Oak Ridge National Laboratory

March 27th, 2025

Bayesian Optimization in Automated Experiments

Rama K. Vasudevan

Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge TN

U.S. DEPARTMENT OF ENERGY

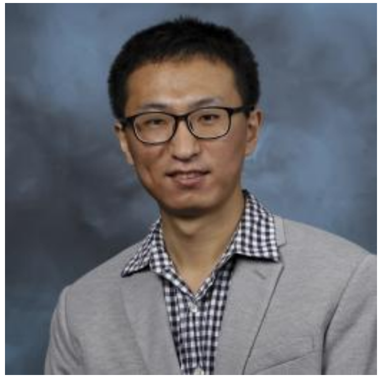
ORNL IS MANAGED BY UT-BATTELLE LLC FOR THE US DEPARTMENT OF ENERGY

Acknowledgements

- Sergei Kalinin (UTK)
- S. Jesse (ORNL)
- P. Ganesh (ORNL)
- Zijie Wu (ORNL)
- G. Duscher (UTK)
- Many other students, postdocs, and staff at CNMS and ORNL

Sumner Harris

Yongtao Liu



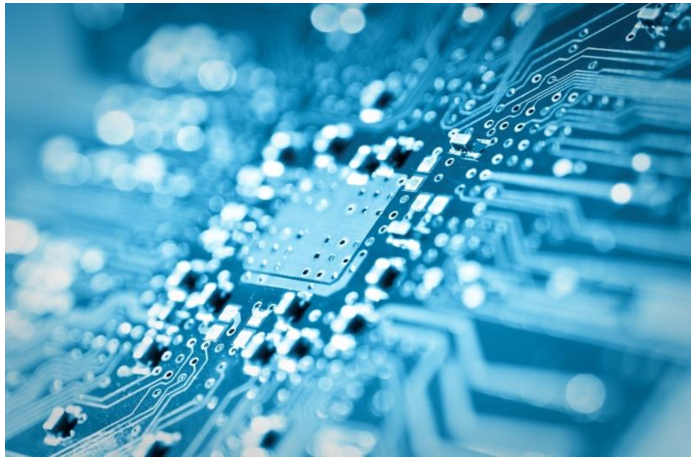
Ganesh Narasimha

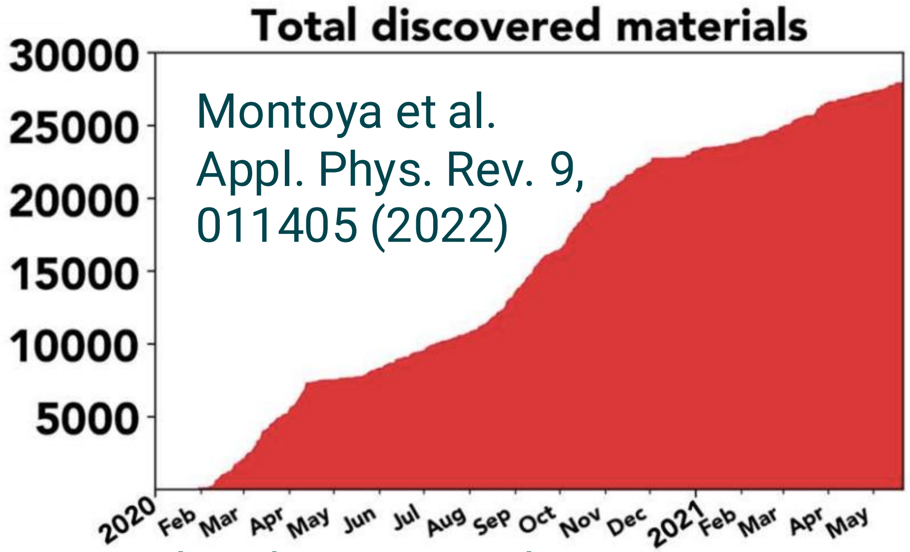
Soumendu Bagchi

OAK RIDGE National Laboratory | CENTER FOR NANOPHASE MATERIALS SCIENCES

OAK RIDGE National Laboratory

Today's materials challenges: autonomous labs?





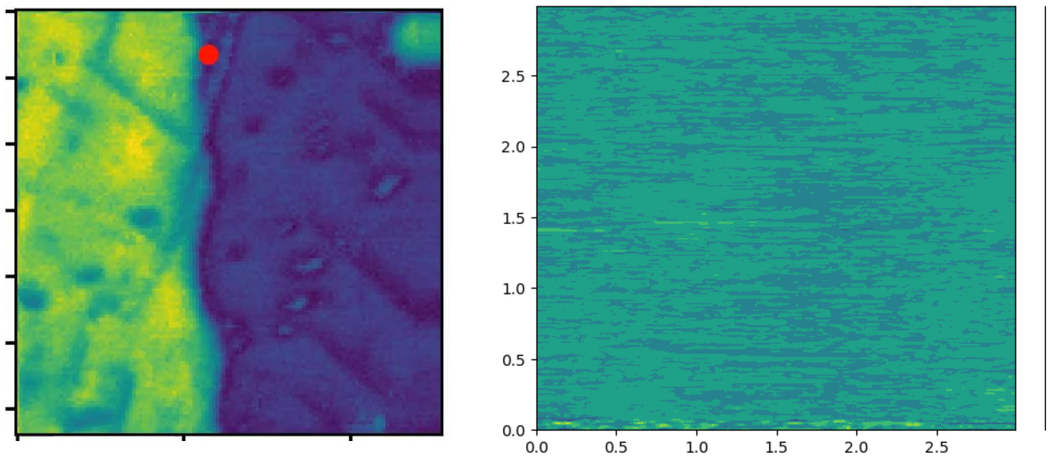
Accumulated new crystal structures in 1374 chemical systems explored autonomously

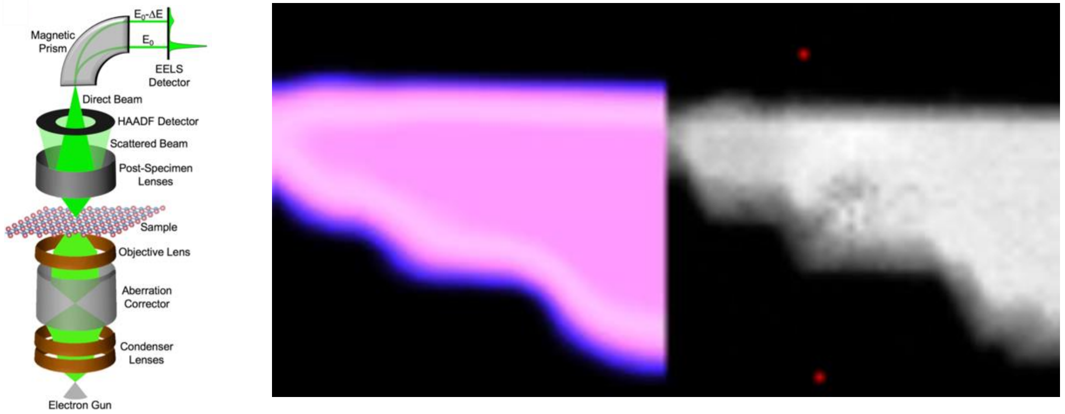
- Renewable energy, self-driving cars, transparent displays, new memory technologies, energy harvesting and generation

- Materials become more complex. We can predict many new materials. But we need to (a) synthesize them and (b) understand them, not just optimize for properties.

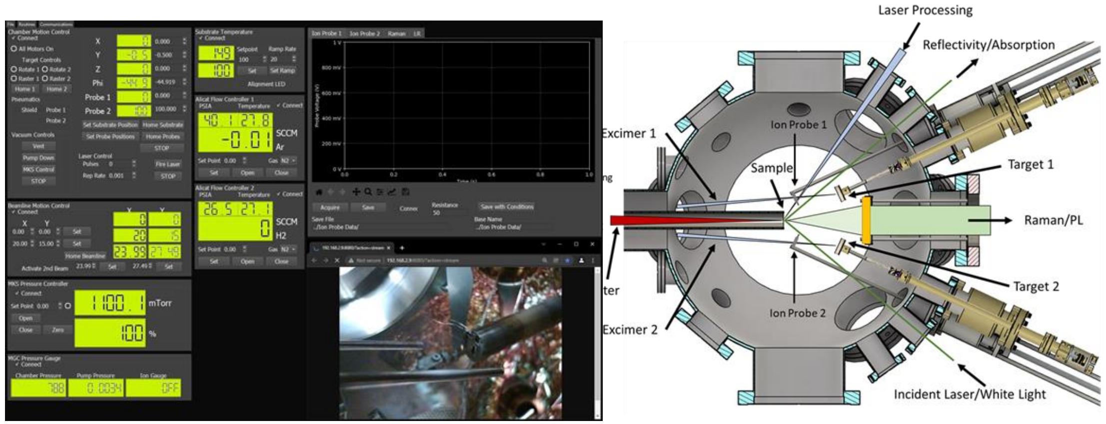
Outline

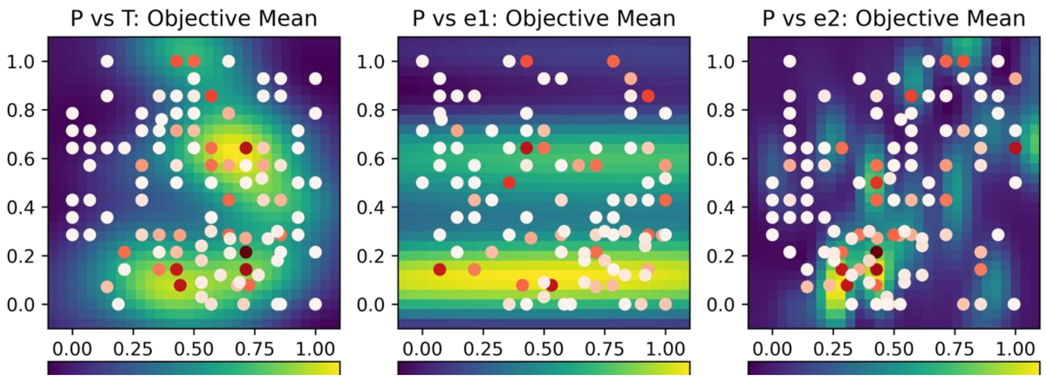
Autonomous Microscopy





Autonomous Synthesis

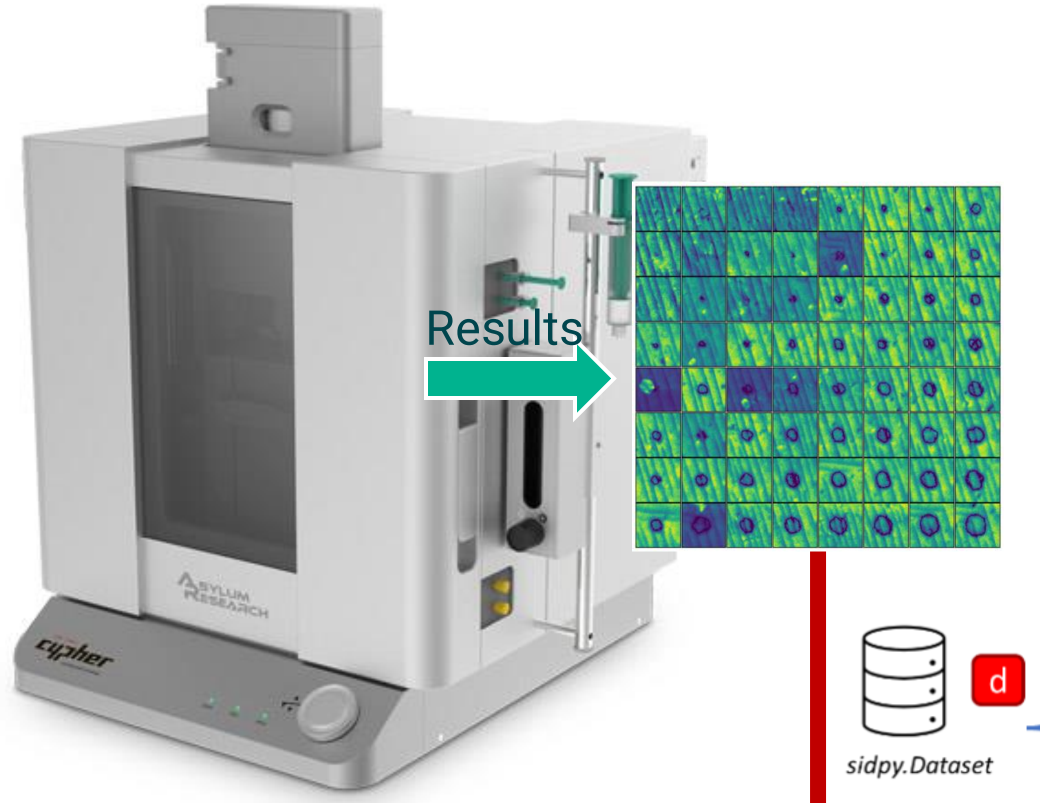




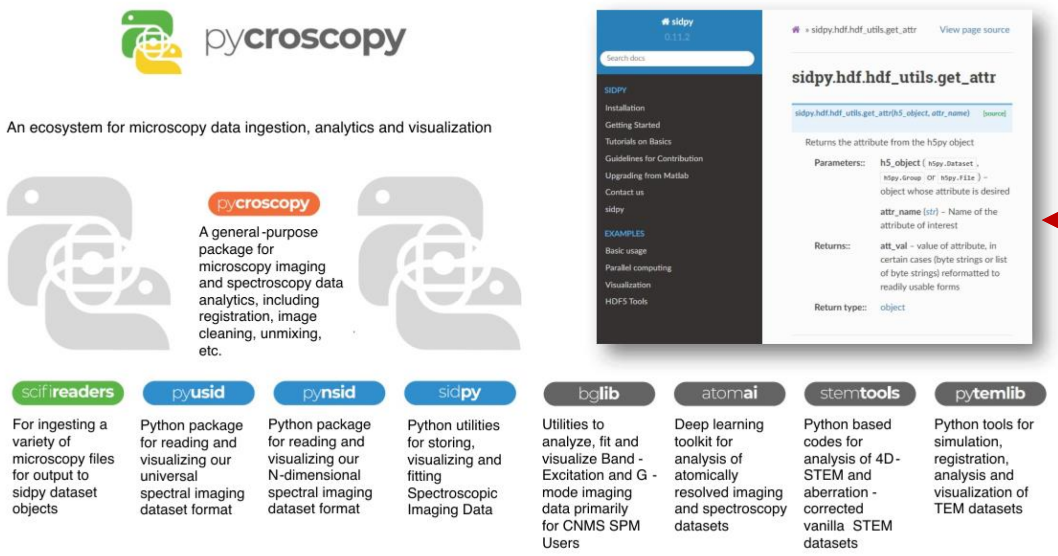
Automating SPM: AEcroscopy

Software Infrastructure

Deploy



Y. Liu et al. Small Methods 2301740 (2024)



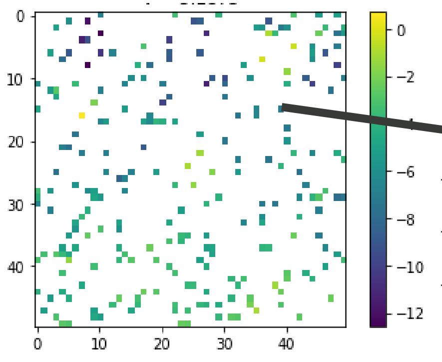
Vasudevan et al. Advanced Theory and Simulations **6**, 2300247 (2023)

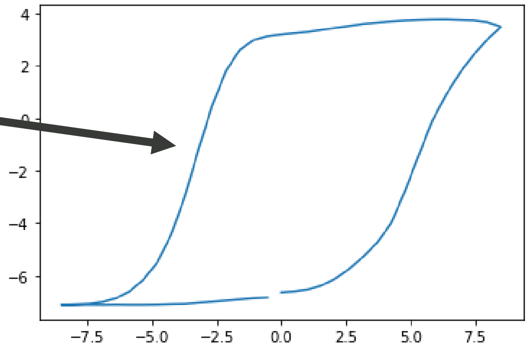
- Standardized data model
- In-built processing and viz utilities

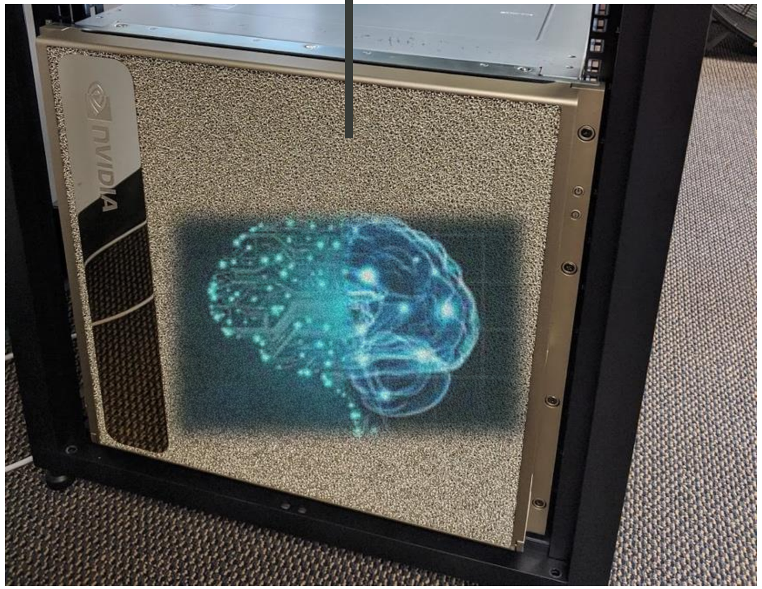
Real-time automated SPM

Provide list of new measurement conditions

Measure a batch of hysteresis loops





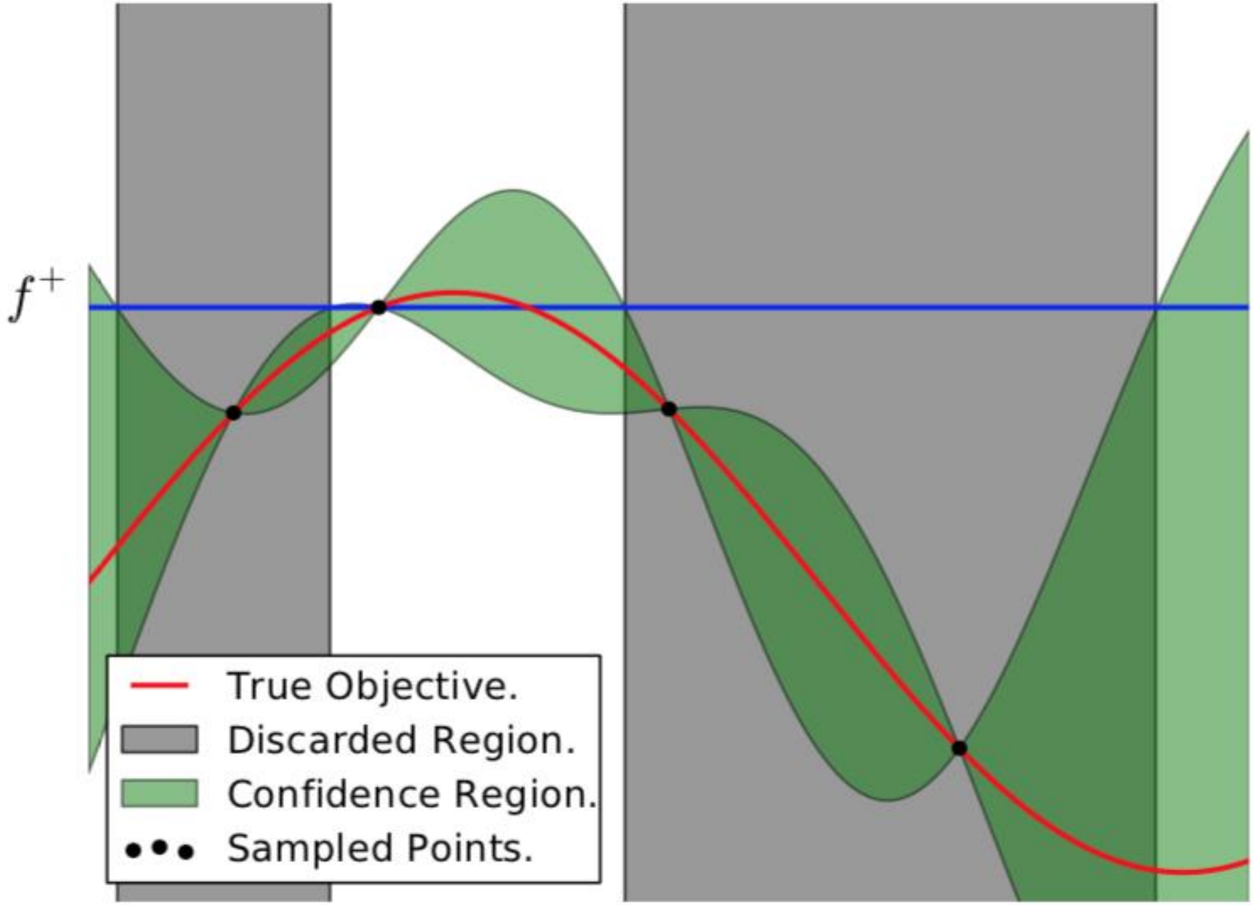


Compute Box

Data preprocessing (scalarizer function)

- For long spectroscopies, measuring every pixel across a grid is not feasible. Bayesian optimization enables sampling only those points that will maximize a property of choice, enabling new experiments that would otherwise not be possible

Bayesian Optimization

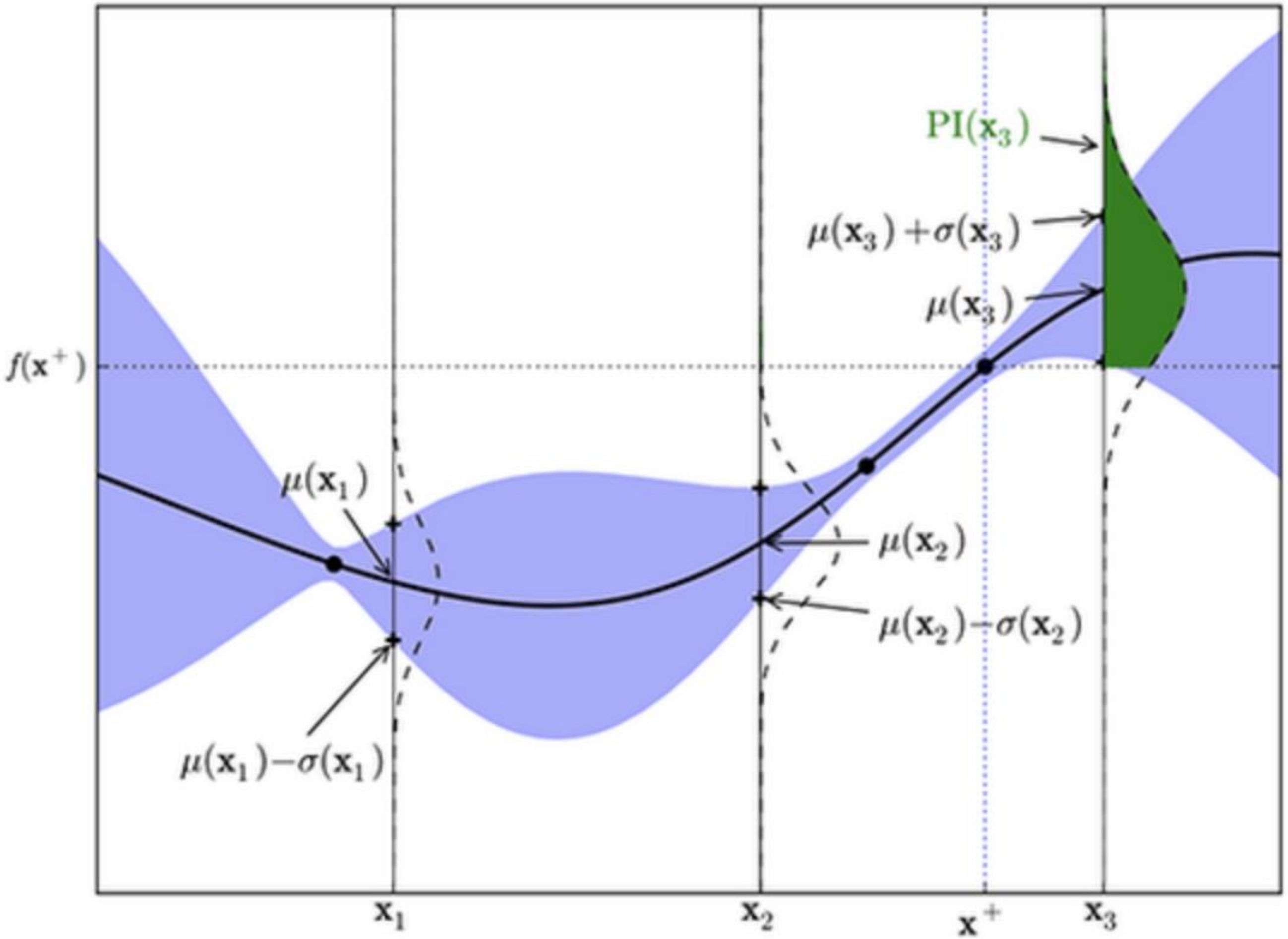


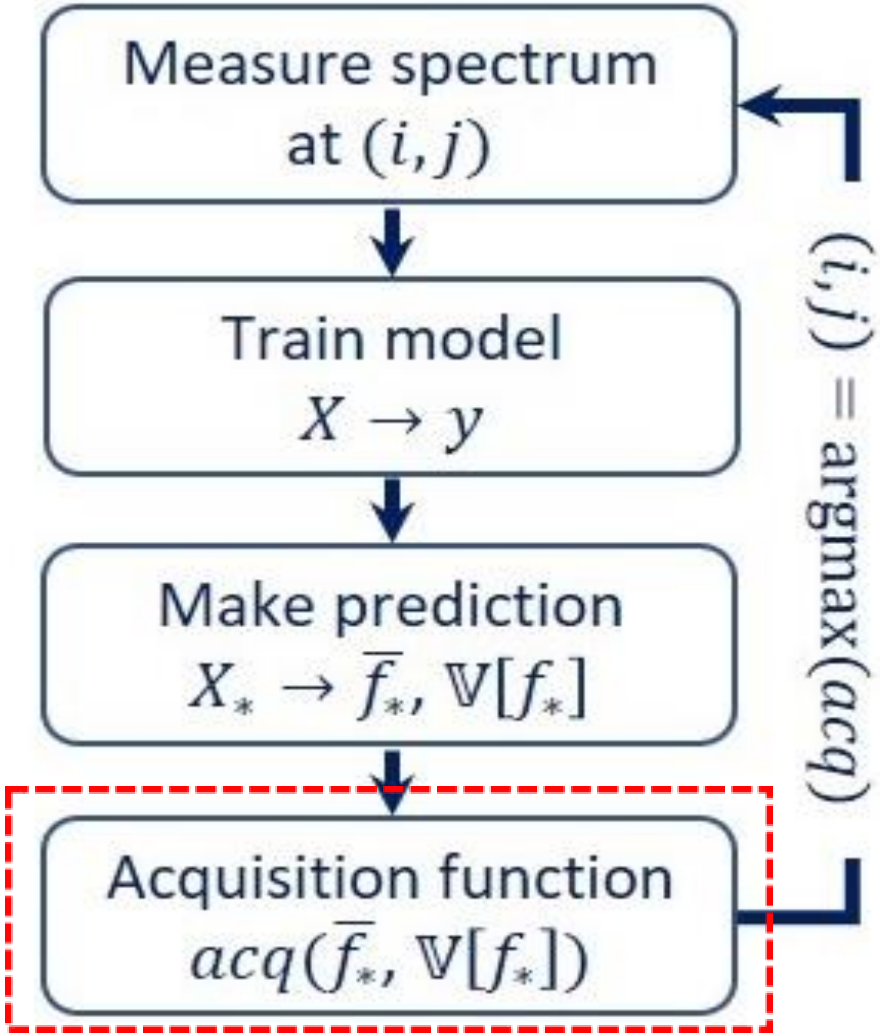
N. de Freitas et al., Taking the Human Out of the Loop: A Review of Bayesian Optimization , *Proceedings of the IEEE* **104**, 148 (2015)

- We have some measurements in space X, and we want to maximize some property f(X).
- Generalizes to higher dimensions
- Recently has become computationally tractable

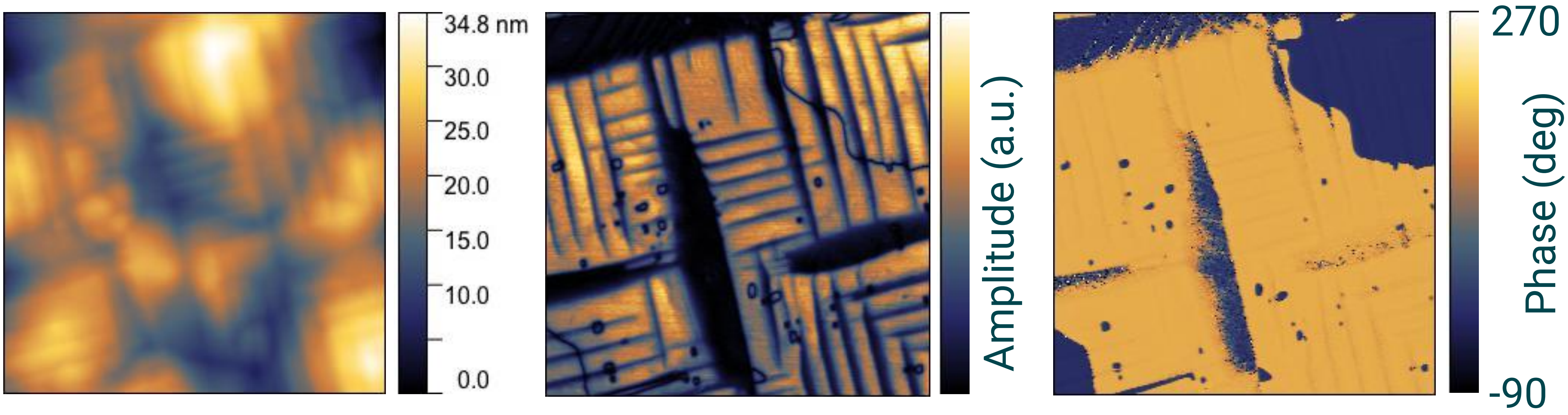
Acquisition Functions

Probability of Improvement
Acquisition Function





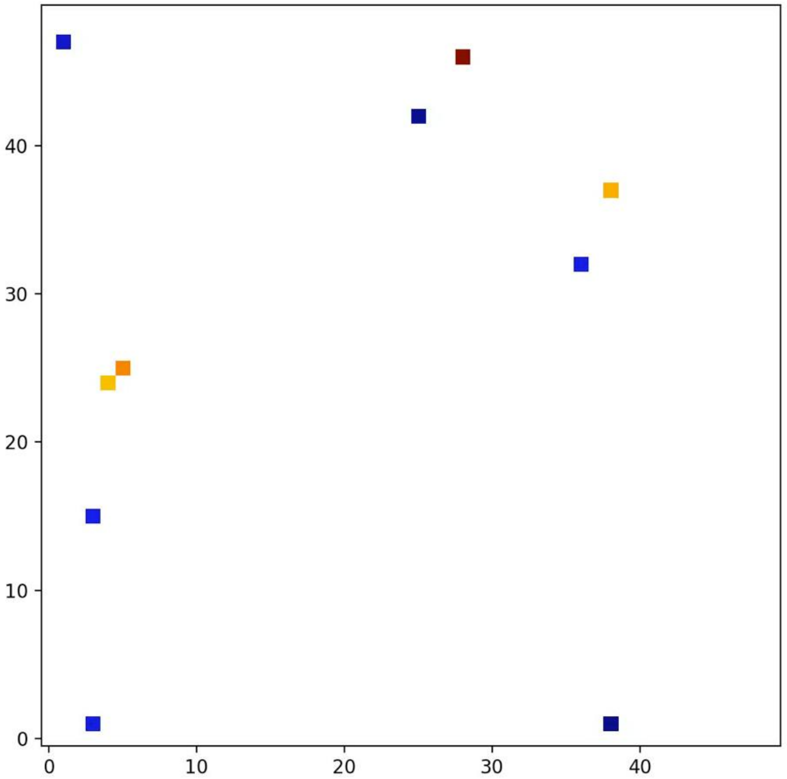
Spectroscopy on a ferroelectric film



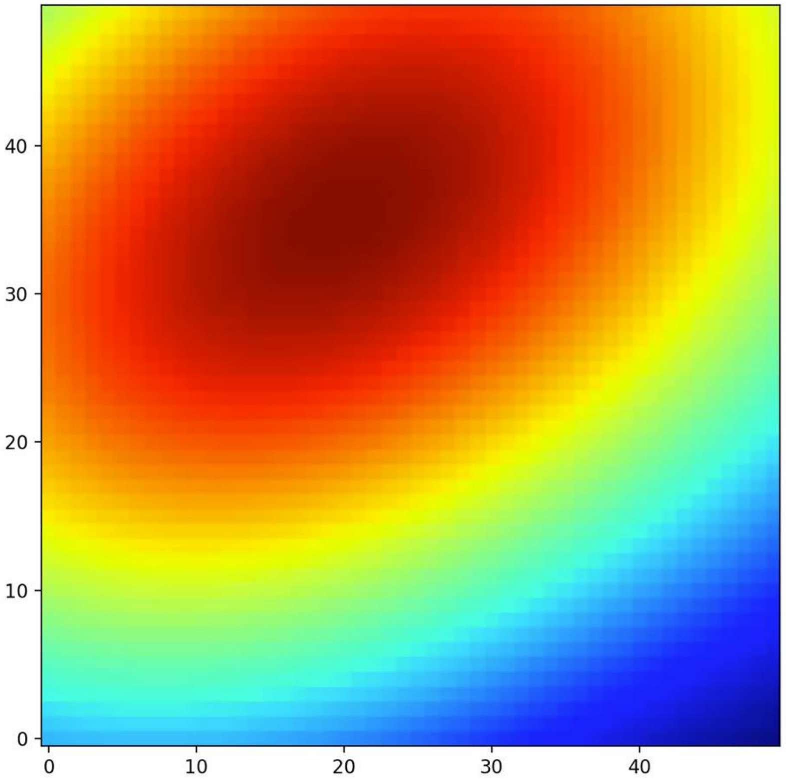
Can image the domain structure with the microscope -> 4 minutes
Spectroscopy – obtaining spectra pixel by pixel – can take 2-24 hours depending on type of measurement.

Automated Experiment example

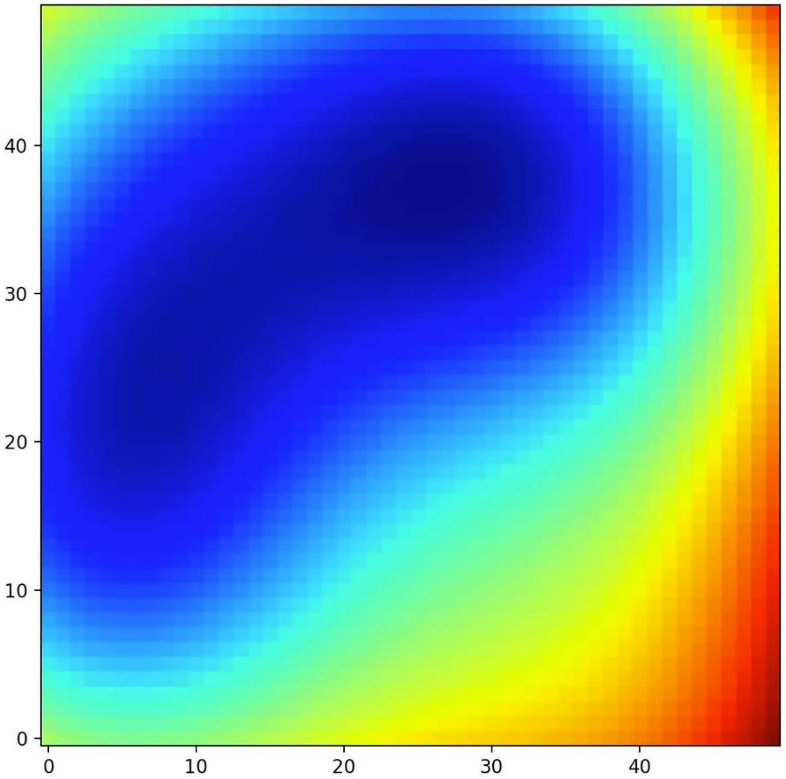
Measured Loop Areas



GP Prediction



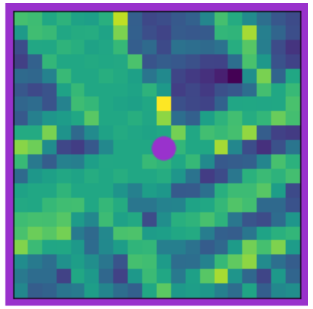
GP Uncertainty

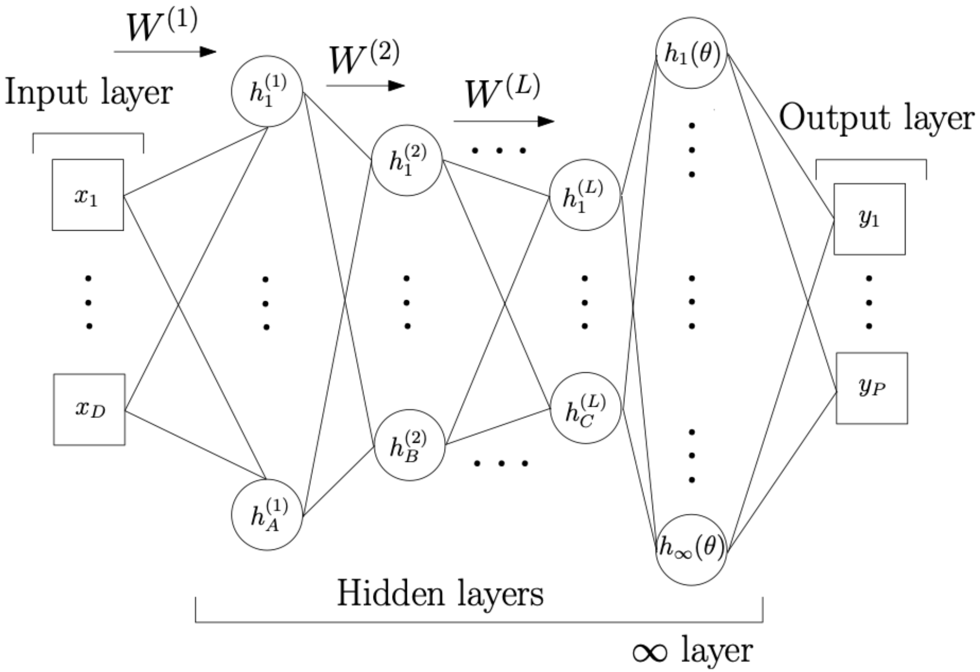


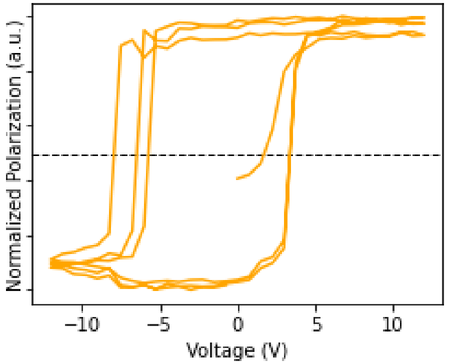
R. Vasudevan et al., ACS Nano 15, 11253 (2021)

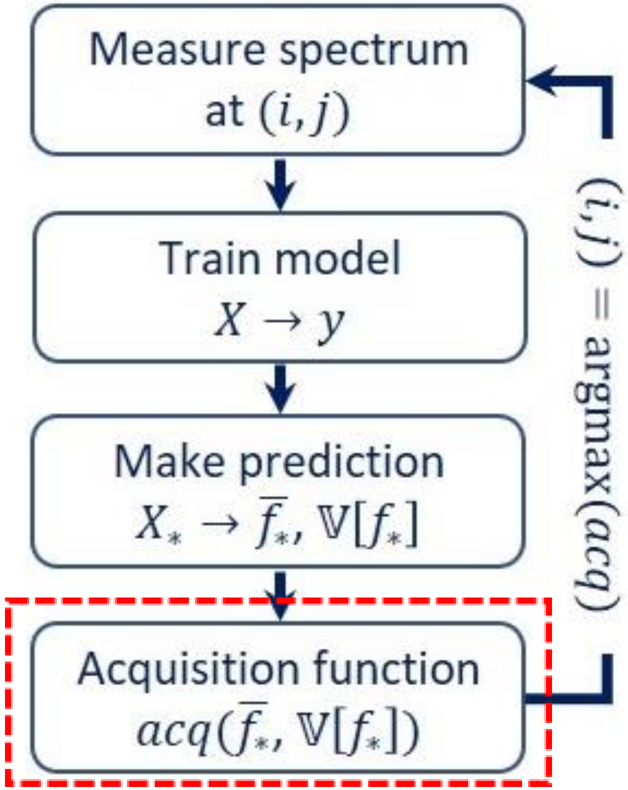
Deep Kernel Learning: More images, better kernels

Deep Kernel Learning (A.G. Wilson, 2015)



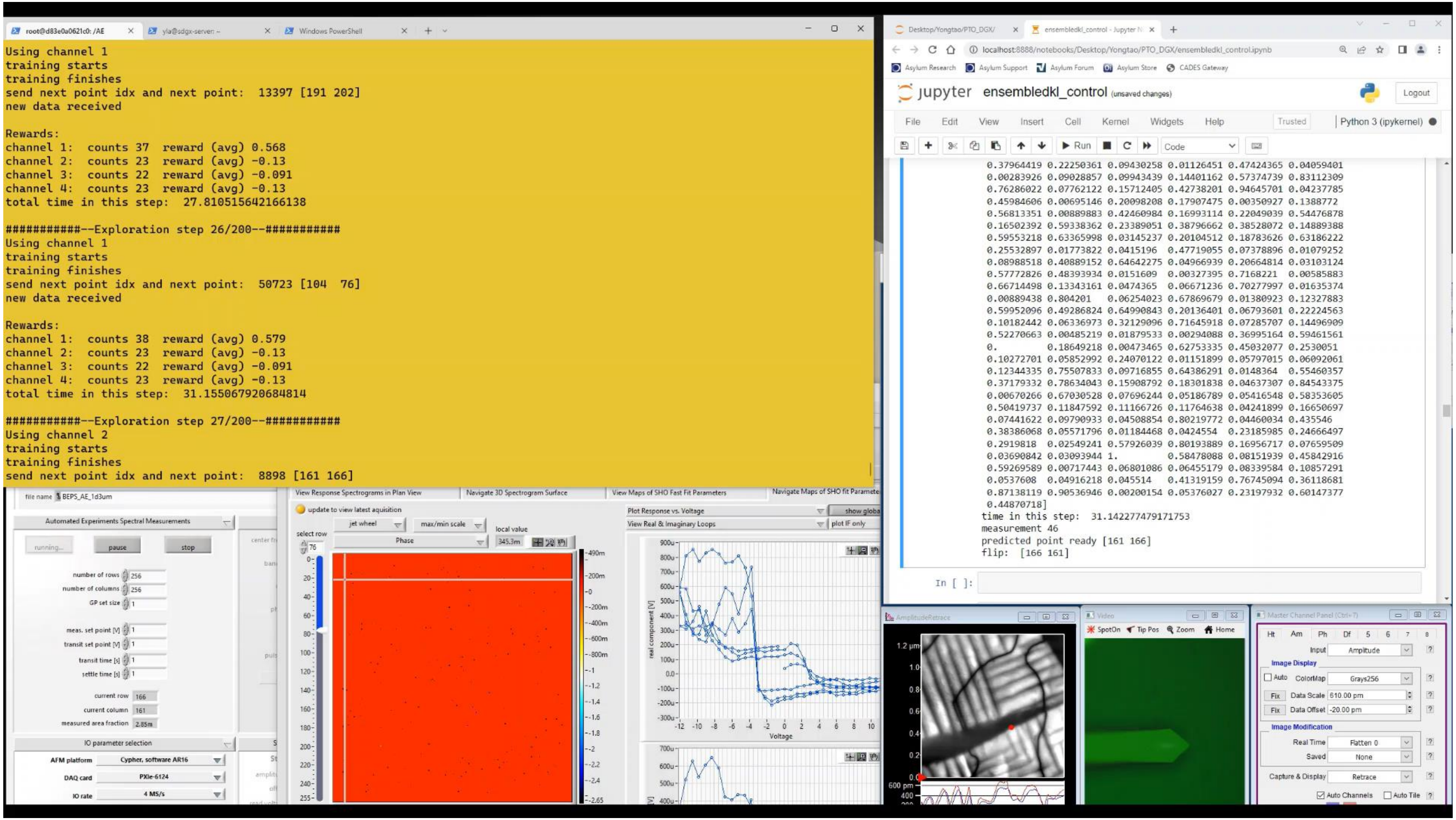






Combine a neural network with GP, learn parameters jointly

Liu, Yongtao, Kyle P. Kelley, Rama K. Vasudevan, Hiroshi Funakubo, Maxim A. Ziatdinov, and Sergei V. Kalinin. Nature Machine Intelligence 4, 4 (2022): 341-350.



Using channel 1
training starts
training finishes
send next point idx and next point: 13397 [191 202]
new data received

Rewards:
channel 1: counts 37 reward (avg) 0.568
channel 2: counts 23 reward (avg) -0.13
channel 3: counts 22 reward (avg) -0.091
channel 4: counts 23 reward (avg) -0.13
total time in this step: 27.810515642166138

###########--Exploration step 26/200--###########
Using channel 1
training starts
training finishes
send next point idx and next point: 50723 [104 76]
new data received

Rewards:
channel 1: counts 38 reward (avg) 0.579
channel 2: counts 23 reward (avg) -0.13
channel 3: counts 22 reward (avg) -0.091
channel 4: counts 23 reward (avg) -0.13
total time in this step: 31.155067920684814

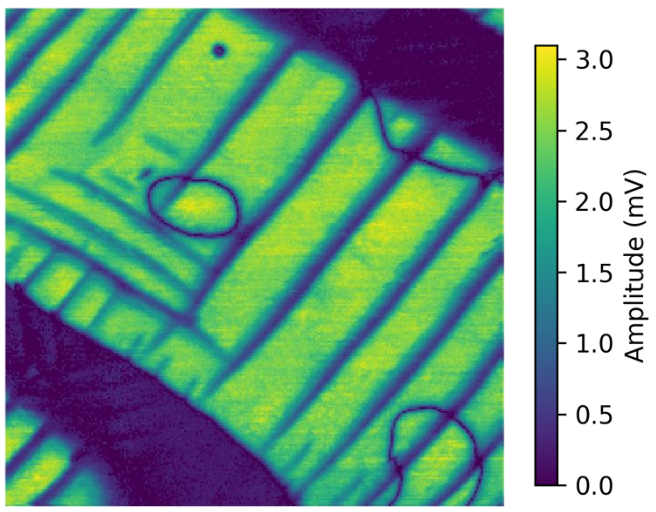
###########--Exploration step 27/200--###########
Using channel 2
training starts
training finishes
send next point idx and next point: 8898 [161 166]

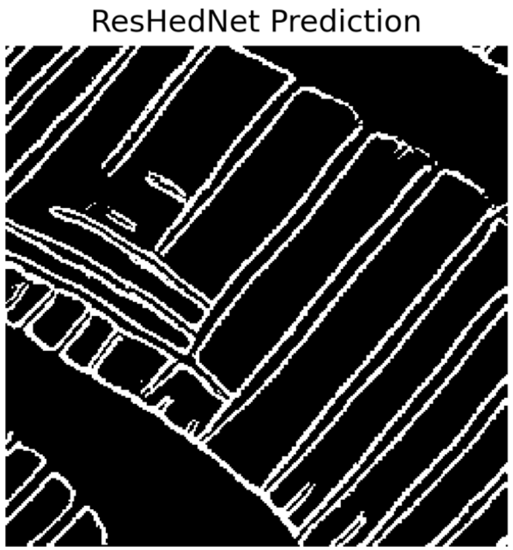
jupyter ensembledkl_control (unsaved changes)
File Edit View Insert Cell Kernel Widgets Help
Trusted
Python 3 (ipykernel)
Logout
0.37964419 0.22250361 0.09430258 0.01126451 0.47424365 0.04059401
0.00283926 0.09028857 0.09943439 0.14401162 0.57374739 0.83112309
0.76286022 0.07762122 0.15712405 0.42738201 0.94645701 0.04237785
0.45984606 0.00695146 0.20098208 0.17907475 0.00350927 0.1388772
0.56813351 0.00889883 0.42460984 0.16993114 0.22049039 0.54476878
0.16502392 0.59338362 0.23389051 0.38796662 0.38528072 0.14889388
0.59553218 0.63365998 0.03145237 0.20104512 0.18783626 0.63186222
0.25532897 0.01773822 0.0415196 0.47719055 0.07378896 0.01079252
0.08988518 0.40889152 0.64642275 0.04966939 0.20664814 0.03103124
0.57772826 0.48393934 0.0151609 0.00327395 0.7168221 0.00585883
0.66714498 0.13343161 0.0474365 0.06671236 0.70277997 0.01635374
0.00889438 0.804201 0.06254023 0.67869679 0.01380923 0.12327883
0.59952096 0.49286824 0.64990843 0.20136401 0.06793601 0.22224563
0.10182442 0.06336973 0.32129096 0.71645918 0.07285707 0.14496909
0.52270663 0.00485219 0.01879533 0.00294088 0.36995164 0.59461561
0. 0.18649218 0.00473465 0.62753335 0.45032077 0.2530051
0.10272701 0.05852992 0.24070122 0.01151899 0.05797015 0.06092061
0.12344335 0.75507833 0.09716855 0.64386291 0.0148364 0.55460357
0.37179332 0.78634043 0.15908792 0.18301838 0.04637307 0.84543375
0.00670266 0.67030528 0.07696244 0.05186789 0.05416548 0.58353605
0.50419737 0.11847592 0.11166726 0.11764638 0.04241899 0.16650697
0.07441622 0.09790933 0.04508854 0.80219772 0.04460034 0.435546
0.38386068 0.05571796 0.01184468 0.0424554 0.23185985 0.24666497
0.2919818 0.02549241 0.57926039 0.80193889 0.16956717 0.07659509
0.03690842 0.03093944 1. 0.58478088 0.08151939 0.45842916
0.59269589 0.00717443 0.06801086 0.06455179 0.08339584 0.10857291
0.0537608 0.04916218 0.045514 0.41319159 0.76745094 0.36118681
0.87138119 0.90536946 0.00200154 0.05376027 0.23197932 0.60147377
0.44870718]
time in this step: 31.142277479171753
measurement 46
predicted point ready [161 166]
flip: [166 161]
In [ ]:

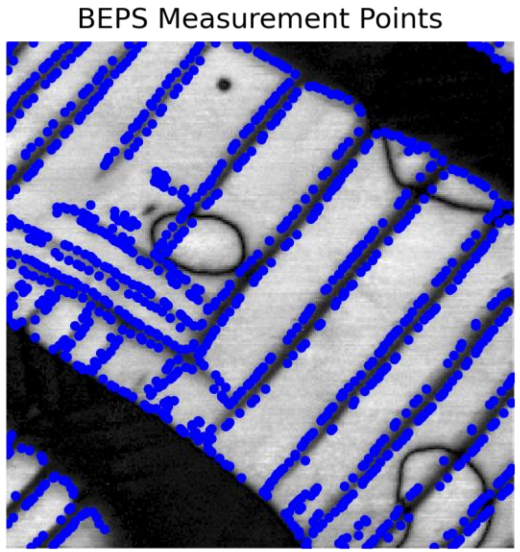
Automated Experiments Spectral Measurements
pause
stop
number of rows 256
number of columns 256
current row 166
current column 161
AFM platform Cypher, software AR16
DAQ card PXIe-6124
IO rate 4 MS/s
Voltage

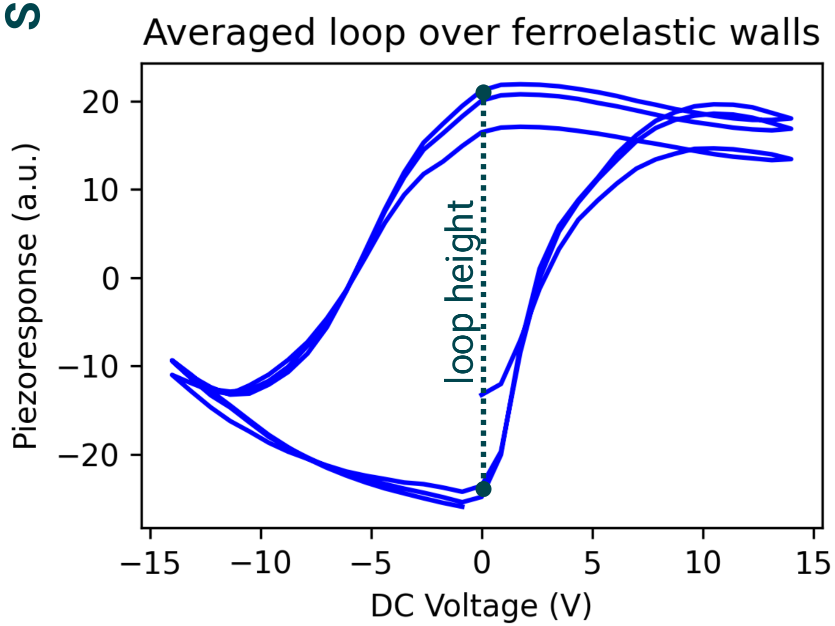
Automated SPM—Objects of Interest

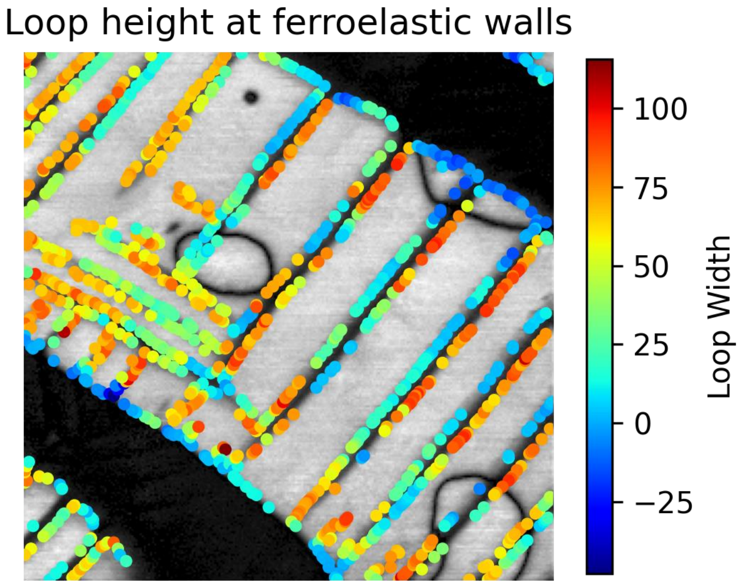
Slide by Yongtao Liu ORNL

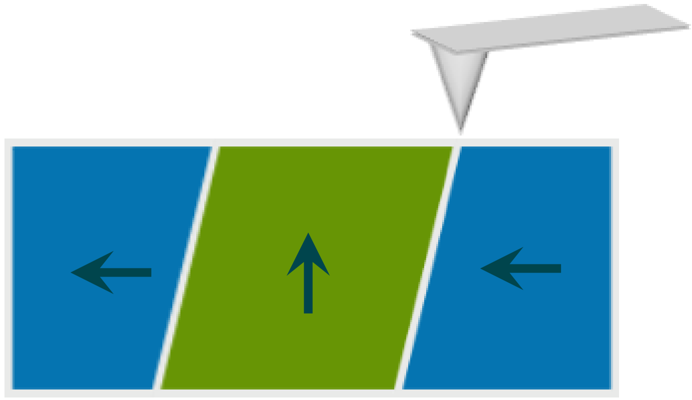






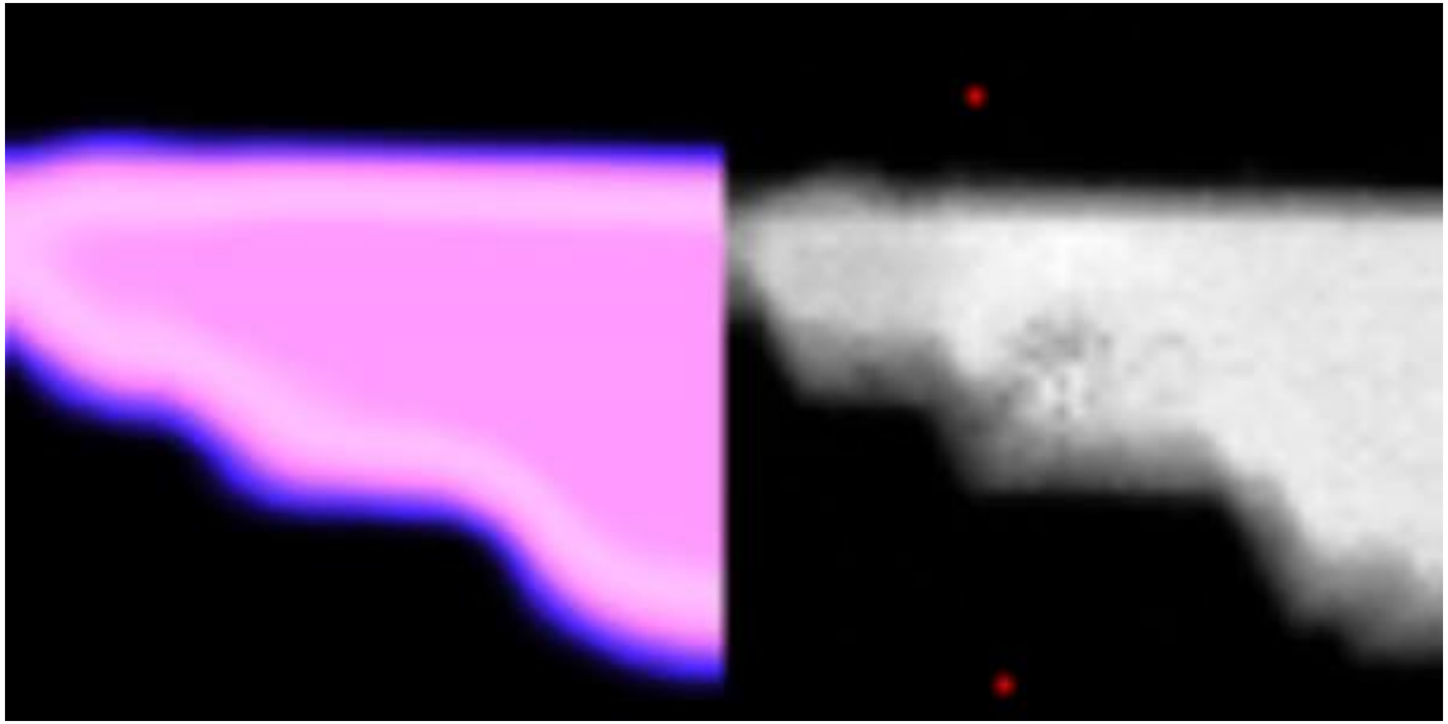






Liu, Y., Kelley, K.P., Funakubo, H., Kalinin, S.V., Ziatdinov, M., Advanced Science, 2022, 2203957

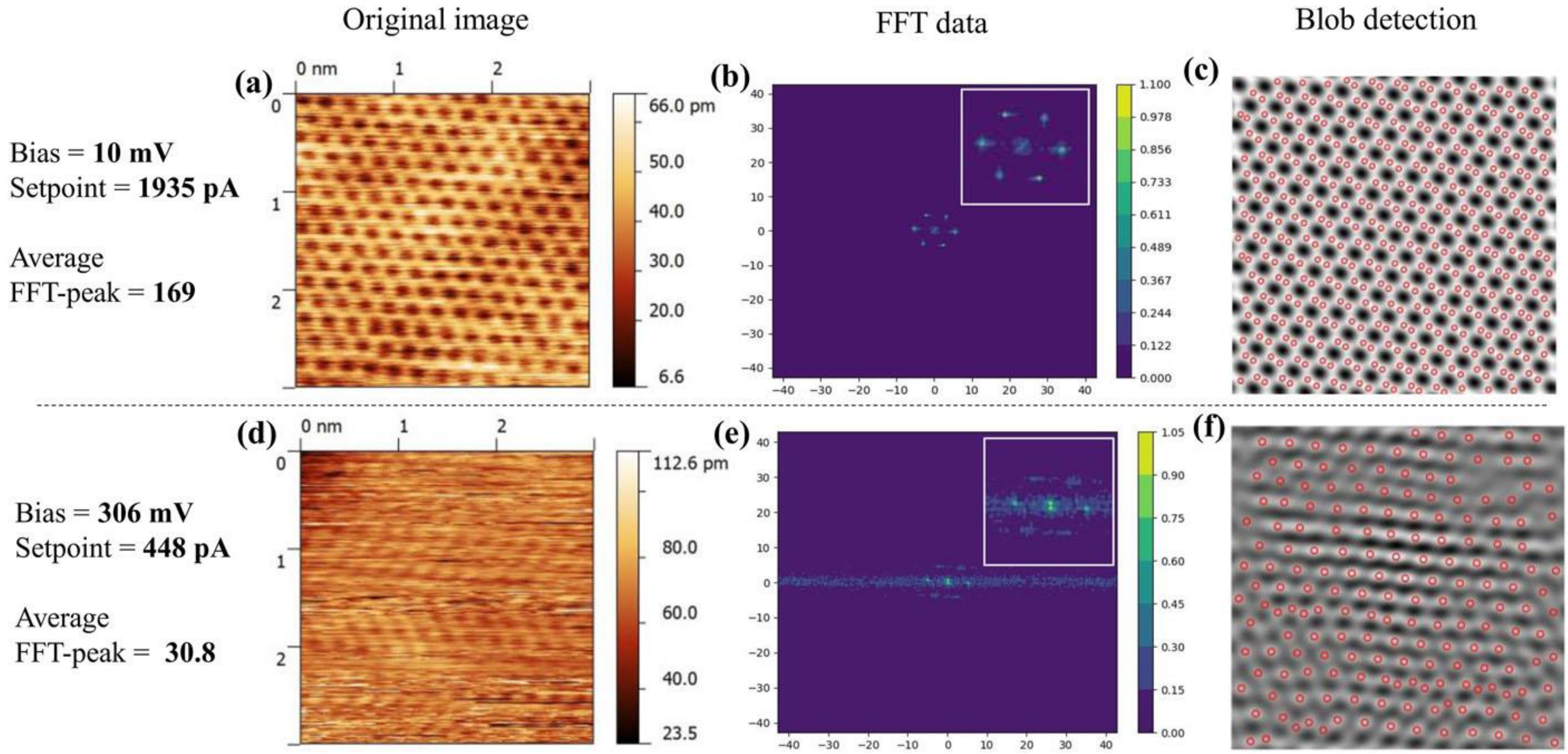
STEM Bayesian Optimization



- Bulk and edge plasmons in $MnSP_3$ investigated via deep kernel active learning
- Optimize for ratio of peaks to find where edge plasmon is strongest
- Automated structure-property relationship determination

K. Roccapriore et al., Adv. Sci. 2203422 (2022)

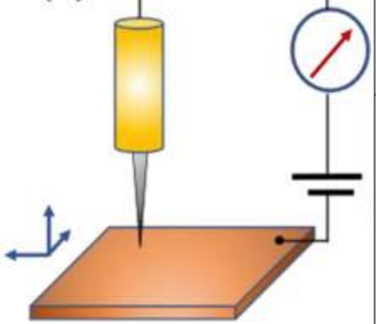
Another example: Image optimization in STM

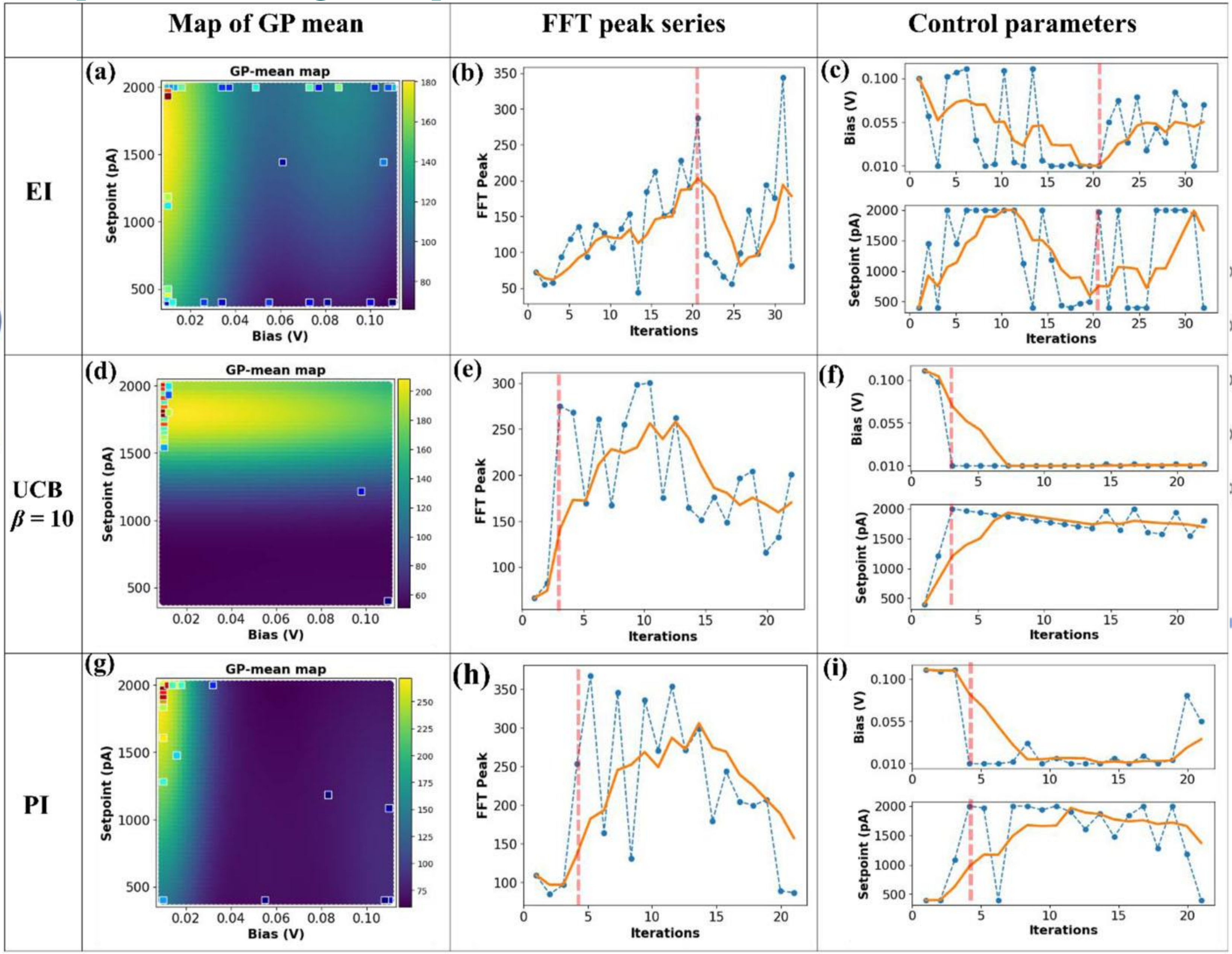


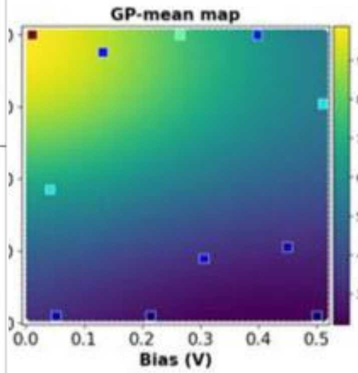
G. Narasimha et al. *APL Mach. Learn.* 2, 016121 (2024)

Another example: Image optimization in STM

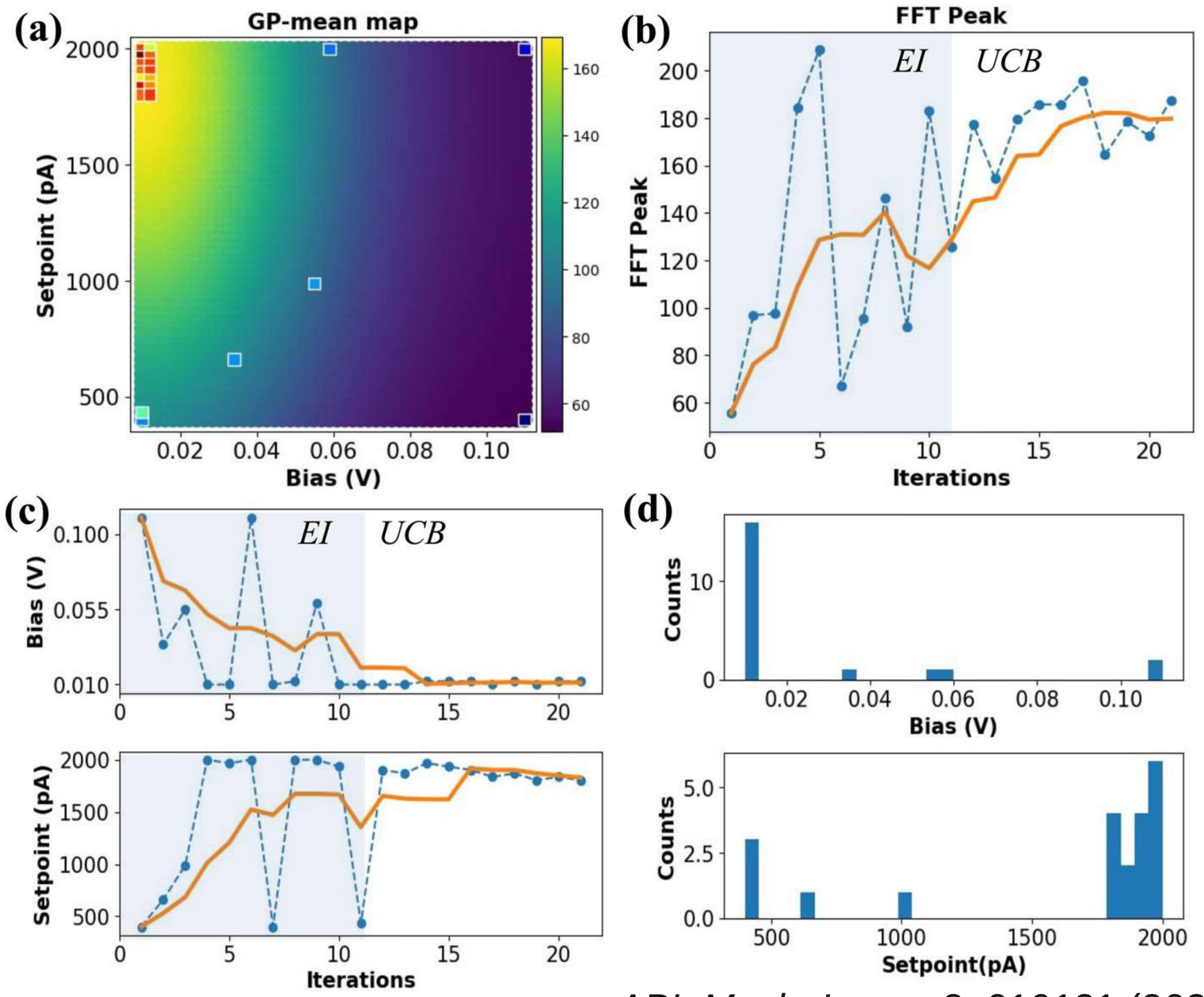
(a)







Another example: Image optimization in STM



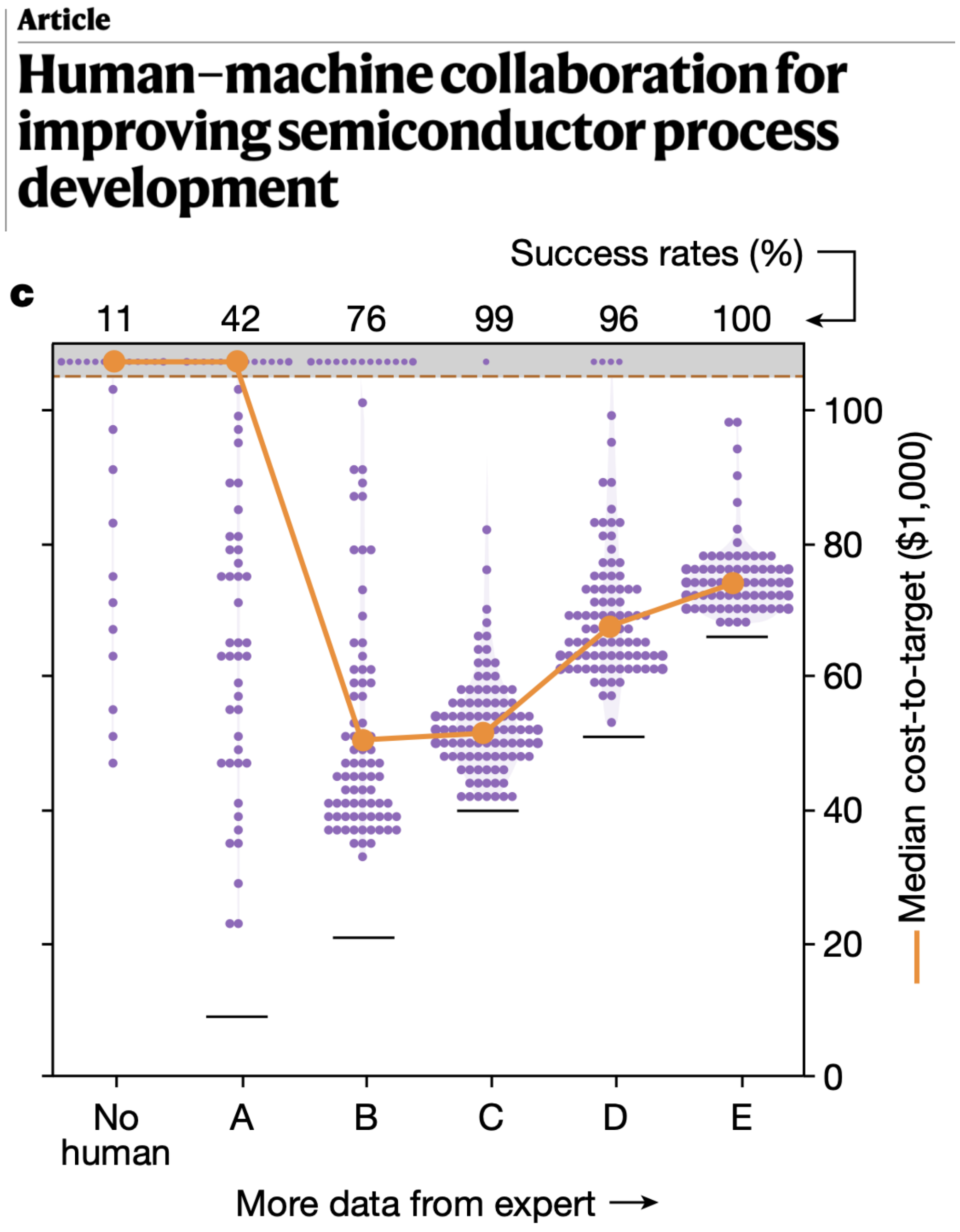
APL Mach. Learn. 2, 016121 (2024)

Beyond Simple Bayesian Optimization

- There is a significant issue with traditional BO based methods
  - We need a scalarizer function to be known *apriori*
- Two solutions
  - Use recommender system to learn the scalarizer
  - Use the notion of curiosity and optimize for it

Algorithms: Human in the Loop

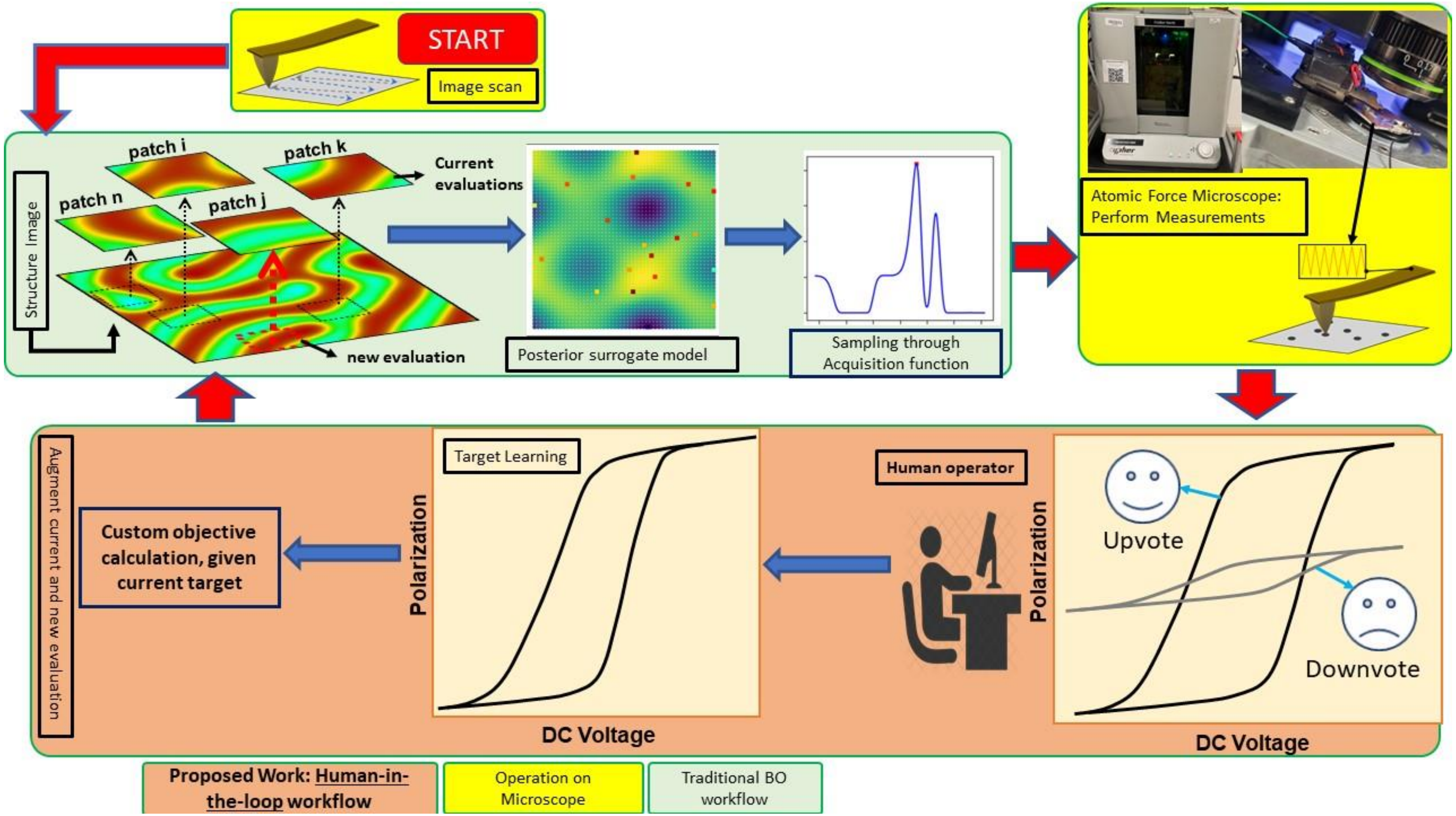
- Human-AI systems work better together than either alone
- Prior knowledge injection can be used to shape targets, reduce wastage, can be informed by simulations
- Increase AI alignment, monitor autonomous progress. Develop better metrics for monitoring autonomous experiments, and bette methods for human input


Article

Human–machine collaboration for improving semiconductor process development

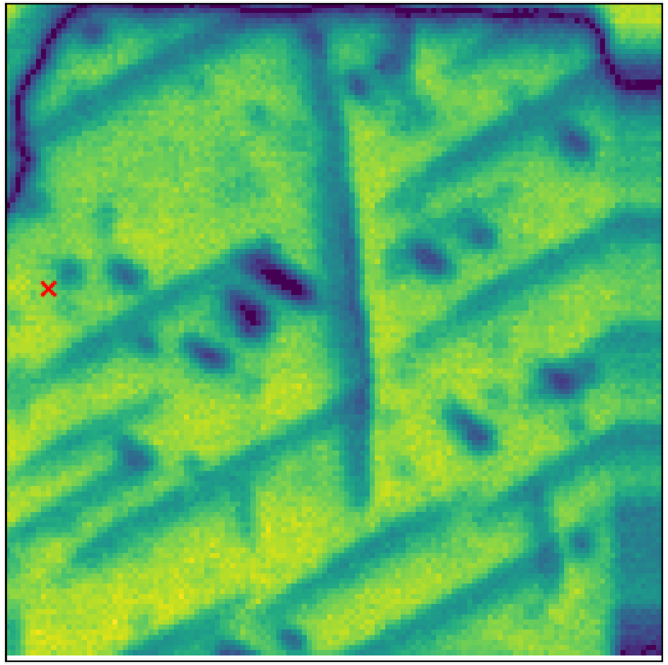
Kanarik et al. Nature **616**, 707 (2023)

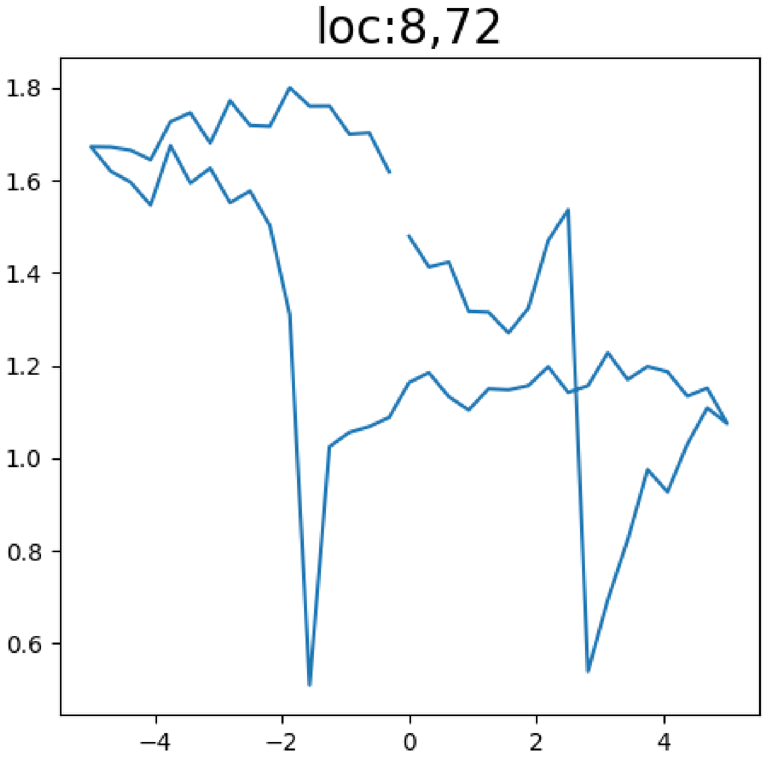
Learning what to optimize: reccomender systems

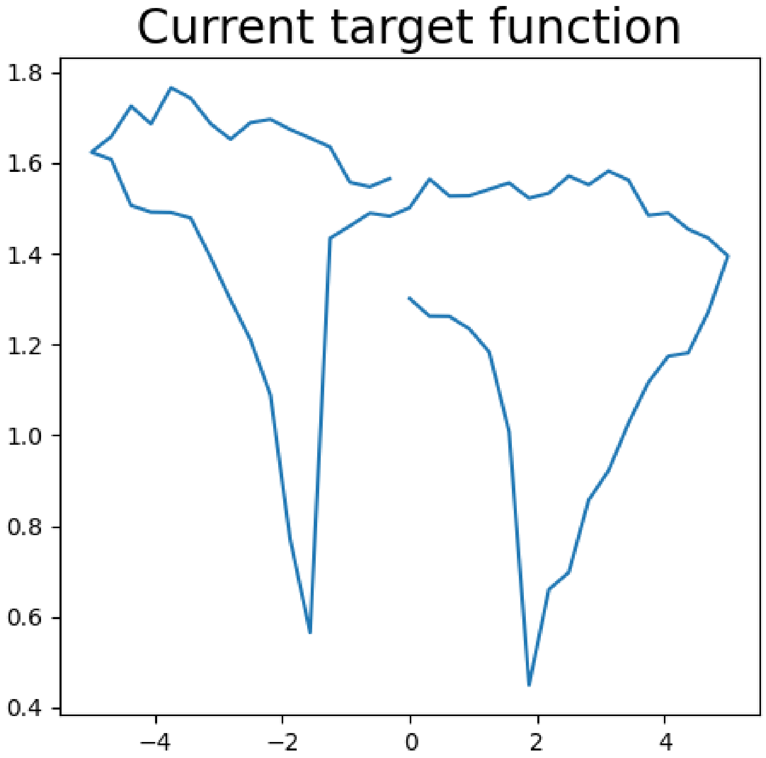


A. Biswas et al. npj Comp. Mater. 10, 29 (2024)

Spectral Recommender: When no Target will suffice







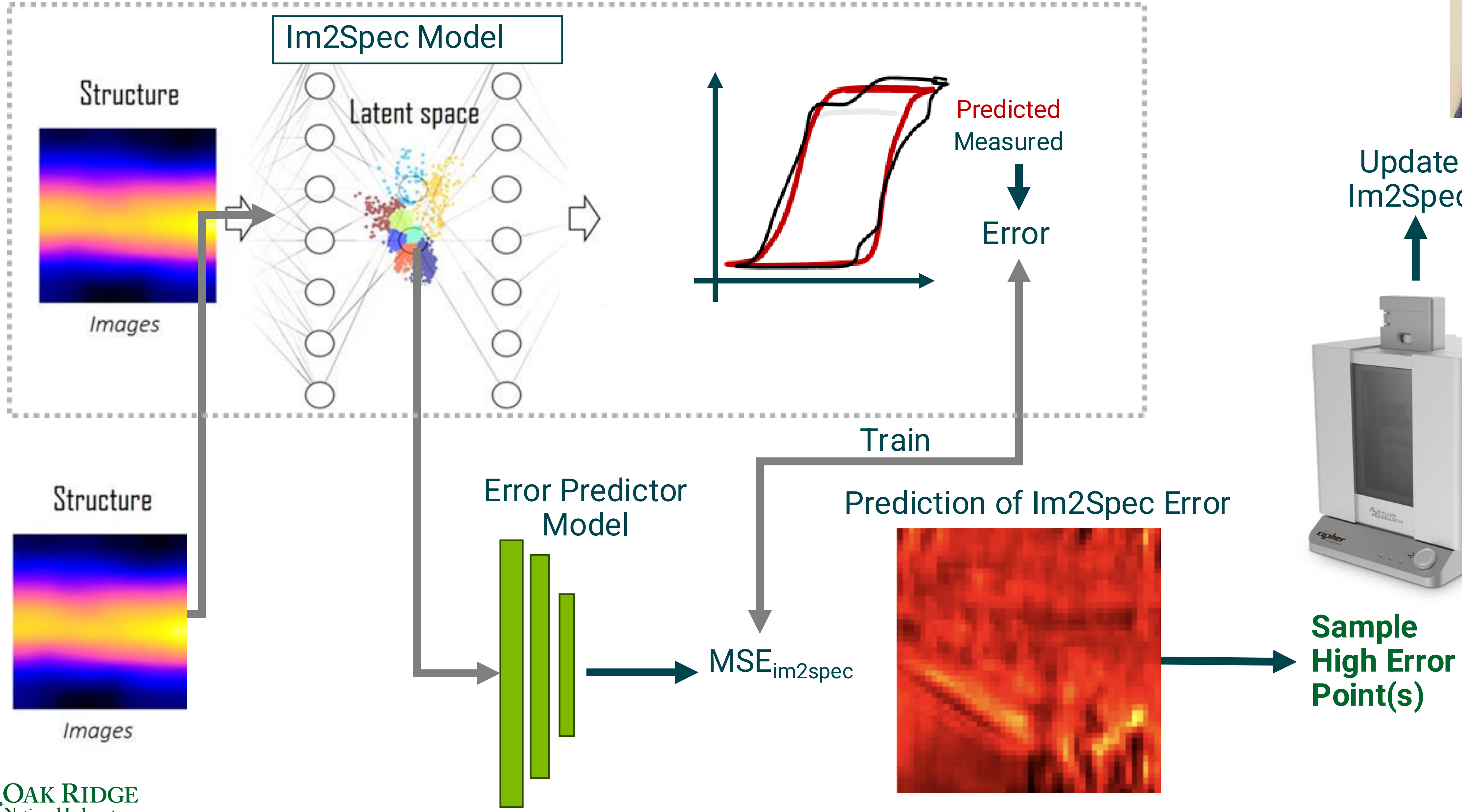
User gets to rate spectra (in this case 15 spectra were rated)

Sample by J. C. Yang (NCKU/Taiwan)

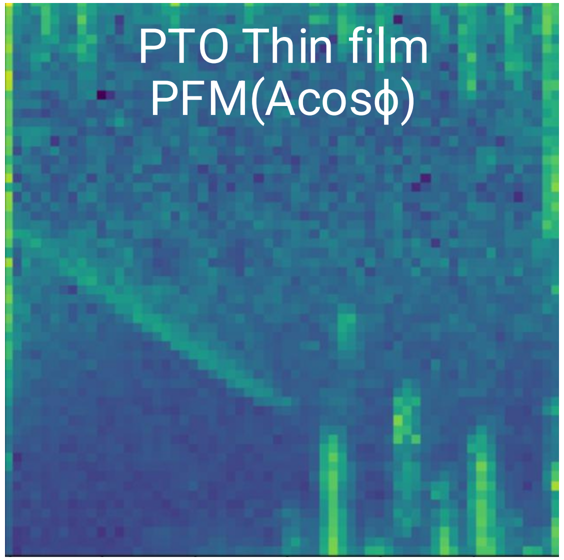
Spectral Recommender system

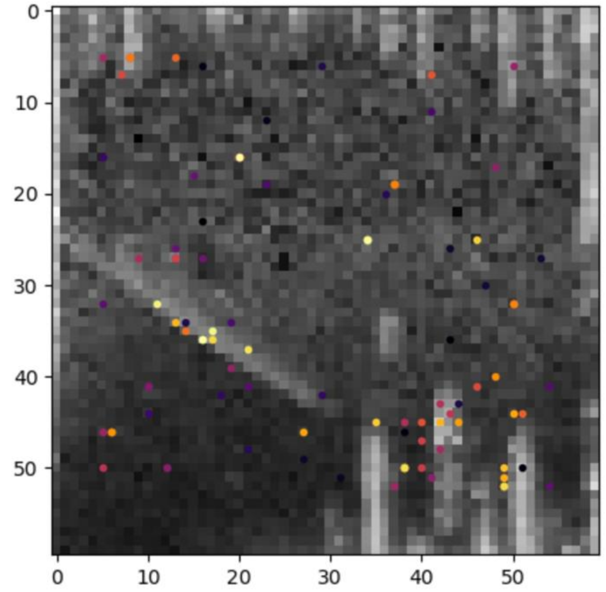
Another alternative: “Curiosity” optimization

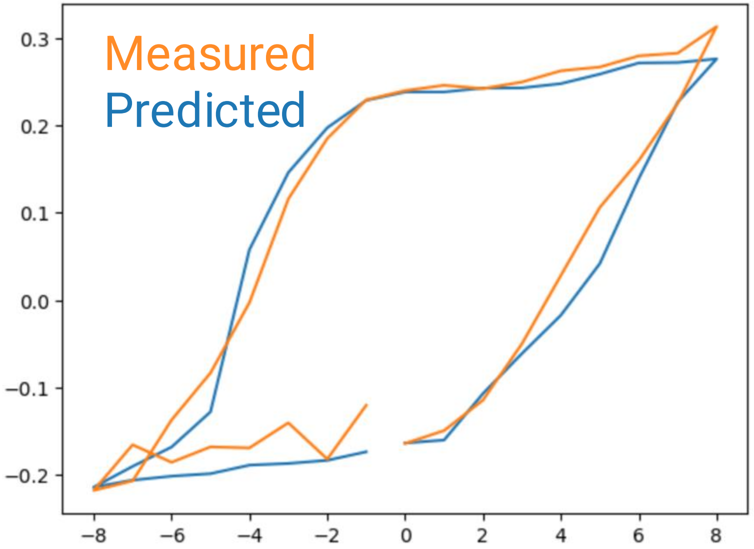
Adi Vatsavai
UNC



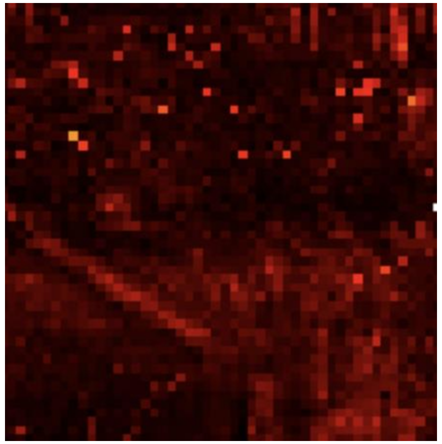
Error prediction model enables efficient sampling!



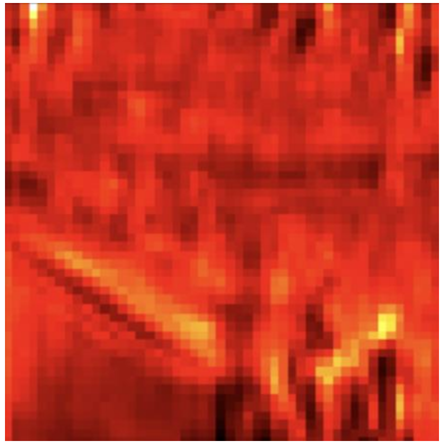




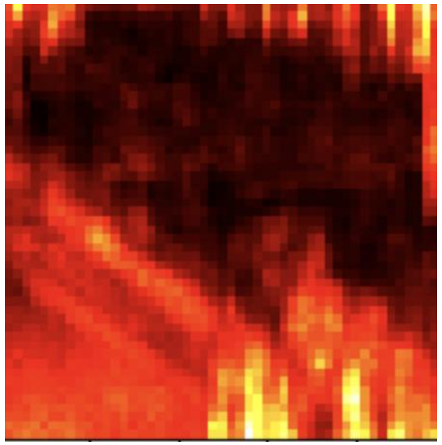
GT Error

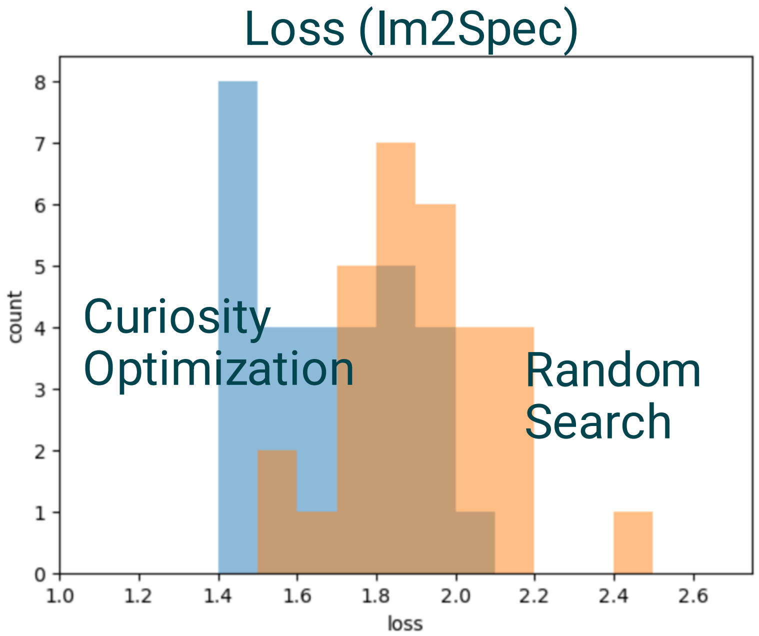


Predicted Error



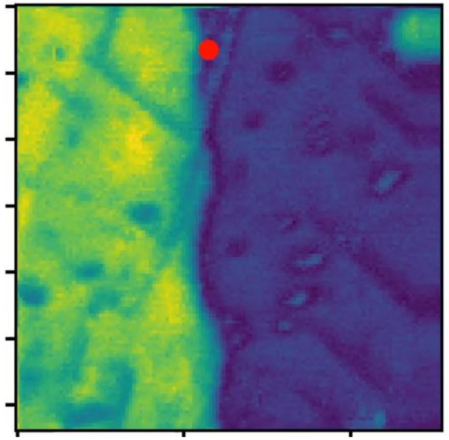
Uncertainty

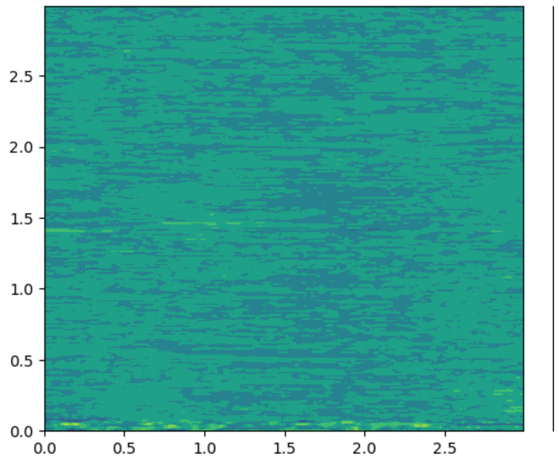


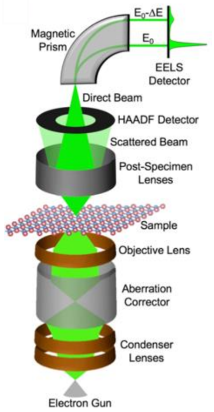


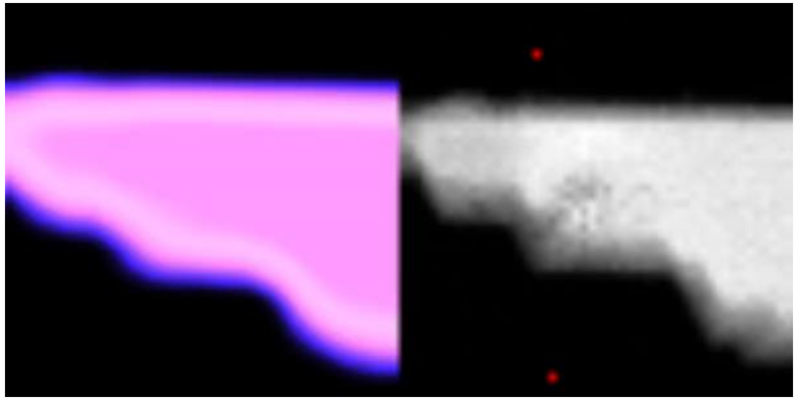
Outline

Autonomous Microscopy

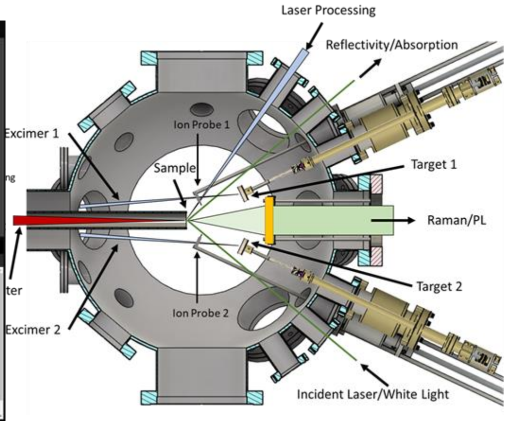


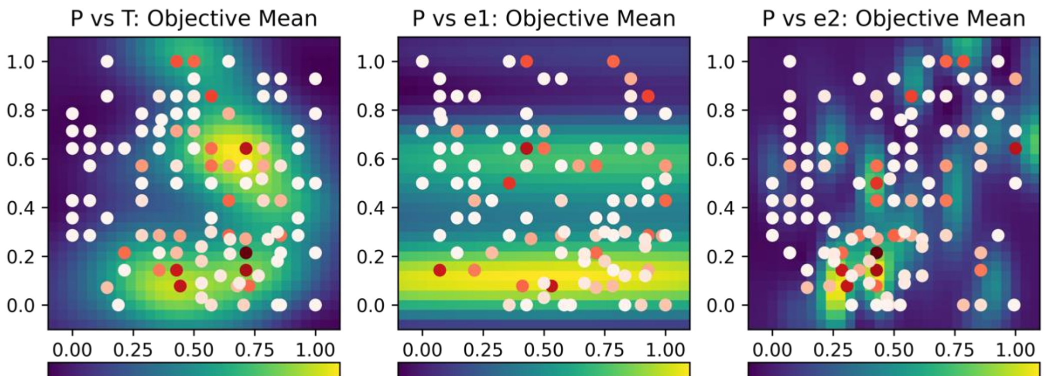






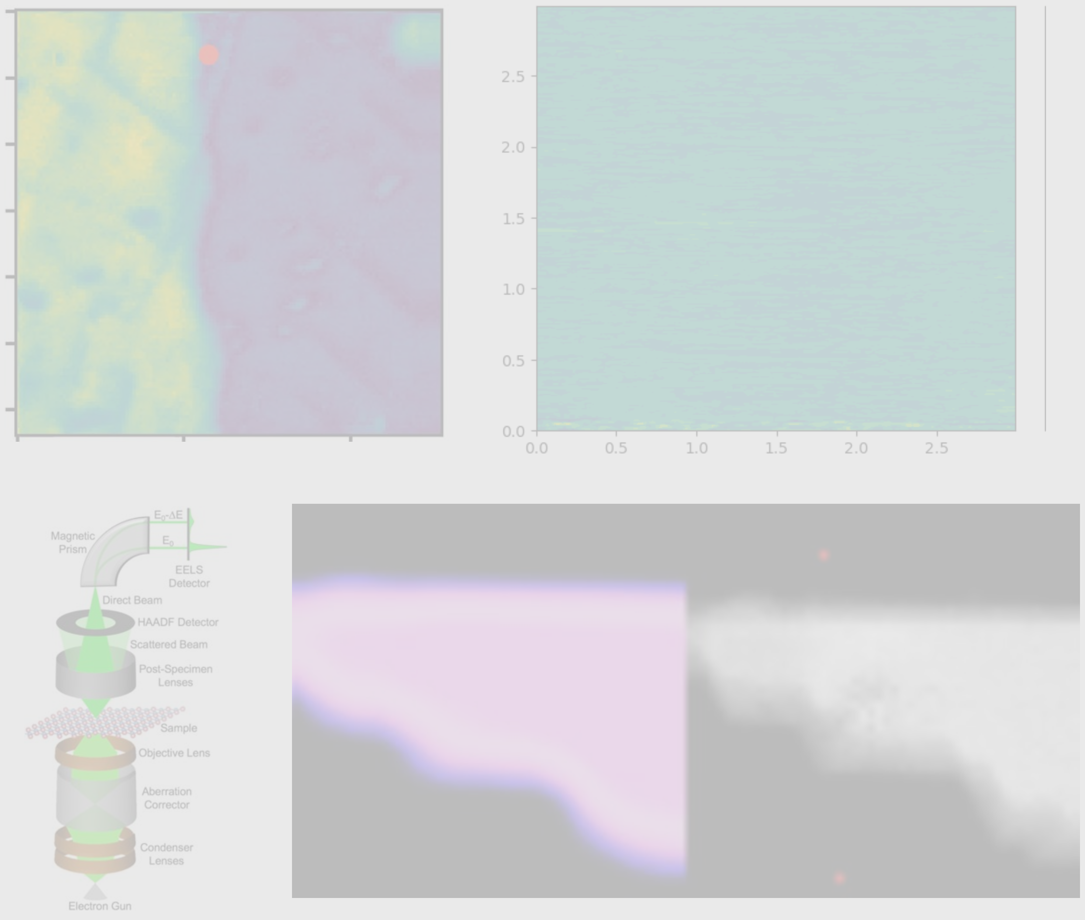
Autonomous Synthesis



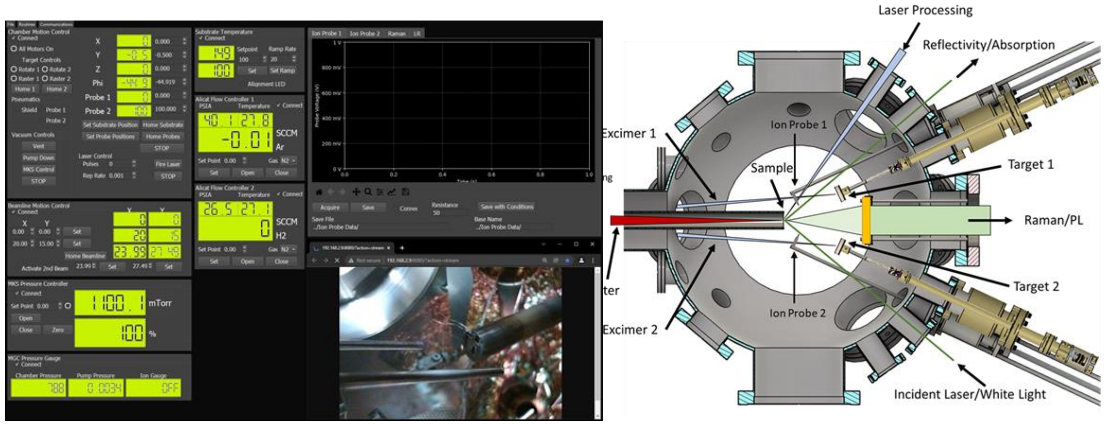


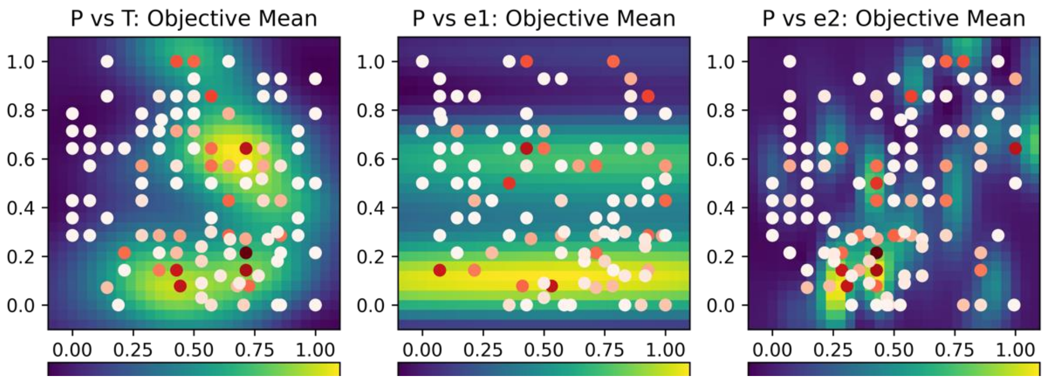
Outline

Autonomous Microscopy



Autonomous Synthesis

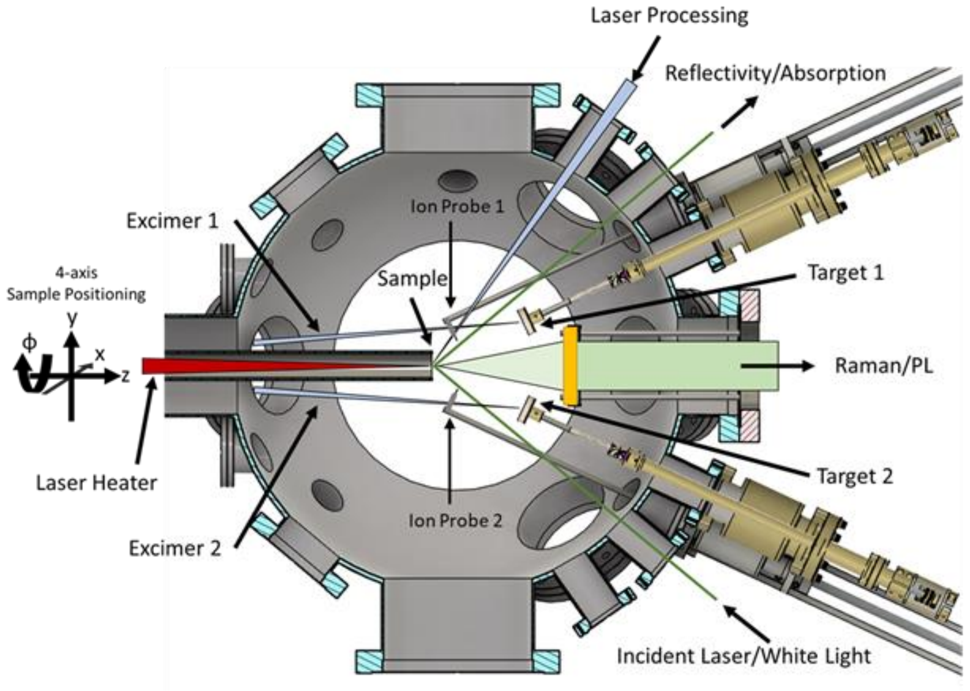




Autonomous Pulsed Laser Deposition

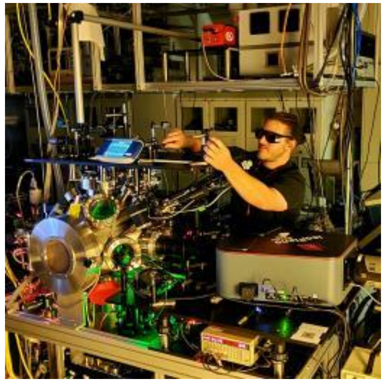
✓ **Full process automation**
✓ **Automated material characterizations**

- **PLD Chamber Design**
  - Two PLD targets enable **simultaneous deposition** of different materials – *tune stoichiometry of deposition*
  - **Plasma diagnostics**: ICCD Imaging, ion probe, gated plume spectroscopy.
  - **Sample diagnostics/characterization**: Raman, PL, laser reflectivity, white light absorbance.
  - **Laser Processing**: x,y nanosecond excimer laser processing of the sample



- **Full automation** of chamber processes
  - Allows in-person and ***remote/programmatic*** control of all PLD processes and diagnostics via Python-> ***enables machine autonomy and seamless integration with ML libraries***
  - **Automated control** of: excimer laser energy, spot size, pulse number, and rep rate; background gas pressure, flow rates, mix of 2 gases; processing laser energy, pulses, rep rate, sample position; etc…

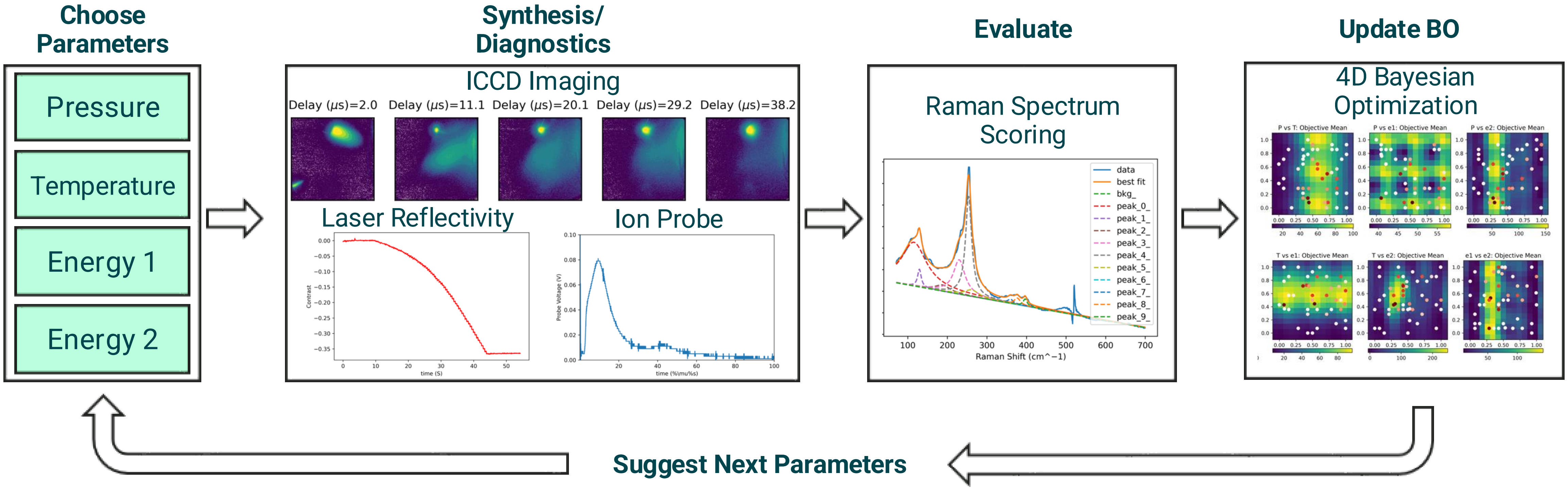
Sumner Harris



Autonomous synthesis of WSe_2

Goal: Autonomously search a broad parameter space to identify good growth conditions for ultrathin (<3 monolayers) WSe_2 by PLD.

- 4D Bayesian optimization to search for optimal Raman "Score"
- Co-deposition PLD, general Workflow:

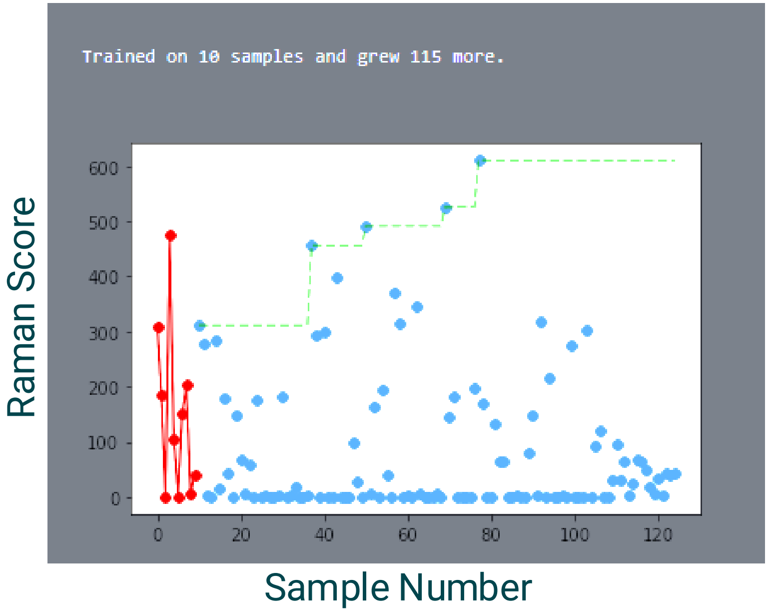


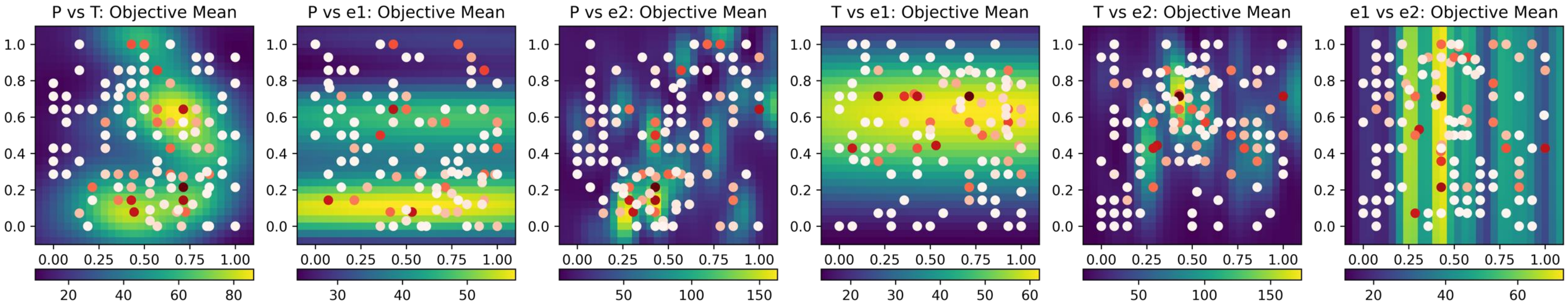
S. B. Harris et al., Small Methods (2024)

Autonomous synthesis of WSe_2

Started with 10 initial samples that were previously grown, autonomously grew 115 more with Bayesian optimization.

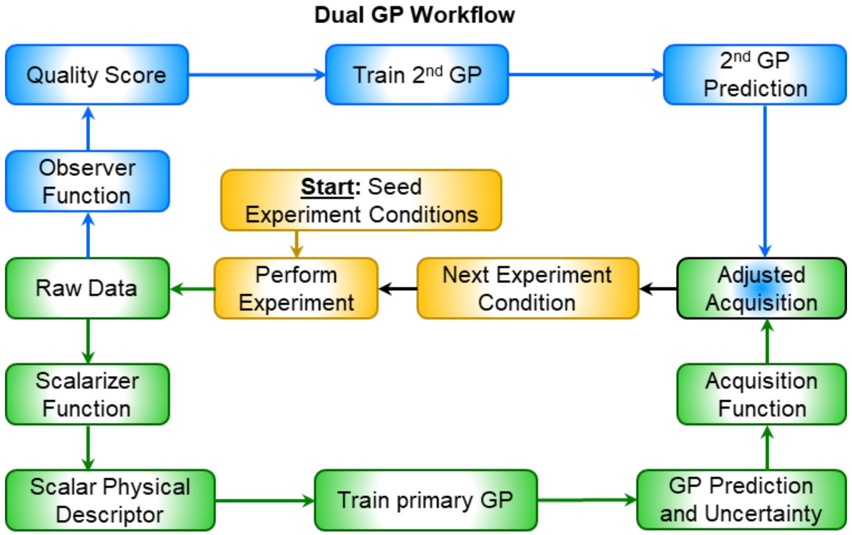
- Result after 115 samples (0.25% of space explored)
  - Showing averaged **surrogate function** projected into each 2D parameter plane
  - Seems **insensitive to WSe_2 laser energy**
  - Se laser energy seems critical -> ***Needs Se compensation***
  - Potential identification of multiple growth regimes



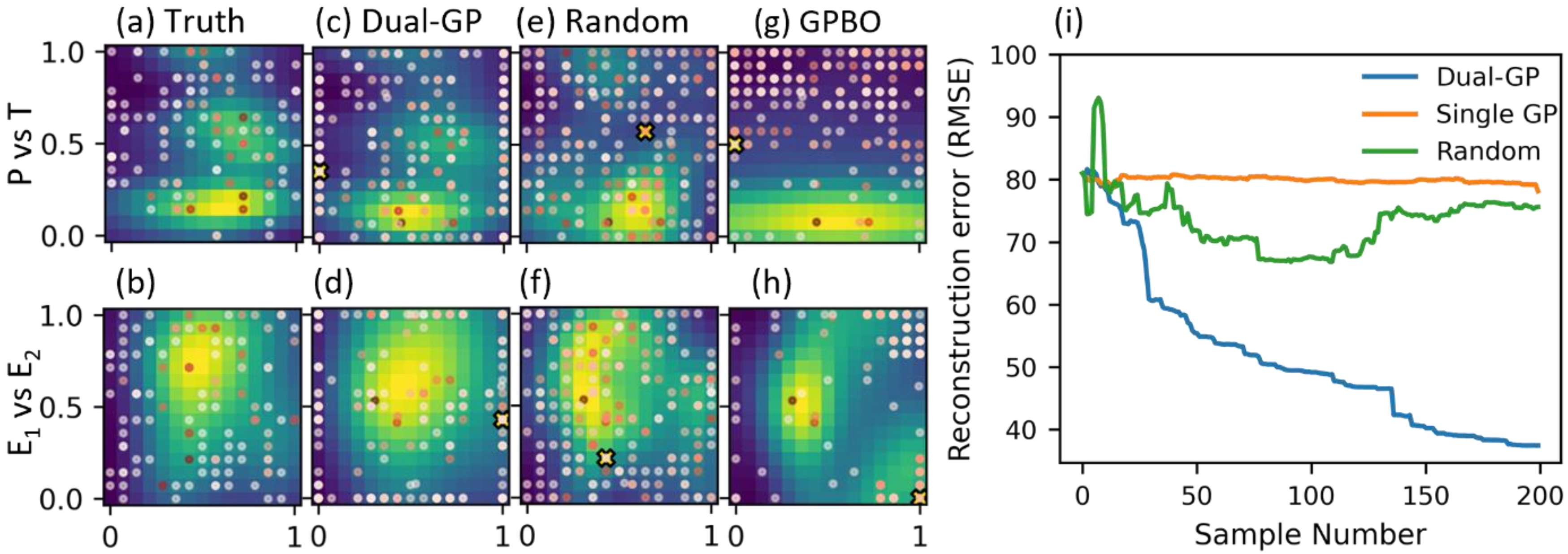


Colormap of predicted Raman spectrum score projected into each 2D plane.

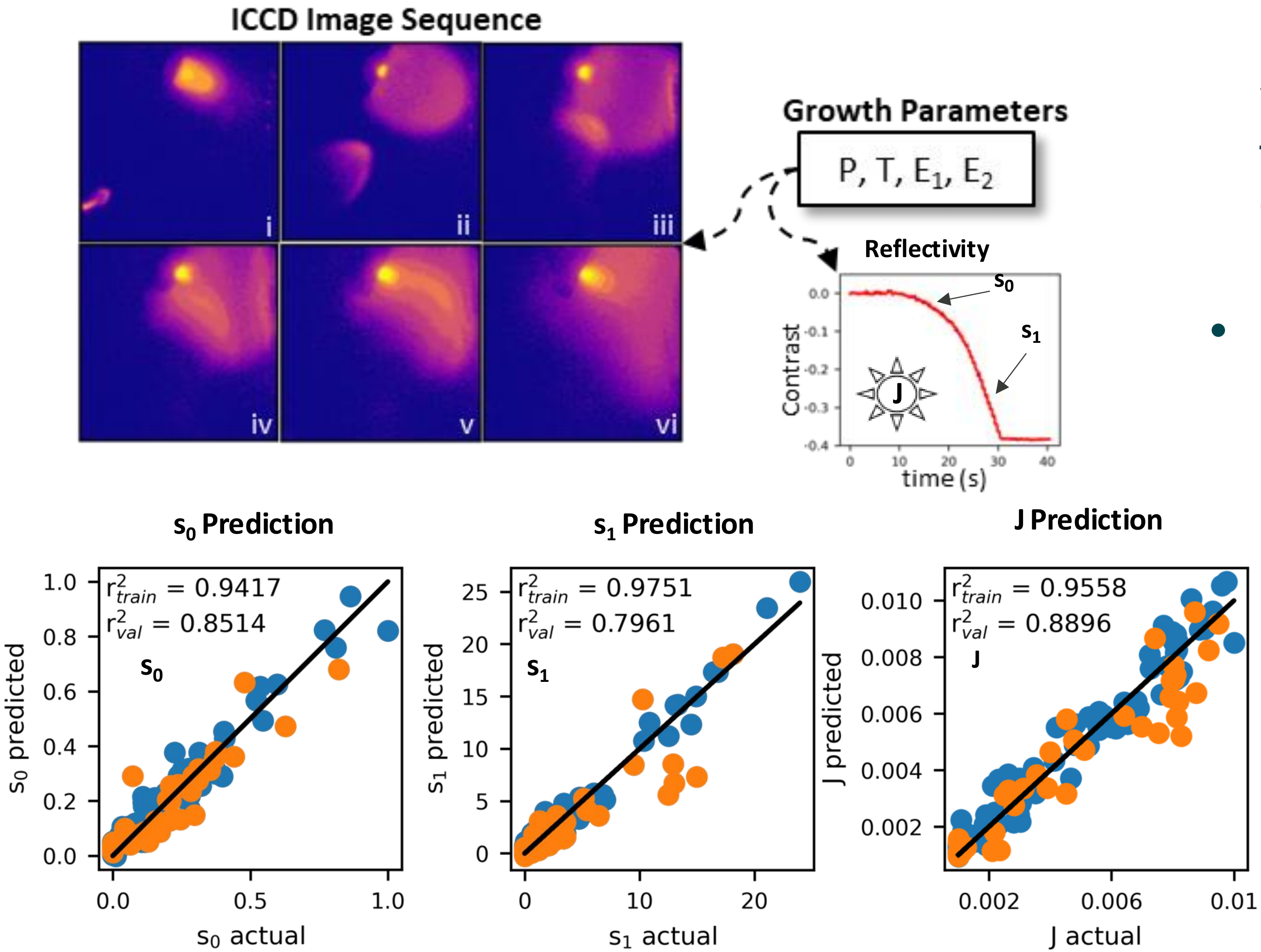
Dual GP method to reduce 'bad' samples



- One of the problems is that samples are wasted if the quality of the spectra is not sufficient for the scalarizer
- A dual GP method is proposed to alleviate this, that forces exploration only in areas where the quality of the score is sufficiently high



In-situ plume imaging



- By imaging the PLD plume, we can predict values of pressure, temperature, and laser fluences
- We can also estimate parameters of a kinetic model!

S. B. Harris, C. M. Rouleau, K. Xiao, and R. K. Vasudevan. *npj Comput. Mater.* **10**, 105 (2024).

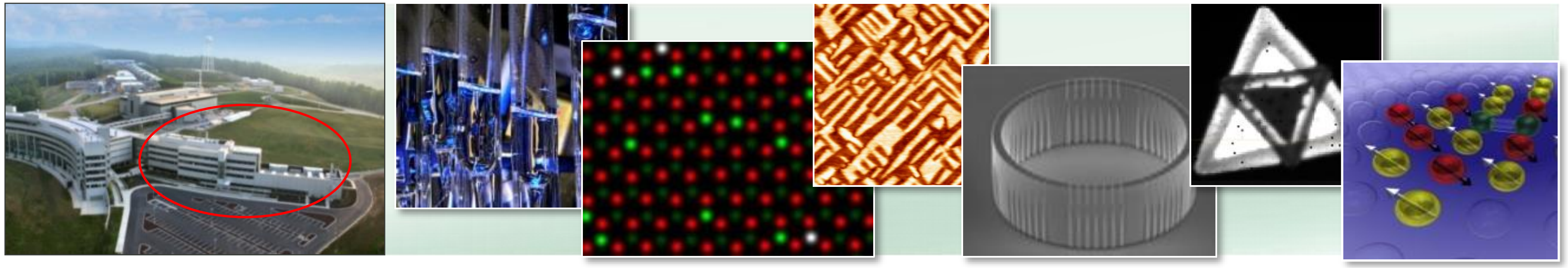
CNMS is a national user facility with a mission to advance nanoscience

- About CNMS:

- **Unlike many user facilities, you don't need to have samples to apply for time**
- **Two calls per year for continuous access; anytime for short-term projects**
- **Simple 2-page proposal**
- **Free access to laboratories, equipment and expertise if you agree to publish**
- **Proposal deadlines: early May and mid-October**
- **Joint proposals with neutron sources (SNS, HFIR)**

- Research areas:

- **Synthesis –** 2D, precision synthesis, selective deuteration
- **Nanofabrication –** direct-write, microfluidics, cleanroom
- **Advanced Microscopy –** AFM, STM, aberration-corrected TEM/STEM, atom-probe tomography
- **Functional Characterization –** laser spectroscopy, transport, magnetism, electromechanics
- **Theory and Modelling –** including gateway to leadership-class high performance computing



CNMS is a Nanoscale Science Research Center supported by the U.S. Department of Energy, Office of Science, Scientific User Facilities Division

OAK RIDGE National Laboratory

Beyond Bayesian Optimization

Simulations and human experts can enable more efficient learning