

Special Topic: Learning Physics From Images

Rama Vasudevan

Virtual summer school on Machine Learning in Electron Microscopy

Center for Nanophase Materials Sciences, Oak Ridge National Laboratory

August 22nd 2023

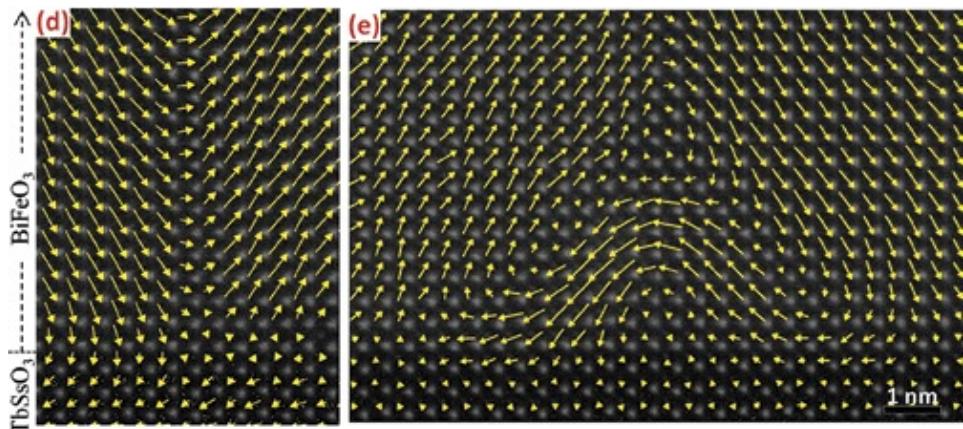


ORNL is managed by UT-Battelle, LLC for the US Department of Energy

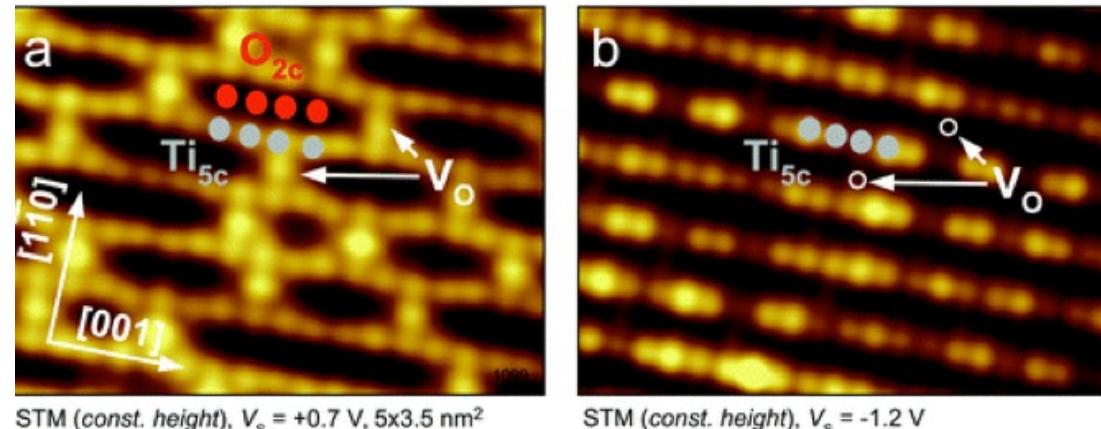


Motivation

Each day, we collect more and more high resolution imaging data through advanced microscopy



Closure states in BiFeO_3
Nano Lett., 2011, 11 (2), pp 828



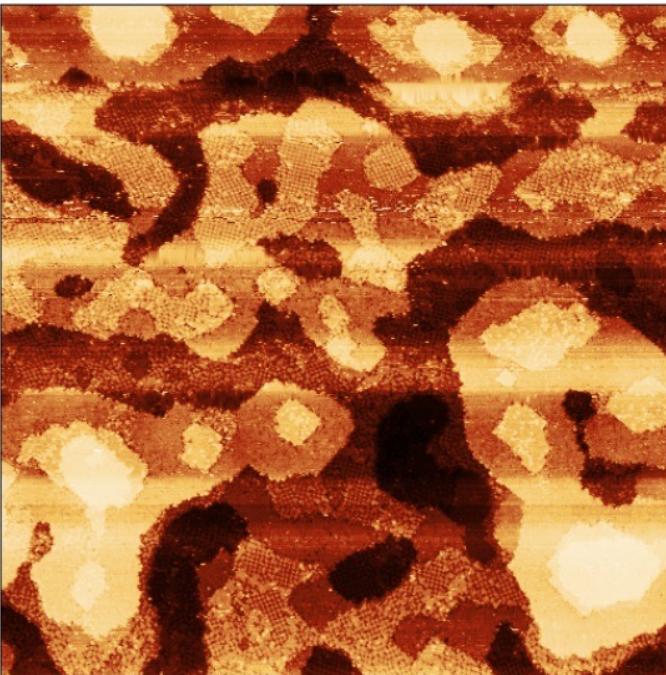
STM rutile $\text{TiO}_2(110)$ surface
Chem. Soc. Rev., 2017, 46, 1772

While image analysis techniques are well developed, only limited amount of transferable materials science knowledge is extracted from atomically resolved images

Data Driven approaches to learning from images

Information from atomic positions: Local Crystallography

Ag-doped LCMO (STM)



Physics is Local

REVIEW

doi:10.1038/nature14453

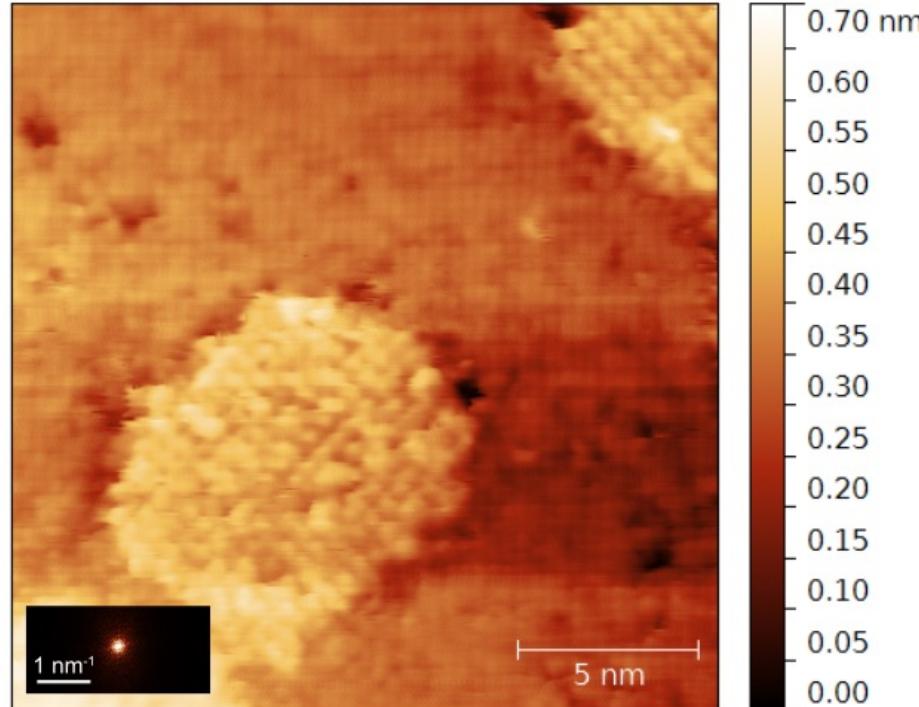
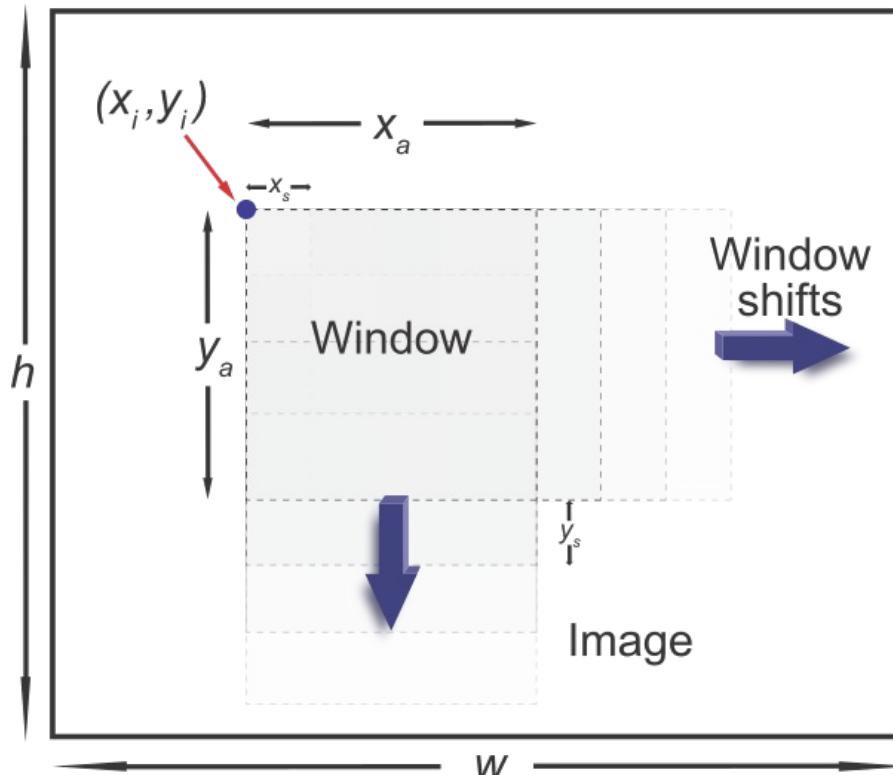
The crystallography of correlated disorder

David A. Keen¹ & Andrew L. Goodwin²

Classical crystallography can determine structures as complicated as multi-component ribosomal assemblies with atomic resolution, but is inadequate for disordered systems—even those as simple as water ice—that occupy the complex middle ground between liquid-like randomness and crystalline periodic order. Correlated disorder nevertheless has clear crystallographic signatures that map to the type of disorder, irrespective of the underlying physical or chemical interactions and material involved. This mapping hints at a common language for disordered states that will help us to understand, control and exploit the disorder responsible for many interesting physical properties.

Need Local Crystallography!

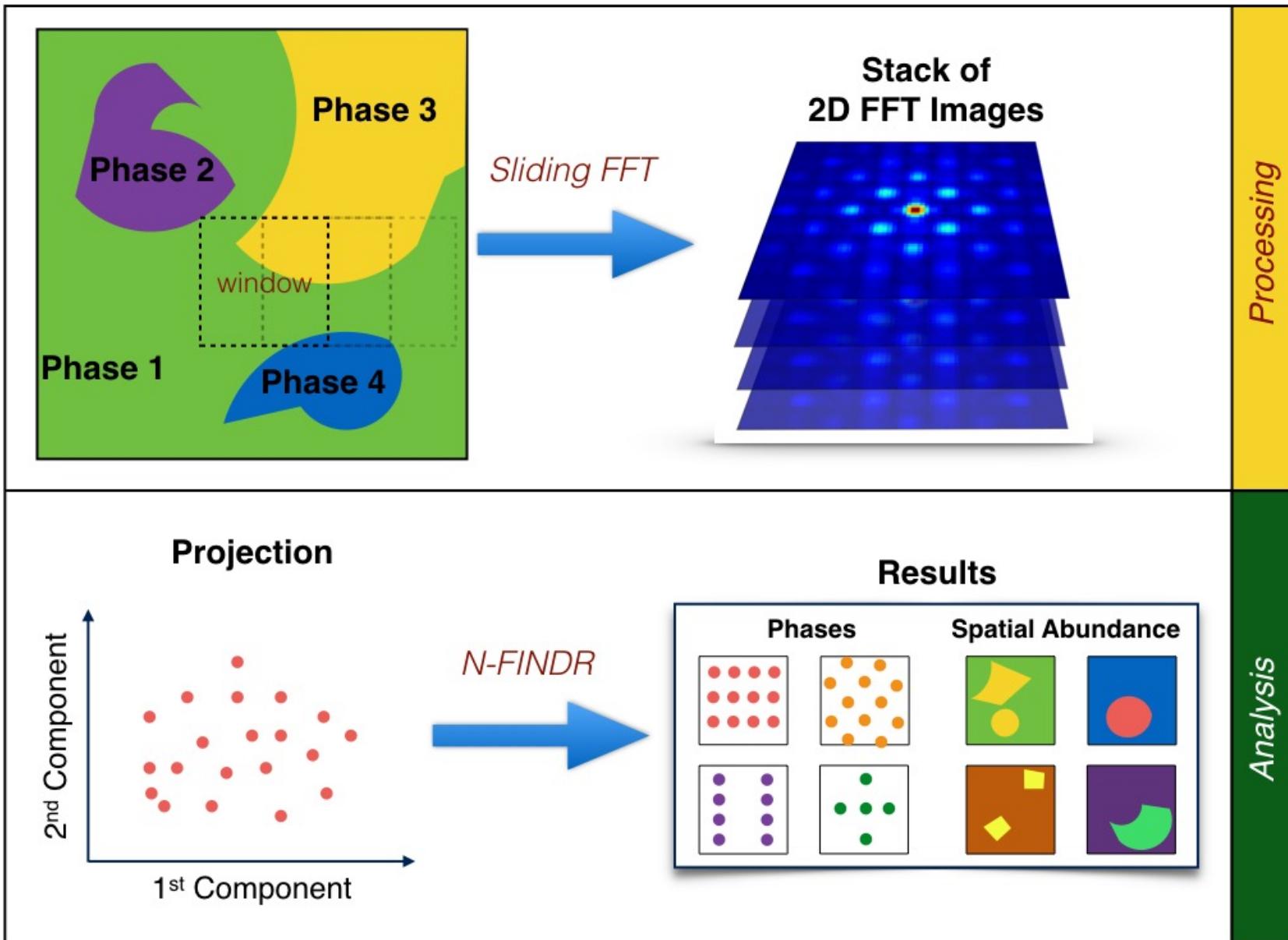
Sliding FFT: More phase identifications



Vasudevan et al., APL 106, 091601 (2015)

- Sliding FFT can be used to determine structure factors which can then be unmixed by appropriate algorithms, such as ICA, N-FINDR or Bayesian Linear Unmixing

Analysis of FFT Stack



Use of N-FINDR

- Spectra for a given pixel is assumed to be a linear combination of the end-member spectra (+ Gaussian noise). The mixing proportions sum to 1

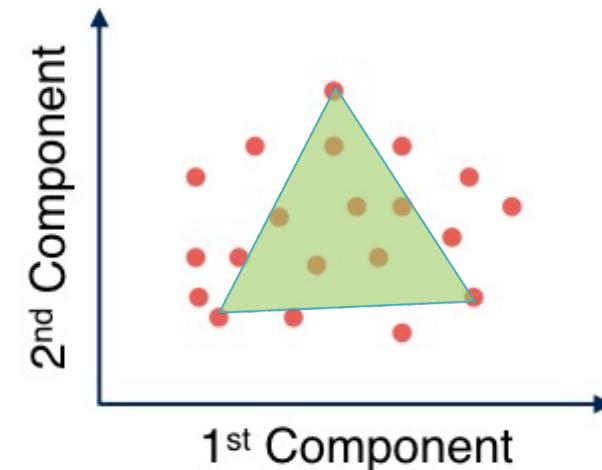
Physics constraint

$$p_{ij} = \sum_k e_{ik} c_{kj} + \varepsilon$$

$$\sum_k c_{kj} = 1$$

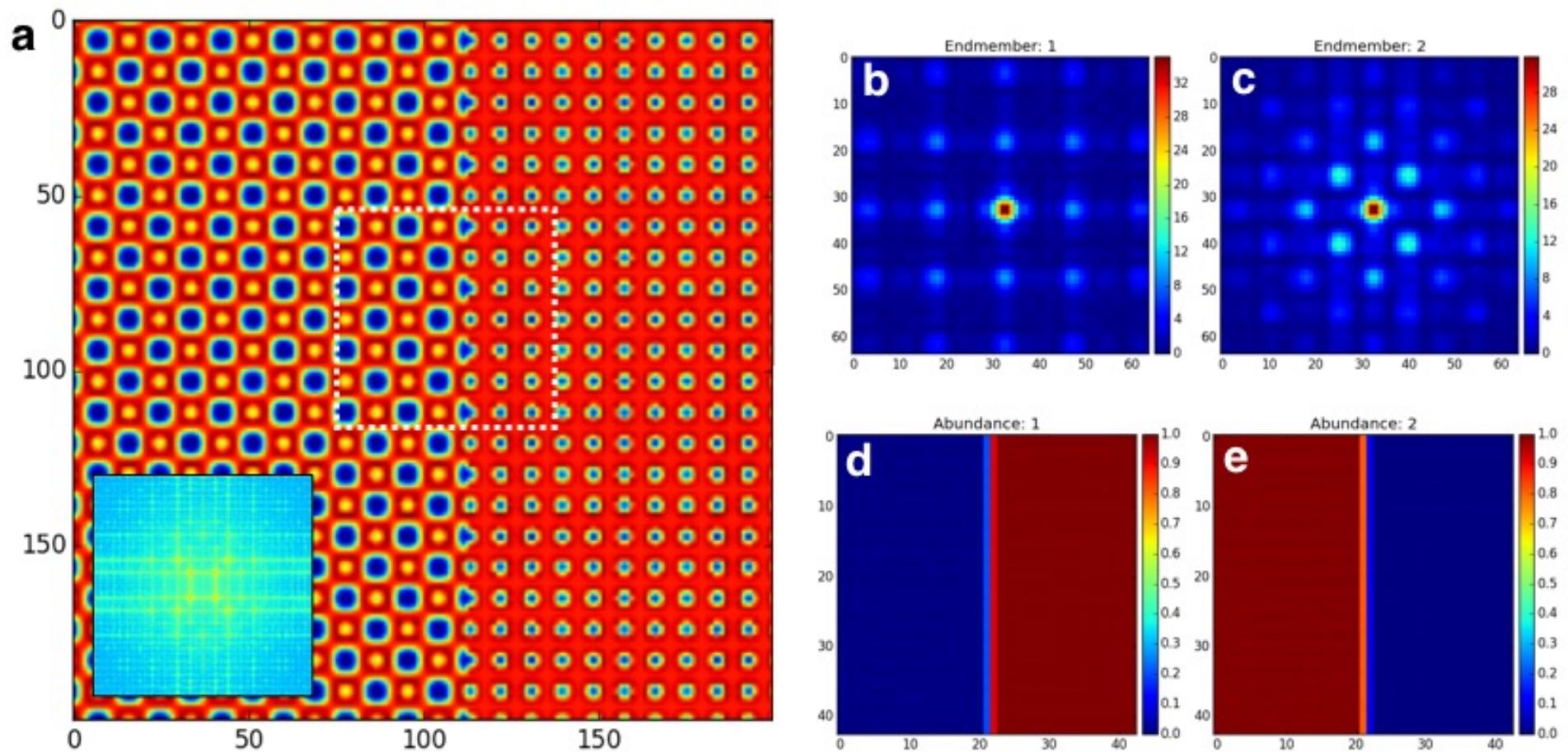
- Let E be the matrix of end-members (here, 3).

$$E = \begin{bmatrix} 1 & 1 & 1 \\ \vec{e_1} & \vec{e_2} & \vec{e_3} \end{bmatrix} \quad V \left(\frac{\mathbf{1}}{(l-1)!} \right) |\det(E)|$$

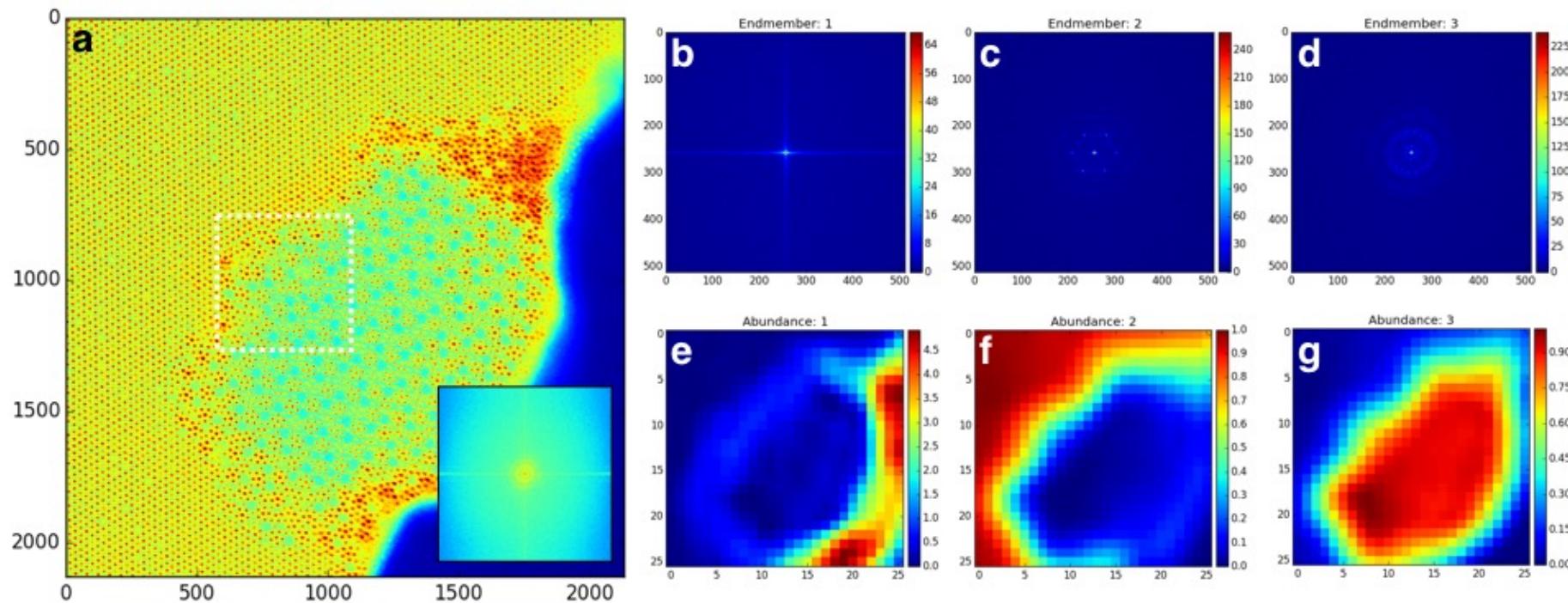


- Iteratively select endmembers, accepting the new selection if the volume increases

Ideal test case



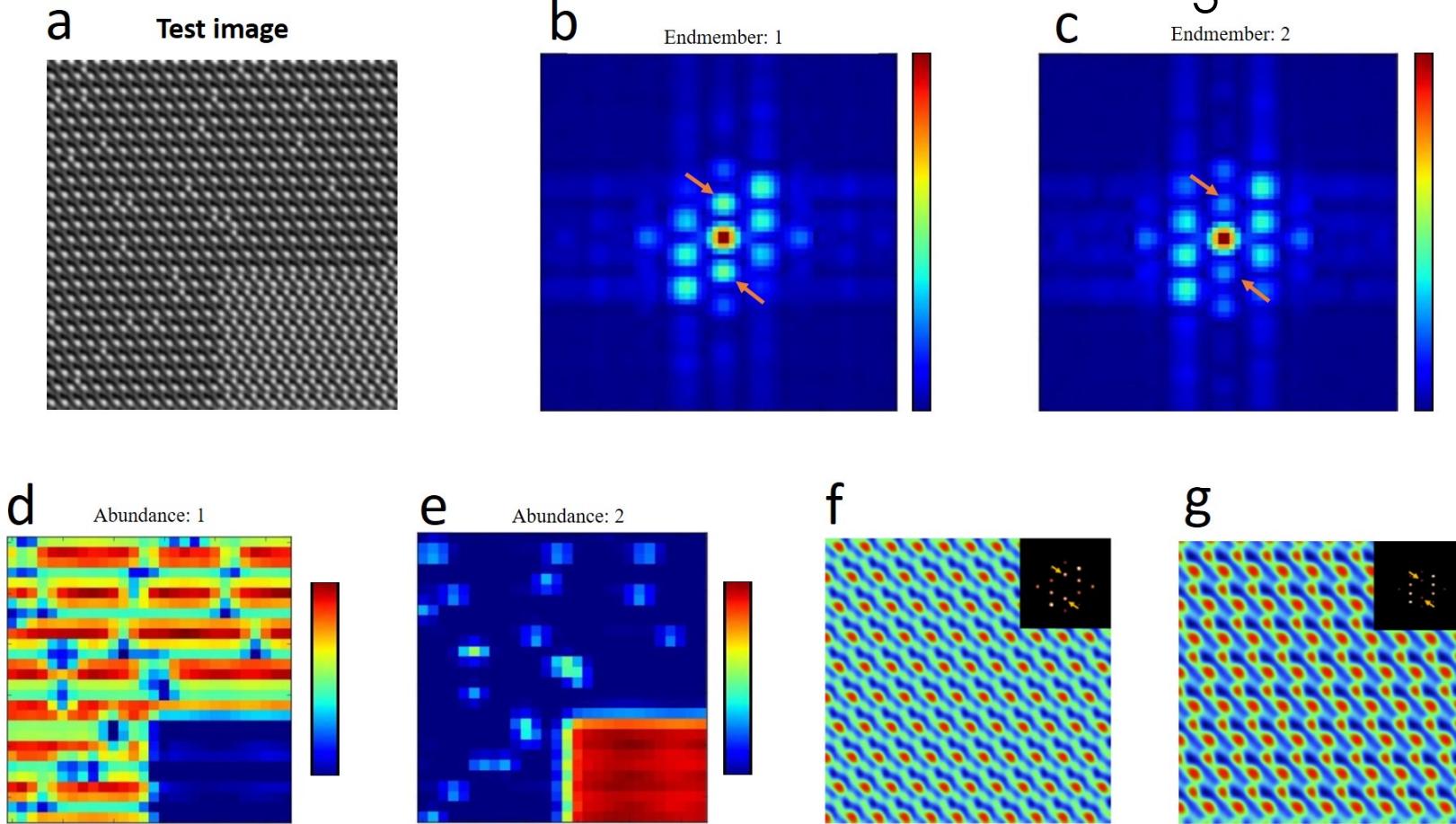
Sliding FFT on an oxide catalyst



- Combination of N-FINDR and Sliding FFT can quickly and automatically determine the two phases present and their spatial abundance.

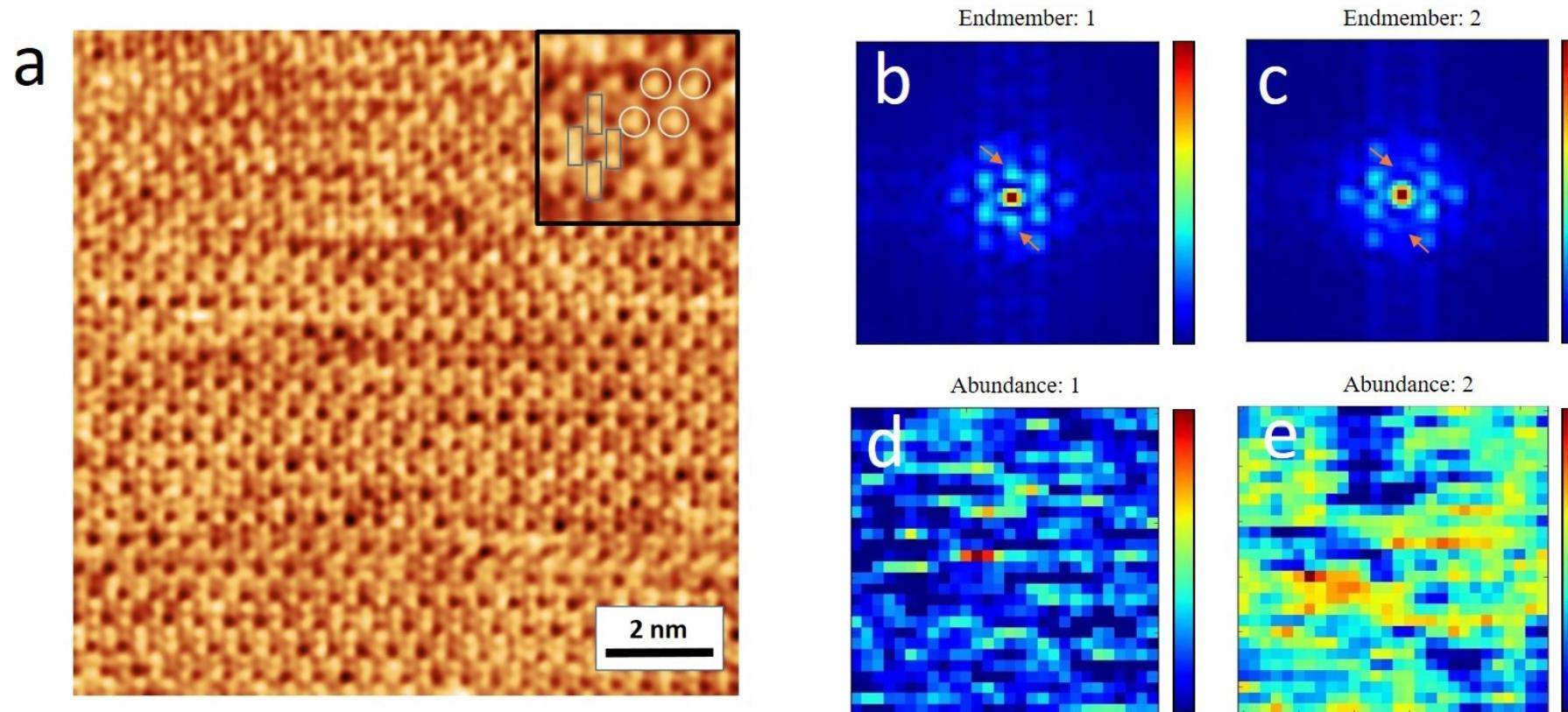
Vasudevan et al. Nano Lett. 16, 5574 (2016)

Fine structures in Electronic Superlattice in a correlated Mott insulator α – RuCl₃



- What would dimers overlayed on hexagonal lattice look like?
- Strong suppression of vertical spots in FFT for dimers
- Scale bar: 0 to 1

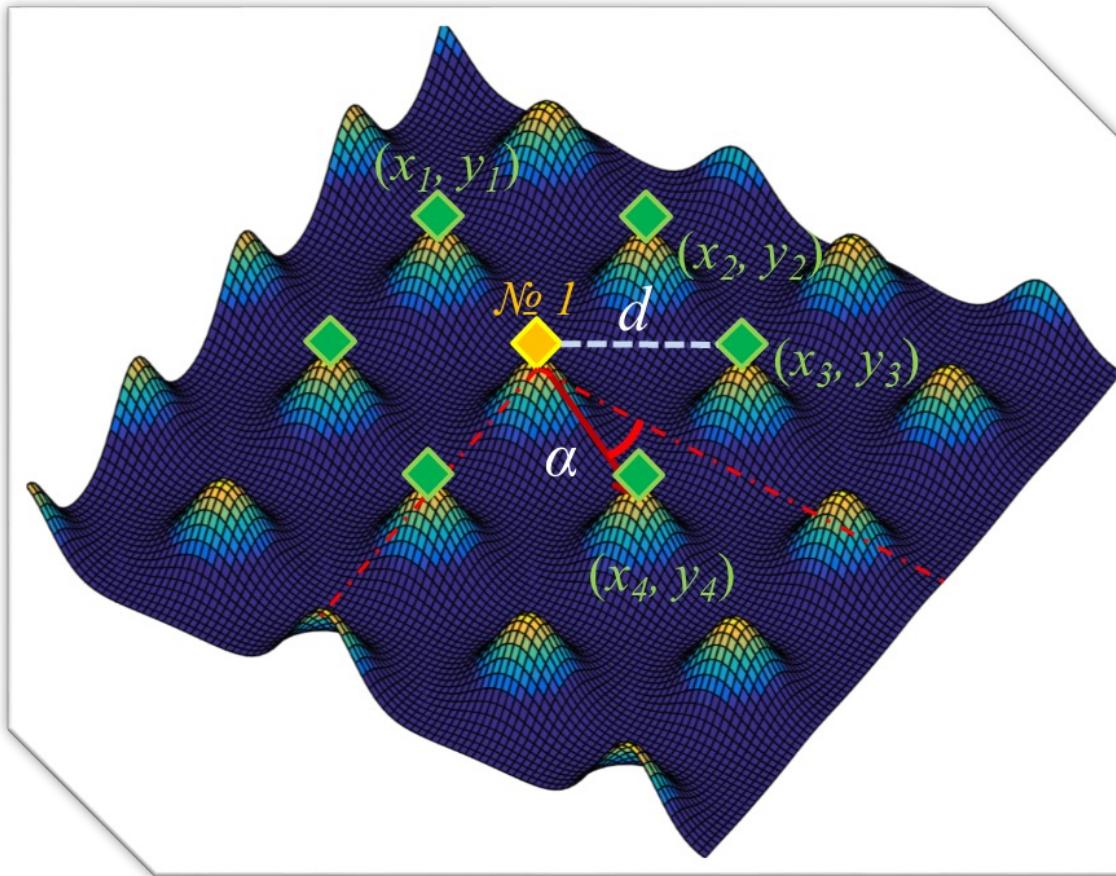
Fine structures in Electronic Superlattice in a correlated Mott insulator



- Clear evidence of dimerization in RuCl_3 through Sliding FFT/N-FINDR

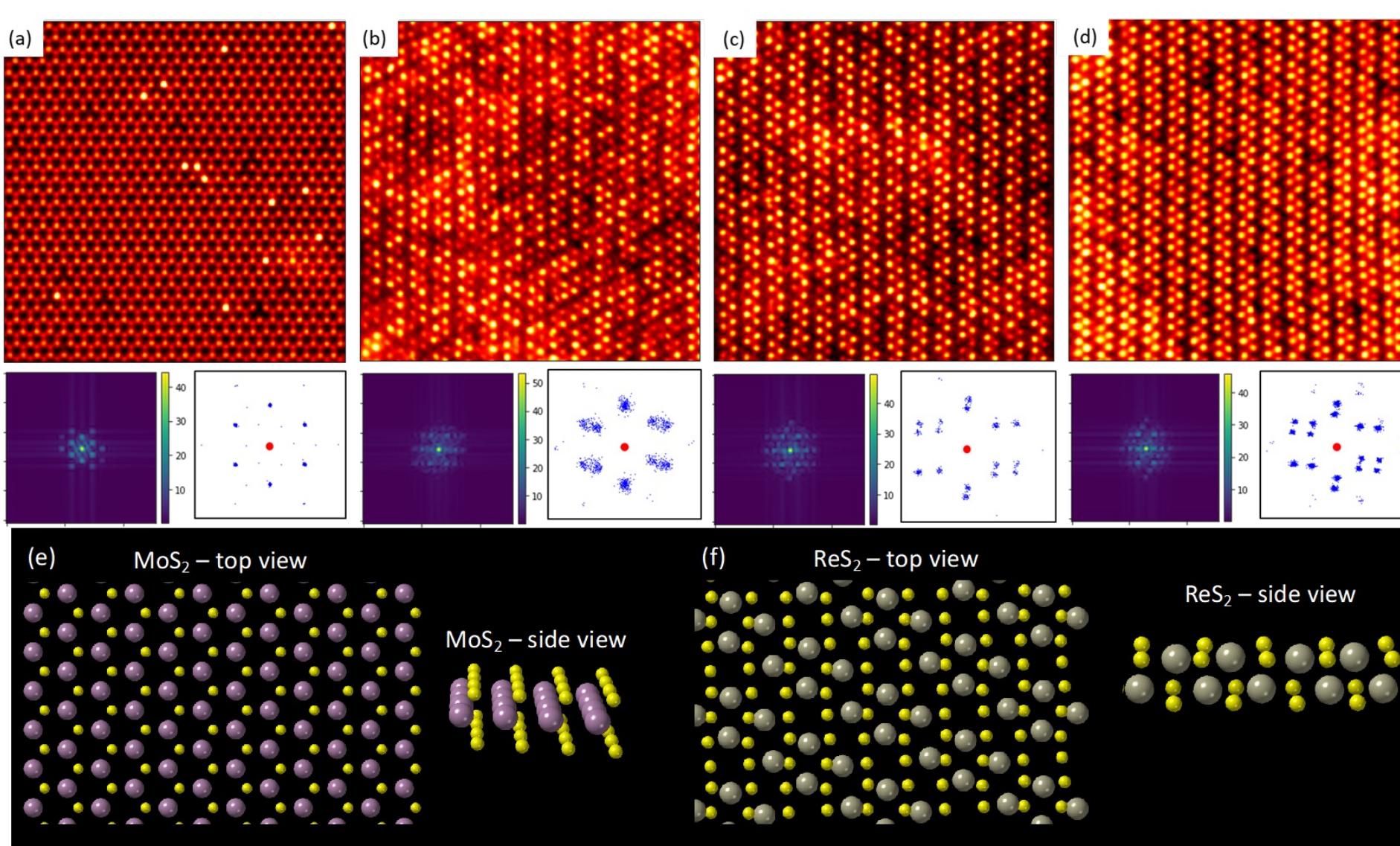
Vasudevan et al. Nano Lett. 16, 5574 (2016)

Atomic positions: Obtaining statistics



- Find atoms using some algorithms
- Determine the distance from each individual atom to the J nearest neighbors, and compute the angles
- Iterate over all N atoms in image to build the matrix of size $N \times J \times 2$
- Compute PCA over $N \times J$ (angle or distance metric)

Example: MoS₂-ReS₂

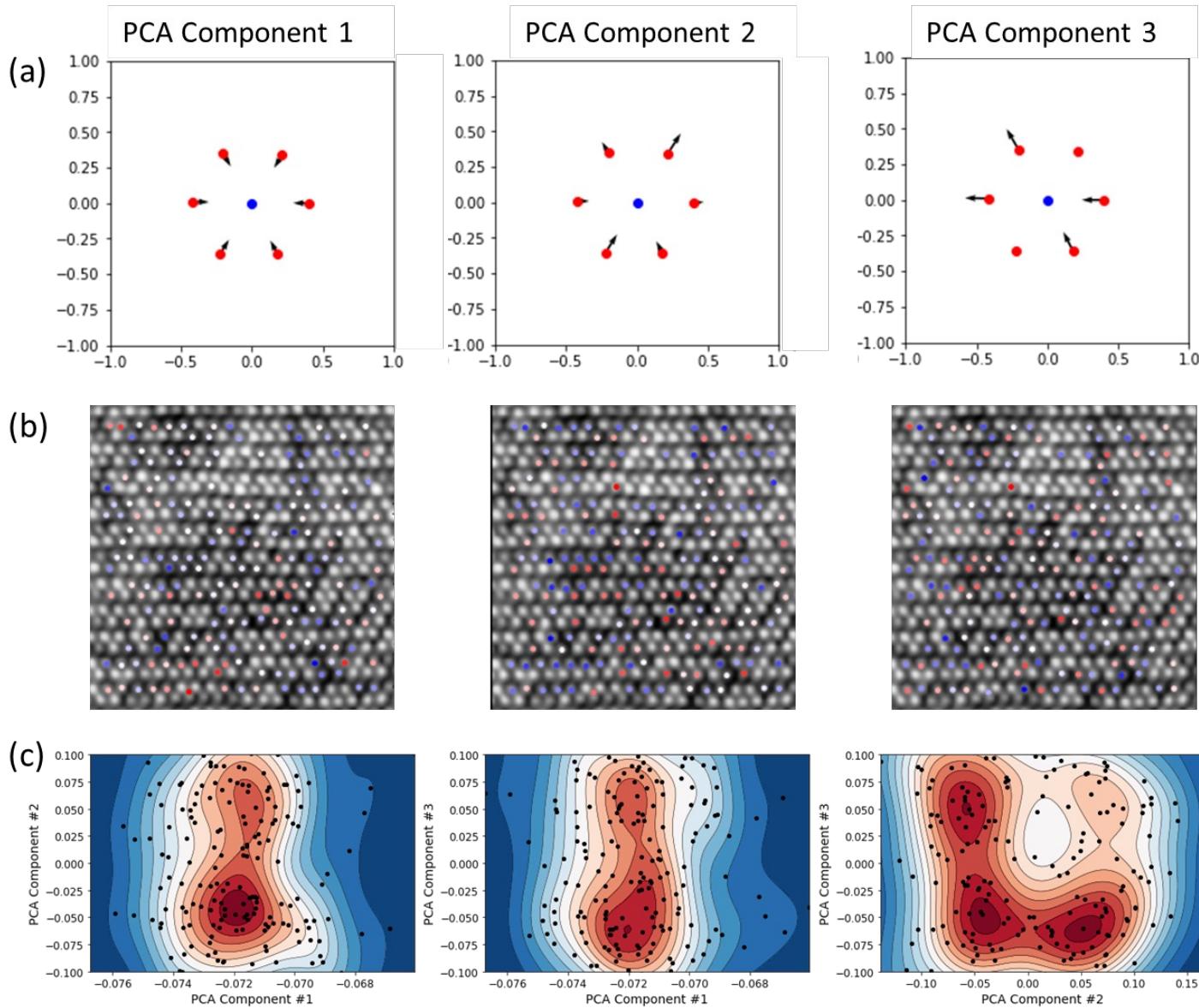


Phase transition
between MoS₂
("H") and ReS₂
("T").

Note that there
must exist
broken
bonds/defects

Question: Does
transformation
occur from
global
composition, or
is it locally
driven?

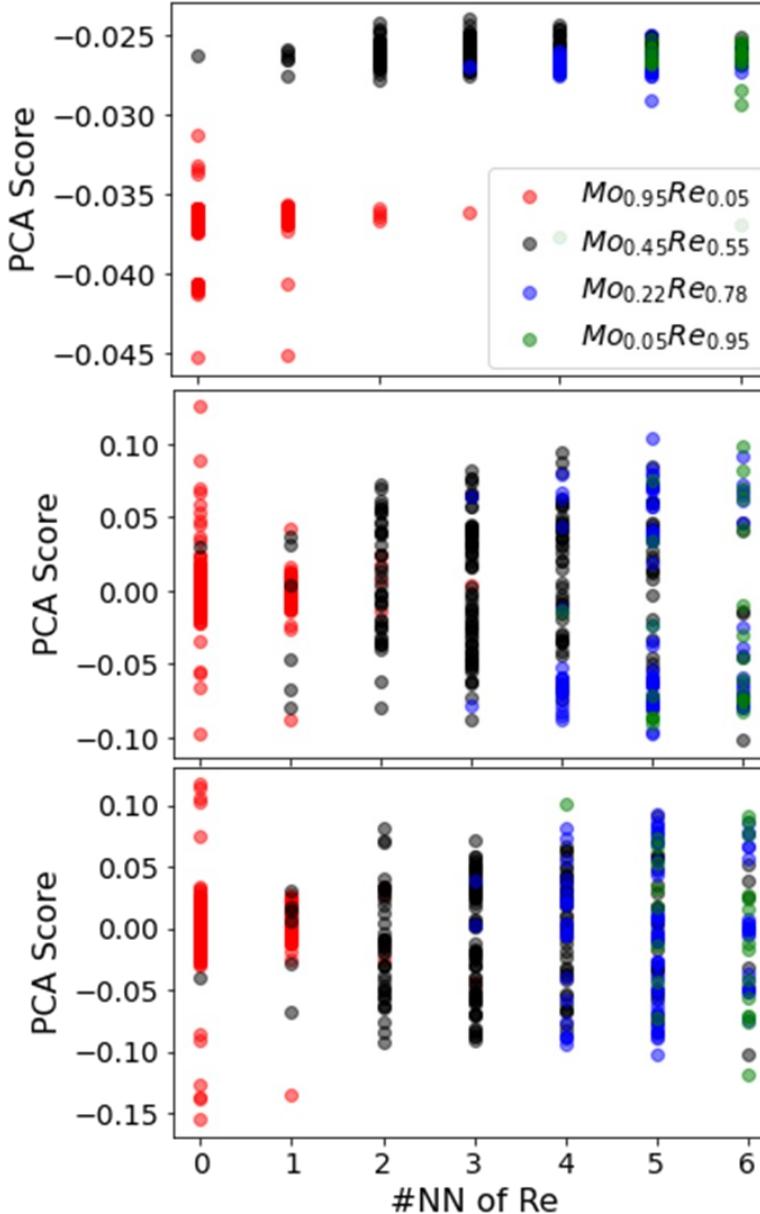
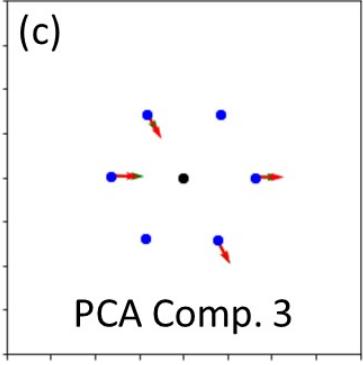
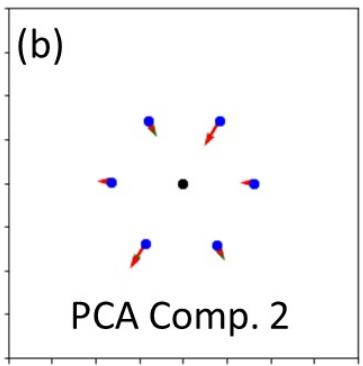
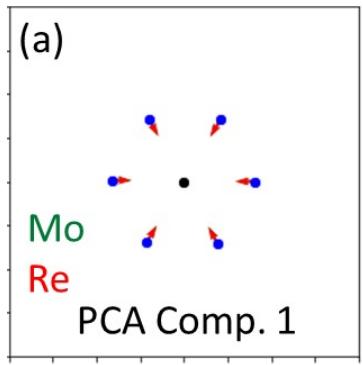
Principal Component Analysis



- PCA of local atomic neighborhoods shows progression of distortions across the phase transition

Vasudevan et al. Appl. Phys. Rev. **8**, 011409 (2021).

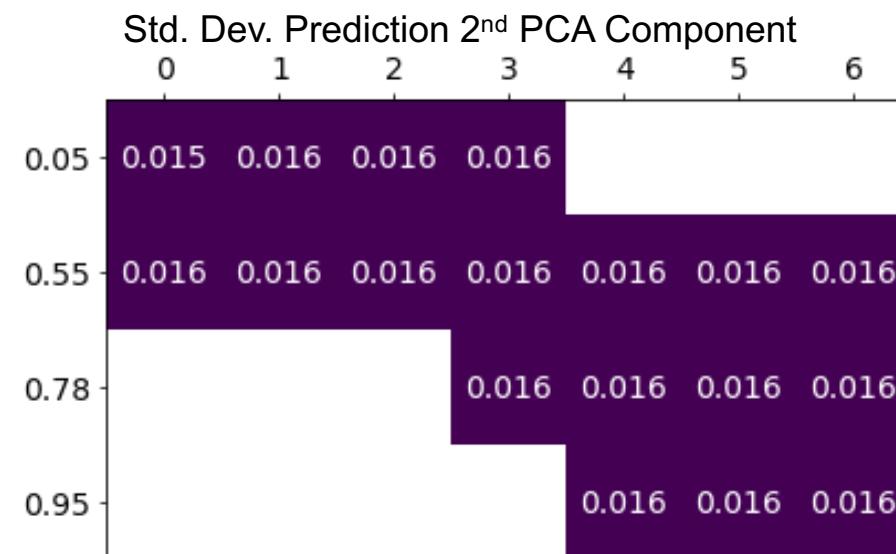
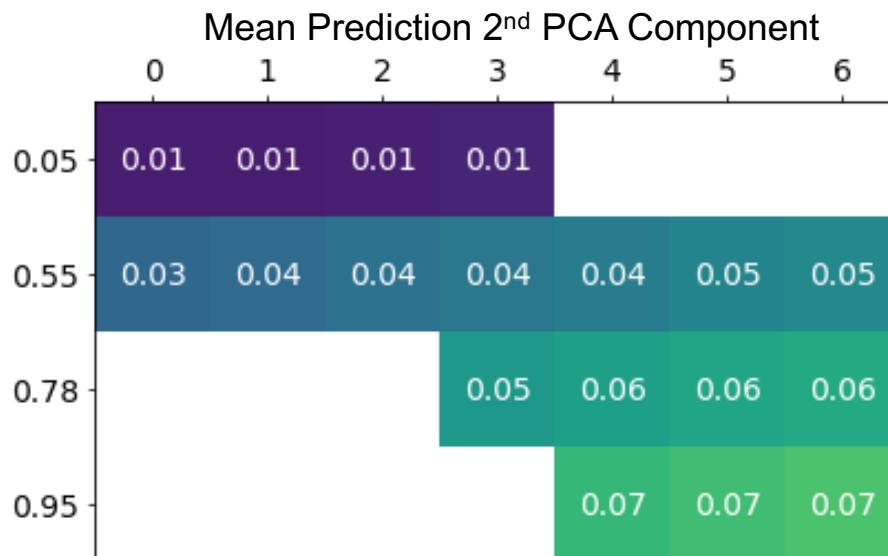
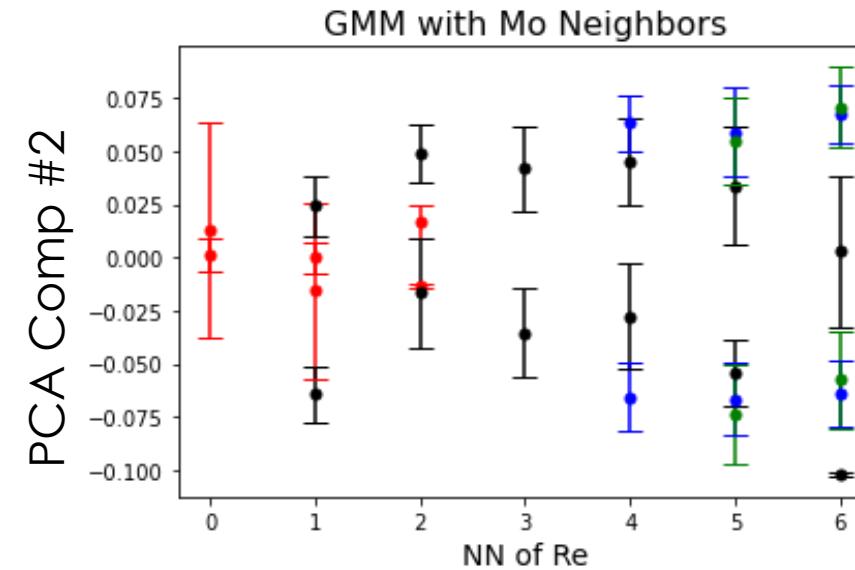
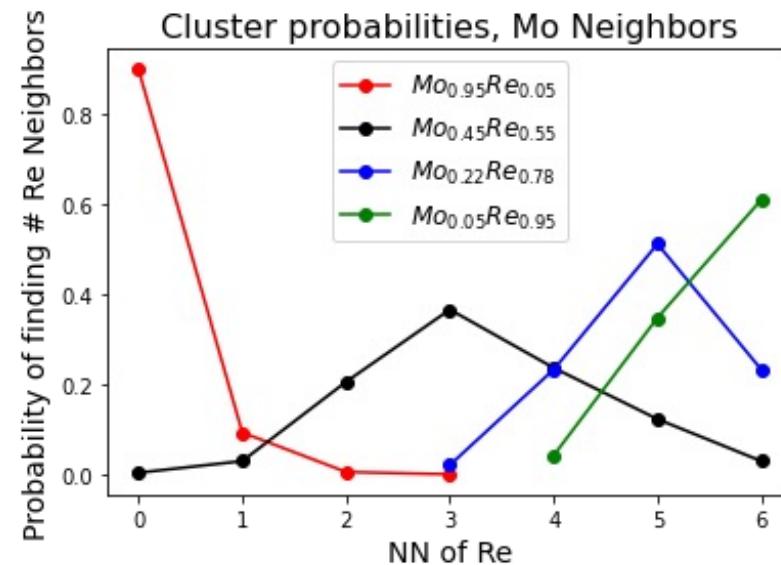
Principal Component Analysis



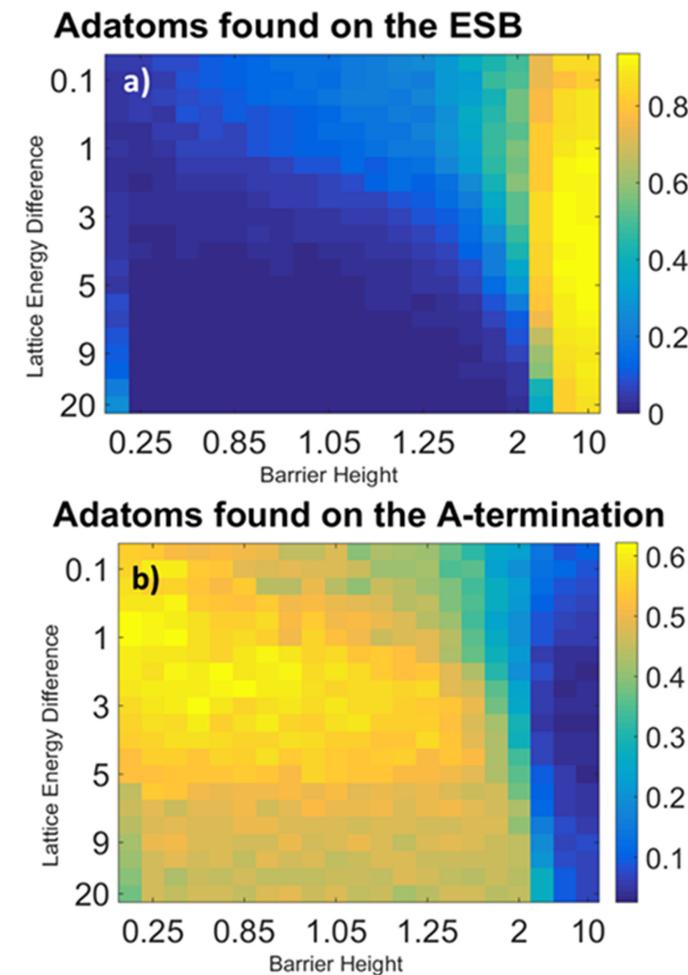
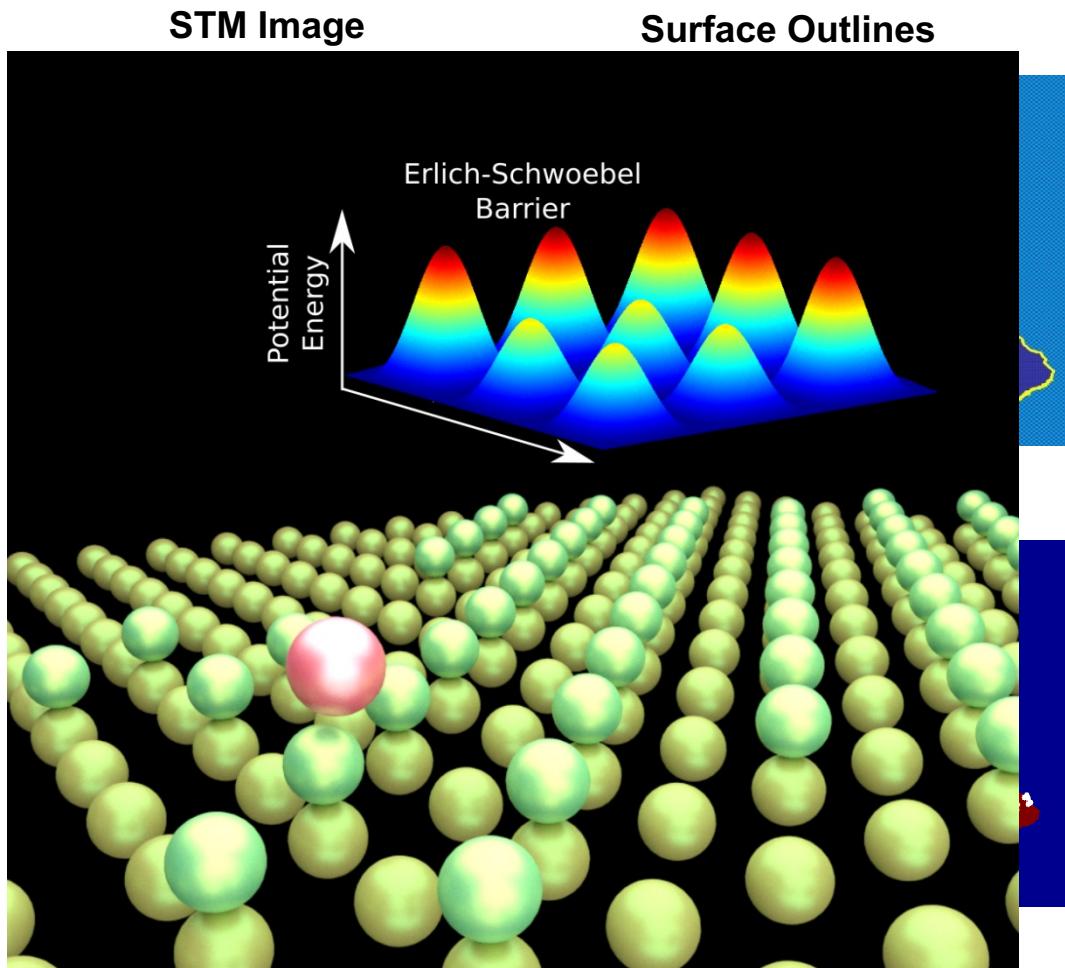
- The second component appears to show a splitting as the number of Re neighbors increases, irrespective of composition

Vasudevan et al. Appl. Phys. Rev. **8**, 011409 (2021).

Quantitative analysis of phase transition



Combined with Modeling: simple Boltzmann statistics model



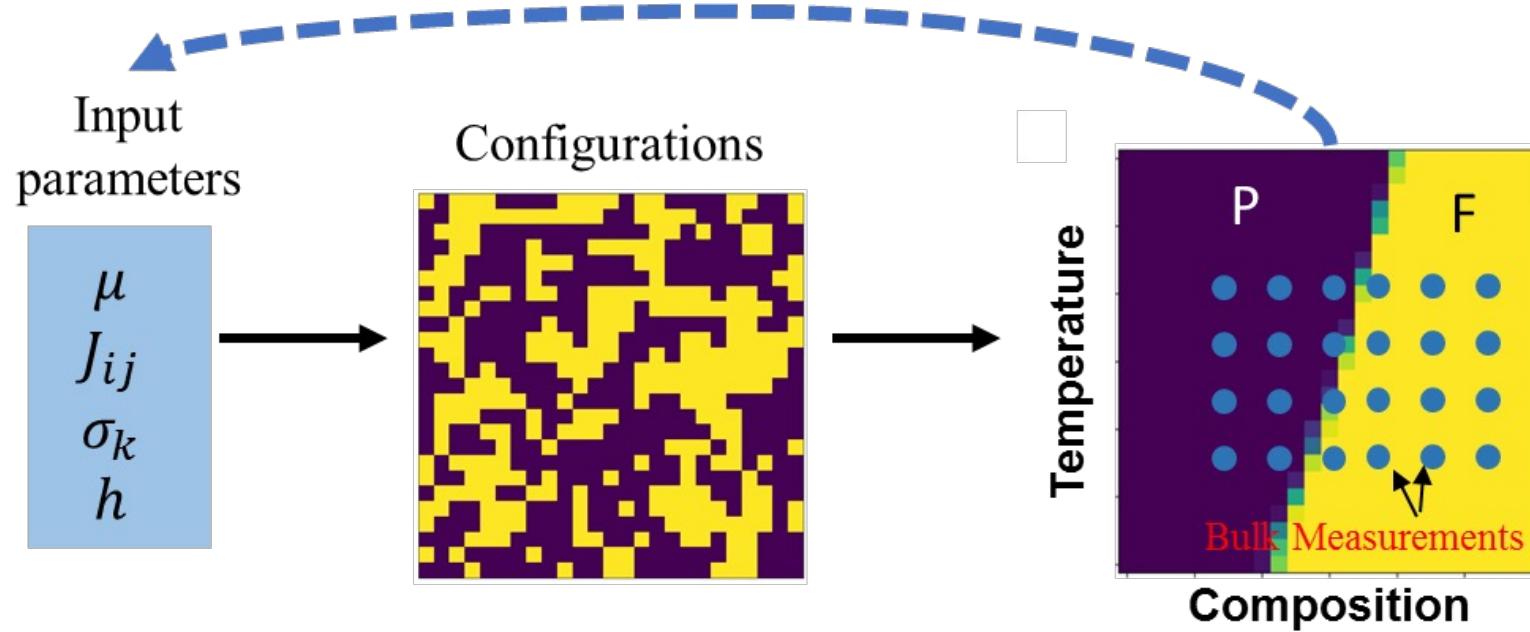
$$P_i = \frac{e^{-\frac{\epsilon_i}{k_B T}}}{\sum_{j=1}^N e^{-\frac{\epsilon_j}{k_B T}}}$$

- Monte-Carlo simulation with Boltzmann statistics to model atomic arrangement and extract surface energies and barrier heights.
- Study adatom dynamics by real space STM images as a function of time and temperature.

Statistical Distance Framework

Phase Diagrams

How we currently do things

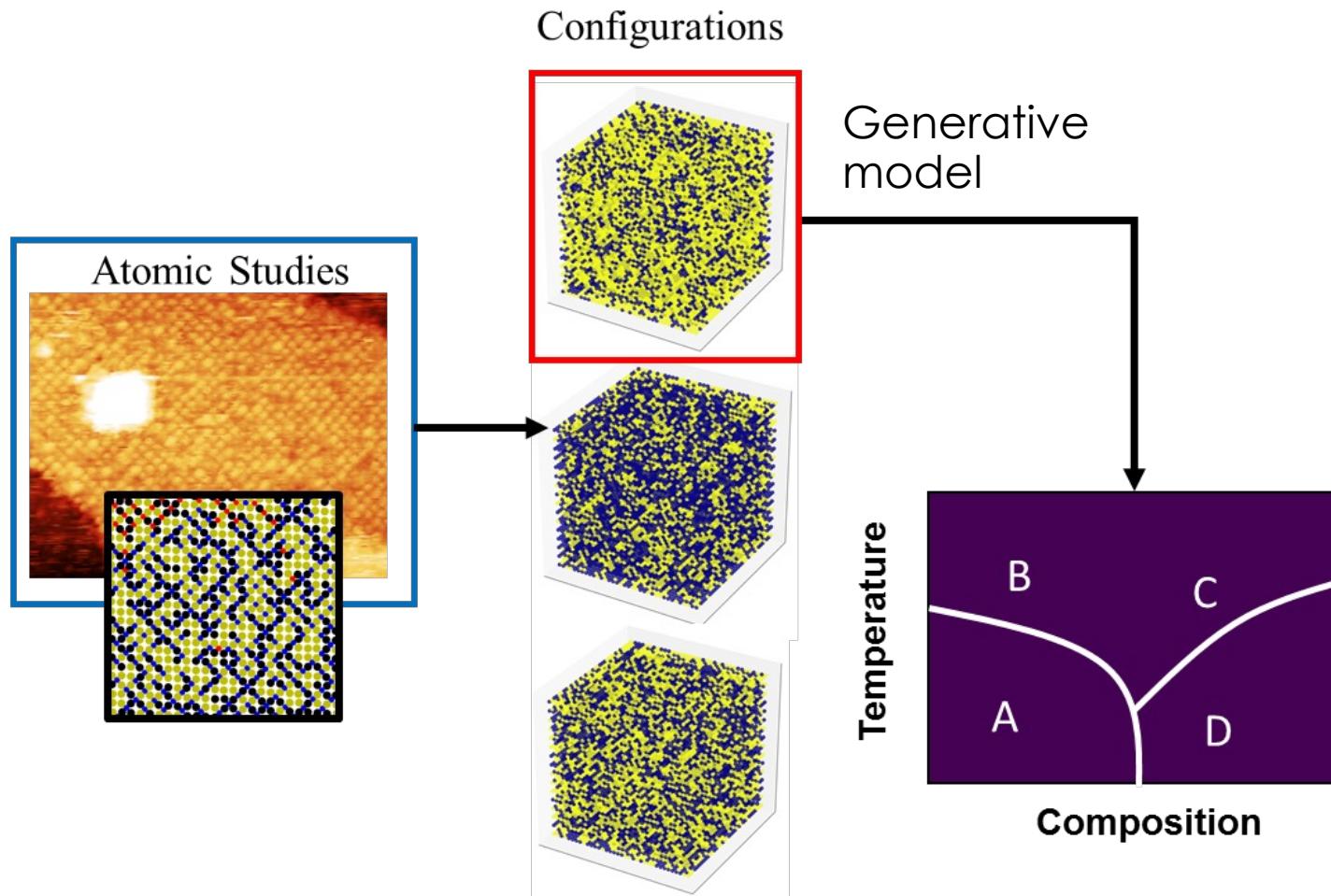


Make many samples, measure properties
Change model parameters to match measured properties

Time consuming, wasteful
Where to sample next?

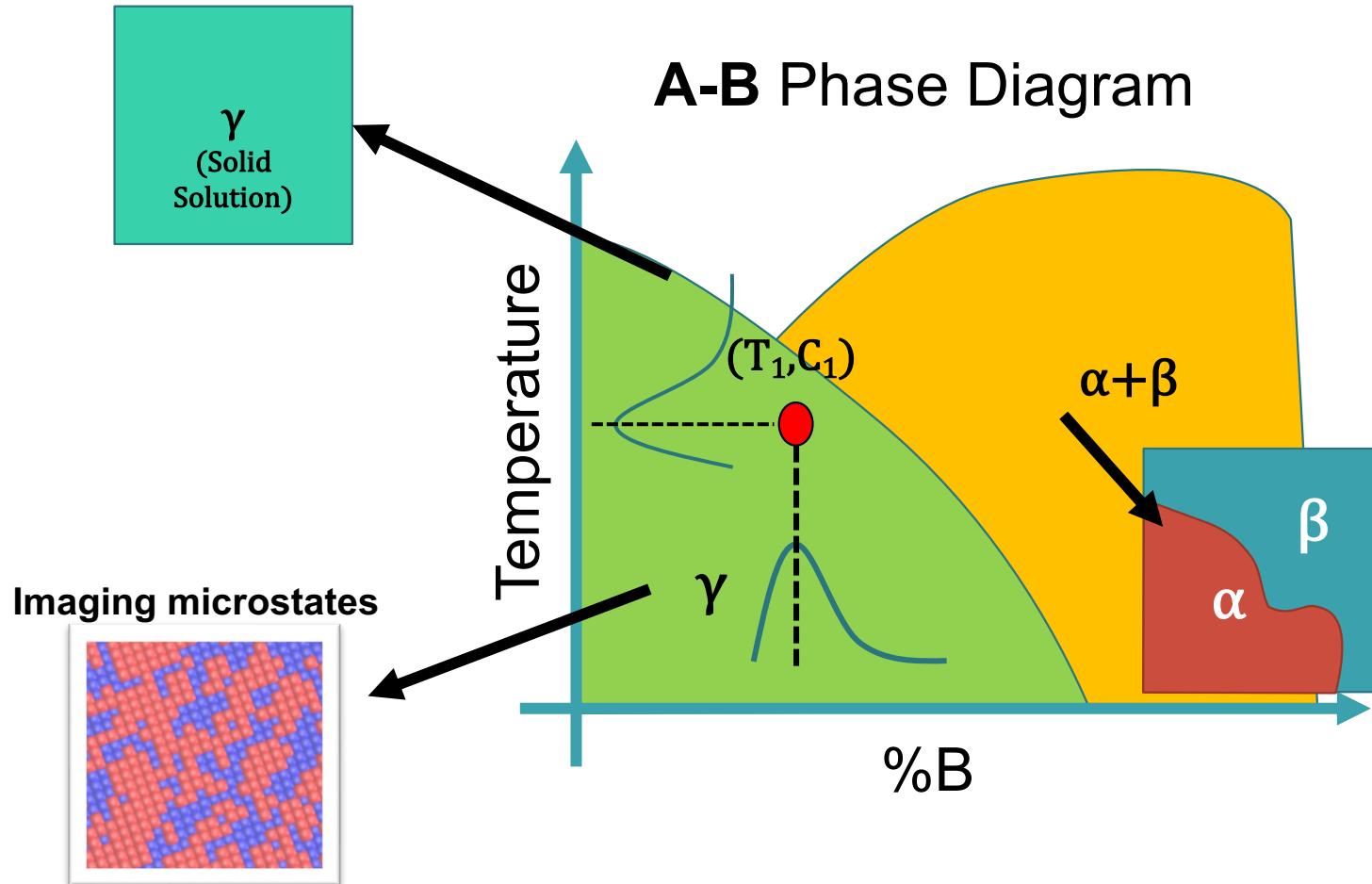
From images to models

New approach



From atomically resolved images, build generative models and sample from them to obtain predictions

What an atomic image really tells us



- Response of system to (T,C) is encoded within the imaged fluctuations!
- If we can image this, then can we use it to predict the phase diagram?

What do we need to make it work?

1. We need a generative model (e.g., lattice model (Ising type))

2. **We need to compare the right statistics:**

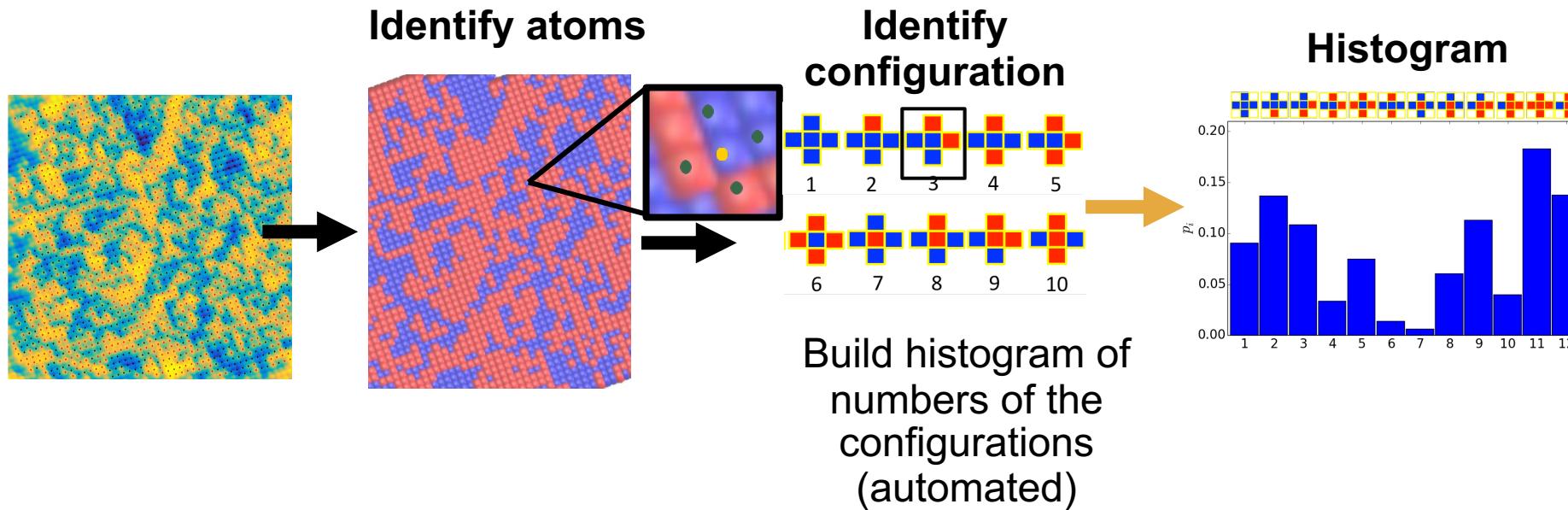
- In the case of atomically resolved images, these will be the atomic configurations
- Can also be augmented with e.g. pair distribution functions, etc.

3. **We need a metric to compare the model to the experimental data, which will correctly capture fluctuations.**

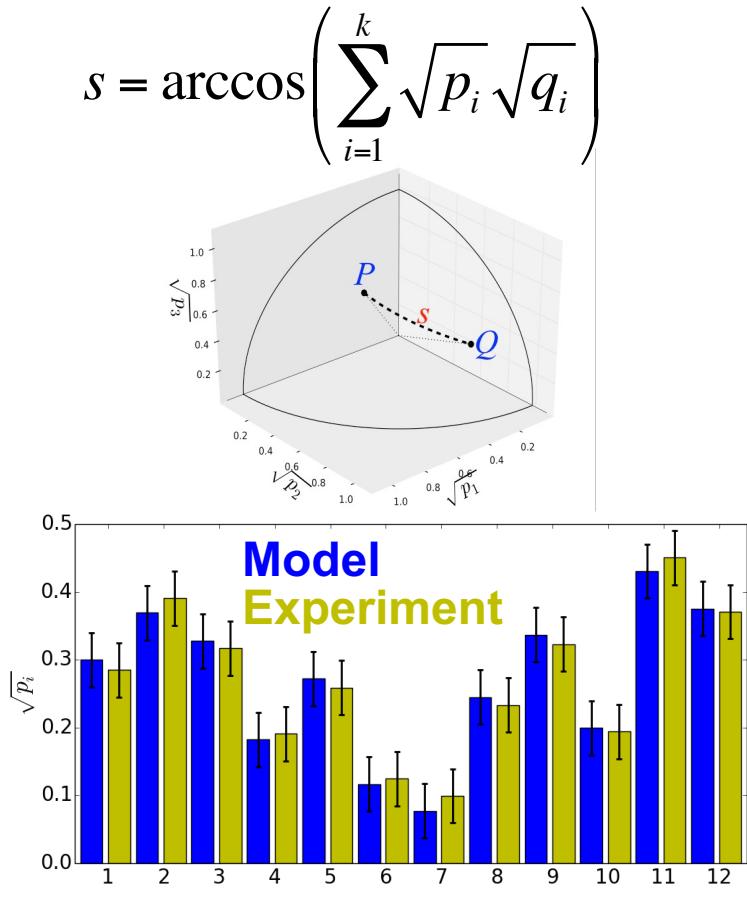
- The appropriate metric in this instance is the statistical distance, which equalizes statistical weights of experimental and statistical data, and enables appropriate comparison
- This is a measure of distinguishability of two systems (e.g., similar to Kullback-Leibler divergence)

Atomic configurations

- For atomically resolved imaging, we need to compare the probability of imaging particular atomic configurations

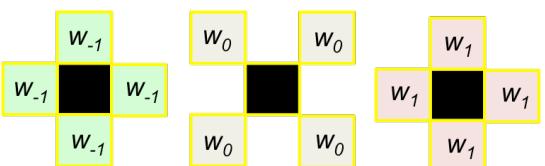
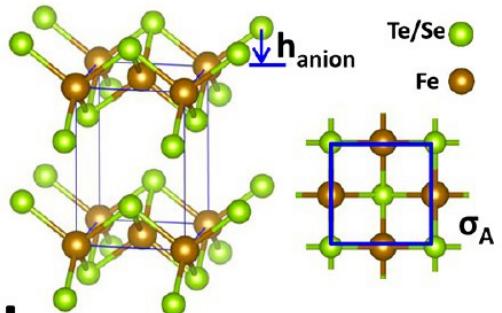
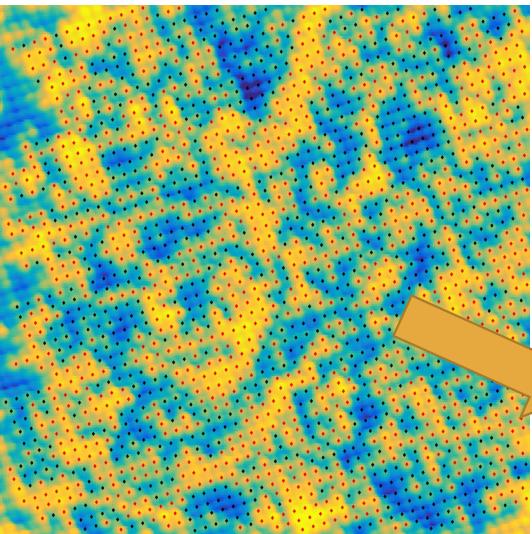


Statistical Distance Metric



- For the generative model to be predictive, it requires that when we sample from it, those samples are indistinguishable from the target system
- For thermodynamic systems, the distinguishability is expressed in terms of statistical distance, which measures the shortest distance in the probability space
- This metric also equalizes statistical weights of the experimental and simulated data
- Correctly captures fluctuations: essential for predictive abilities

Stochastic model matching

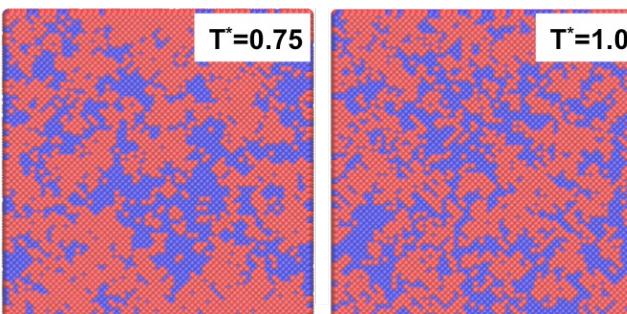
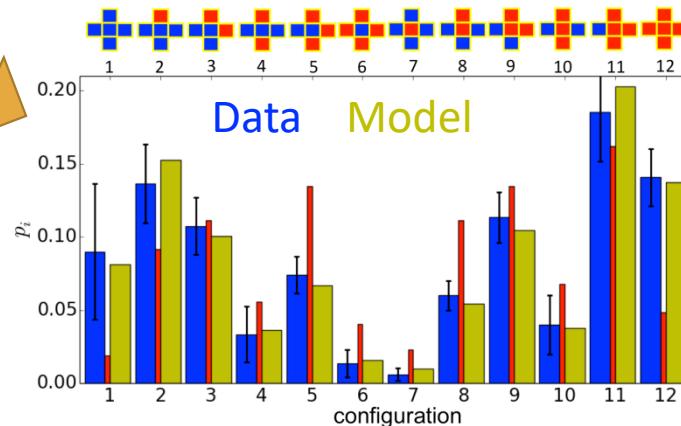


Vlcek et al., ACS Nano **11**, 10313 (2017)

Interacting solid solution model

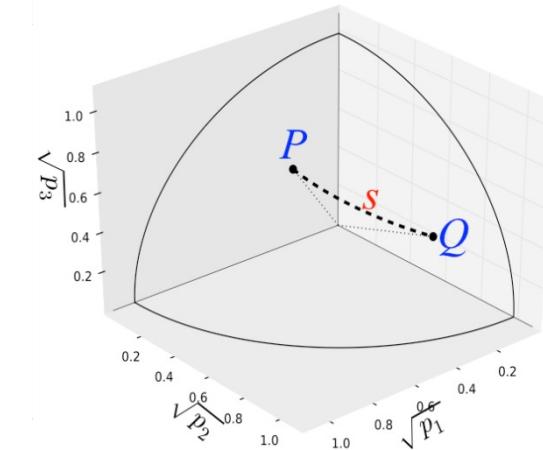
$$u_i = w_0^p a_i^0 + w_{-1}^p a_i^{-1} + w_1^p a_i^1$$

$$a_i^0 = \sum_{NN} \delta^0; \quad a_i^{-1} = \sum_{NN} \delta^{-1}; \quad a_i^1 = \sum_{NN} \delta^1$$



Minimize statistical distance between histograms

$$s = \arccos \left(\sum_{i=1}^k \sqrt{p_i} \sqrt{q_i} \right)$$



- Utilize all available statistical information in the image
- Using generative model, infer parameters of the model via the experiment.

- 1) Wootters W.K. *Phys. Rev. B*. **(23)2**, 357-362 (1981).
- 2) Braunstein S.L., Caves S.M.. *Phys. Rev. Lett.* **72(22)**, 3439-3443 (1994).

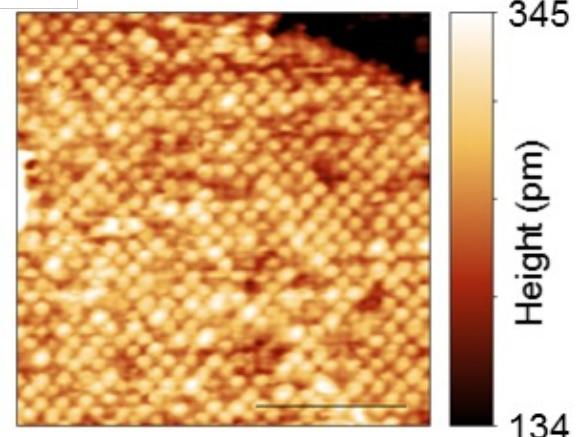
Test Case: Segregation in $\text{La}_{5/8}\text{Ca}_{3/8}\text{MnO}_3$ film

- Segregation of the divalent cation in manganite thin films is well known, and can present problems for situations requiring ideal interfaces such as tunnel junctions

Available Data

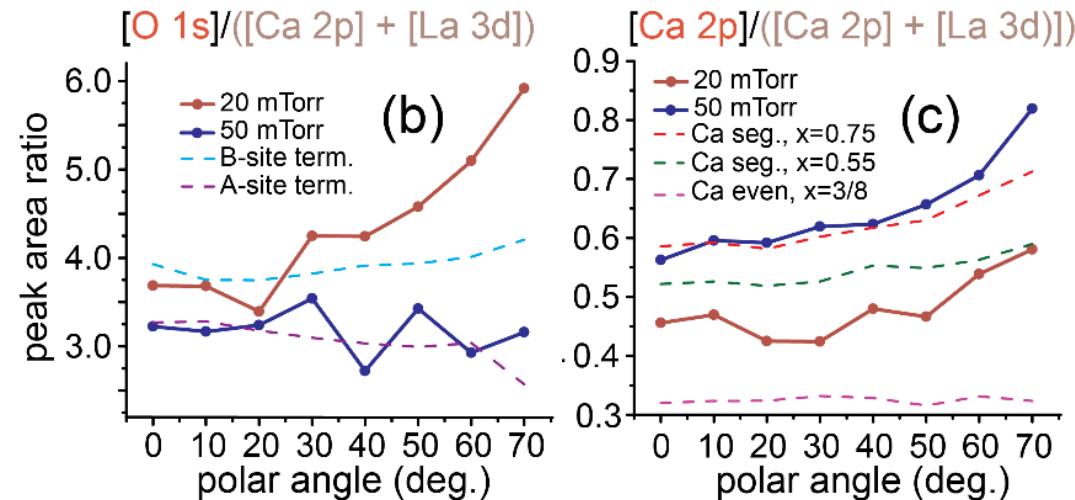
Microscopic

STM Images
(A-site terminated)



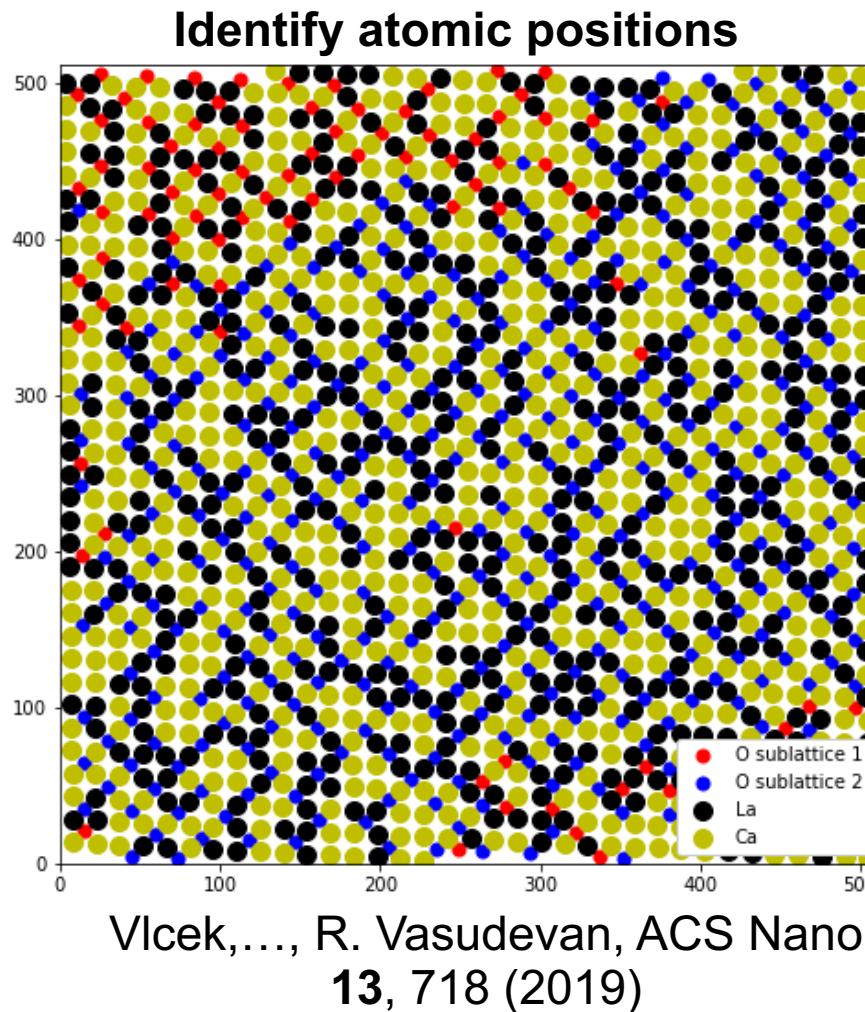
Macroscopic

Angle-resolved XPS

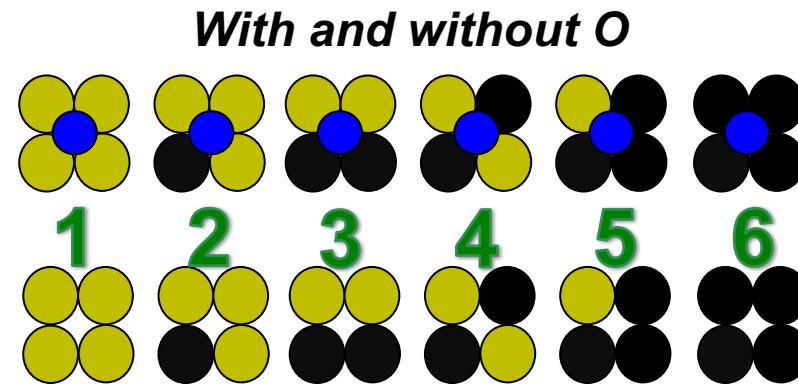


Tselev et al., ACS Nano, 9, 4316 (2015)

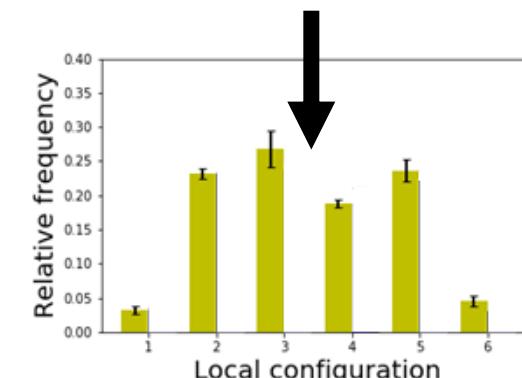
Image analysis and model setup



After image analysis, we can compute the histogram of configurations present



Compute histograms



Generative model

Energy of configuration

$$u_i = w_{CL} \sum_{\substack{\{C,L\} \\ \text{Bulk}}} \delta_i^{b,CL} + w_{LL} \sum_{\substack{\{L,L\} \\ \text{Surface}}} \delta_i^{s,LL} + w_{LO} \sum_{\substack{\{L,O\} \\ \text{Surface}}} \delta_i^{s,LO}$$

Interaction terms w_{CL} , w_{LL} and w_{LO}

Where C = Calcium, L = Lanthanum, O = Oxygen

Loss function

$$S^2 = c_{SO} S_{SO}^2 + c_{SN} S_{SN}^2 + c_{PR} S_{PR}^2 ; \quad s = \arccos \left(\sum_{i=1}^k \sqrt{p_i} \sqrt{q_i} \right)$$

↑
Surface configurations with oxygen

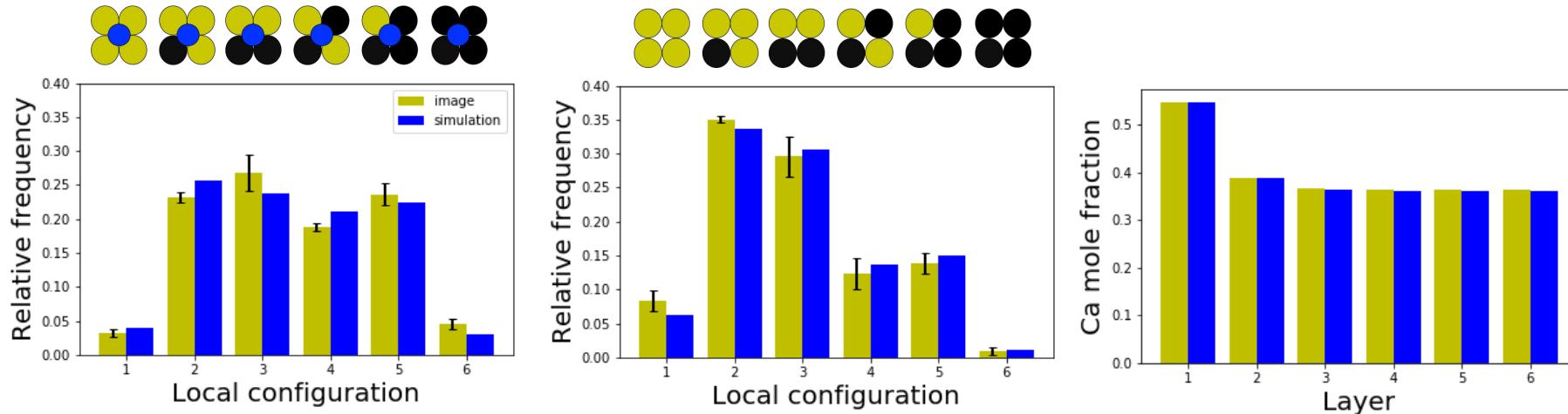
↑
Surface configurations with no oxygen

Ca concentration profile

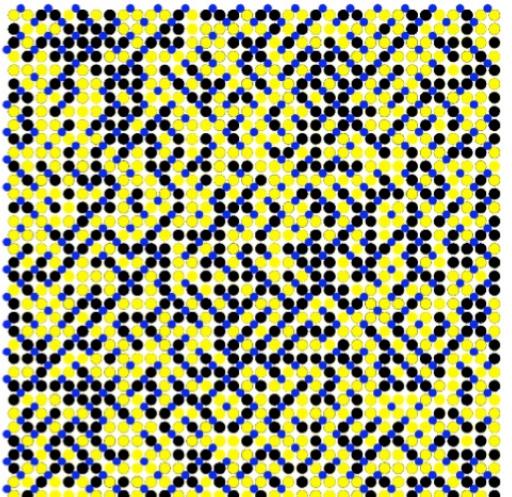
Loss function accounts for both microscopic and macroscopic data

Results

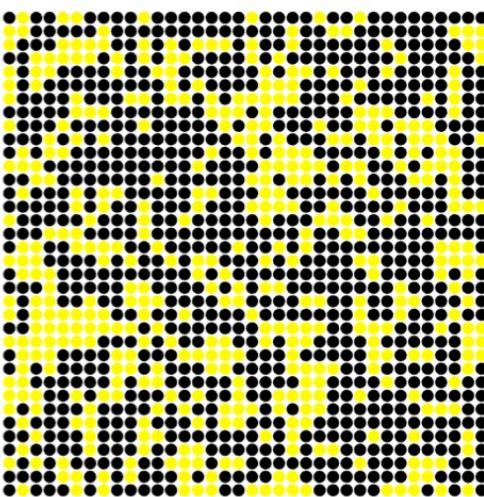
$$w_{LC} = 1.65 \text{ kJ/mol}, w_{LL} = 6.28 \text{ kJ/mol}, w_{LO} = -7.57 \text{ kJ/mol}$$



Surface

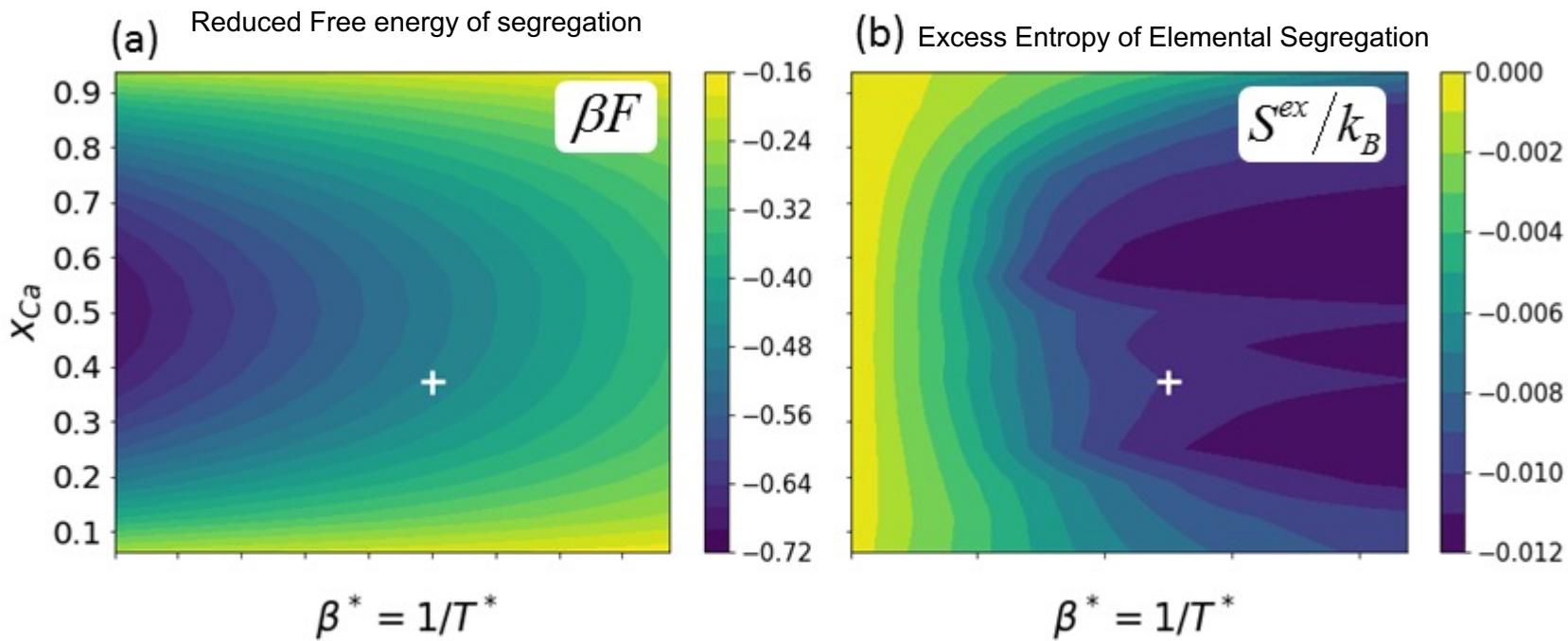


Bulk



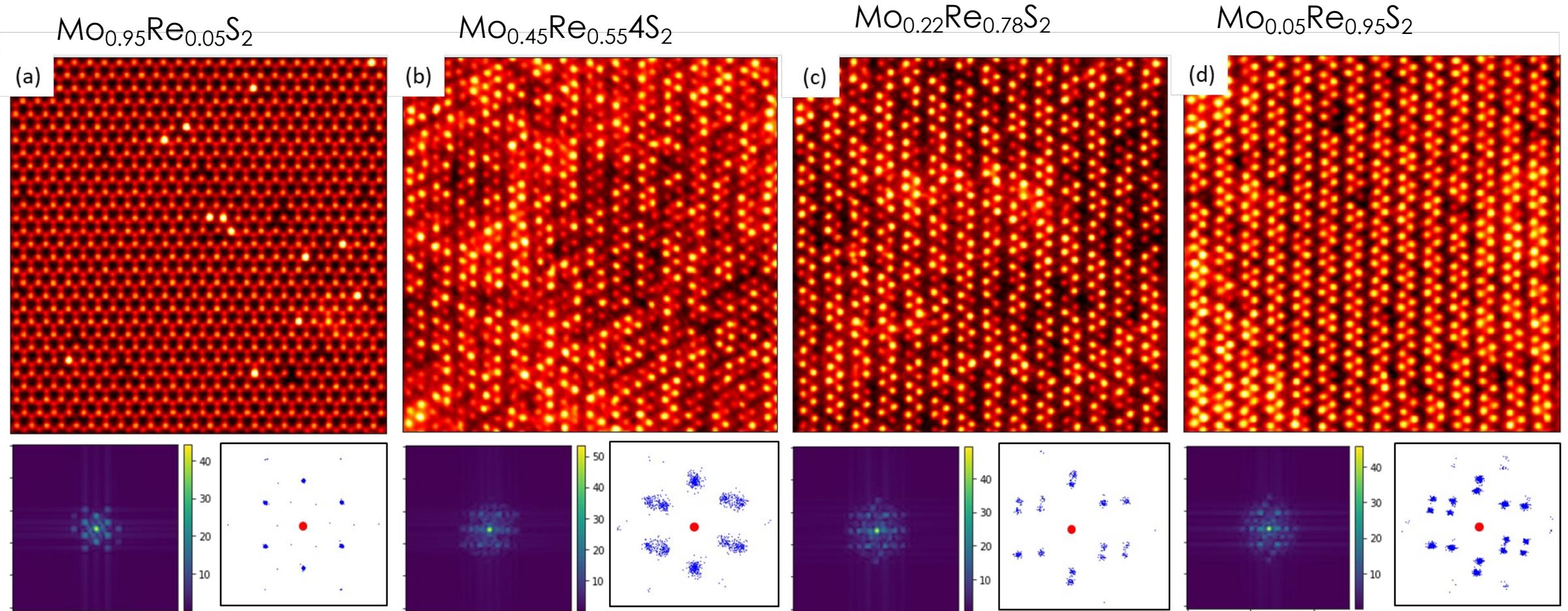
- Good match between experiment and model
- Unfavorable electrostatic interactions between like cation pairs, and favorable elastic interactions in deeper layers lead to increased tendency to segregation

Results



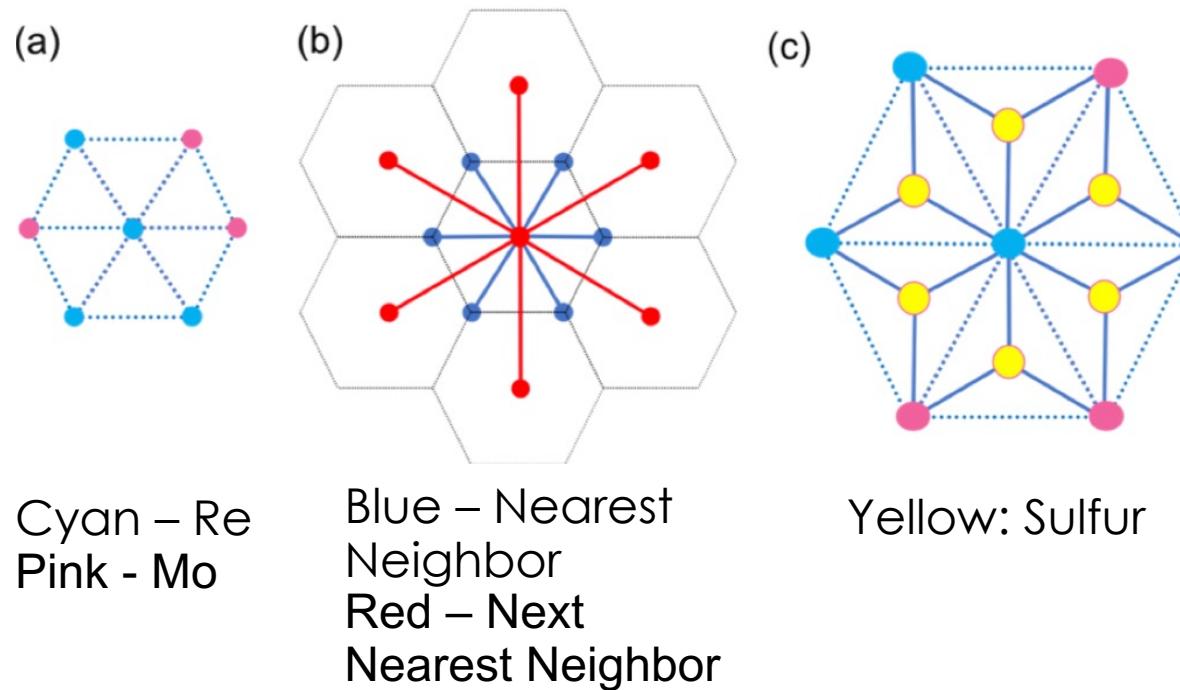
By combination of STM and spectroscopy, and with basic modeling, we could construct a generative model with predictive capability

Another: MoS₂-ReS₂ system



Vasudevan et al. Appl. Phys. Rev. **8**,
011409 (2021).

Application: MoS₂-ReS₂ system



1. Pair-additive Model

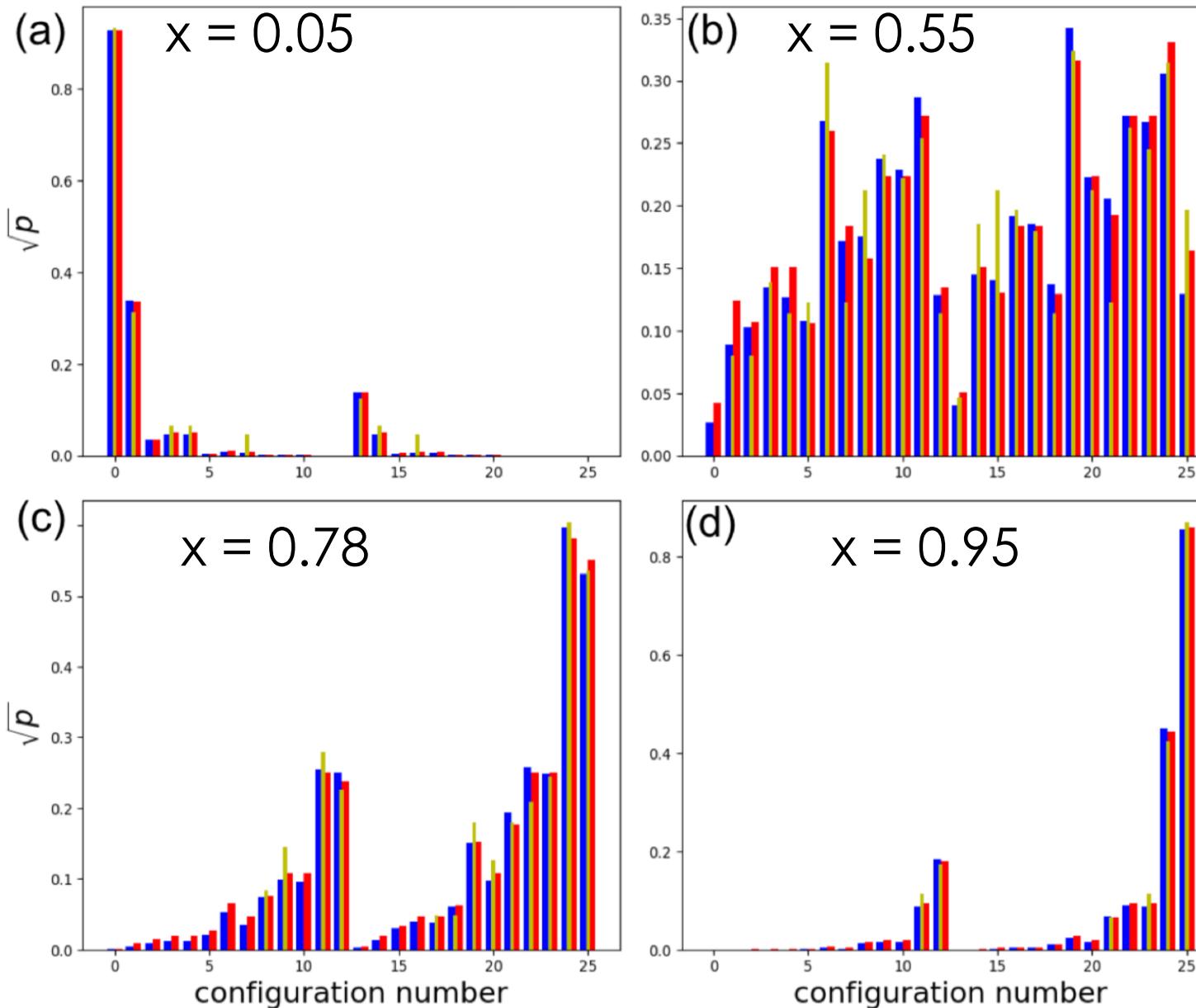
$$u_i = w_1 \sum_{\{NN\}} \delta_{MoRe} + w_2 \sum_{\{NNN\}} \delta_{MoRe}$$

2. Many-body Model

$$u_i = w_1 \sum_{\{S\}} \delta_{MoMoRe} + w_2 \sum_{\{S\}} \delta_{MoReRe}$$

Vlcek et al. npj Computational Materials, 7, 42 (2021)

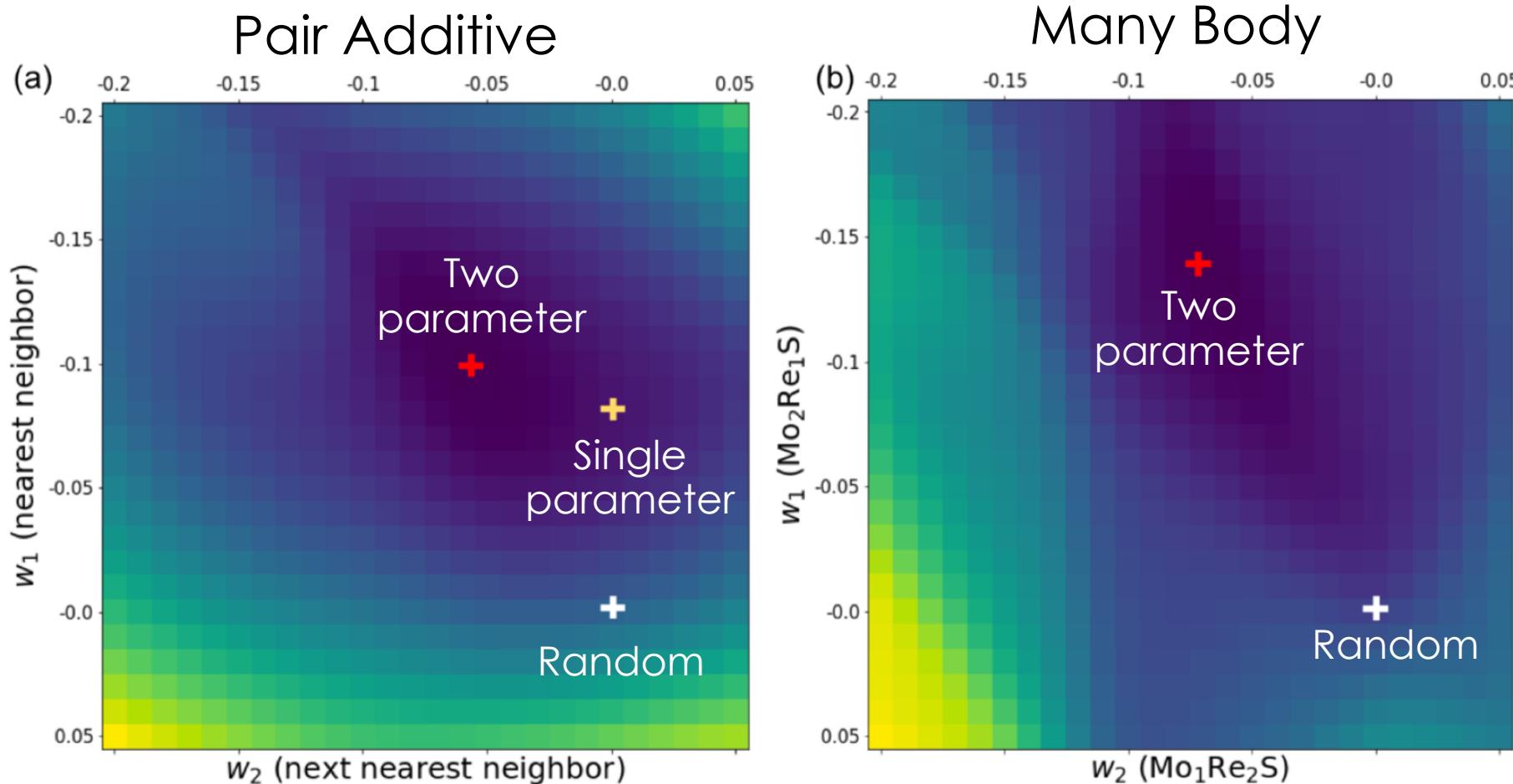
Application: MoS₂-ReS₂ system



Red – Random
Blue – Pairwise Model
Teal - Experiment

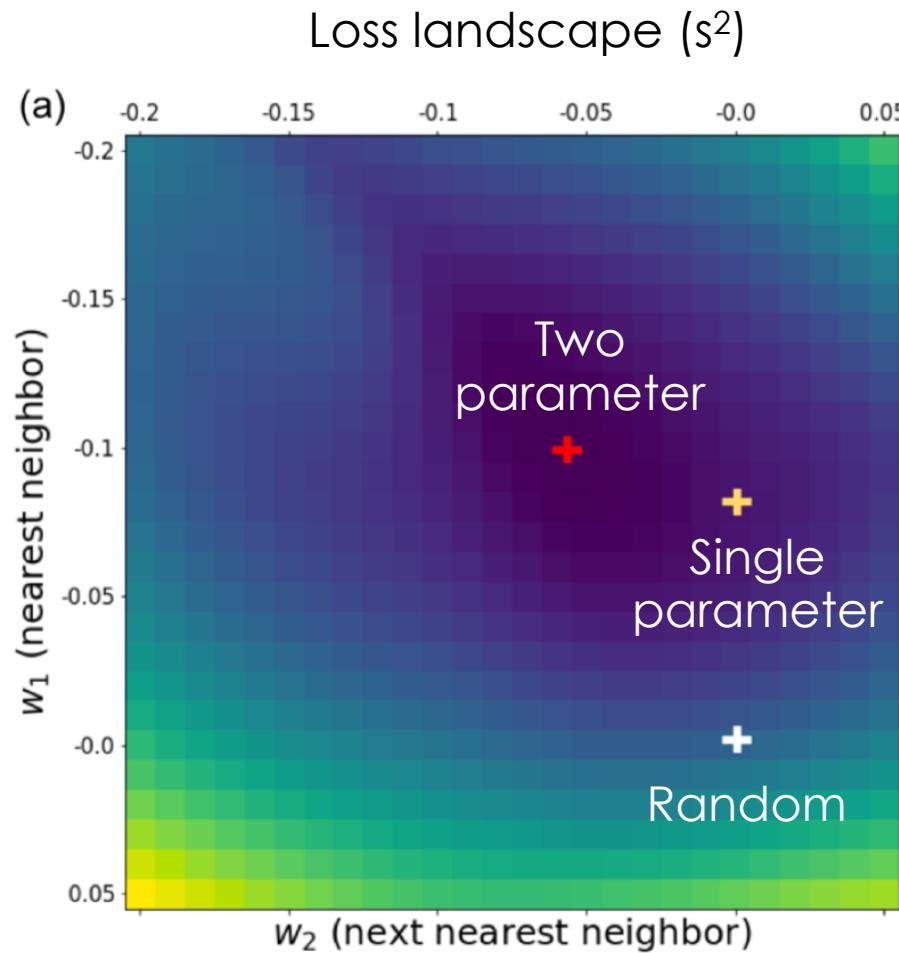
1. Random model good at low and high concentrations, but not so great in between.
2. Equilibrium model better, but still not great.

Application: MoS₂-ReS₂ system



Statistical
distance
landscapes
for two
distinct
choices of
the model

Extension: UQ for interactions



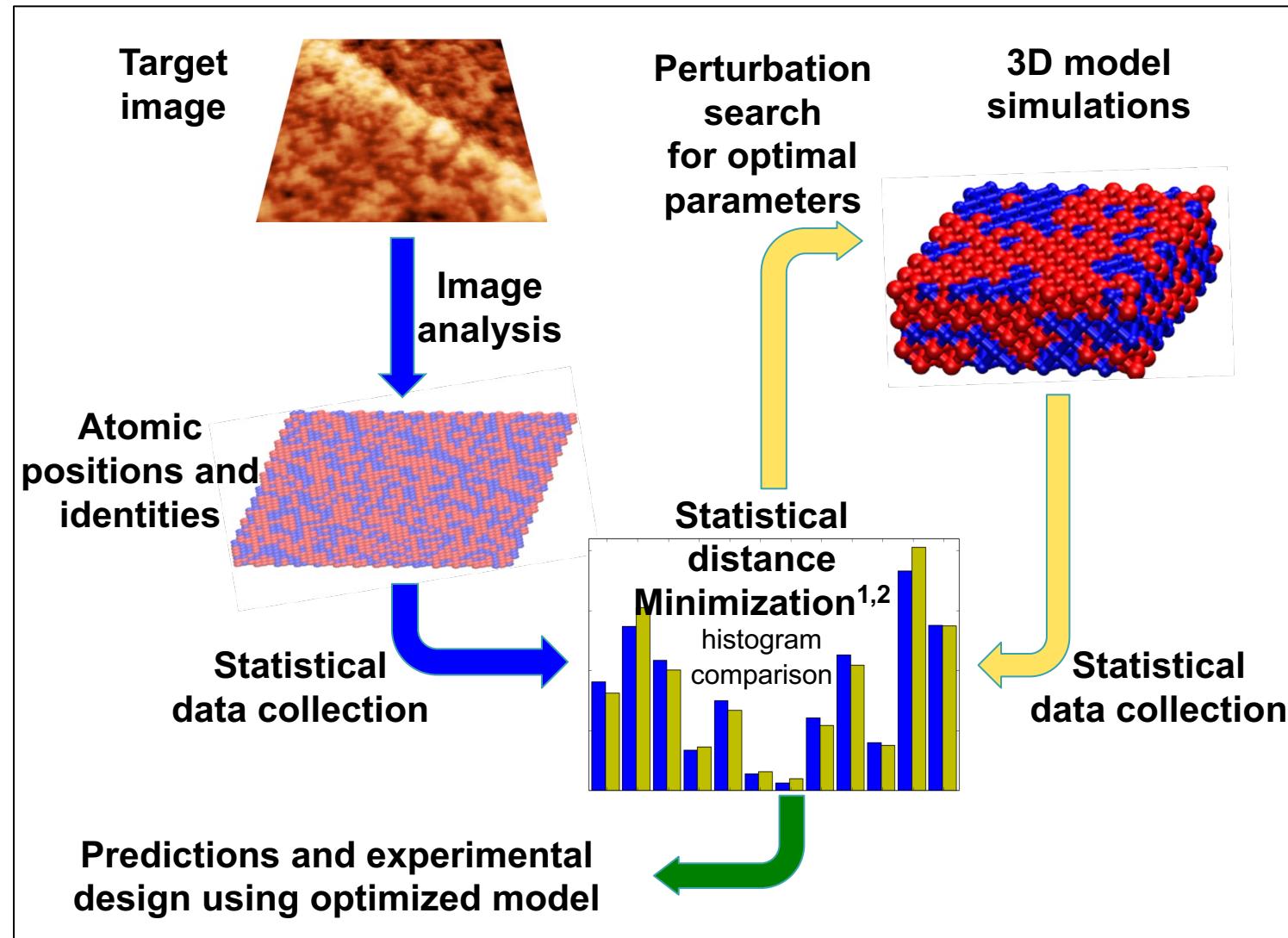
- We wish to determine the form of $p(\mathbf{w} | \mathbf{d})$
- We can obtain estimates of the mean and variance via bootstrapping:
 - obtain M images;
 - optimize model on a subset
 - Repeat for different subsets
 - Provides uncertainty estimate
- Alternatively, simply use MC sampling to determine $p(\mathbf{w}|\mathbf{d})$ with log-likelihood given by s^2
- But what about model uncertainty?

Model selection: Information Criterion

dataset	N	$S^2(R)$	PV (R)	$S^2(P1)$	PV (P1)	$S^2(P2)$	PV (P2)	$S^2(M)$	PV (M)
$x=0.05$	464	0.0062	0.9899	0.0067	0.9829	0.0070	0.9771	0.0073	0.9680
$x=0.55$	466	0.0325	0.0001	0.0283	0.0009	0.0288	0.0007	0.0284	0.0009
$x=0.78$	434	0.0298	0.0013	0.0263	0.0069	0.0247	0.0143	0.0275	0.0040
$x=0.95$	471	0.0124	0.5602	0.0129	0.4976	0.0121	0.5852	0.0121	0.5864
Total	1835	0.0200	0.0015	0.0184	0.0107	0.0180	0.0168	0.0187	0.0079
BIC		73.4		71.4		73.6		76.1	

Statistical hypothesis testing for uncertainty

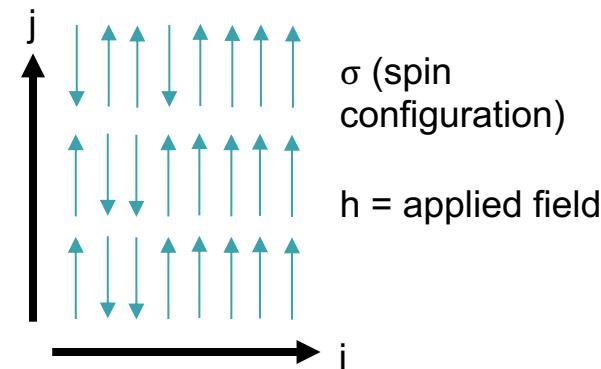
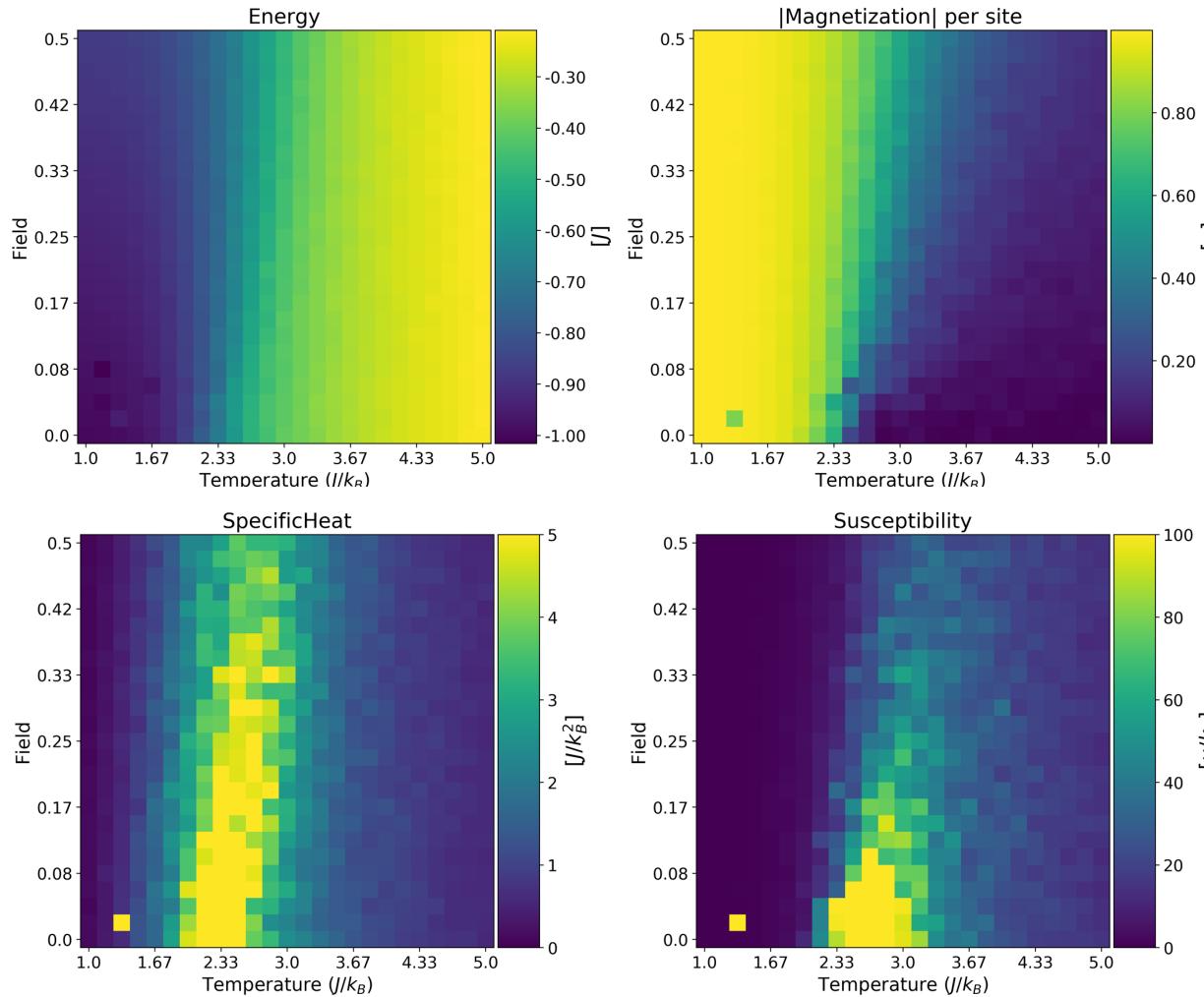
Summary



- We present a framework for analysis of atomically resolved imaging data that captures the physics inherent in the stochastic variations observed
- We combine this with model selection via information criterion approaches
- If applied in real-time, this can help to answer the question of 'how much data is required'

Colab: 2D Ising Model on a square lattice

2D Ising Model as a function of H, T



$$H(\sigma) = - \sum_{\langle i j \rangle} J_{ij} \sigma_i \sigma_j - \mu \sum_j h_j \sigma_j$$

$$\chi = \frac{1}{T} (\langle M^2 \rangle - \langle M \rangle^2) \quad \sigma_k = \pm 1$$

$$C_v = \frac{1}{T^2} (\langle E^2 \rangle - \langle E \rangle^2) \quad \mu = 0.67$$

$$J_{ij} = 1 \quad J_{ij} = 1$$

Let's see how we can use the IsingSim to optimize a model based on a 'simulated' experiment...