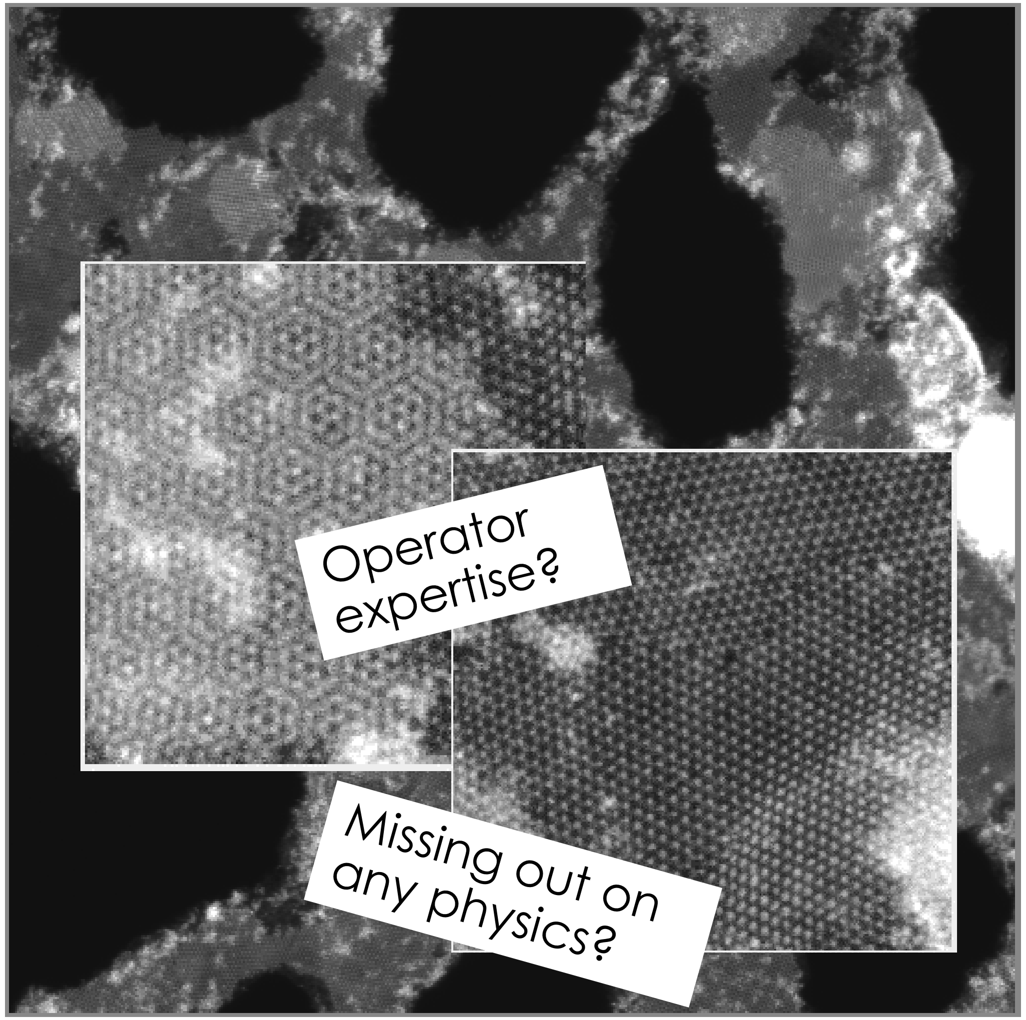


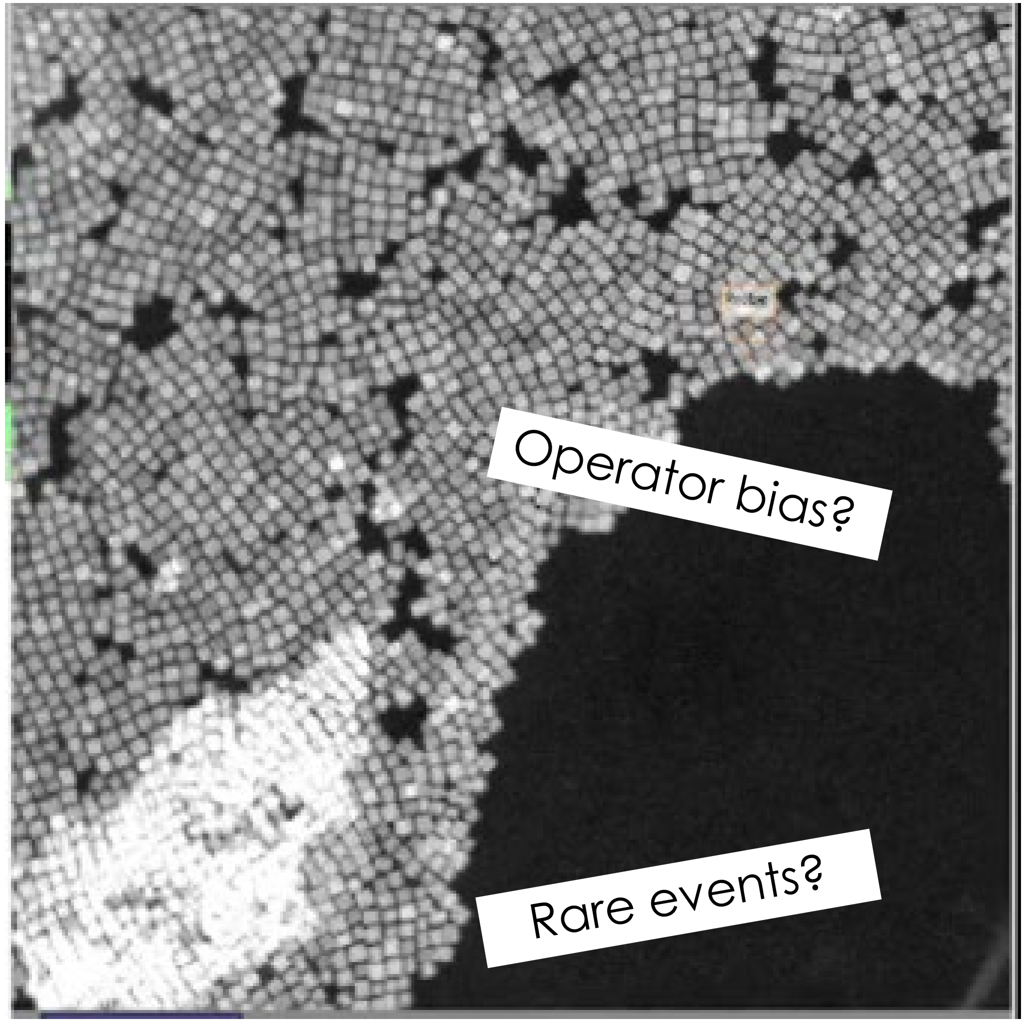

**Kevin Roccapriore**[1,2] and Sergei V. Kalinin[3,4]

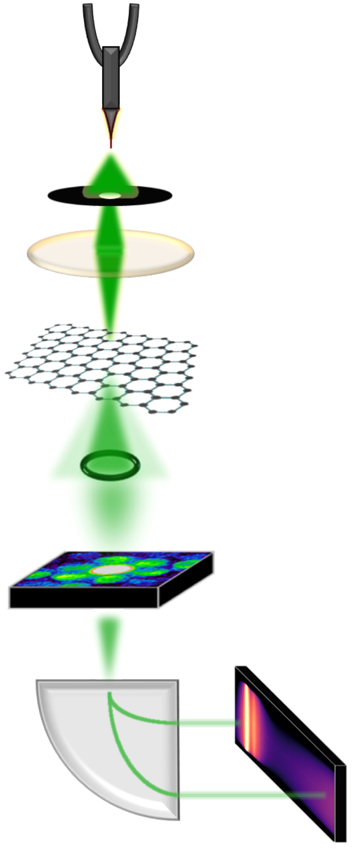
[1]Center for Nanophase Materials Sciences, Oak Ridge National Laboratory
[2]AtomQ
[3]University of Tennessee Knoxville
[4]Pacific Northwest National Laboratory

ORNL is managed by UT-Battelle LLC for the US Department of Energy

# How do we understand structure property relationships

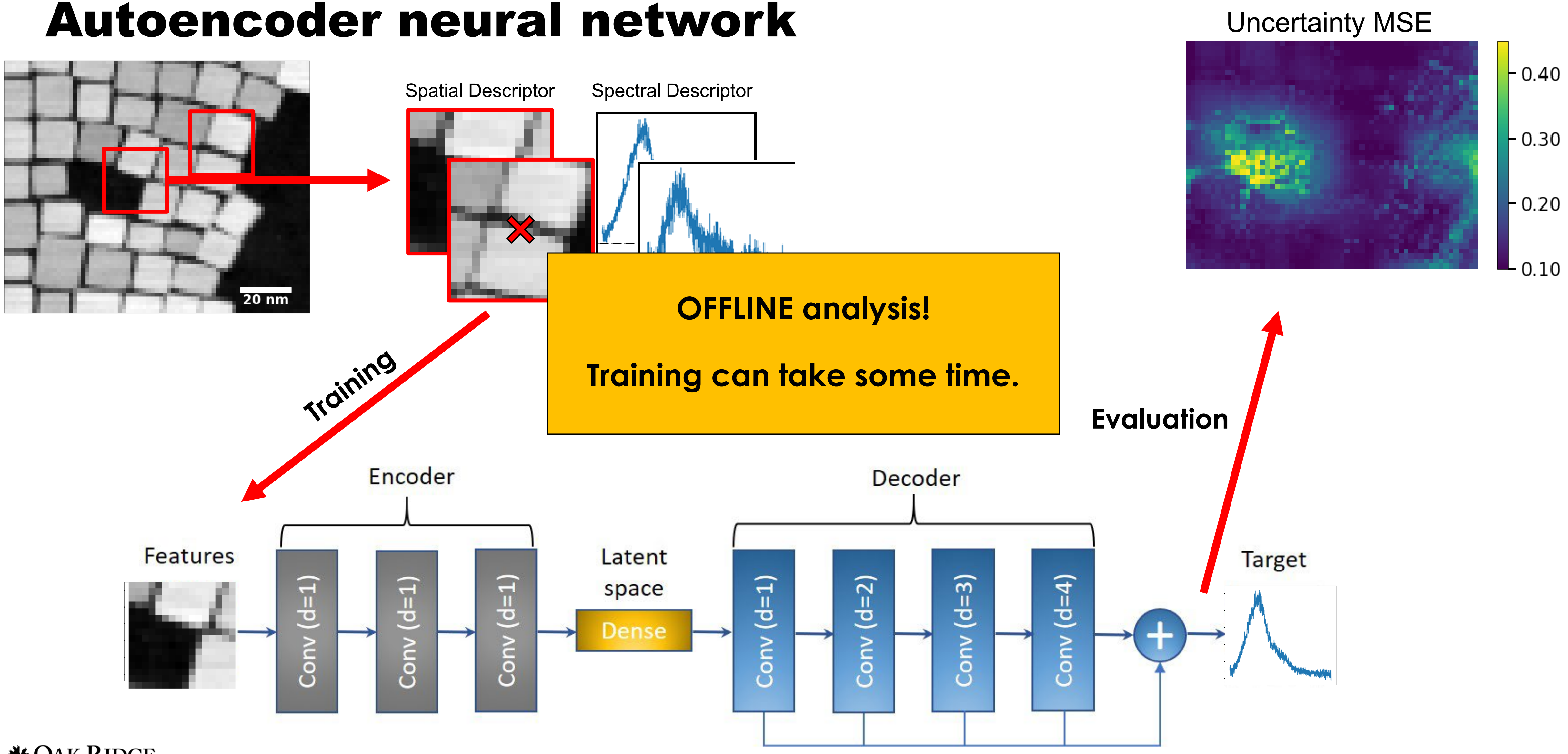


**Q**: Where should (would) you acquire analytical measurements? EELS spectrum image / 4D STEM, point spectra, where? **You can't say all points!**

# Recall:
# Structure-property correlations:
# Autoencoder neural network




**OFFLINE analysis!**

**Training can take some time.**

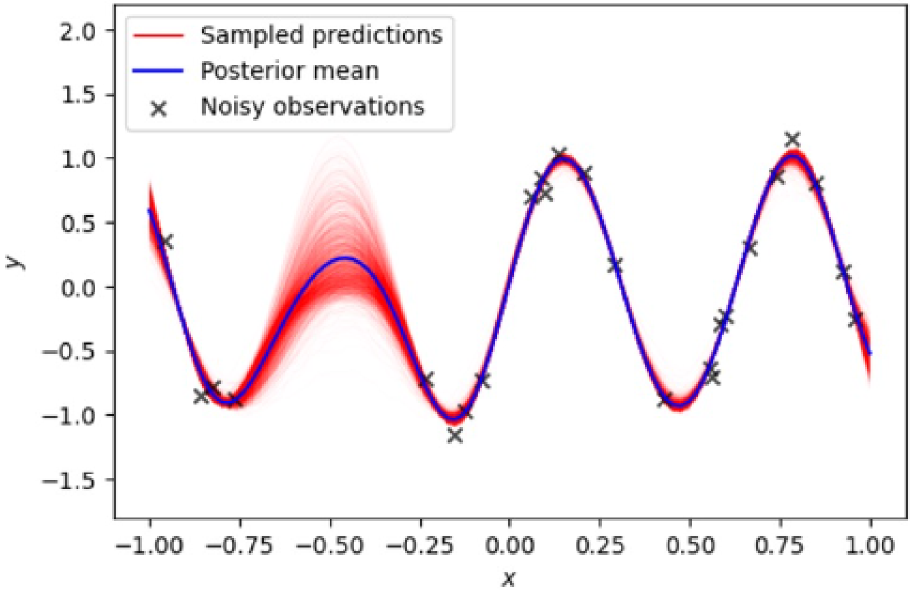
# Automated Experiment

## 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**)

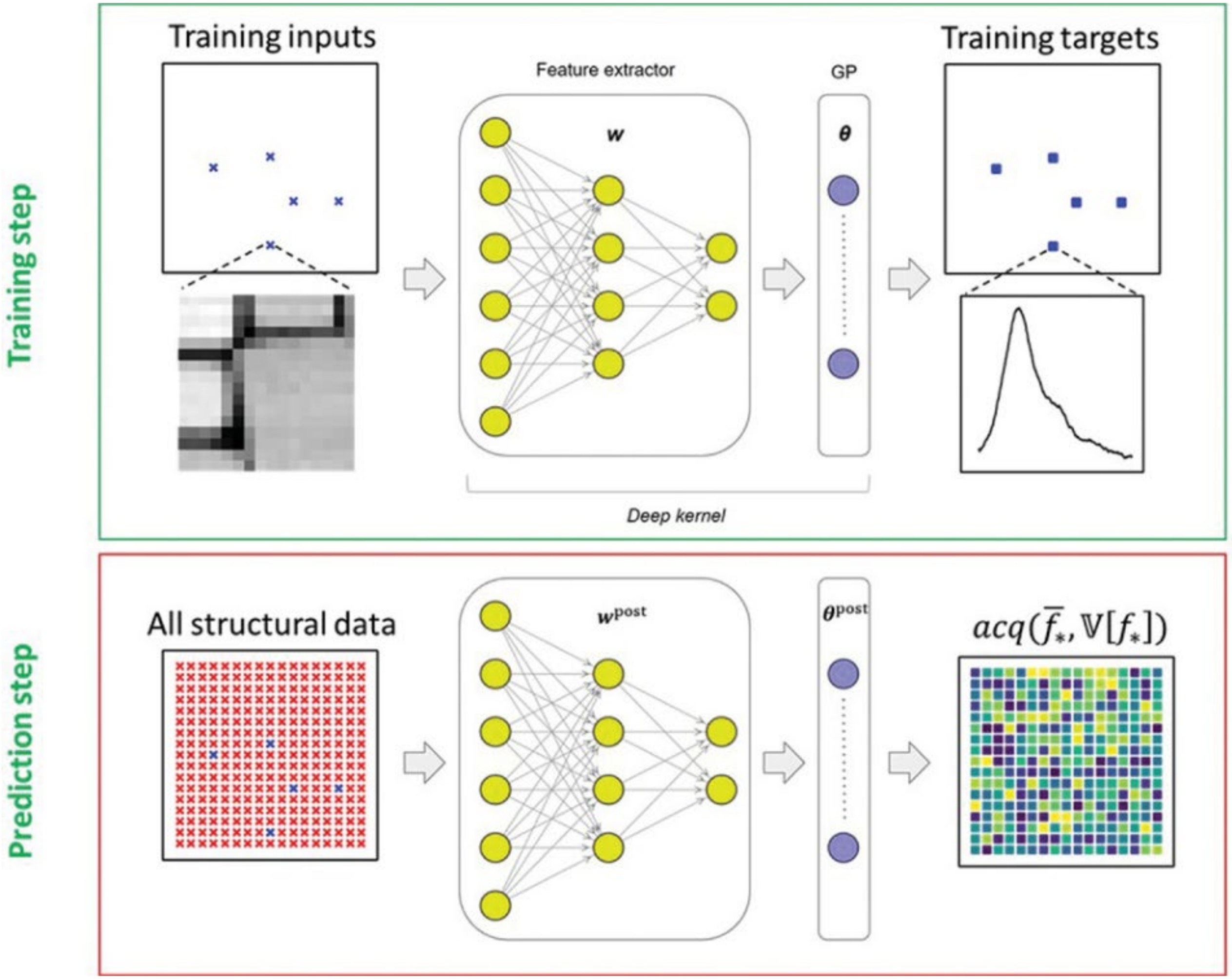
- Gaussian Process → requires a kernel for reconstruction.
-
  - What kernel do we use?
  - Kernel length?

  - 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!

# Deep kernel learning (DKL)




**Tuning knobs**

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

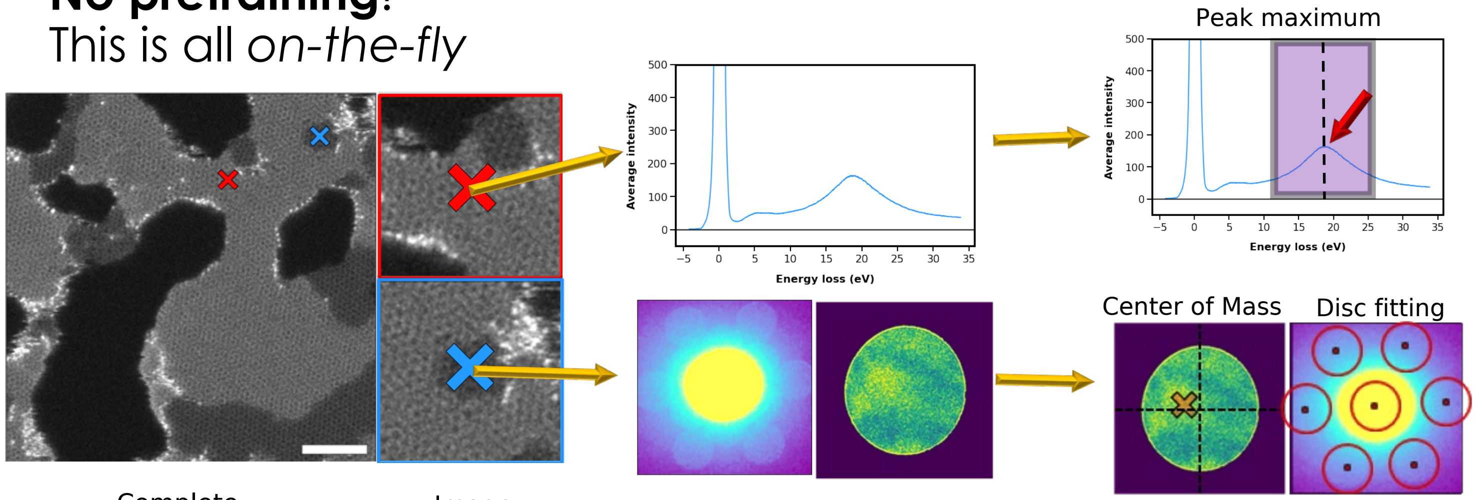
- Acquisition function
  - How to navigate given the predicted space

  - 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.

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

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

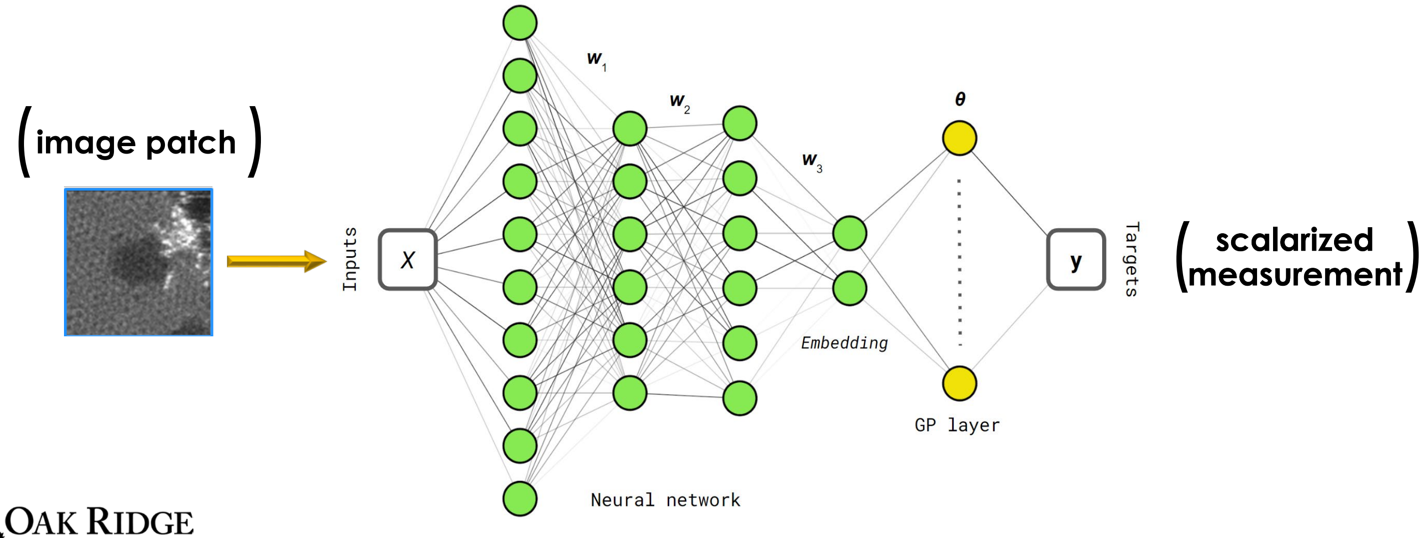



Complete **structural** image

Image patch

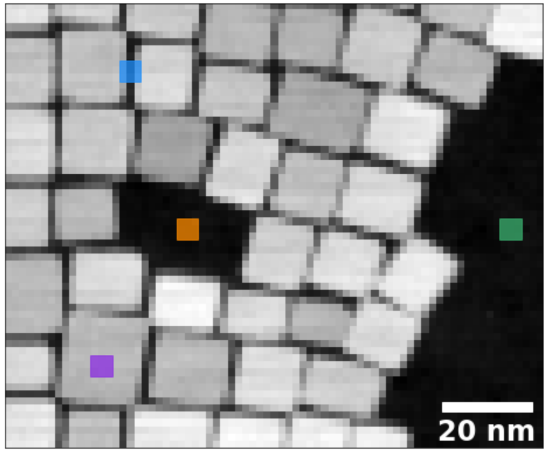
Measurement (1D or 2D)

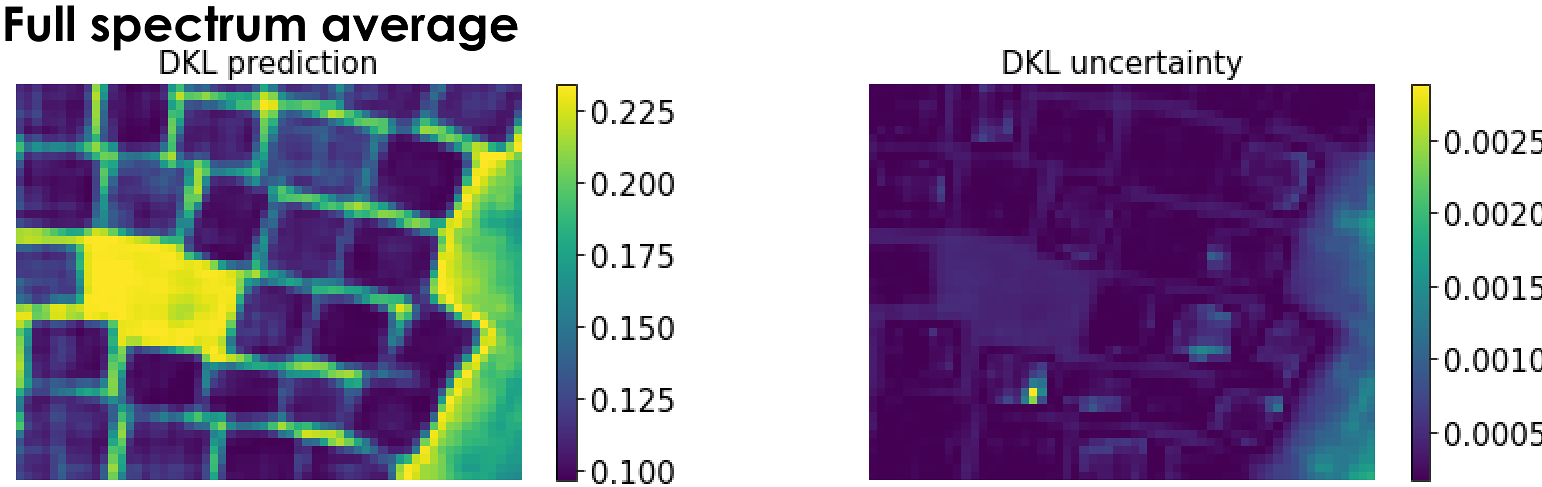
Scalarize (**physics enters here**)

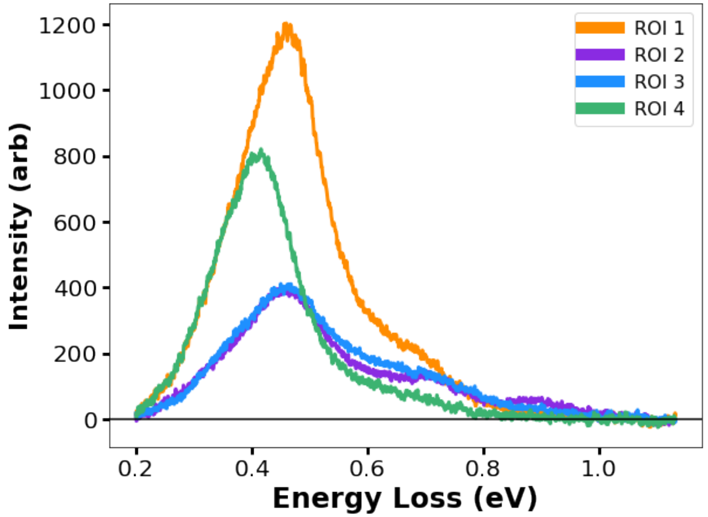


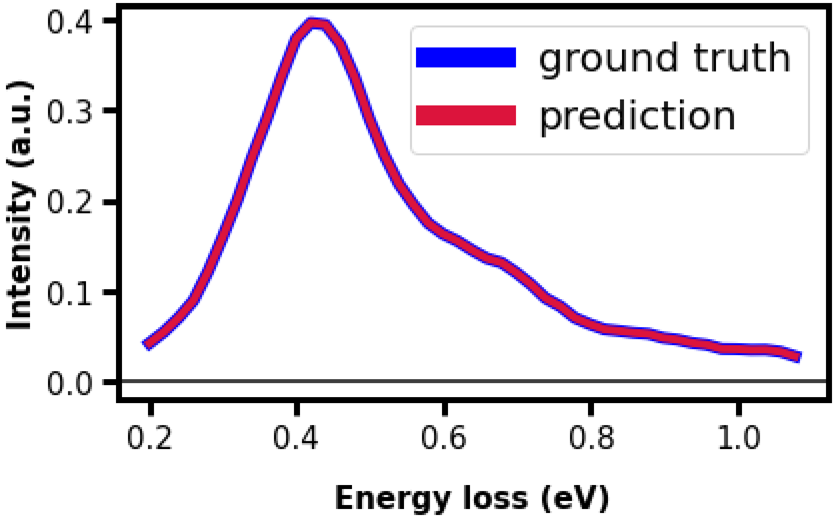

1. A complete **structural image**
2. Image **patches**
3. Perform a **measurement** @ (x1,y1)
4. "**Scalarize**" (**physics** incorporation)
5. Learn **correlation** between image signal and measurement signal (training step)
6. **Next measurement** @ (x2,y2) based on learned correlation
7. Repeat

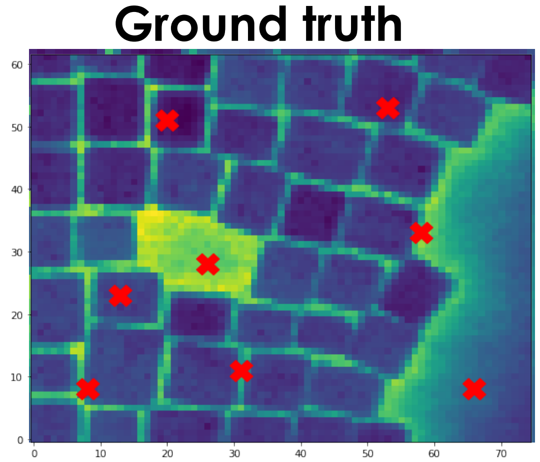
# DKL with FULL pre-acquired data (no active learning or BO)



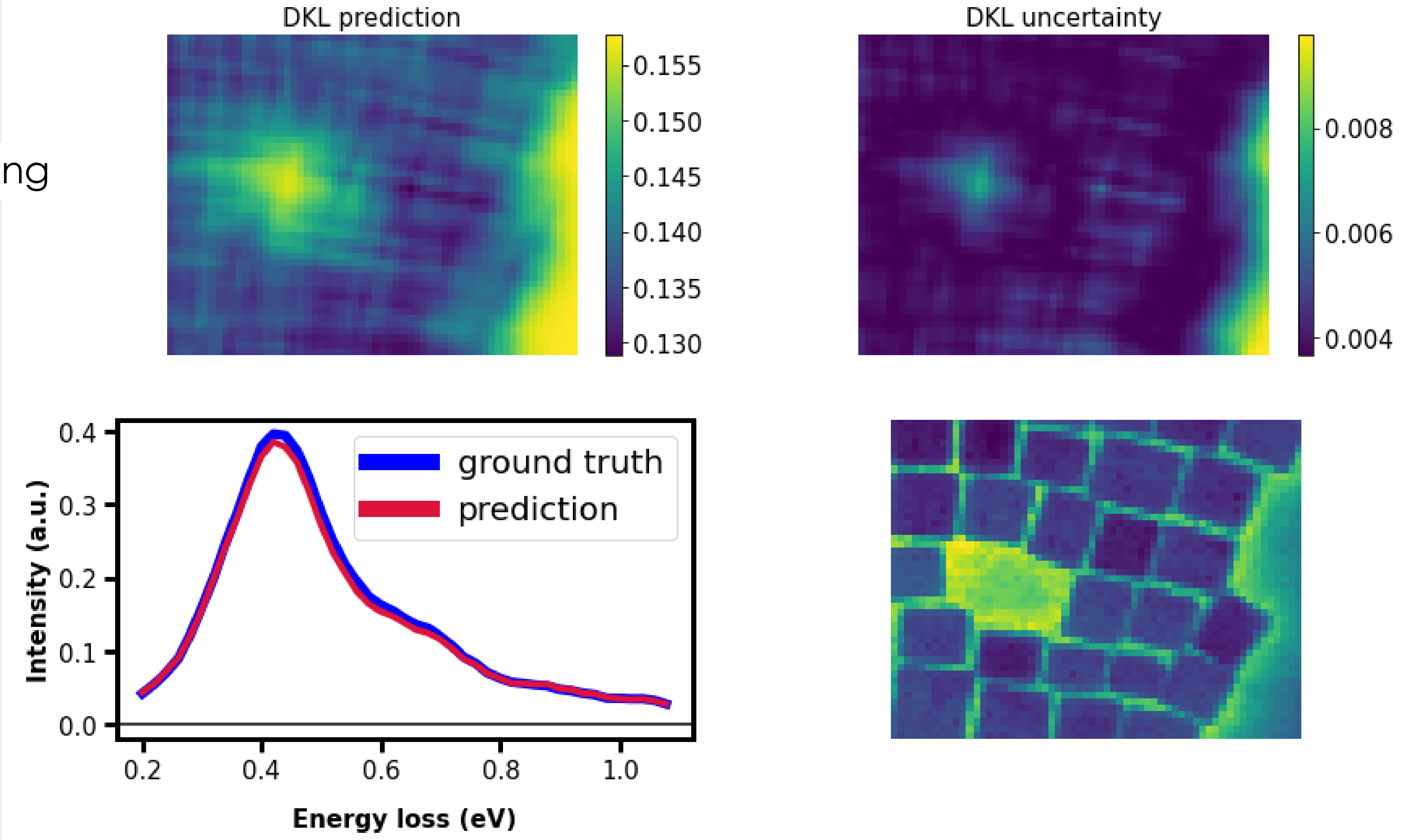




- This is actually "**vector**" DKL which **reconstructs entire 1D spectrum (or 2D image)**
- Is not so amenable for automated experiment, however


Oak Ridge National Laboratory

# 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

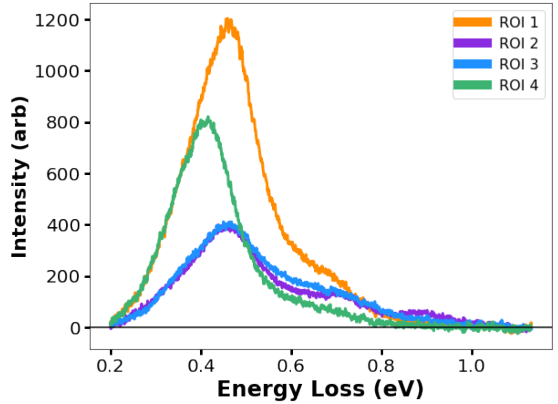
- Expected improvement
- Upper confidence bound
- Uncertainty-based
- Thompson sampling

➢ **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

```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
```



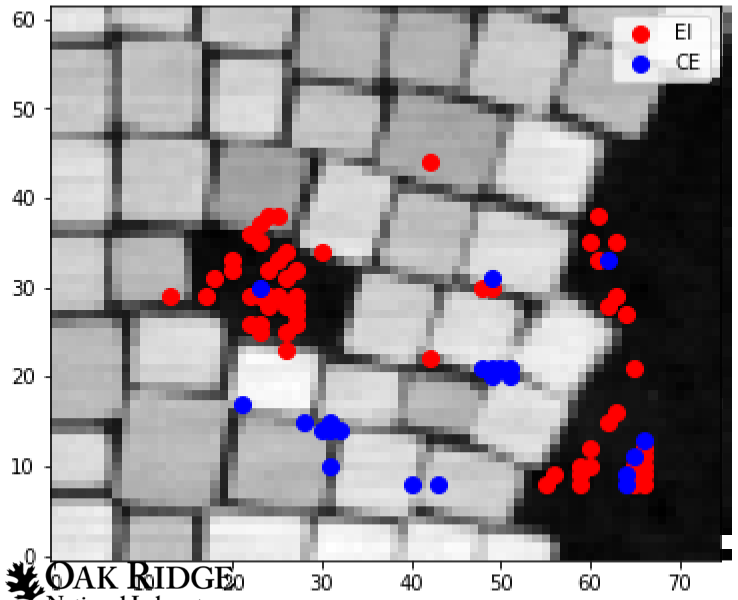

Measure & scalarize

Train

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

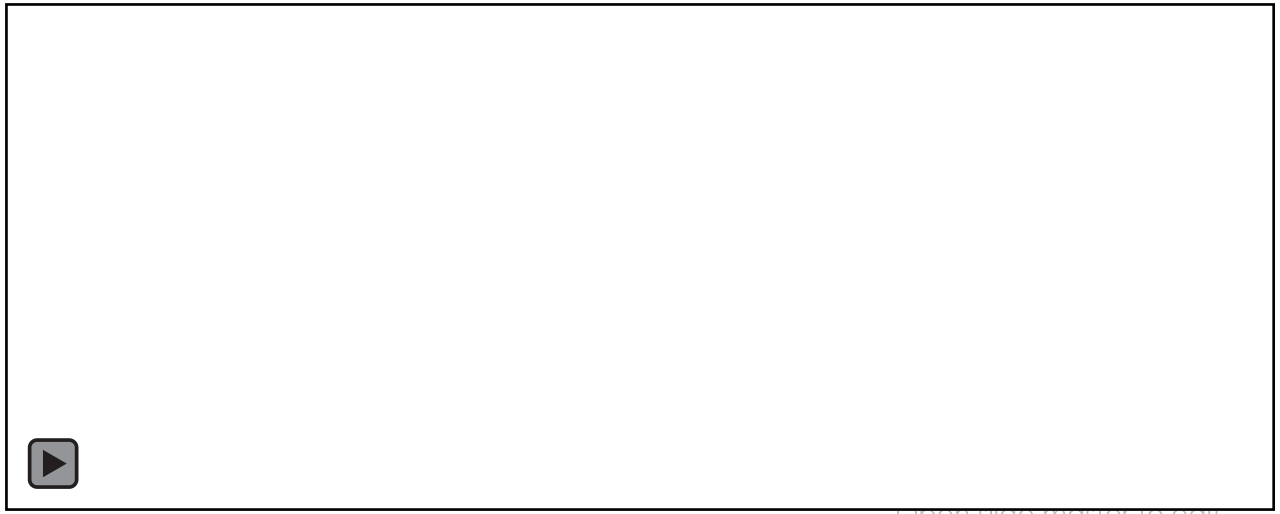
**Next measurement point:** $argmax()$

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



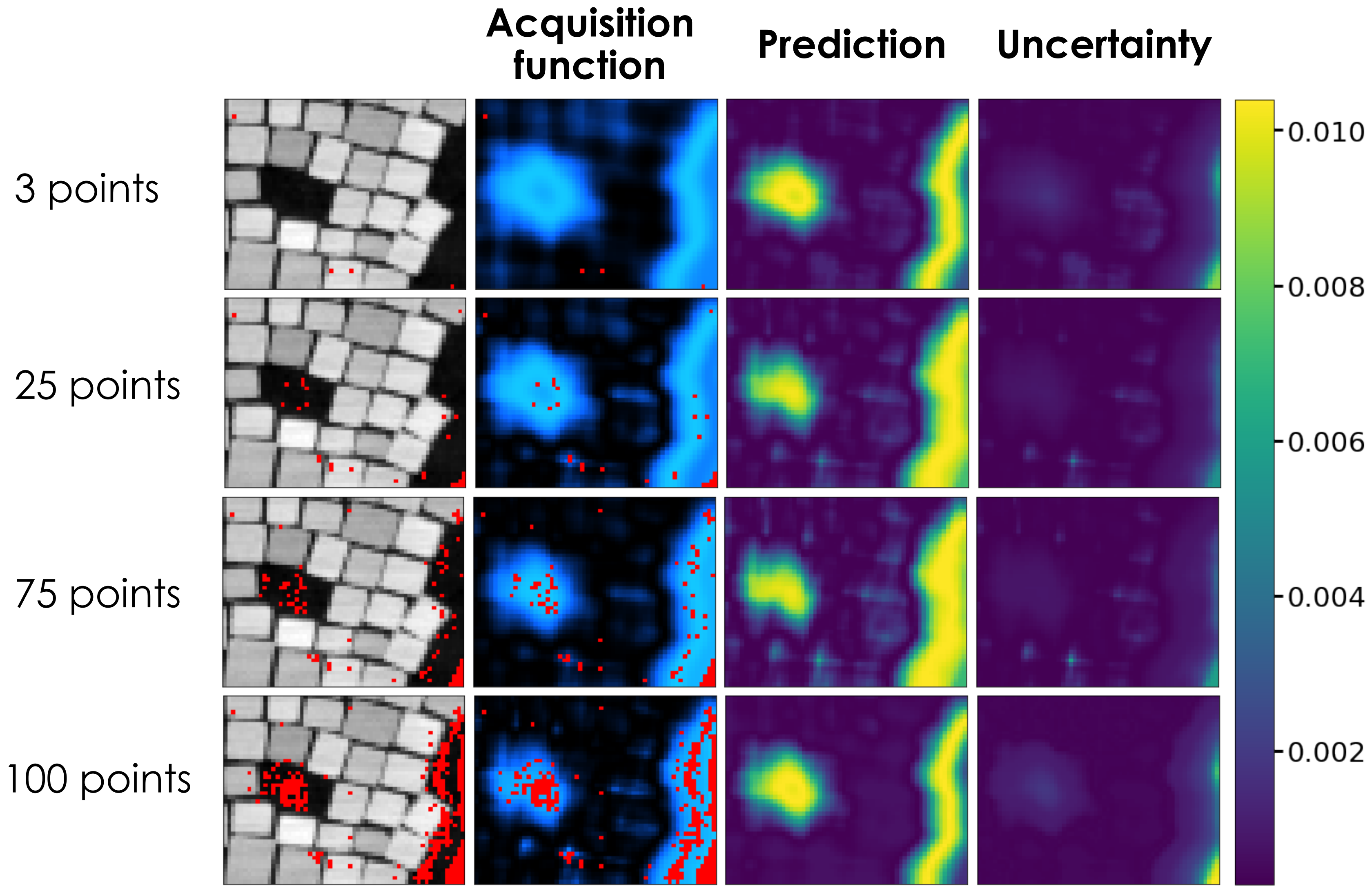

Choice of **acquisition function** affects pathway

**Acquisition function** (Video)



# 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?

(screen record video)

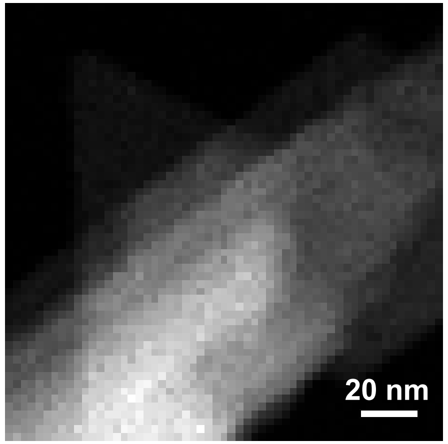
With **access to microscope controls** (Python API), this is deployed on an instrument

- **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 $MnPS_3$**



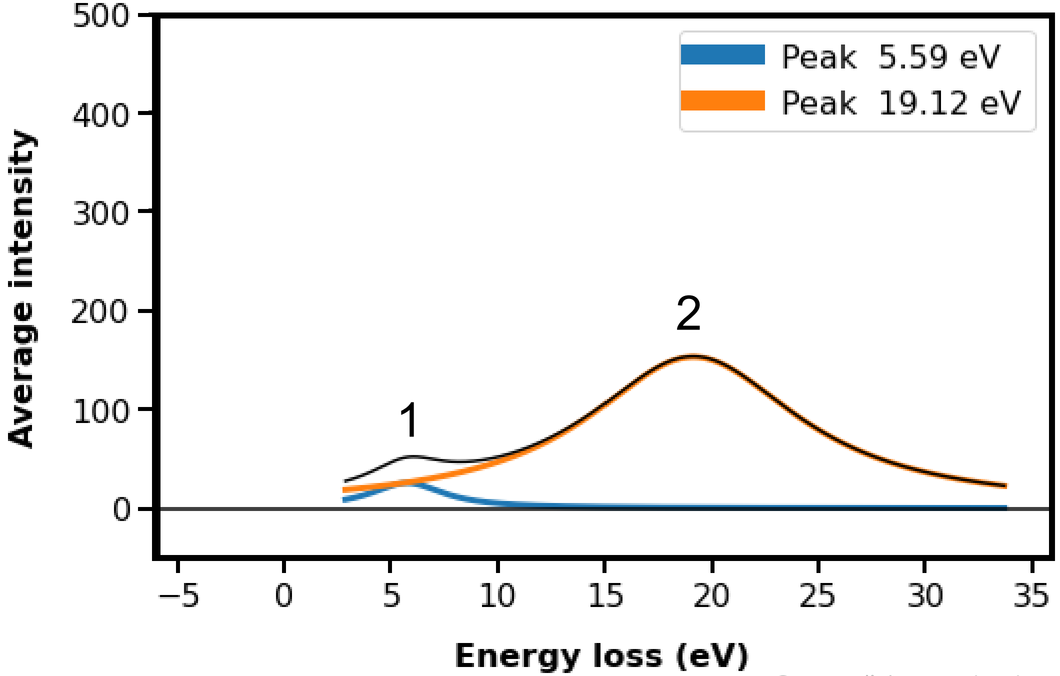

"Acquisition function"

(Video)

HAADF-STEM
+ points visited

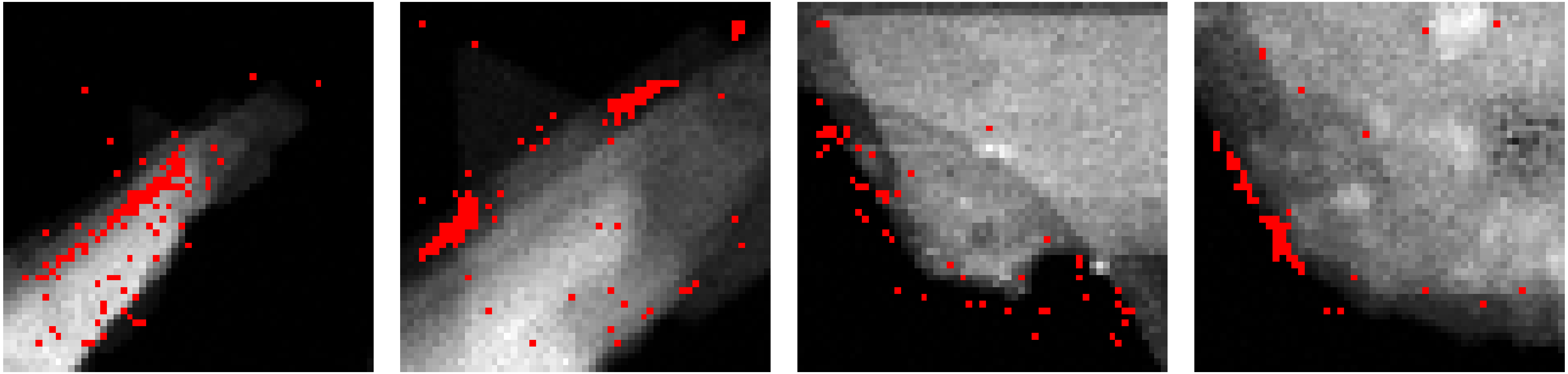
Physics search criteria:

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



# 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

"Acquisition function"

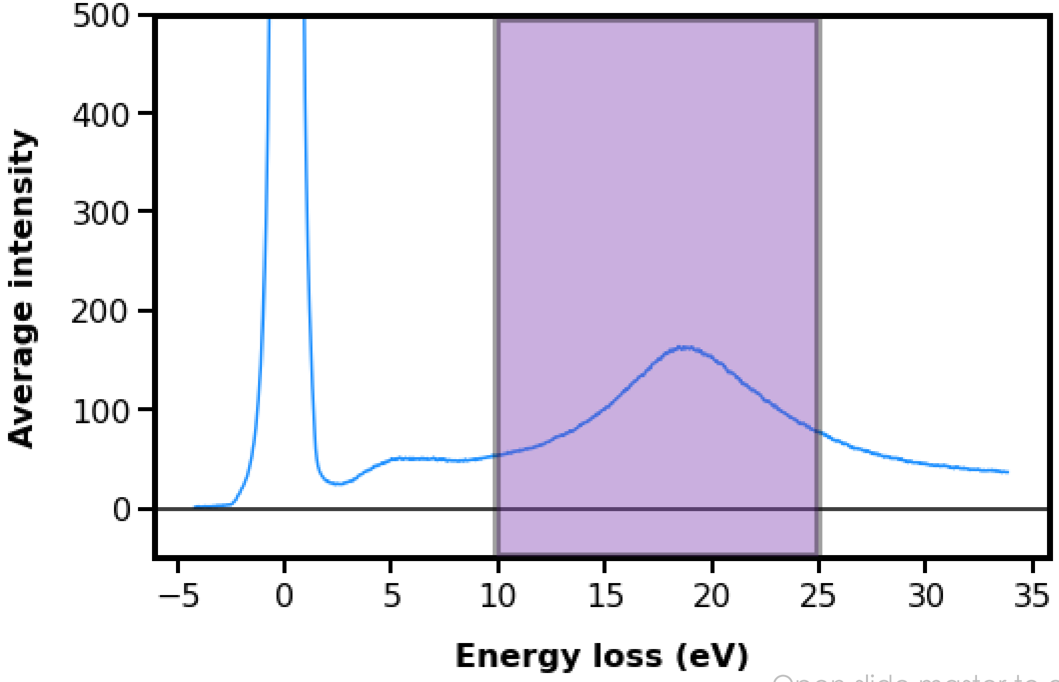
(Video)

HAADF-STEM
+ points visited

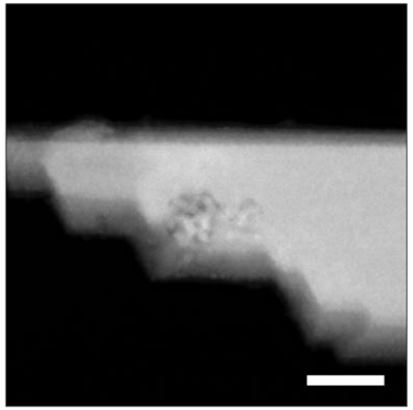
Physics search criteria:

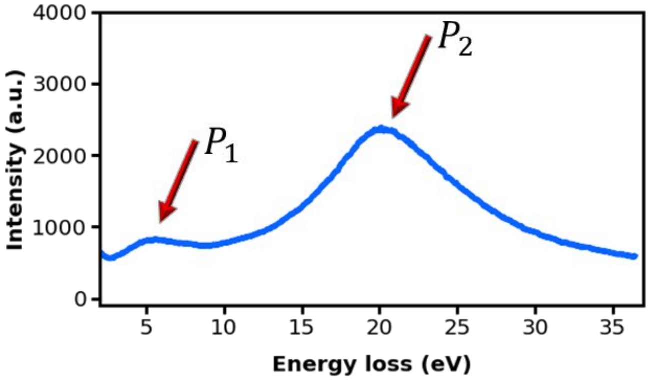
$$\boldsymbol{Maximize(f)}$$

**(Specific peak intensity)**

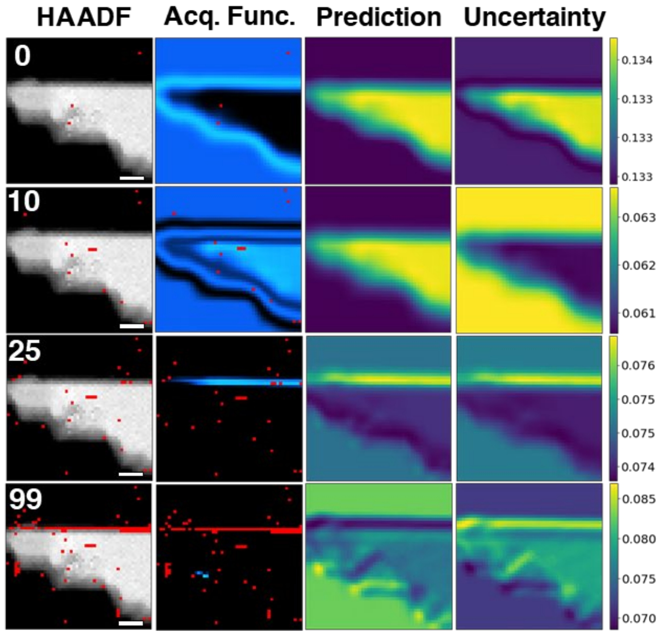


# Summarize these automated experiments..

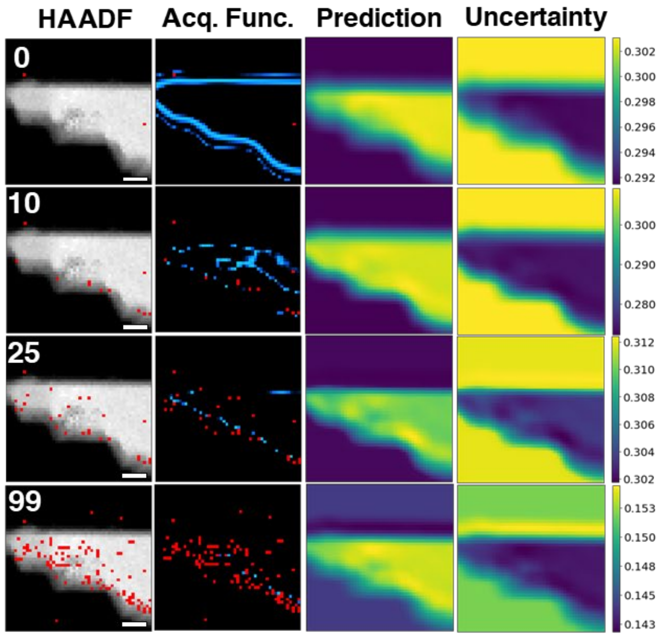





$\max(P_1/P_2)$

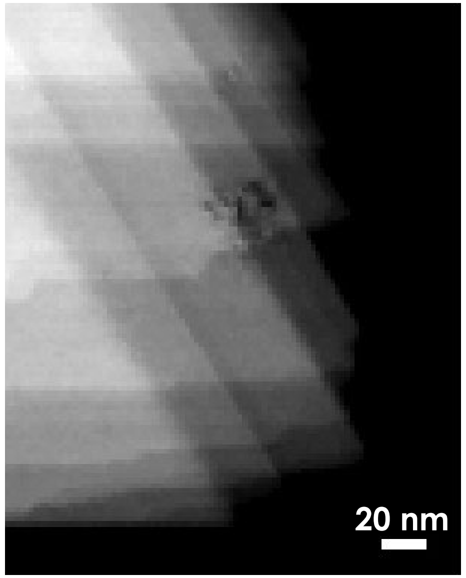


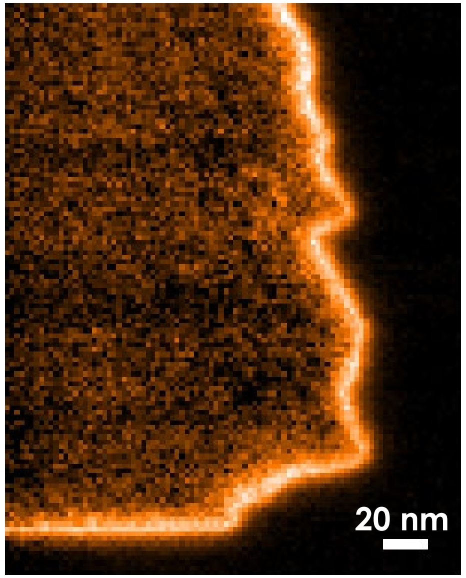

$\max(P_2)$

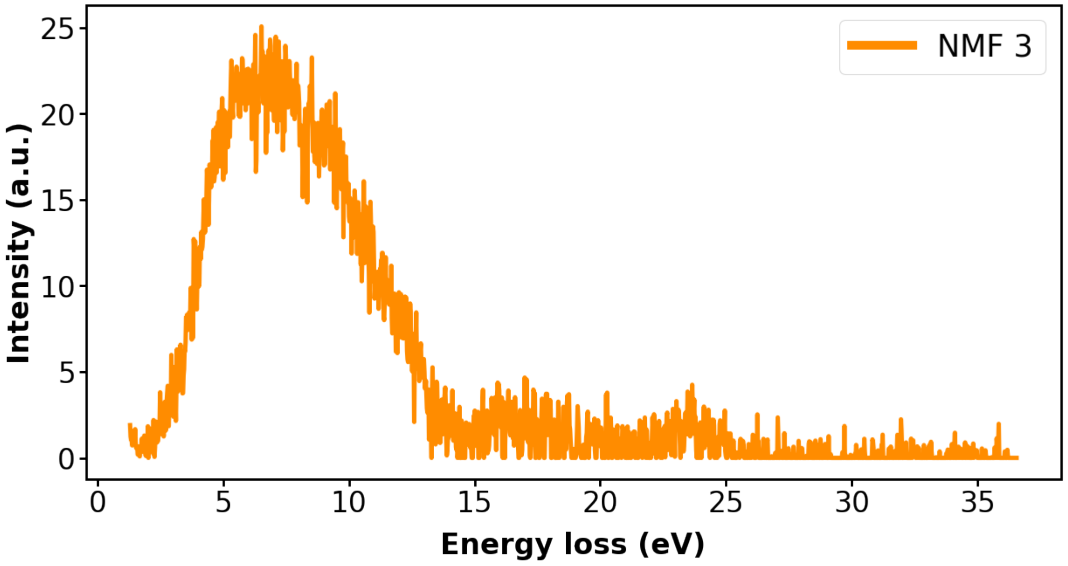


# Confirmation

- 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**

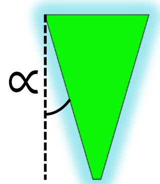
- Electric field
- Potential
- charge density

**NBED disc fitting**

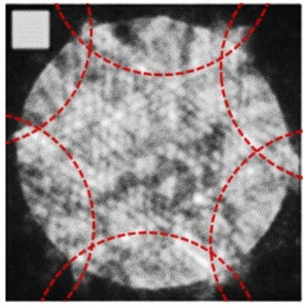
- Strain
- Grains

Require different probe conditions

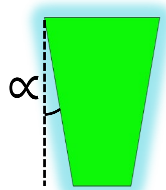
Overlapping diffraction discs



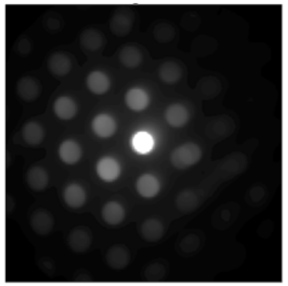

$\propto \approx 30\, mrad$



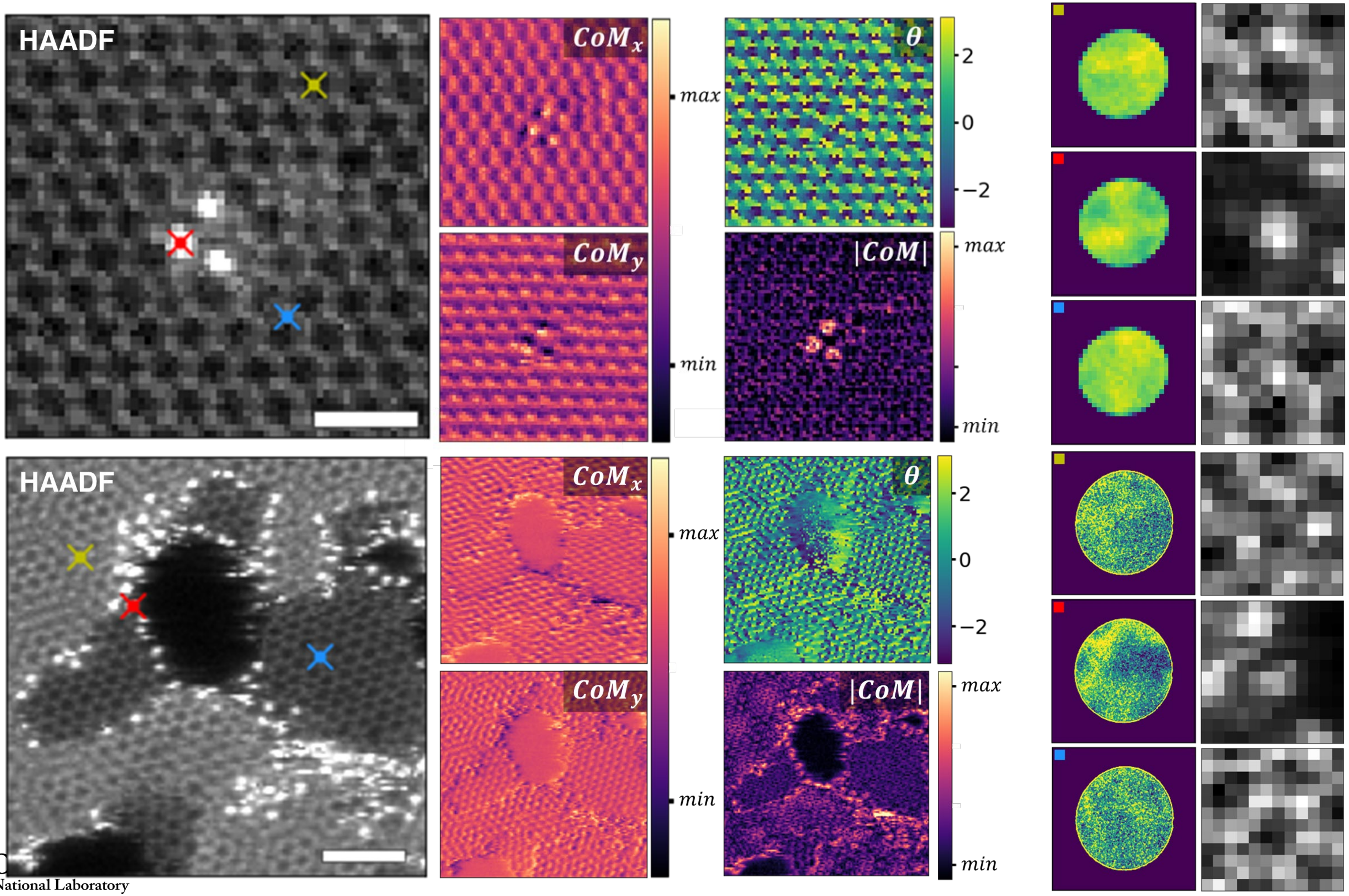
Non-overlapping diffraction discs




$\approx 5\, mrad$



# DKL: 4D STEM – center of mass (DPC)




Do not normally care about the local image information

But again, let's make use of it!



National Laboratory

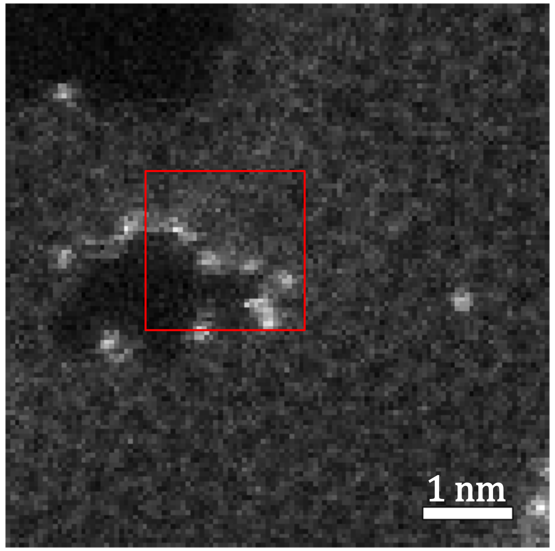
# DKL: 4D STEM – center of mass

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

Acquisition function

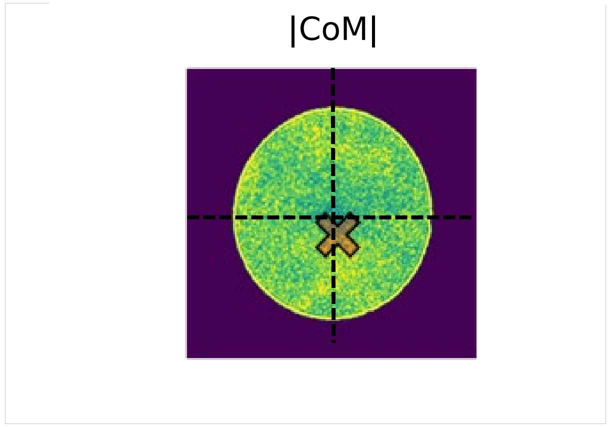
Full structural image

Measured quantity




Prediction map

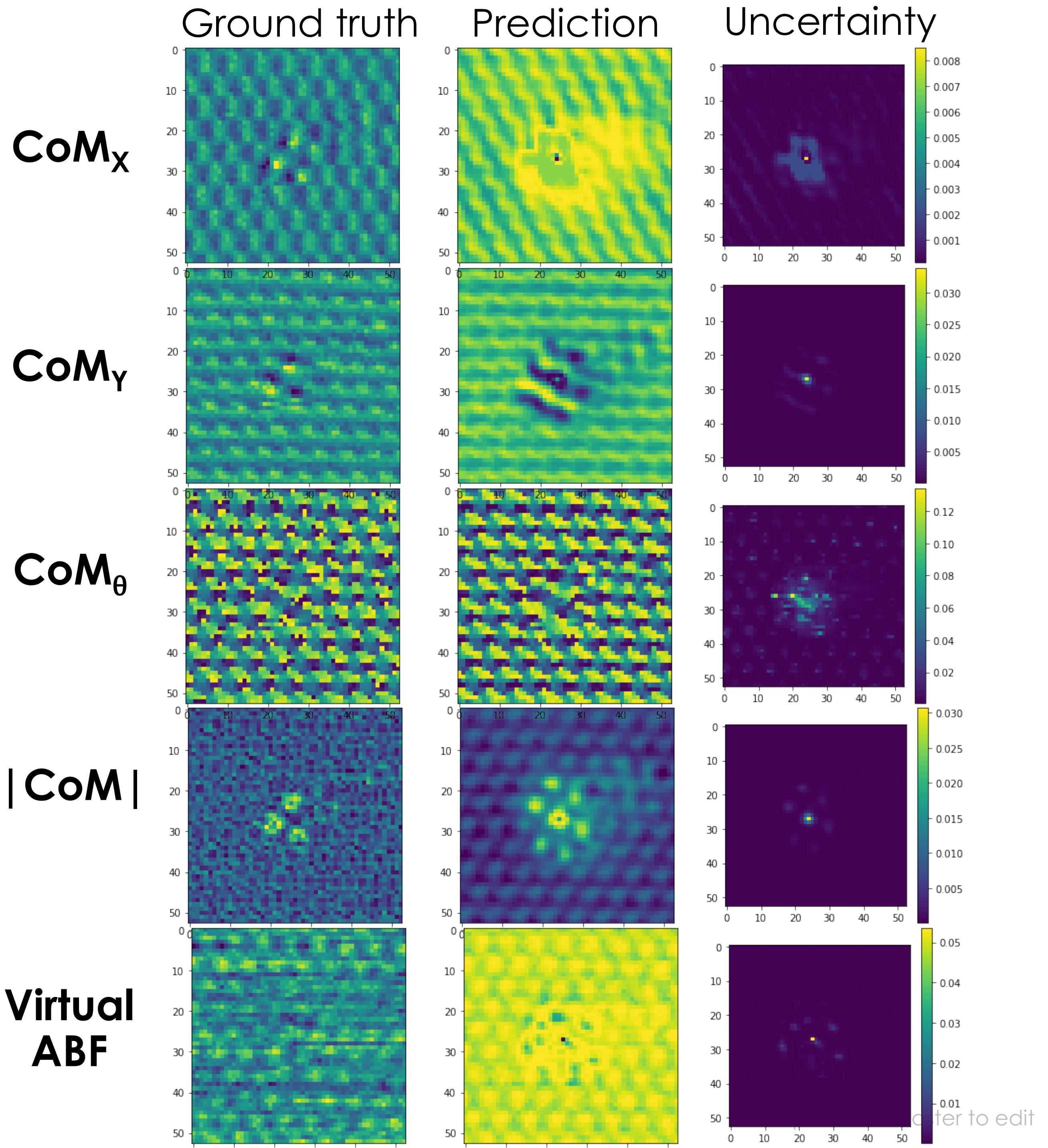
**Optimize physics**
(center of mass, strain, etc.)




Uncertainty map

# DKL: 4D STEM (pre-acquired data)

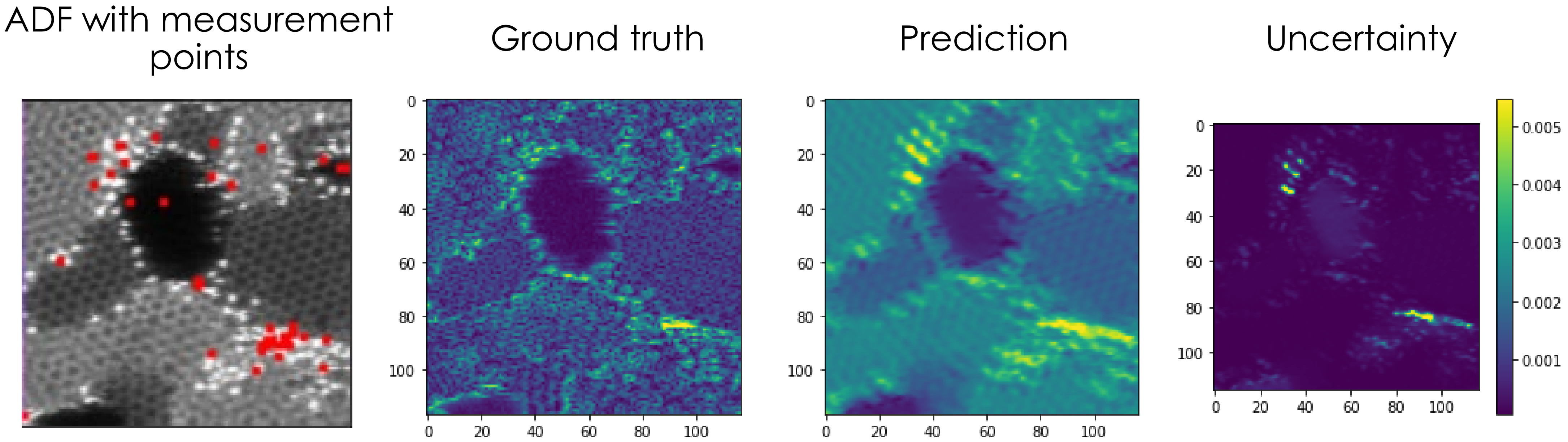
- Choosing **different scalarizers**, does one perform "better"?
- Some interesting (real) artifacts appear in some predictions

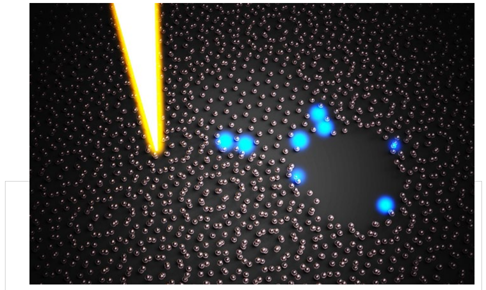


# DKL: 4D STEM active learning (pre-acquired data)

Scalarizer: |CoM|

After 25 measurements



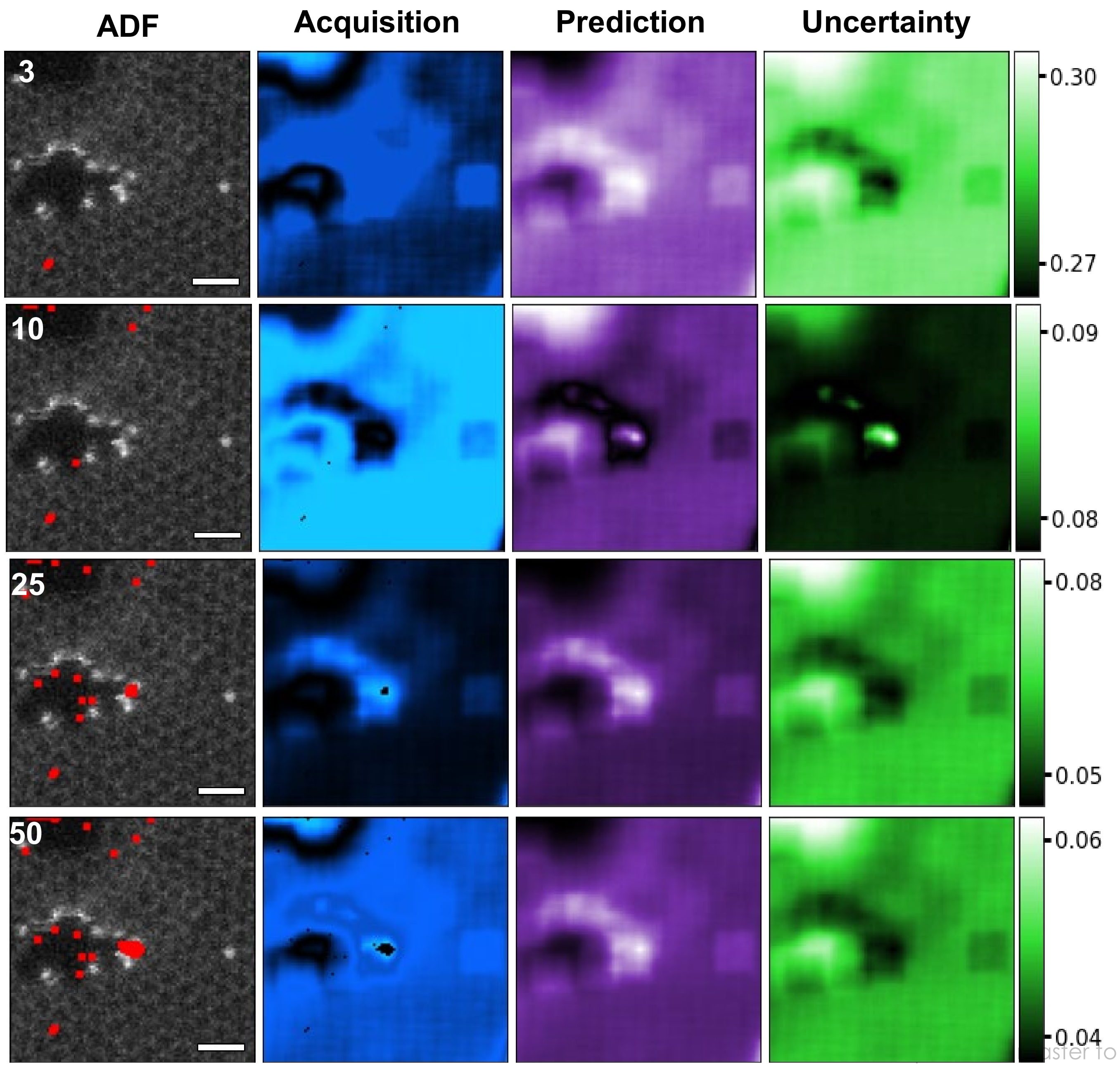


OAK RIDGE National Laboratory

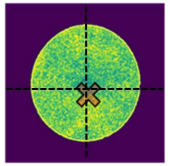
# Active learning: real automated experiment

## Scalarizer: $CoM$ magnitude

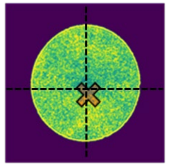
- High uncertainty @ start, but fairly quickly reduces
- Prediction actually doesn't drastically change throughout experiment
    - Structure-property relationship here is fairly rapidly learned
- Note the training can be halted after some criterion is met, making remainder of experiment go much quicker



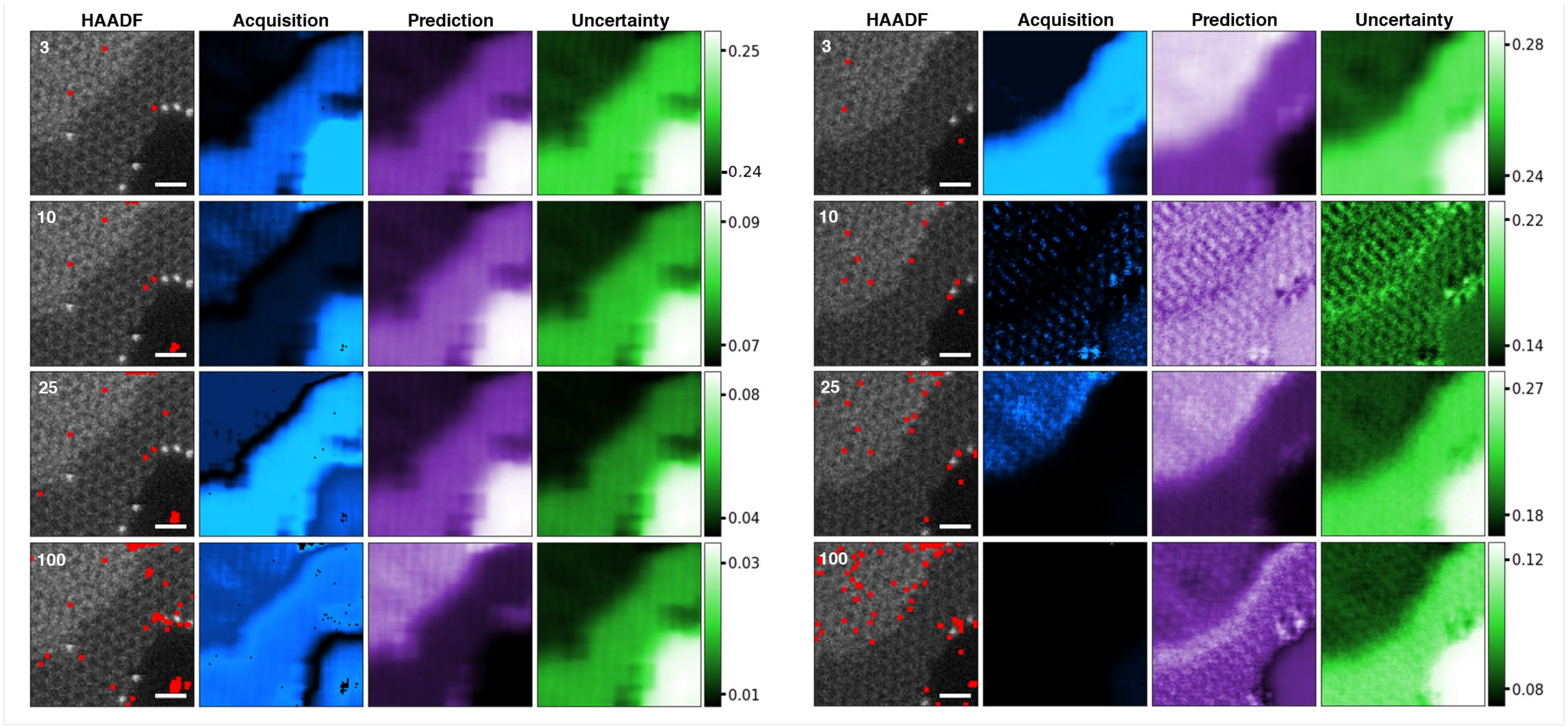
# Effect of physics choice (scalarizer) on acquisition



CoM magnitude



CoM angle






OAK RIDGE National Laboratory

# 4D STEM

**Center of mass**

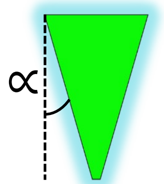
- Electric field
- Potential
- charge density

**NBED disc fitting**

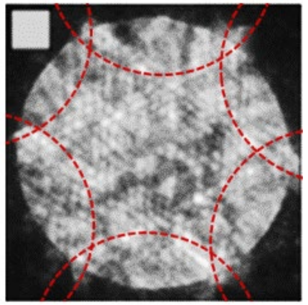
- Strain
- Grains

Require different probe conditions

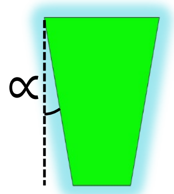
Overlapping diffraction discs



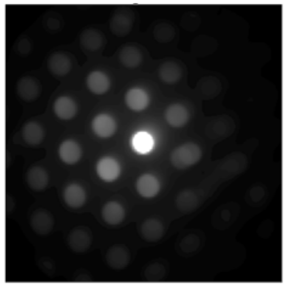

$\propto \approx 30\, mrad$



Non-overlapping diffraction discs

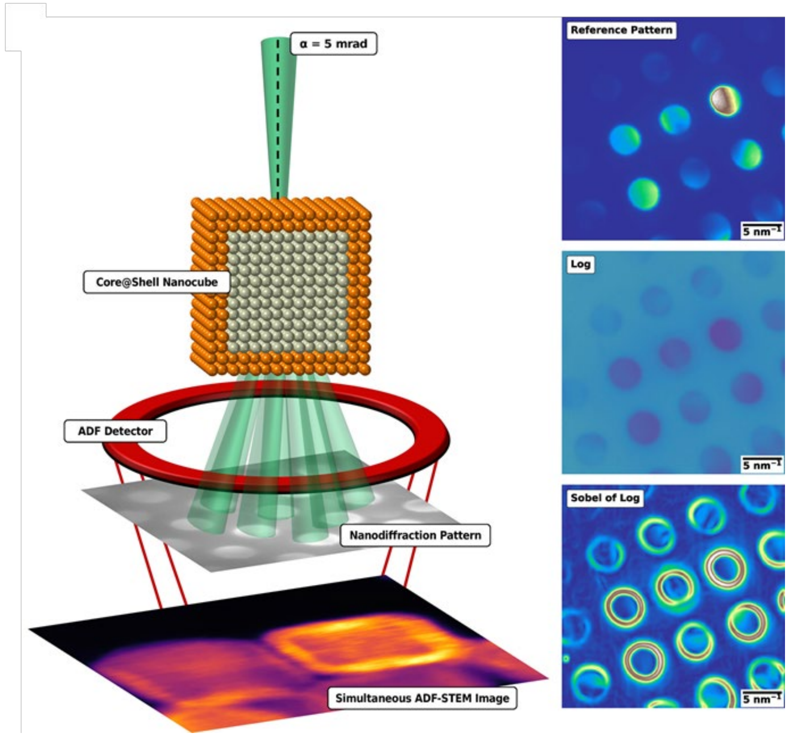



$\approx 5\, mrad$

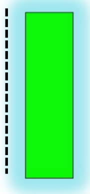


# Measuring strain in the STEM

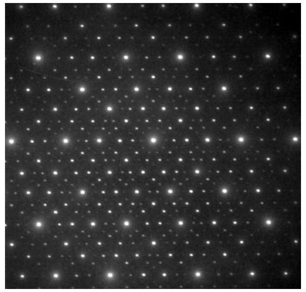
## Using local diffraction




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



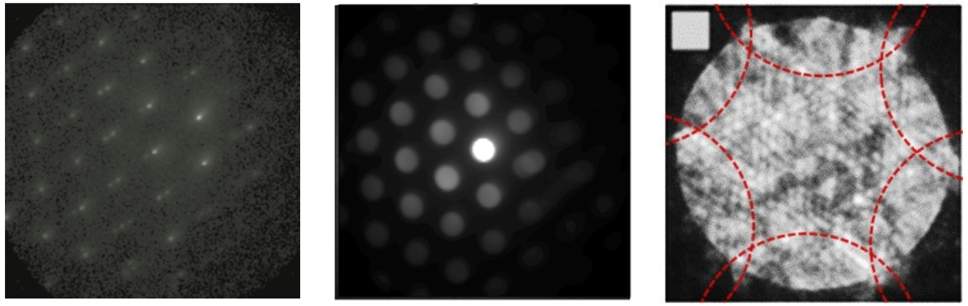
Parallel illumination



(Conventional TEM conditions)

$\approx urads\ to\ few\ mrad$ $\quad$ $\propto \approx 30\ mrad$

Convergent beam



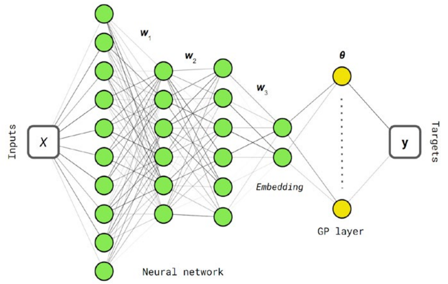
Strain, differential phase contrast (DPC), electric field, atomic potential maps

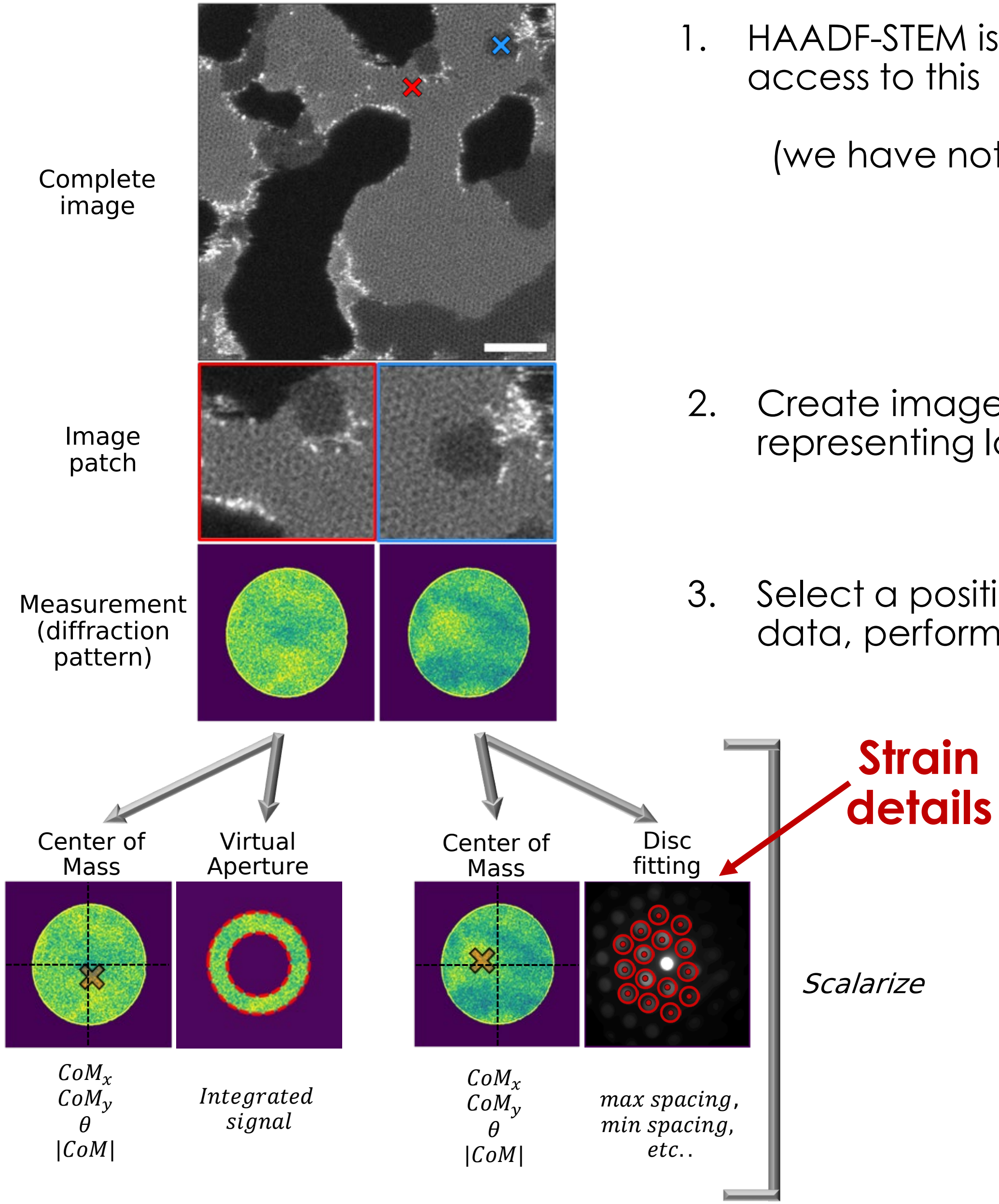
Requirements:

- 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)

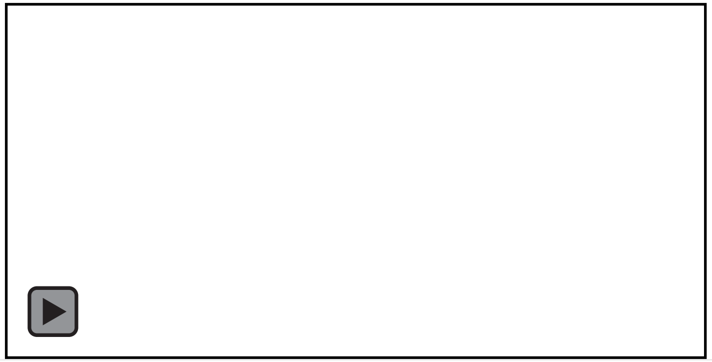
2. Create image patches at each pixel, representing local geometry/structure

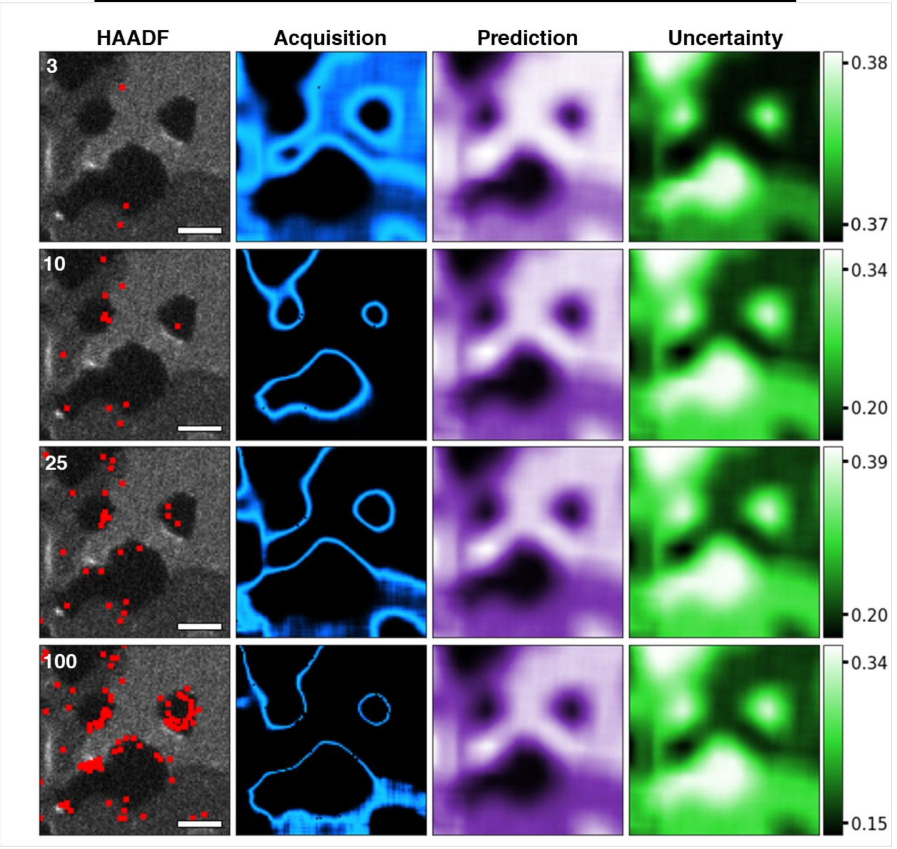
3. Select a position in space to acquire diffraction data, perform computation (scalarize).

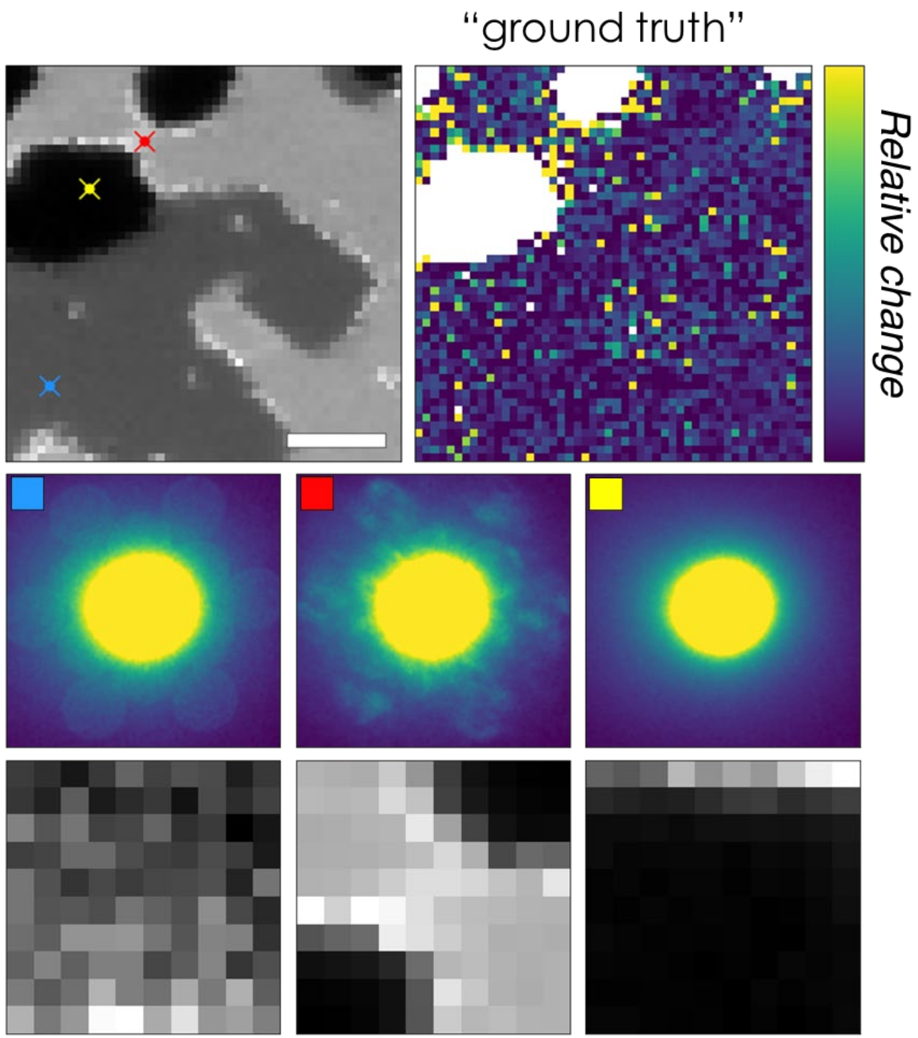
4. Train / update network with **image patch** and **scalarized quantity PAIR**. Repeat 3&4 (automated)

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

# Automated Experiments in 4D STEM: strain





# Hands-on notebook!

# Tips for speed / memory problems

- Change to GPU runtime if not already
- spectralavg = **128**
- dklgp = aoi.models.dklGPR(data_dim, embedim=2, precision="=**single**")