THE UNIVERSITY of TENNESSEE KNOXVILLE

Deep Kernel Learning - I

Sergei V. Kalinin

What have we learned from lectures on GP/BO

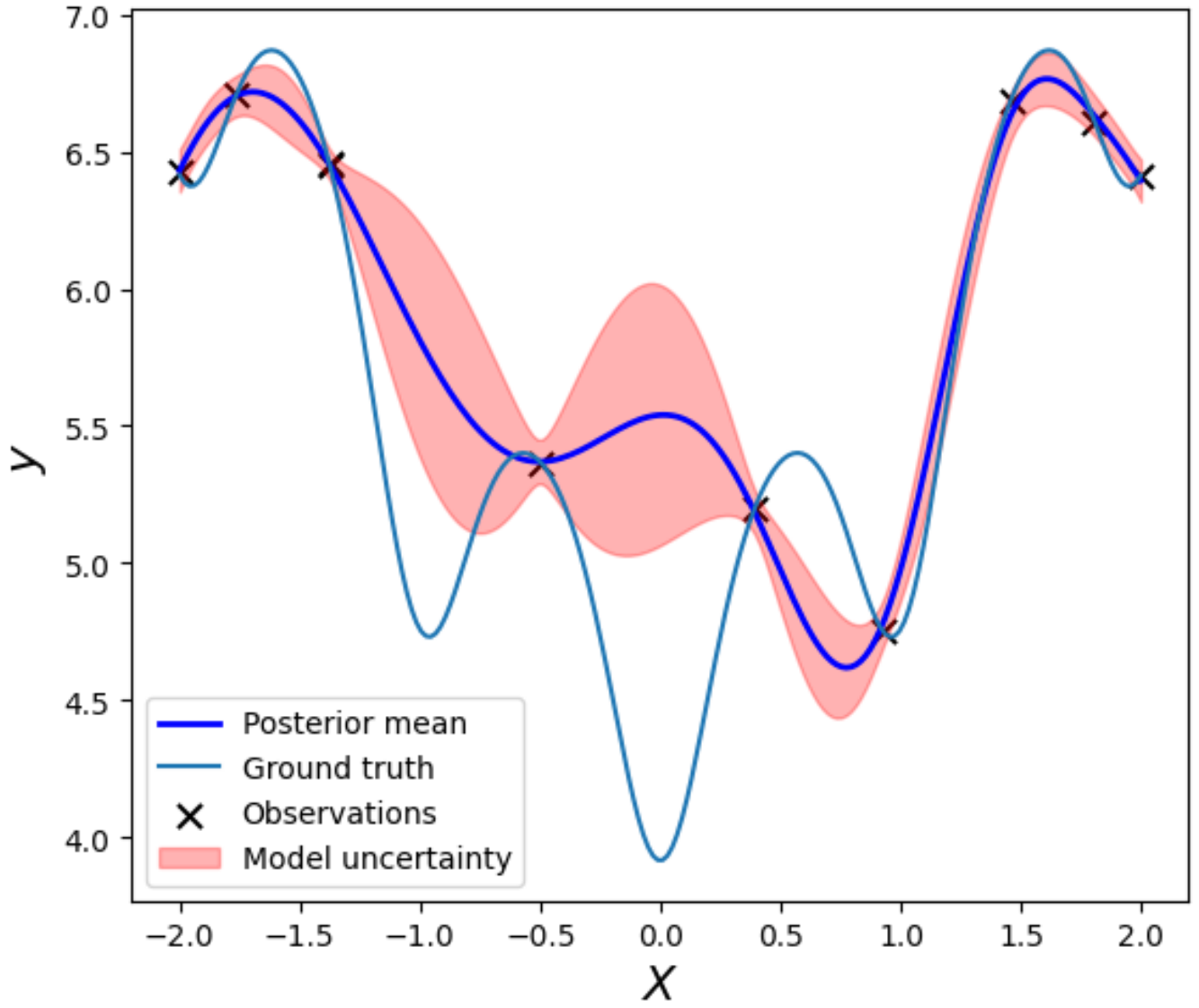
- Gaussian Process

- Kernel and kernel parameters
- Kernel Priors
- Noise Priors
- Mean function and priors

- Posteriors
- Bayesian Inference

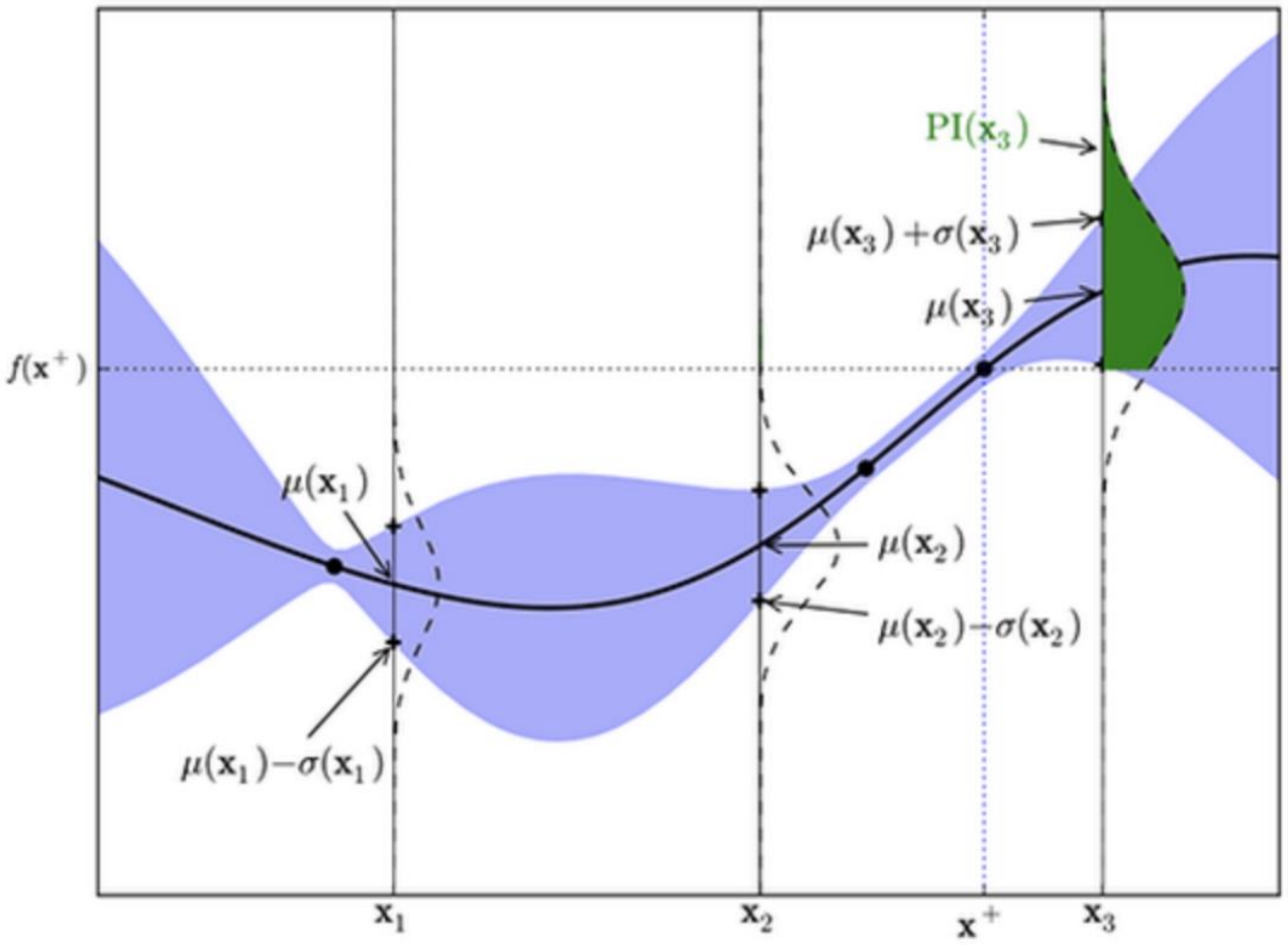
- Bayesian optimization
- Acquisition function

Gaussian Process and Bayesian Optimization



- We have some measurements in space X, and we want to maximize some property f(X).
- We create surrogate model: function and uncertainty based on measurements
- Gaussian Process: purely data driven
- Bayesian Inference: known model and some idea on parameters
- Structured Gaussian Process: physics-derived mean function

Acquisition Functions (Policies)



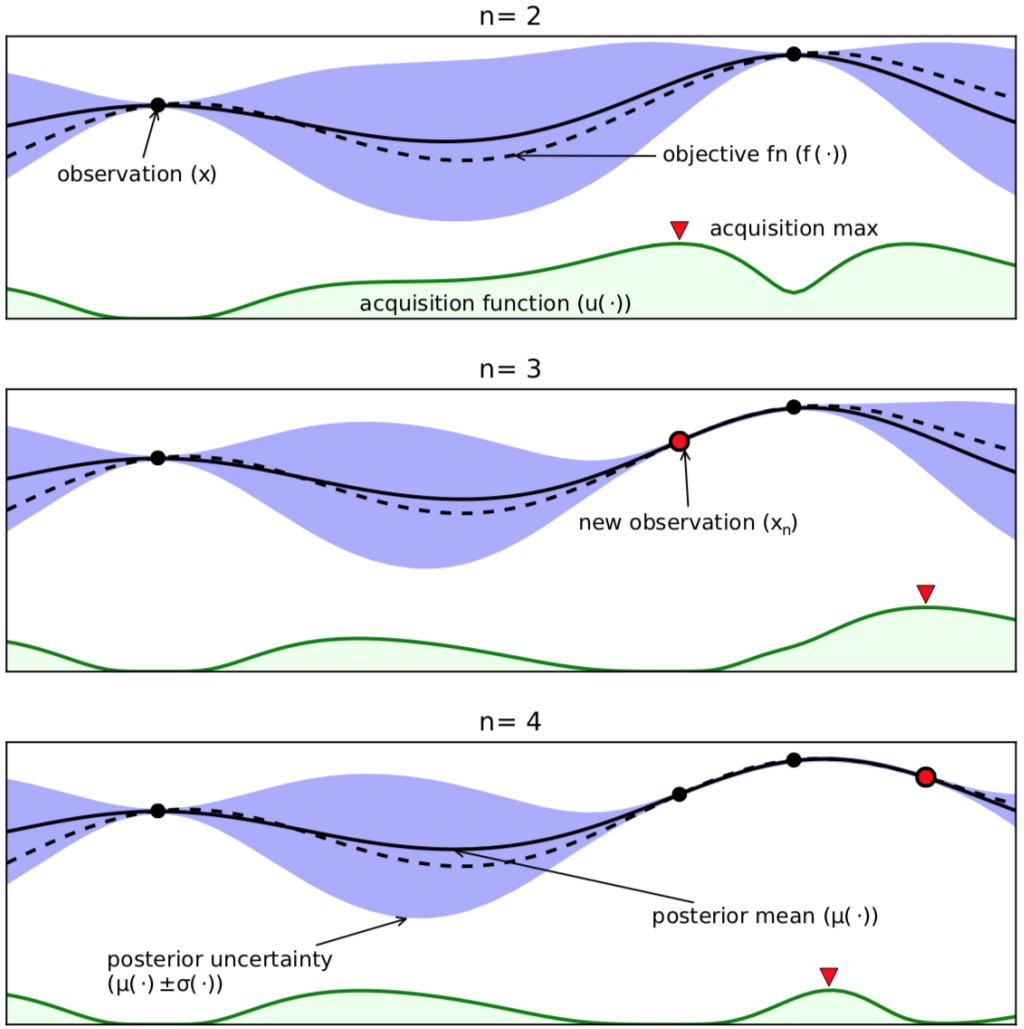
1. **Upper confidence bound:** simplest possible - just take the upper confidence bound from the prediction

2. **Probability of Improvement:** Integral from current functional maximum to upper limit of distribution as test point

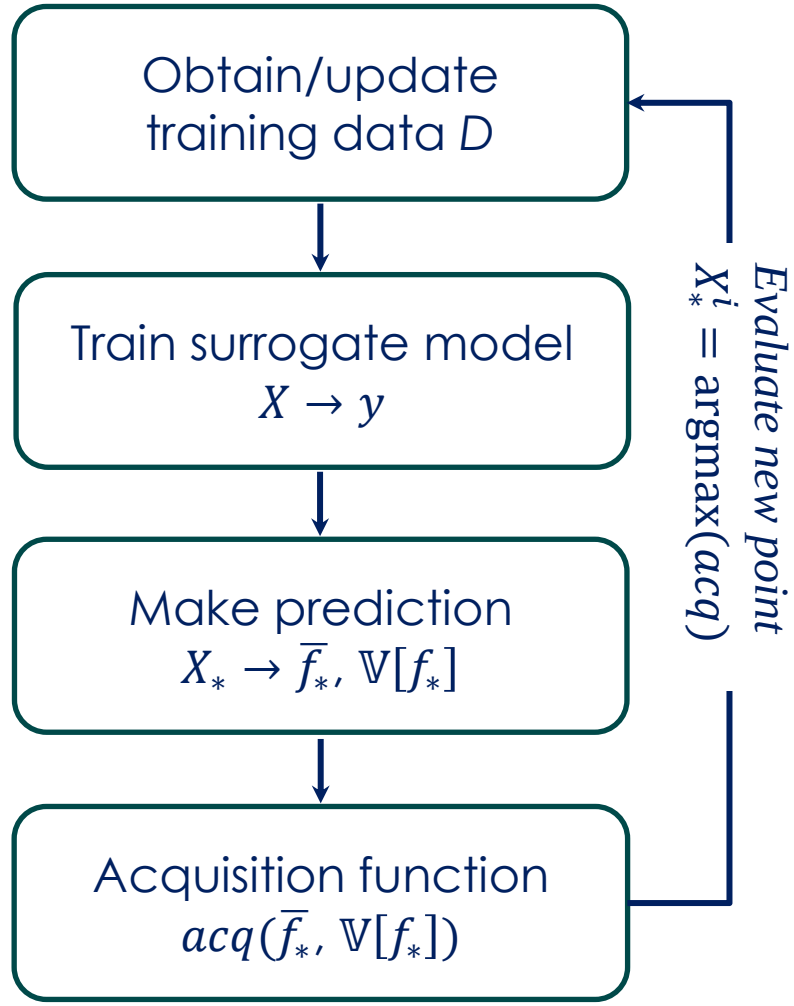
3. **Expected Improvement:** Instead of probability of improvement, we want to maximize the expected increase in the function value

4. **There are (always) more…**

Bayesian Optimization



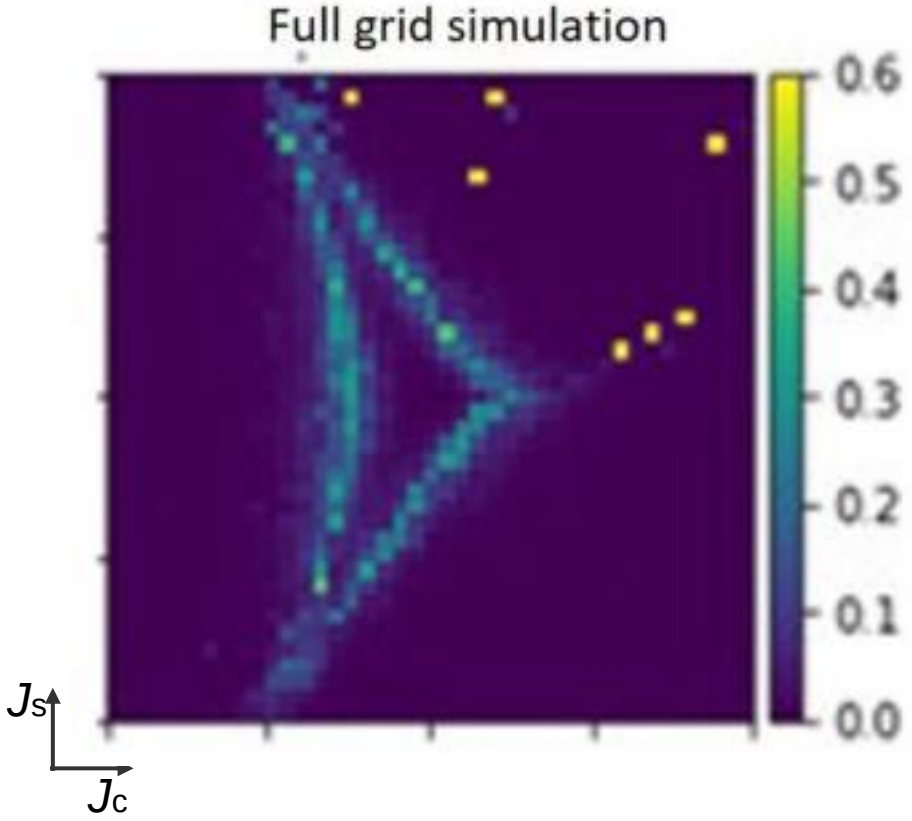
X, y: (sparse) Training data
X_*: New (not yet evaluated) points

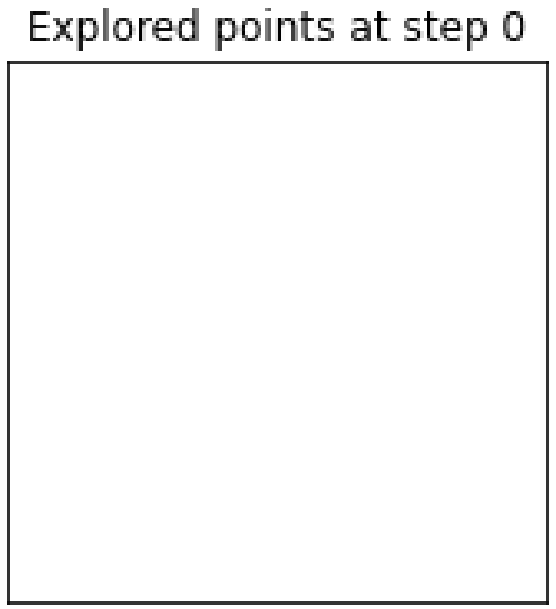


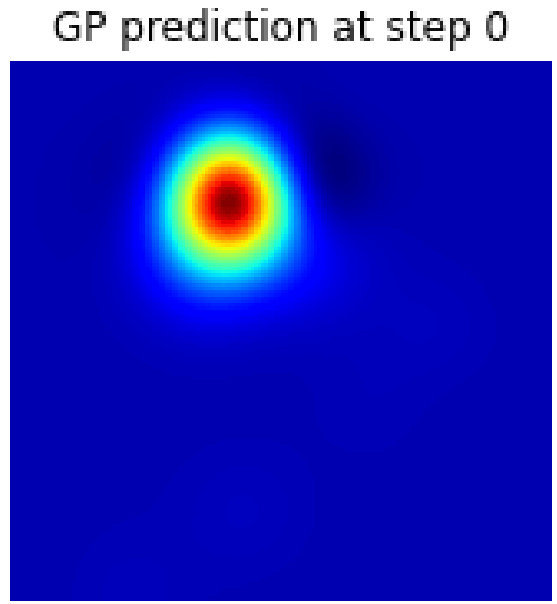
N. de Freitas et al., Taking the Human Out of the Loop: A Review of Bayesian Optimization , *Proceedings of the IEEE* **104**, 148 (2015)

Bayesian Optimization for Physical Discovery

Discovering regions in which the heat capacity is maximized in NNN Ising model

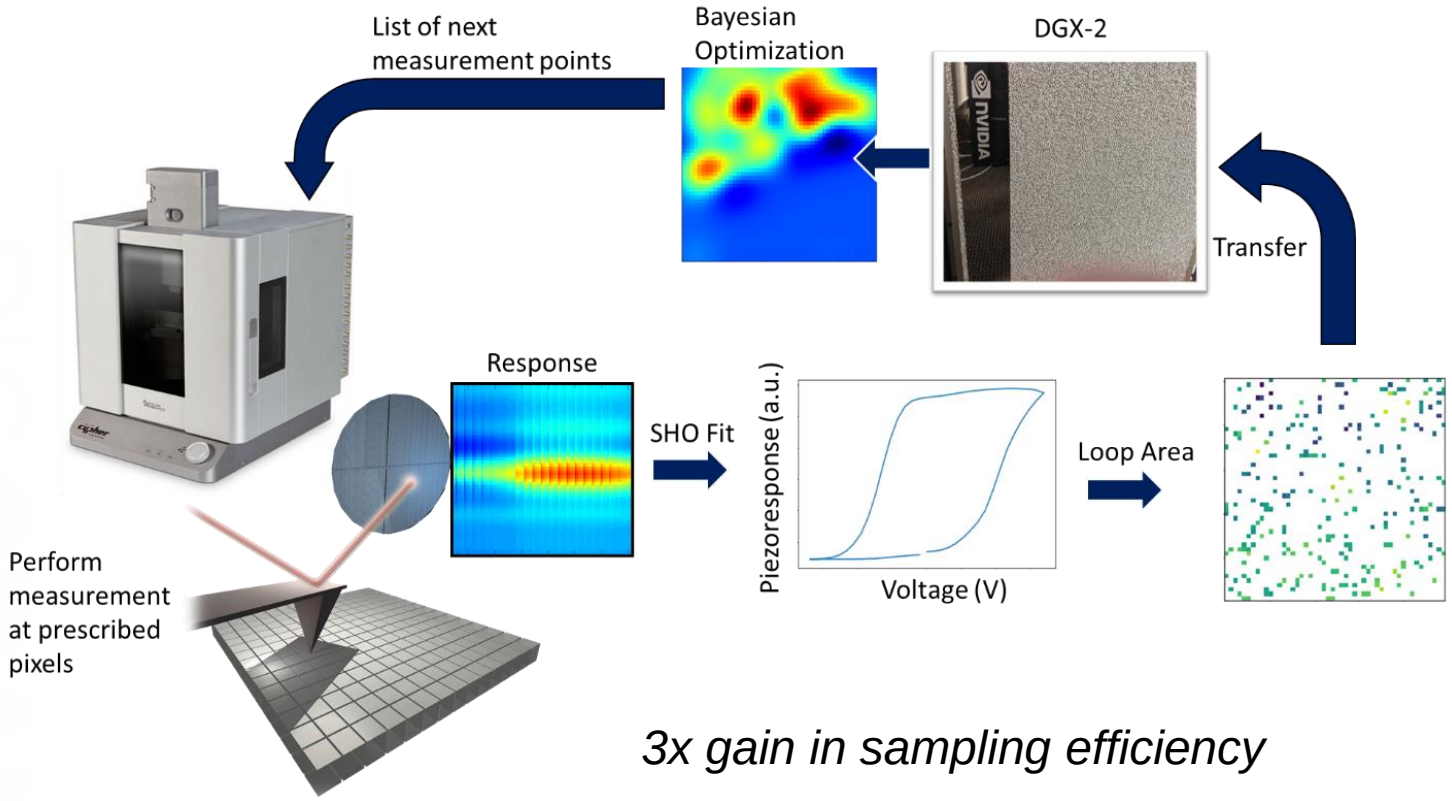




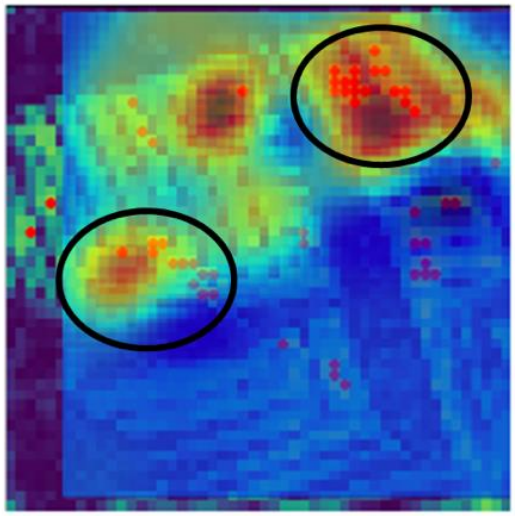


BO for Self-Driving Microscope

First implementation of self-driving microscope: 2020



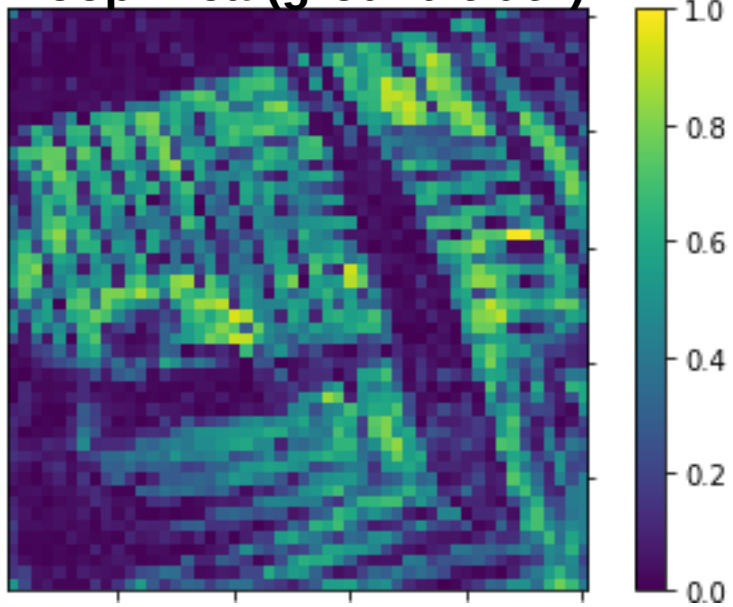
Comparison with "ground truth"



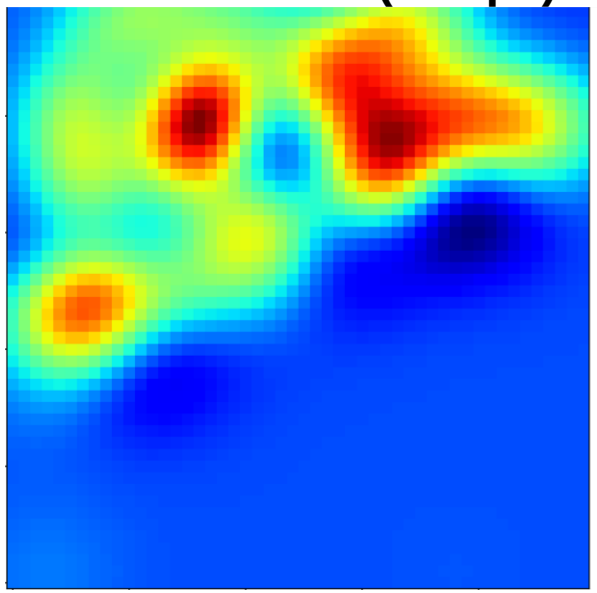
3x gain in sampling efficiency

R. K. Vasudevan, K. Kelley, H. Funakubo, S. Jesse, S. V. Kalinin, M. Ziatdinov,
ACS Nano *(2021) https://doi.org/10.1021/acsnano.0c10239*

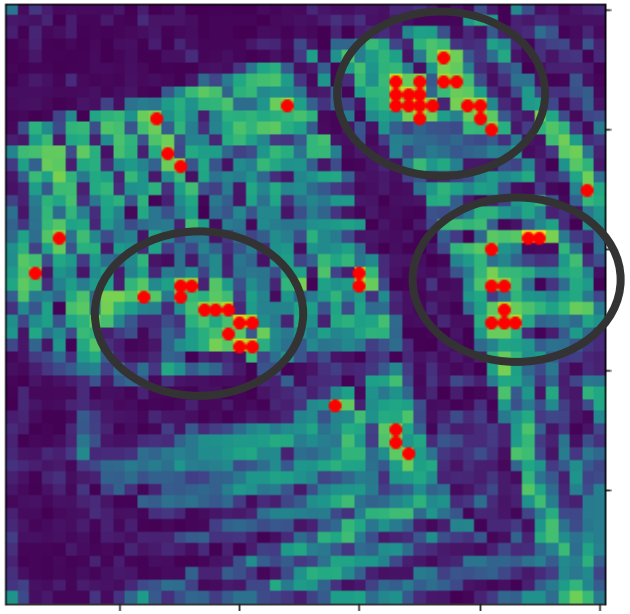
Loop Area (ground truth)



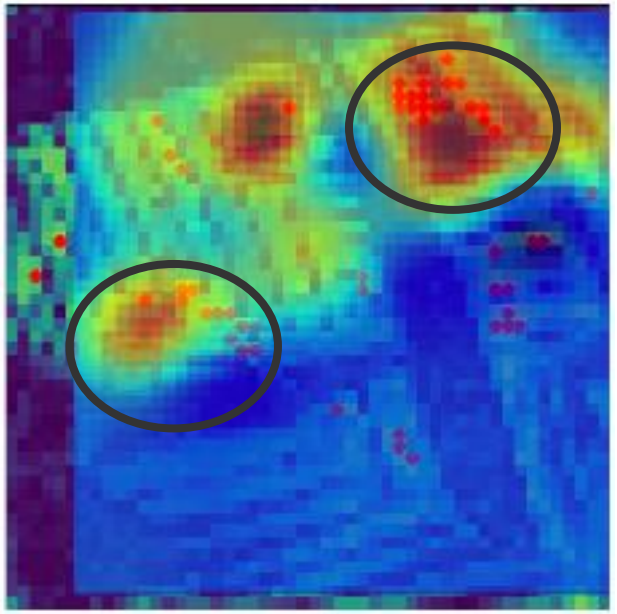
GP Prediction (400 px)



Loop Area >0.8



Overlaid



arXiv:2103.12165

arXiv:2011.13050

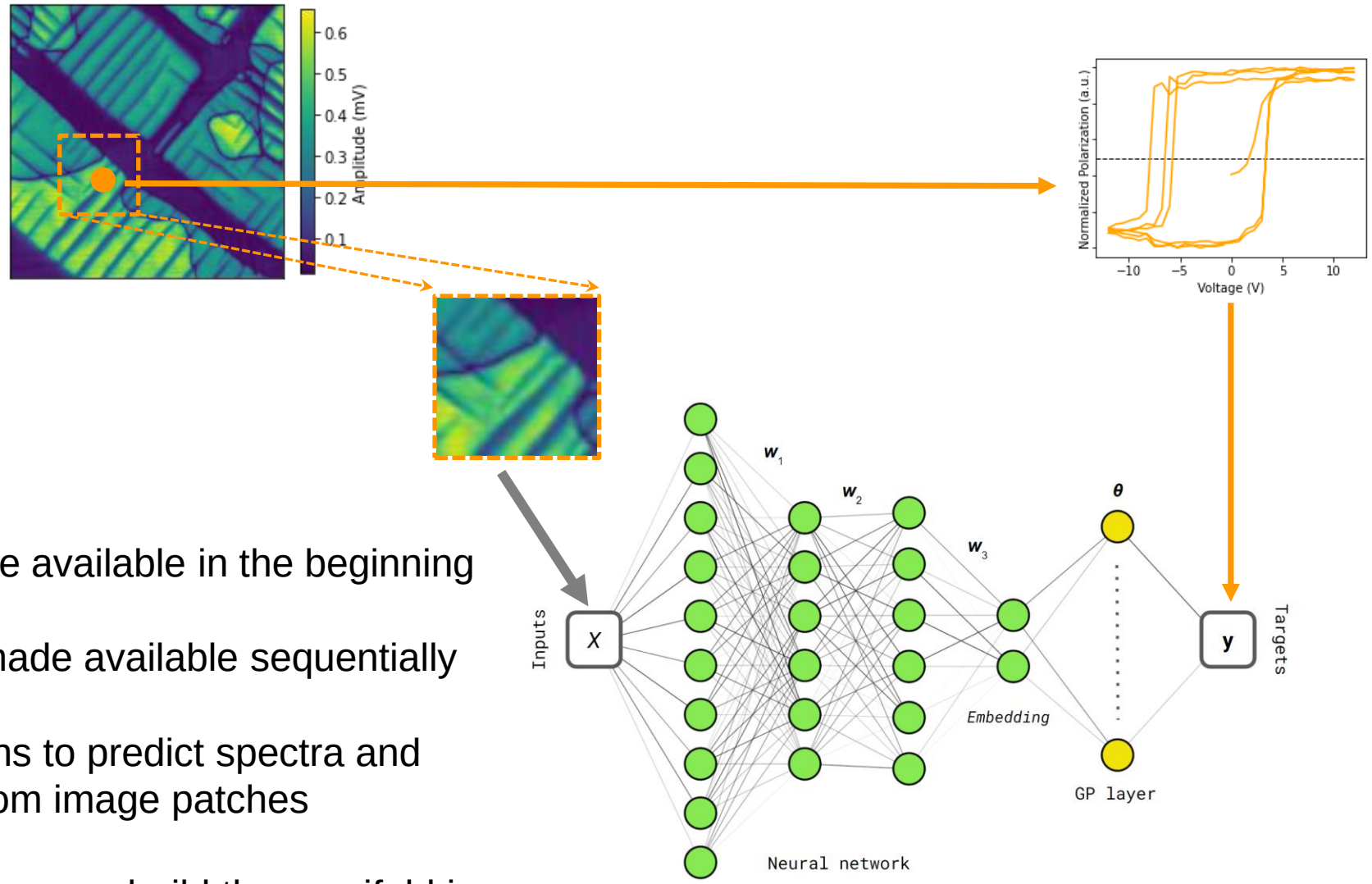
The application of simple data-driven GP for real world scenarios did not work particularly well.

What is the limitation of the GP/BO?

1. Works only in low-dimensional spaces
2. The correlations are defined by the kernel function (very limiting)
3. We do not use any knowledge about physics of the system
4. We do not use cheap information available during the experiment (proxies)

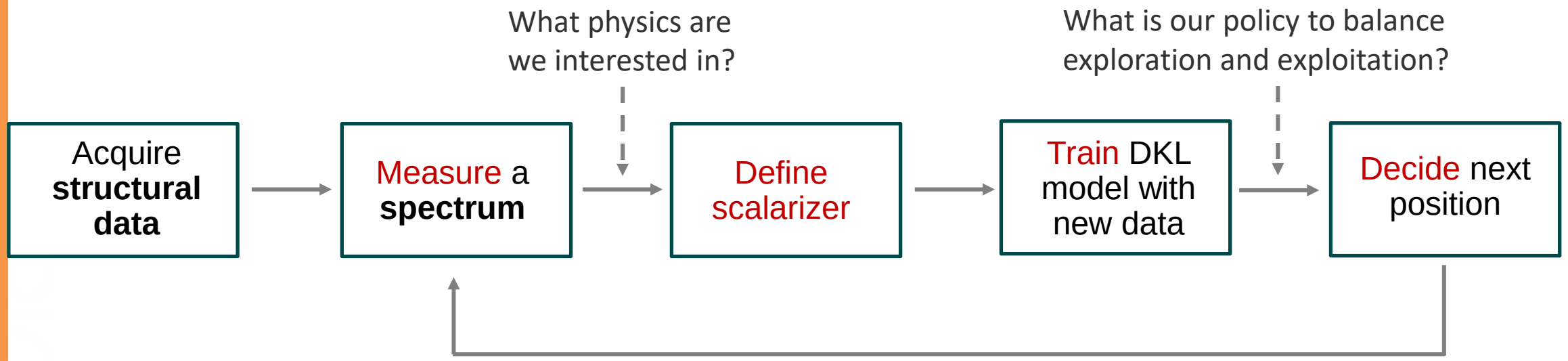
Can we somehow make high dimensional space low-D?

Deep Kernel Learning



- All patches are available in the beginning
- Spectra are made available sequentially
- The DKL learns to predict spectra and uncertainty from image patches
- Key aspect here: we build the manifold in latent space ***dynamically*** (unlike VAE)

Deep Kernel Learning based BO



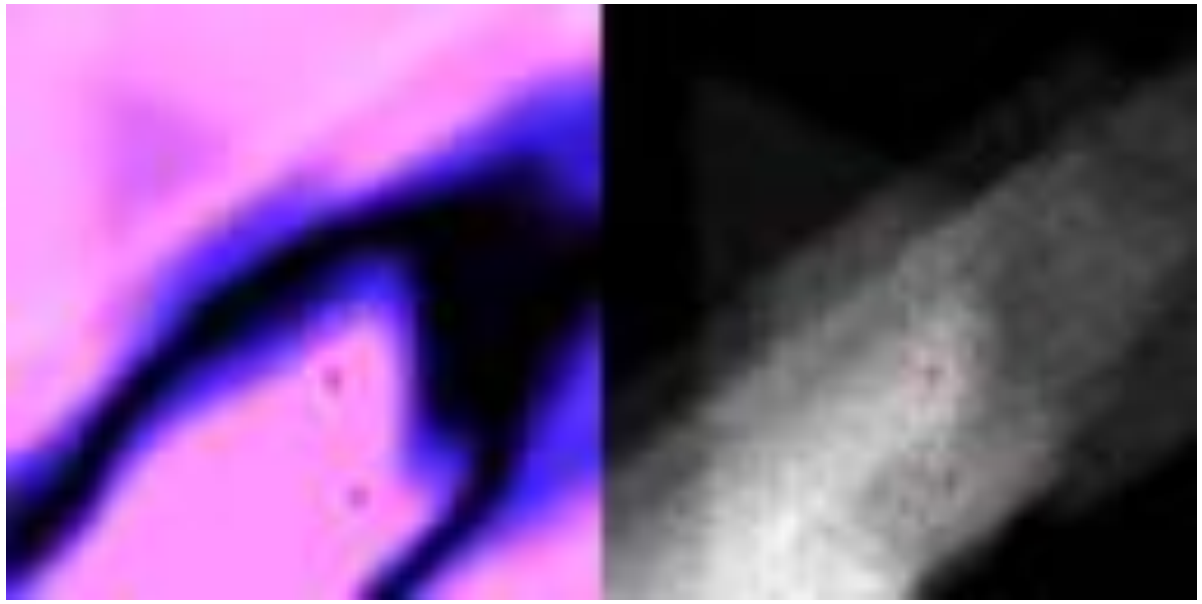
Key concepts:

- **Scalarizer:** (any) function that transforms spectrum into measure of interest. Can be integration over interval, parameters of a peak fit, ration of peaks, or more complex analysis
- **Experimental trace:** collection of image patches and associated spectra acquired during experiment. Note that we collect spectra, not only scalarizers

Discovering Regions with Interesting Physics

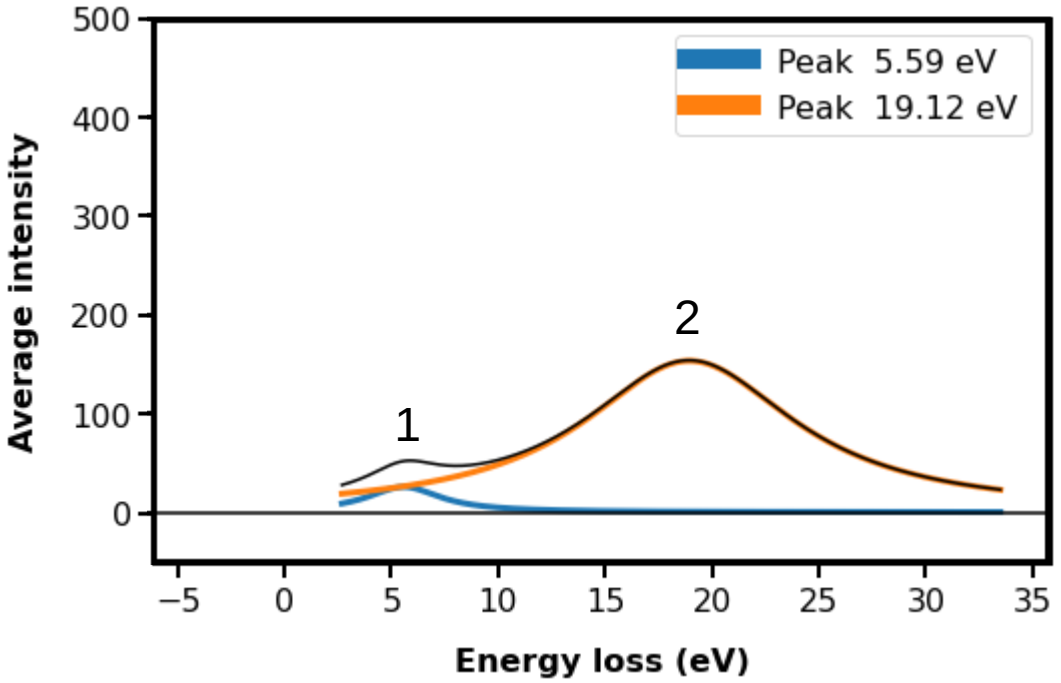
- **Discovering physics in a "new" material $MnPS_3$**
- **Curve fitting** to help enforce physical processes

"Acquisition function" HAADF-STEM

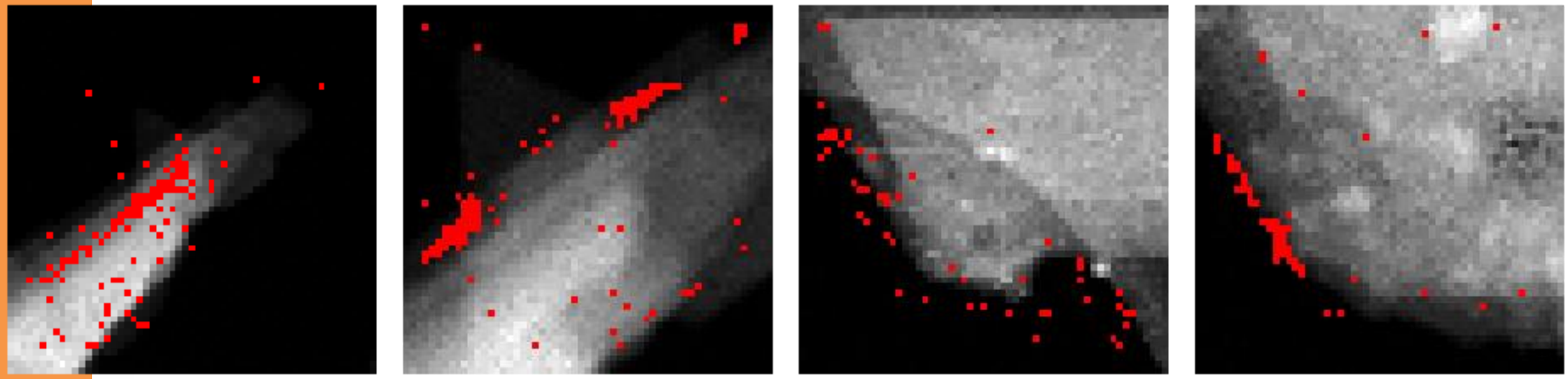


Physics search criteria:

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



More Examples of Physics Discovery



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

Changing the Criterion

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

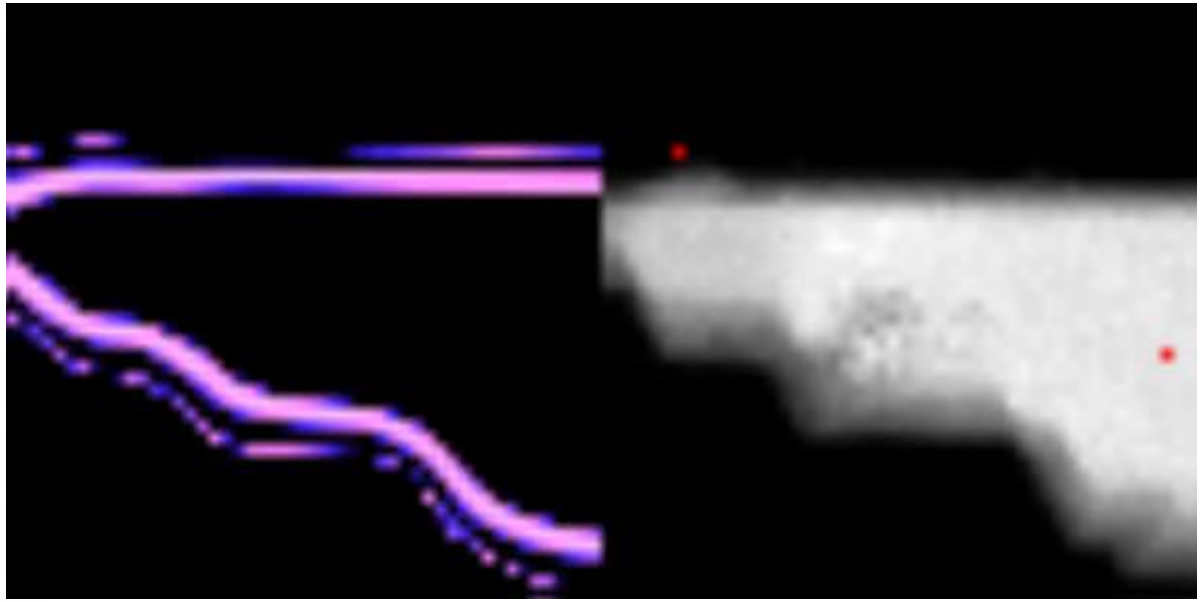
Physics search criteria:

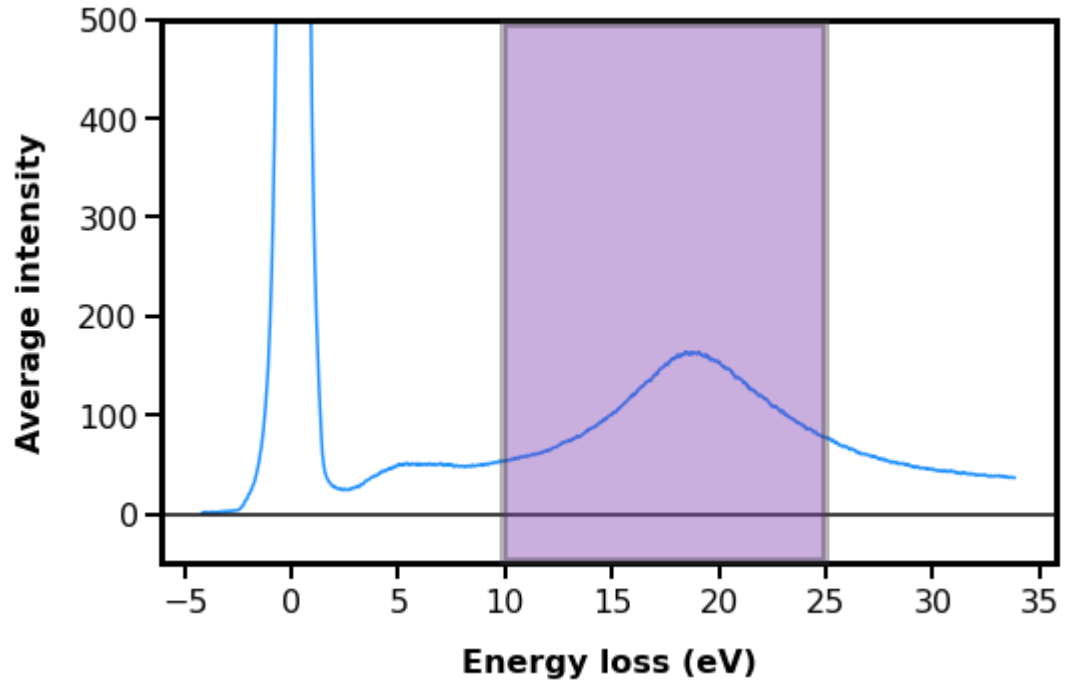
$$\boldsymbol{Maximize(f)}$$

(Specific peak intensity)

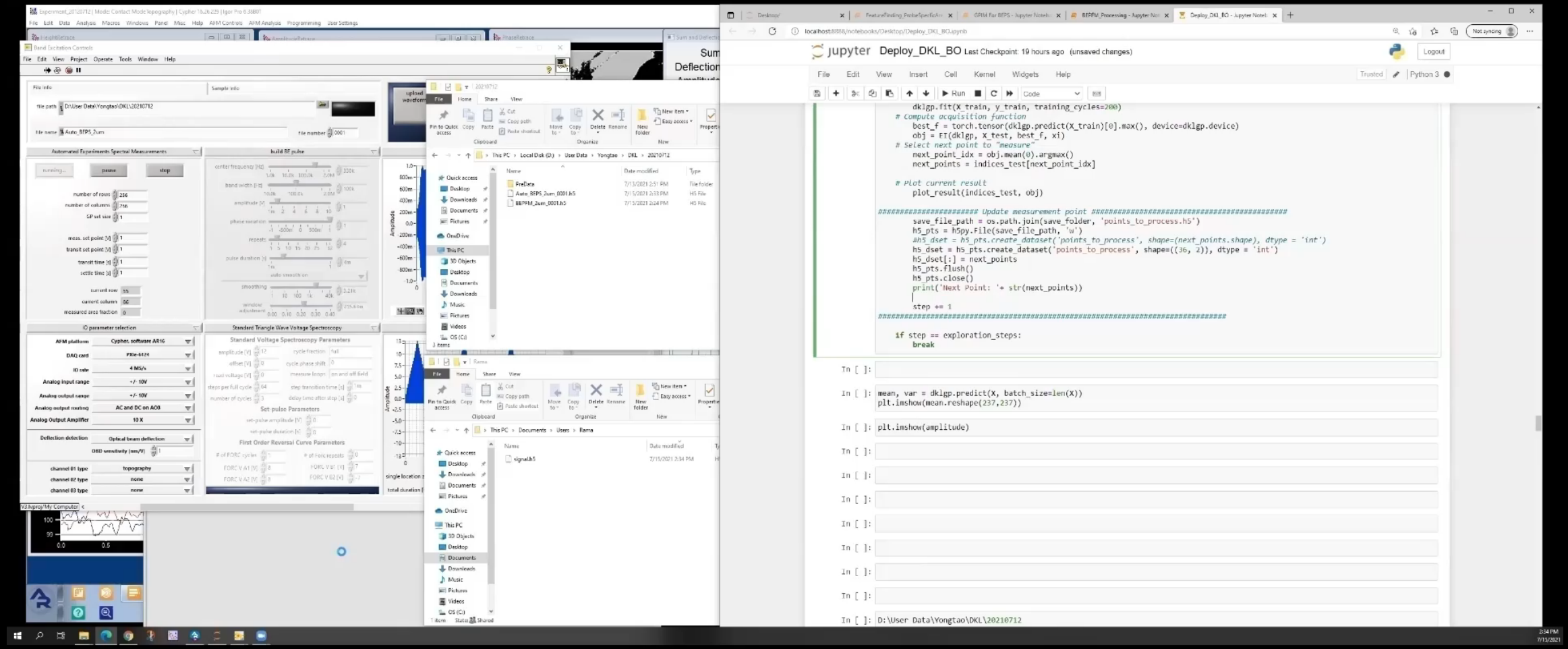
"Acquisition function"

HAADF-STEM
+ points visited



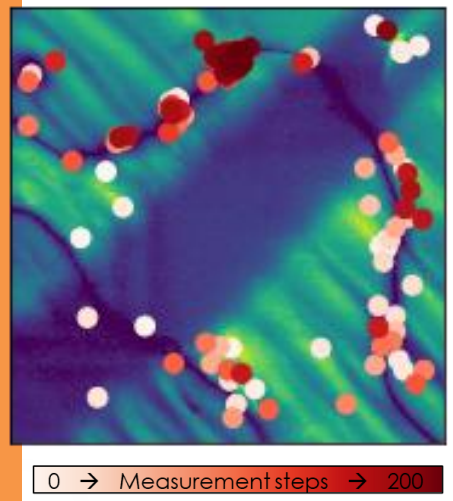


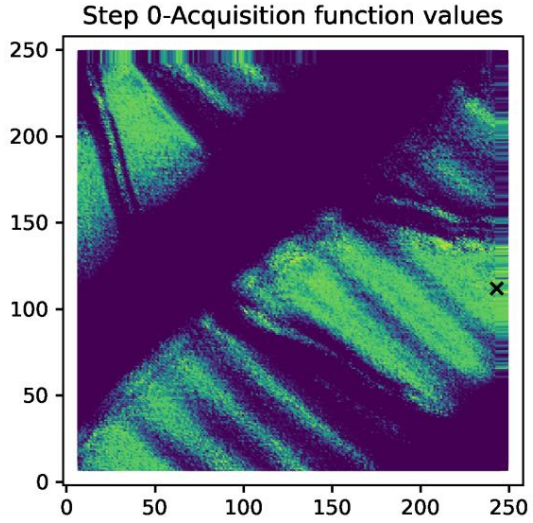
Deep Kernel Learning AE SPM

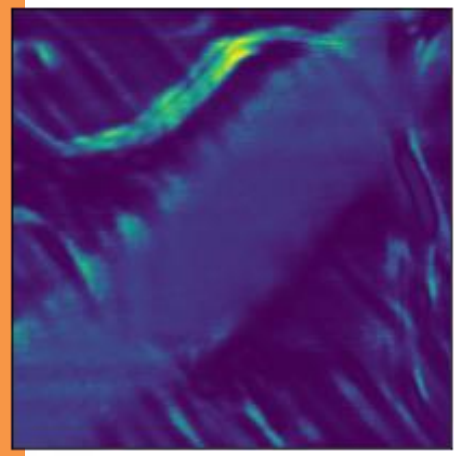


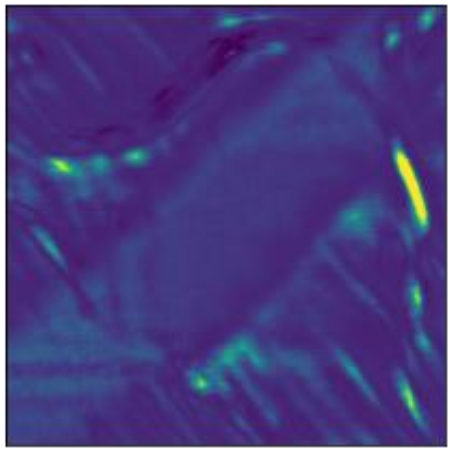
DKL SPM

Guided by: On field loop area

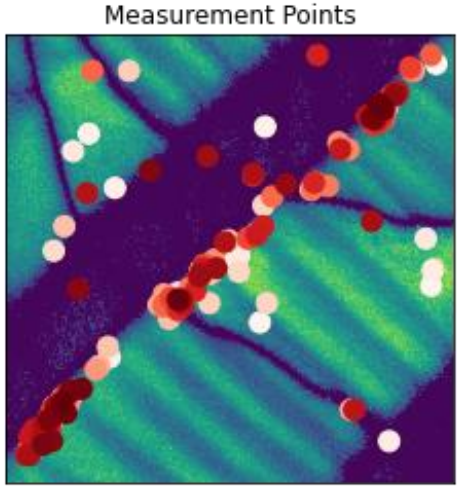


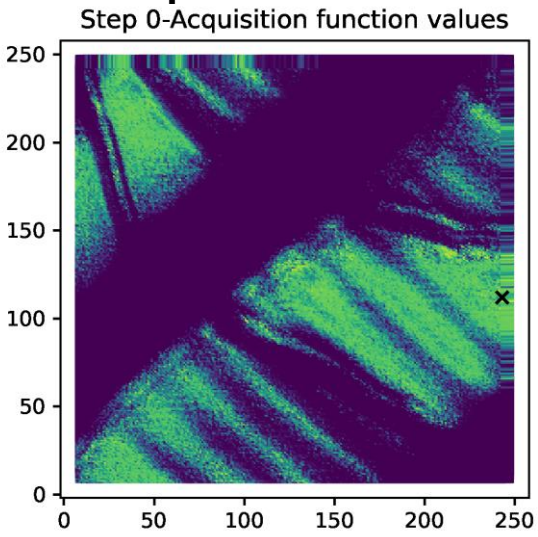


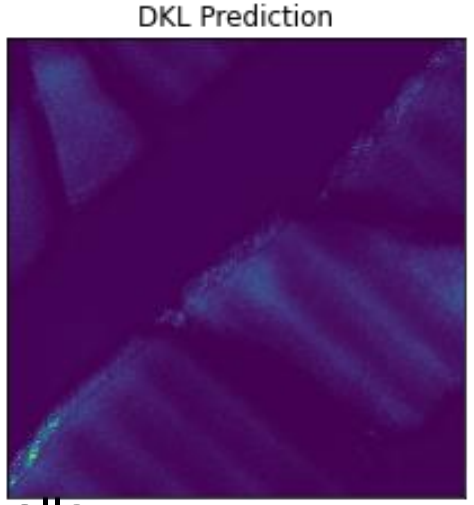


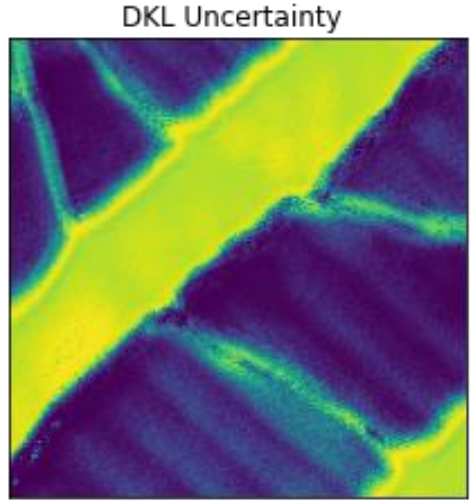


Guided by: Off field loop area





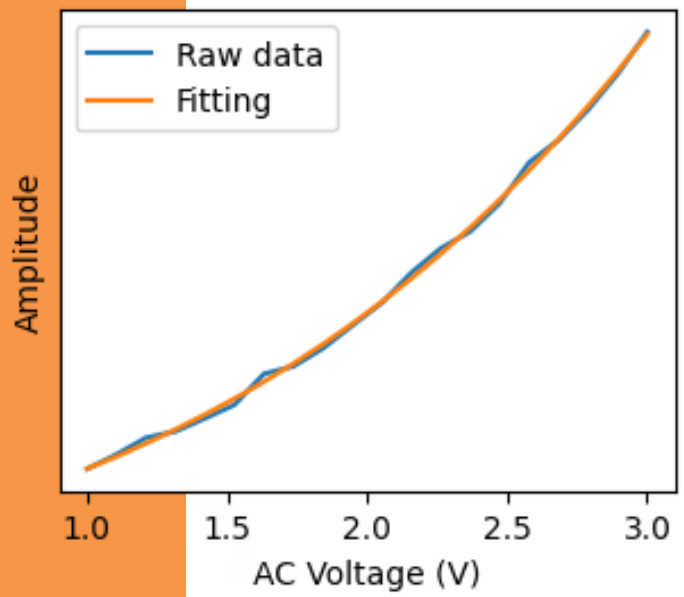




- Large loop opening corresponding 180º domain walls
- This behavior can be attributed to the large polarization mobility of 180º walls

Liu, Yongtao, et al, Nature Machine Intelligence 4, 4 (2022): 341-350.

Exploring Non-Linearity



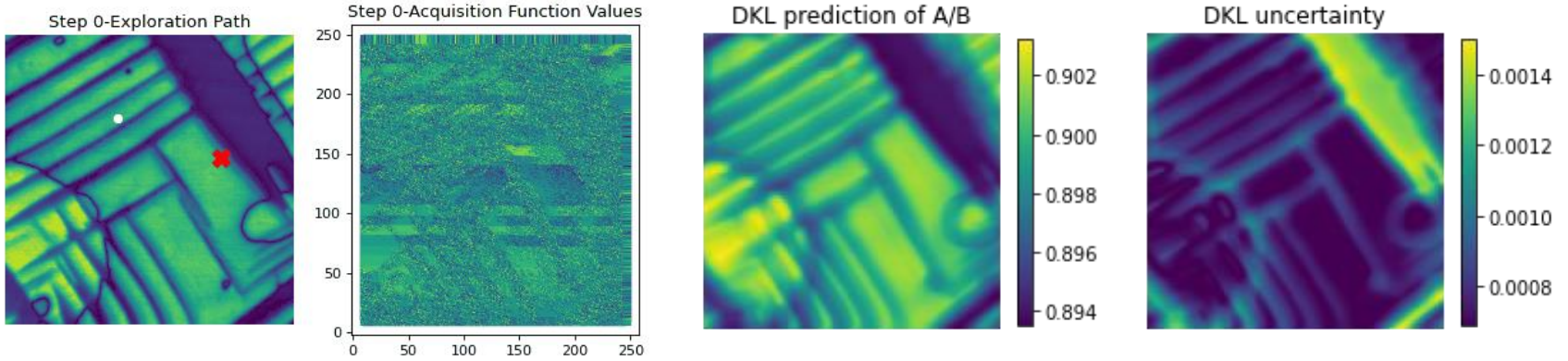
V_{AC} sweep curve at each location was fitted as $y = Ax^3 + Bx^2 + Cx$

A, B, C, and A/B were us as the target function to guide DKL-V_{AC} measurement.

PTO and HZO thin films were studied.

- Shown are 200-step measurements of PTO and HZO thin films
- PFM amplitude was used as structure image; A/B was used to guide the measurement.

PTO experiment process and results



HZO experiment process and results