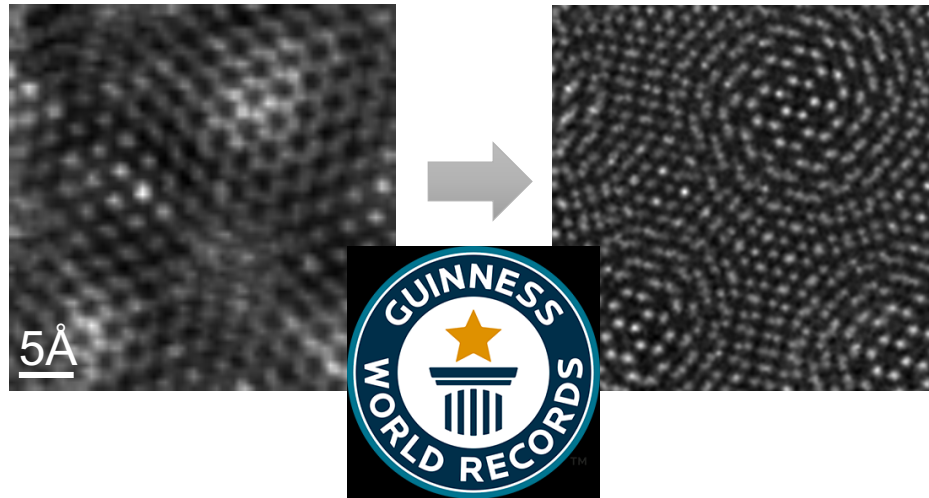


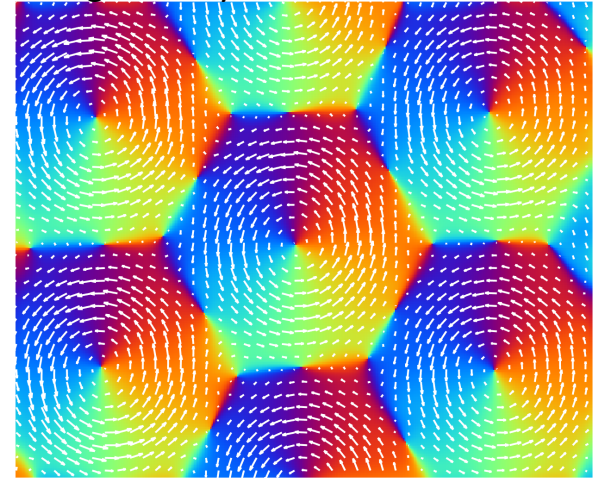
Phase Retrieval for Petavoxel Electron Imaging as a big and deep data problem

David Muller, Cornell University

Michael Cao, Michael Matty, Zhen Chen, Eun-Ah Kim, Li Li (Google)



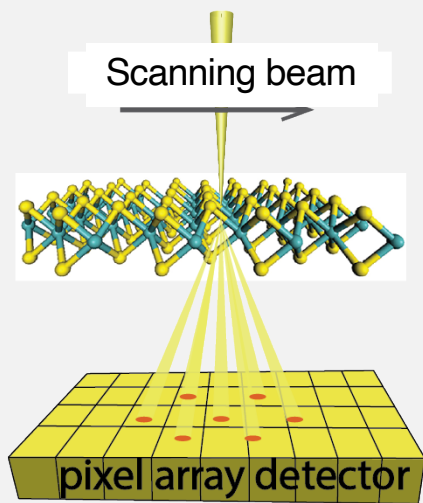
Magnetic Skyrmions in FeGe at 100 K



An inverse problem for Structure and Physics

Electron Microscope Pixel Array Detector (EMPAD)

Records every electron that passes through the sample, at every probe position!



Unique Cornell Design:

High Speed: 0.86 ms/frame

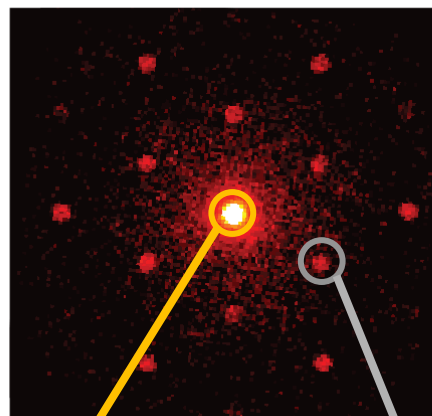
Single Electron Sensitivity (noise is $1/50 e^-$)

e^- Dynamic Range > 1,000,000:1

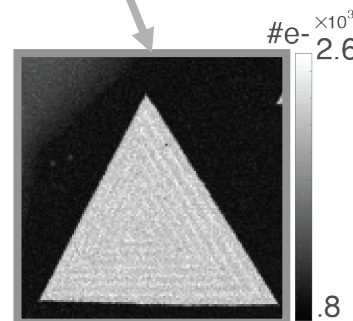
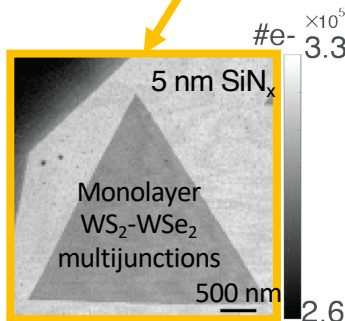
Sol Gruner:

M. W. Tate, et al. *Microsc. & Microanal.* **22**, 237 (2016)

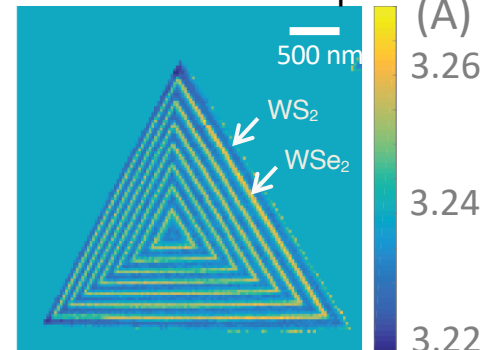
Diffraction pattern in 1 ms



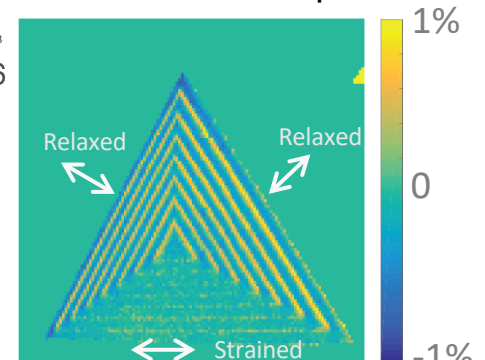
#e-
 10^4
 10^3
 10^2
 10



Sub-pm precision
 Lattice constant map



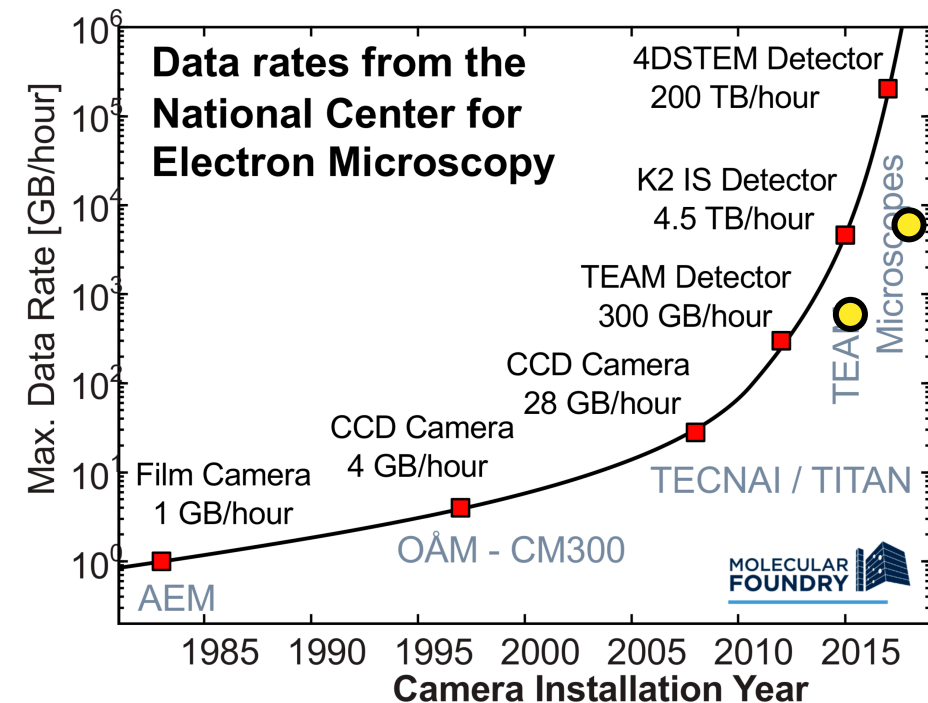
Uniaxial strain map



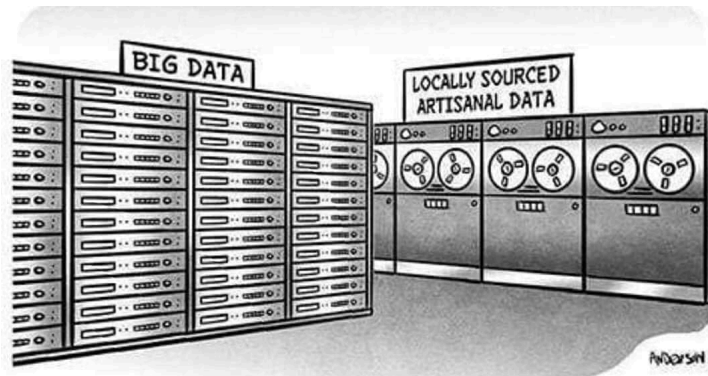
Y. Han et al. *Nano Lett.* (2018), **18**, 3746

Science **359**. (2018). 1131.

A Moore's Law for Detectors



EMPAD ↗



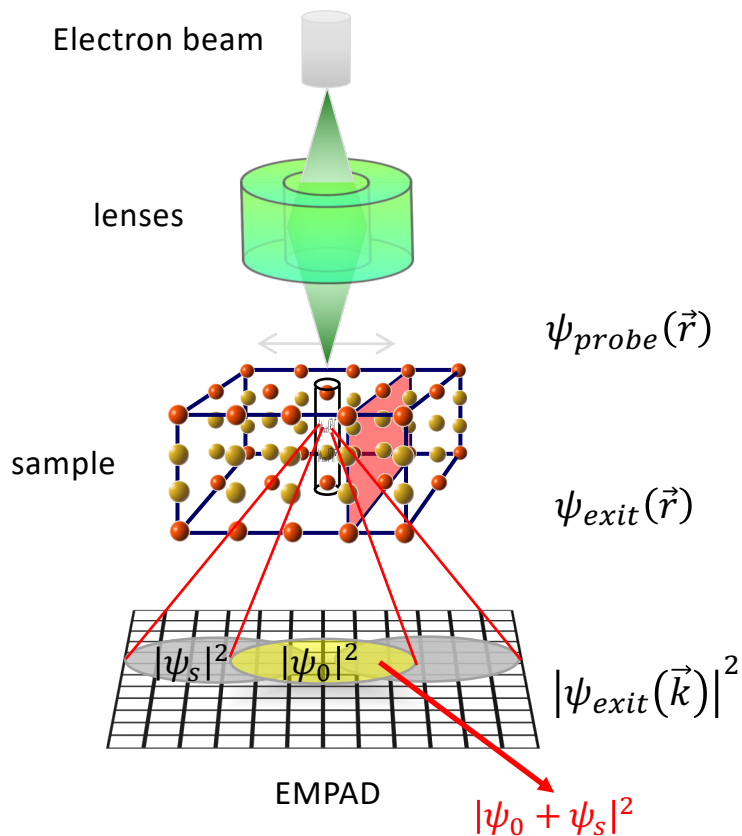
Today:

Probe positions (x,y) : 512×512

Diffraction pattern (p_x, p_y) at each (x,y) : 128×128

$\rightarrow I(x, y, p_x, p_y)$ at 32 bits is 16 GB in 30 s

X10 data rate every 5 years



Goal: reconstruct the phase of exit wave $\psi_{exit}(\vec{r})$

Requirement:

- Strong phase approximation

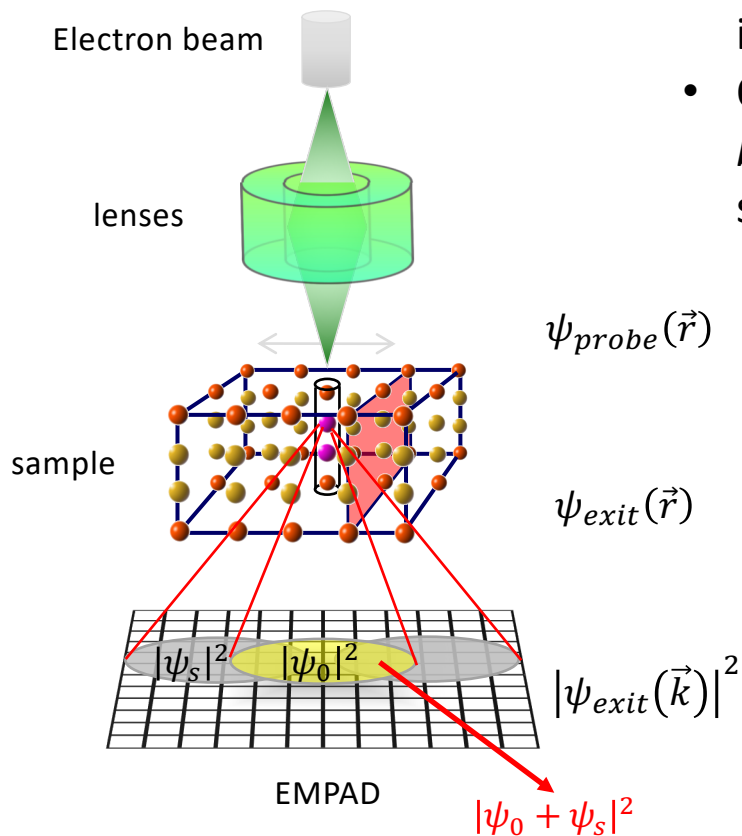
$$\psi_{exit}(\vec{r}) = \psi_{probe}(\vec{r}) \cdot \underbrace{\exp[i\sigma V(\vec{r})]}_{\text{transmission function}}$$

transmission function

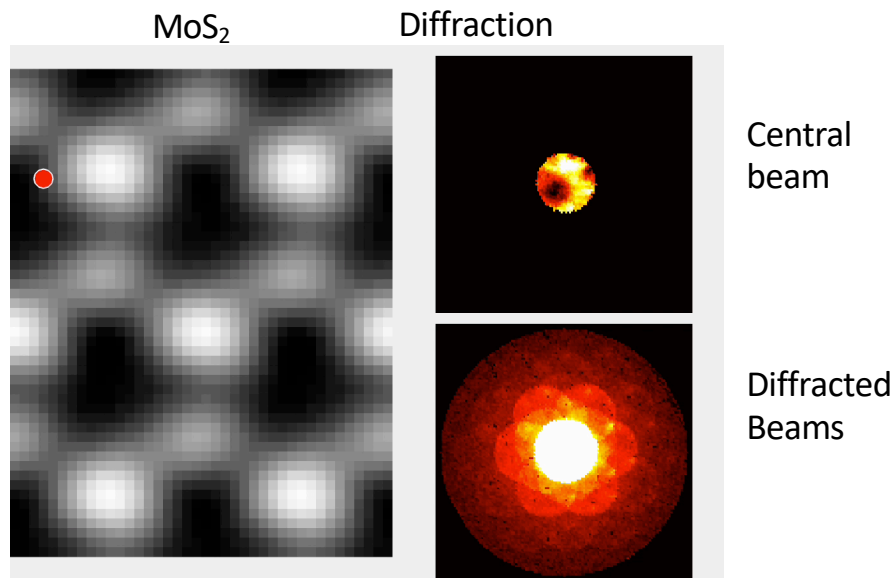
- Diffraction patterns recorded at adjacent scan positions – 4D dataset
- Periodic structure: overlapping diffraction disks

$$\psi_0 = |\psi_0|e^{i\phi_0} \quad \psi_s = |\psi_s|e^{i\phi_s}$$

$$|\psi_0 + \psi_s|^2 = |\psi_0|^2 + |\psi_s|^2 + 2|\psi_0||\psi_s|\cos(\phi_0 - \phi_s)$$



- Diffraction from an atomic-sized beam contains phase information in the overlapping beams.
- Our EMPAD detector is sensitive enough to measure *both* central and higher order beams thus increasing spatial resolution beyond the limitations of the lens



Ptychographic reconstruction methods

- Wigner distribution deconvolution (WDD)

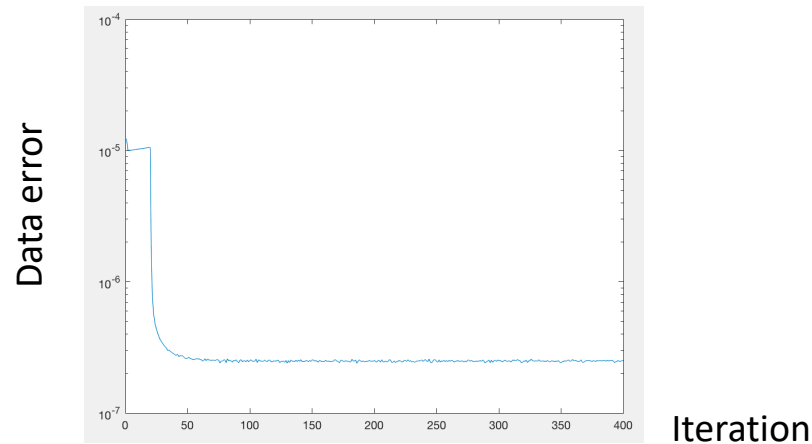
- Non-iterative
- Need to know probe function

$$\psi_{exit}(\vec{r}) = \psi_{probe}(\vec{r}) \cdot \underbrace{\exp[i\sigma V(\vec{r})]}_{\text{transmission function}}$$

- Iterative methods

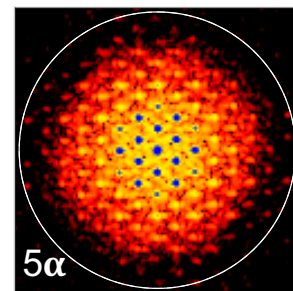
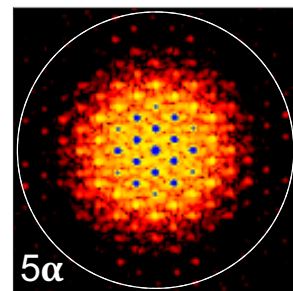
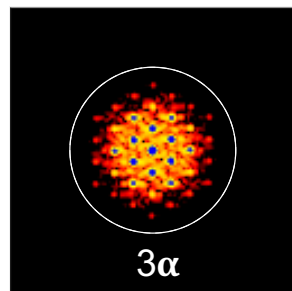
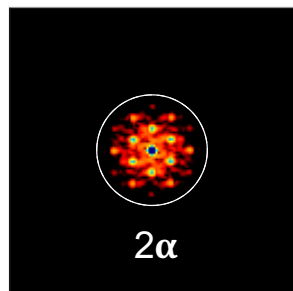
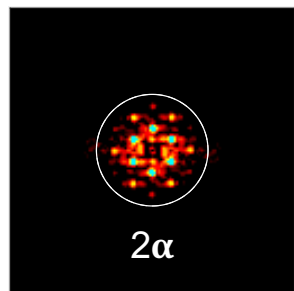
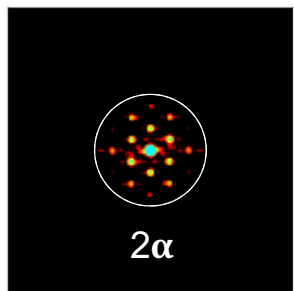
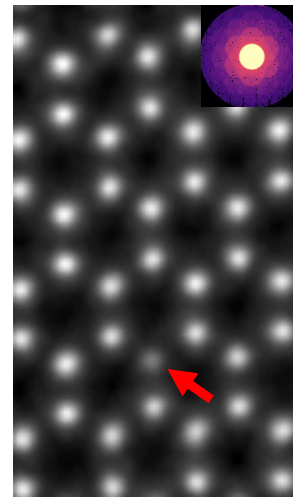
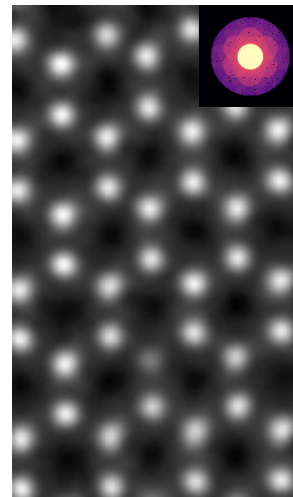
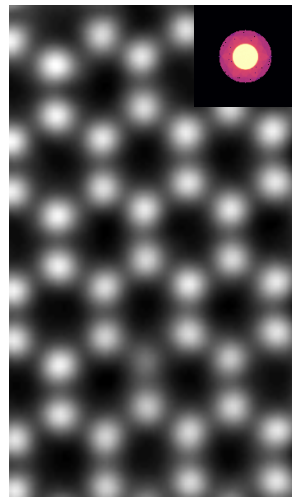
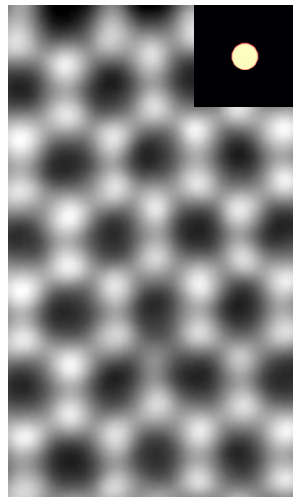
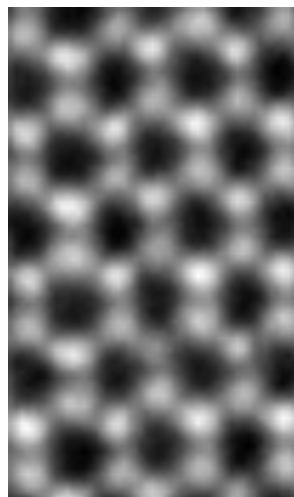
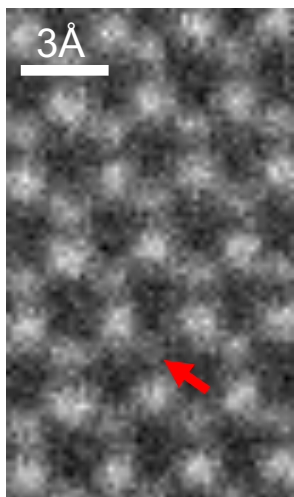
- PIE, difference map, maximum likelihood
- More flexible (include multislice)
- Can use out-of-focus probe
- Bayesian & dictionary
- Artificial neural network ?

$$\min \sum_i \|\sqrt{I_i} - |\mathcal{F}[\psi_{probe}(\vec{r} - \vec{r}_i) \cdot T(\vec{r})]|\|^2$$



EMPAD uses all electrons: improved resolution and contrast

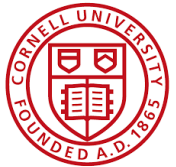
ADF Integrated CoM $\sim V(\vec{r}) \otimes |\psi_{probe}(\vec{r})|^2$ BF ptychography $\longrightarrow$ Full Field Ptychography



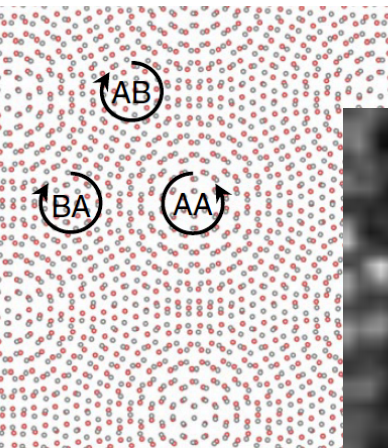
0.98 Å

Yi Jiang, et al., Nature 559, 343–349 (2018)

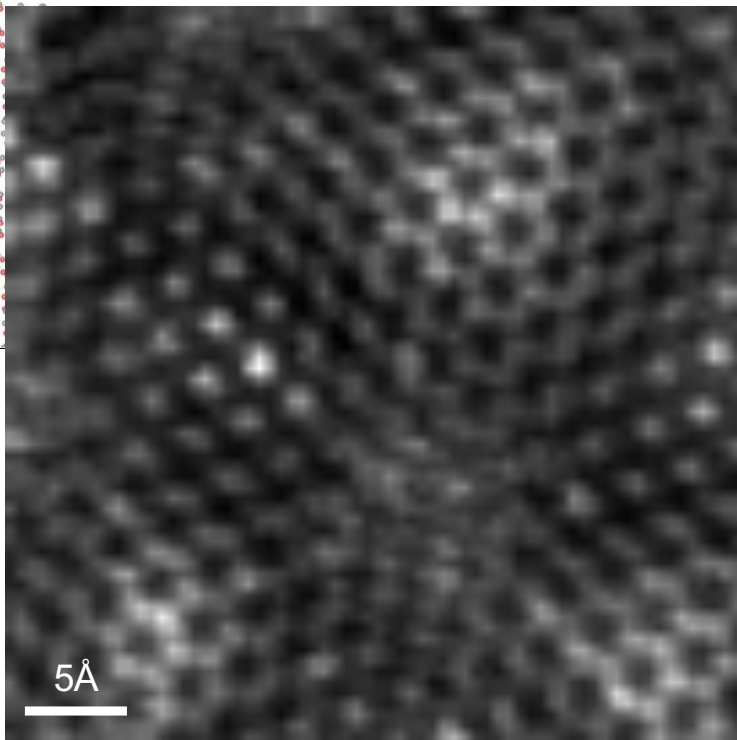
0.39 Å



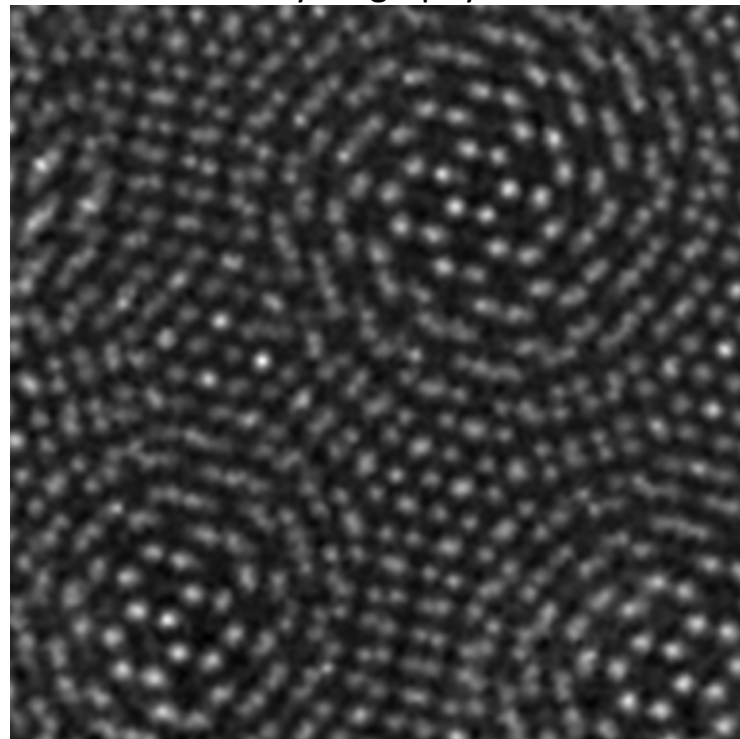
How to test resolution in real-space? ---Twisted bilayer MoS₂

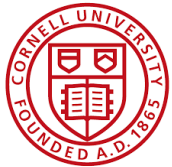


ADF

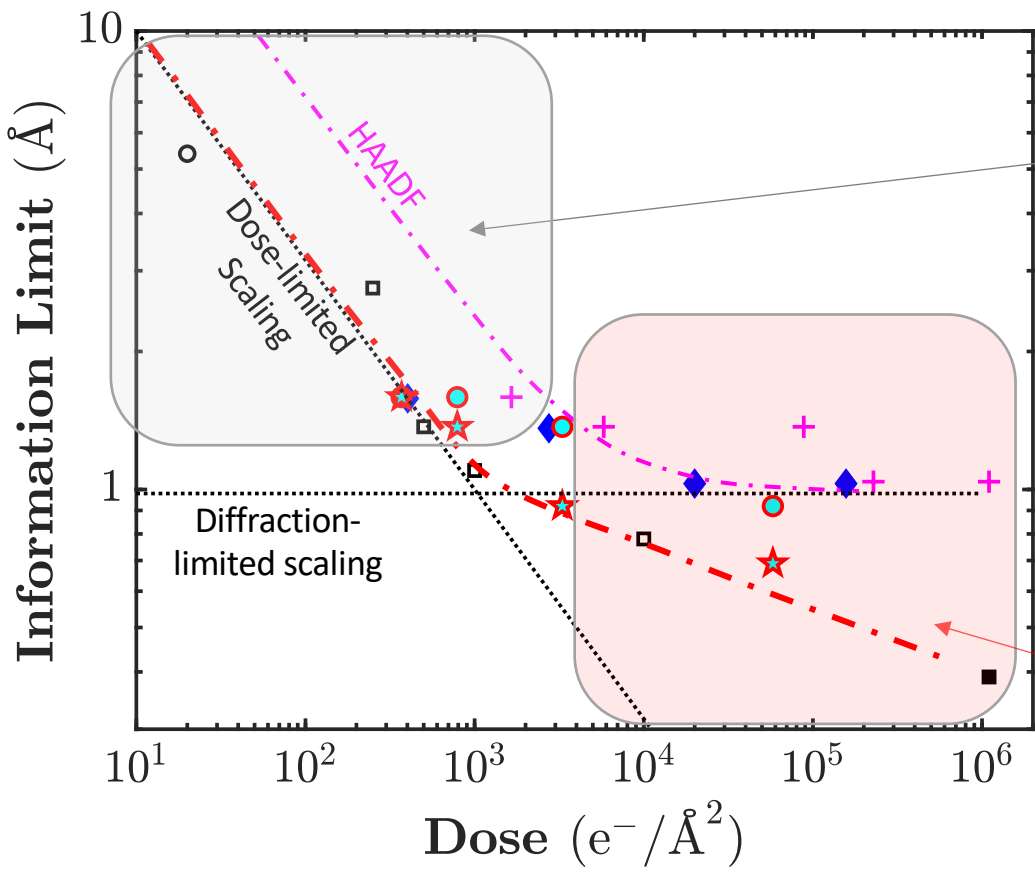


Ptychography





Dose vs. Resolution: 2 Regimes

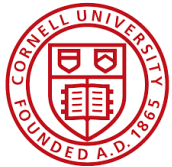


Low Dose:

- Poisson Uncertainty $\gg$ Probe Size
- $d \propto 1/\sqrt{N}$
- Ptychography vs ADF:
double info limit for the same dose

High Dose:

- Probe size set by diffraction limit
- ADF resolution saturates
- Ptychography resolution continues to improve with dose (2-3 x).



Measuring Probability Currents

$$\langle \vec{p} \rangle = \int \Psi^*(\vec{p}, \vec{r}_p) \vec{p} \Psi(\vec{p}, \vec{r}_p) d\vec{p} = \int \vec{p} |\Psi(\vec{p}, \vec{r}_p)|^2 d\vec{p}$$

Center of mass (CoM) of the diffraction pattern is the net probability current flow:

$$\langle \vec{p} \rangle = \frac{\hbar}{2i} \int \Psi^*(\vec{r}, \vec{r}_p) \vec{\nabla} \Psi(\vec{r}, \vec{r}_p) - \Psi(\vec{r}, \vec{r}_p) \vec{\nabla} \Psi^*(\vec{r}, \vec{r}_p) d\vec{r} = 2m \langle \vec{j} \rangle$$

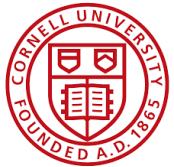
K. Muller (2014), A. Lubk (2015)

From Ehrenfest's Theorem the CoM measures the Lorentz Force:

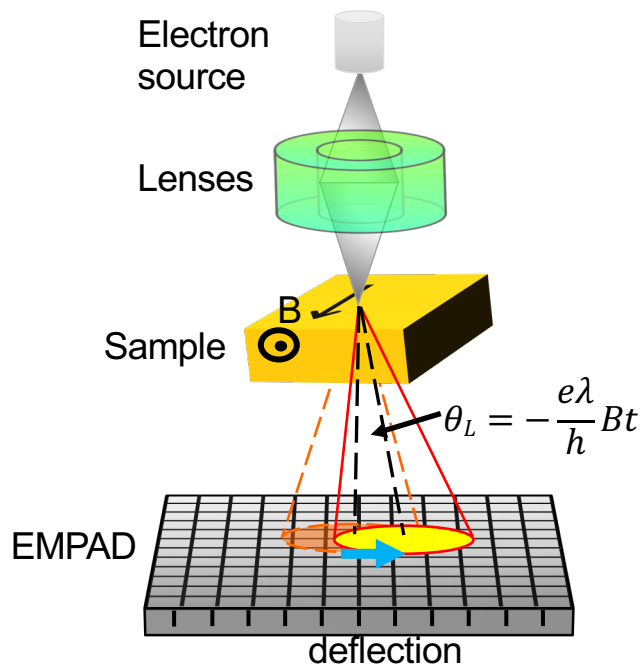
$$\frac{d\langle \vec{p} \rangle}{dt} = \langle [H, \vec{p}] \rangle = \langle q(\vec{E} + \vec{v} \times \vec{B}) \rangle$$

If we know the wavefunction, we can construct expectation values of general operators:

$$\text{Orbital Angular Momentum} \quad \langle \mathbf{L} \rangle = \langle \mathbf{r} \times \mathbf{p} \rangle$$

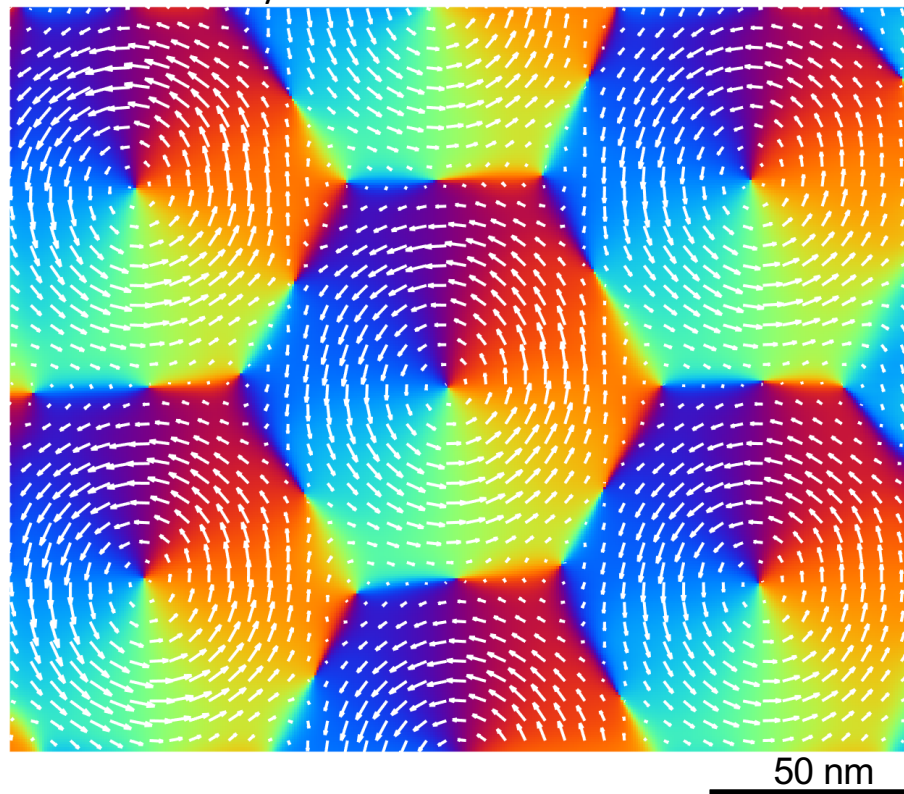


Lorentz Scanning Transmission Electron Microscopy with the Cornell Electron Microscopy Pixel Array Detector (EMPAD)

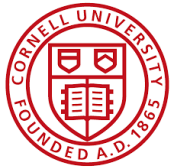


e^- Dynamic Range > 1,000,000:1

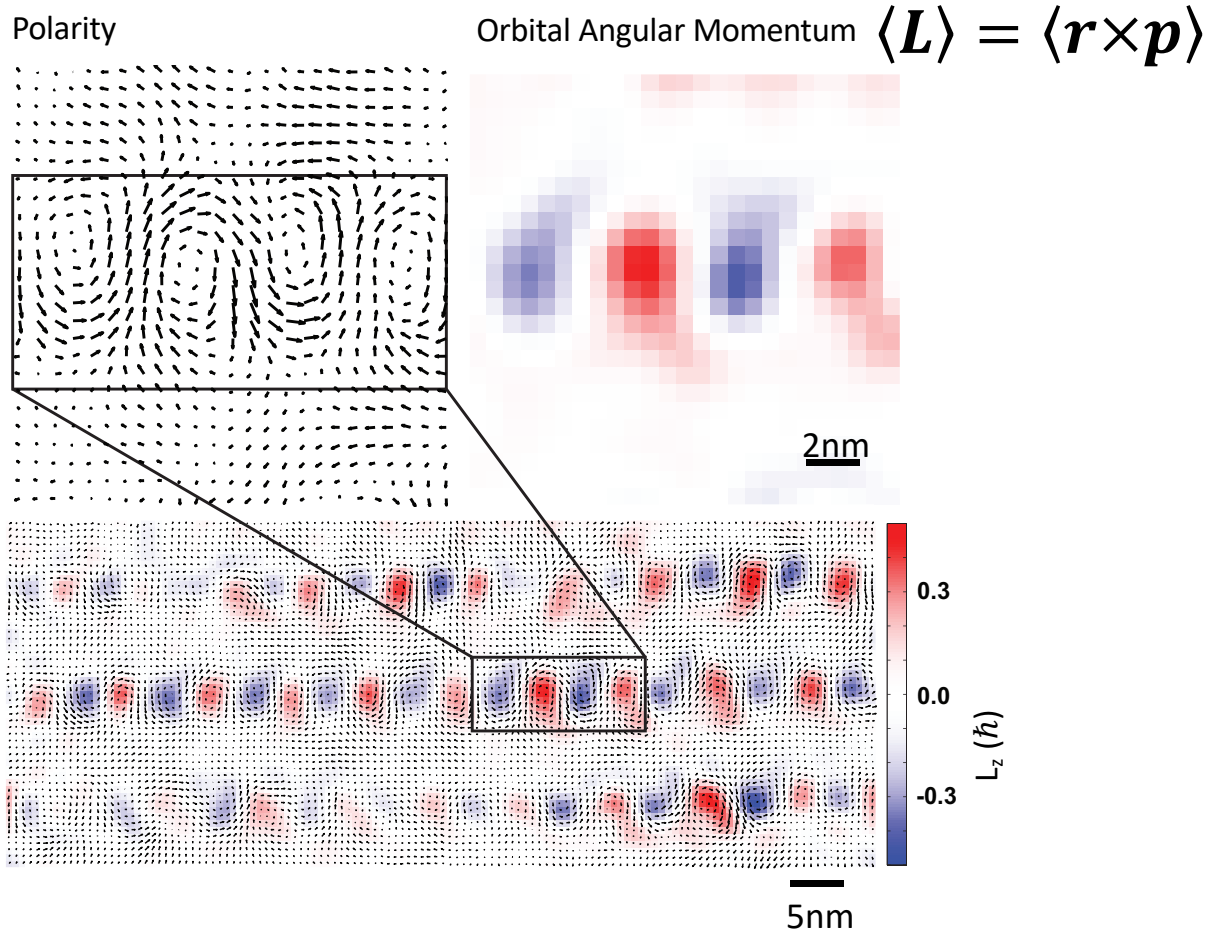
Skyrmions in FeGe at 100 K

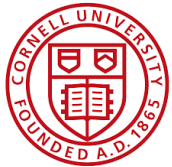


How to see the internal structure of Skyrmions



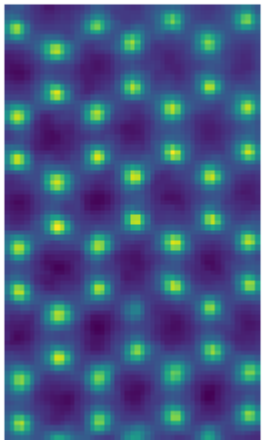
Polarization Vortices in $\text{PbTiO}_3/\text{SrTiO}_3$ Superlattices





Ptychography

Ptychography(x,y)



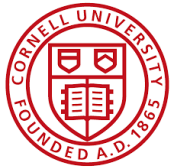
Find scattering potential $V(x, y)$ such that

Diffraction Pattern Model

Diffraction Data

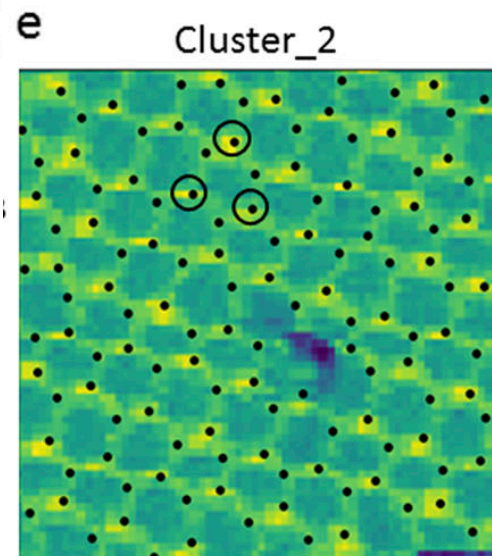
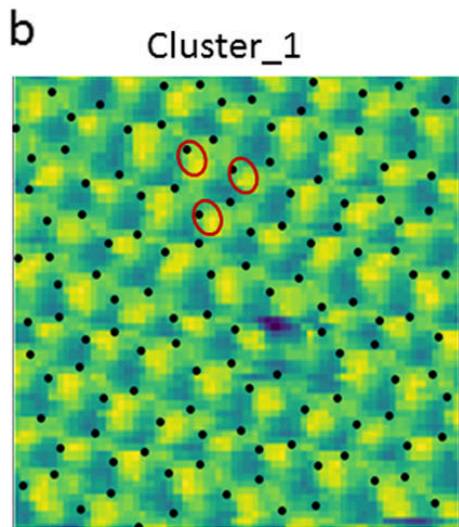
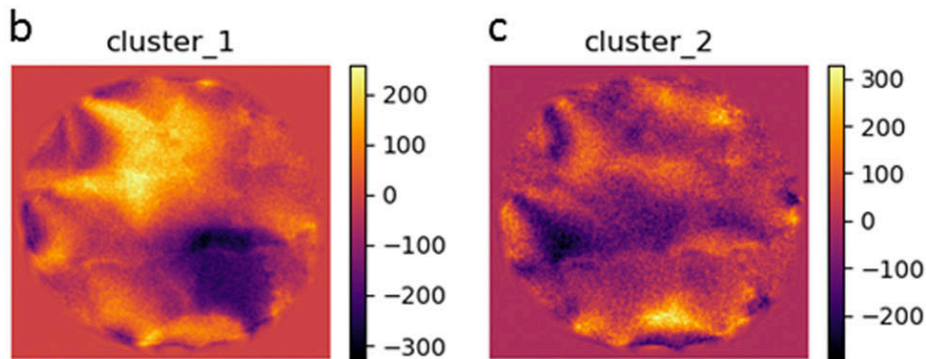
$$\left\| |F[\Psi_0(x' - x, y' - y) \exp i\sigma V(x', y')]|^2 - DP(x, y, k_x, k_y) \right\|_{\min}$$

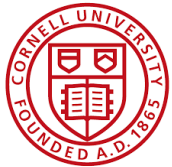
- Ptychography algorithm generates a higher quality image with better resolution and contrast
- Similar to neural network in that it minimizes a loss, but it doesn't learn a filter that can be immediately applied to other data sets (non-transferability).
- Other datasets must go through entirely separate ptychography process, analogous to re-training a network
- Ptychography doesn't work for all datasets (fails for thick samples as we lack a fast model)



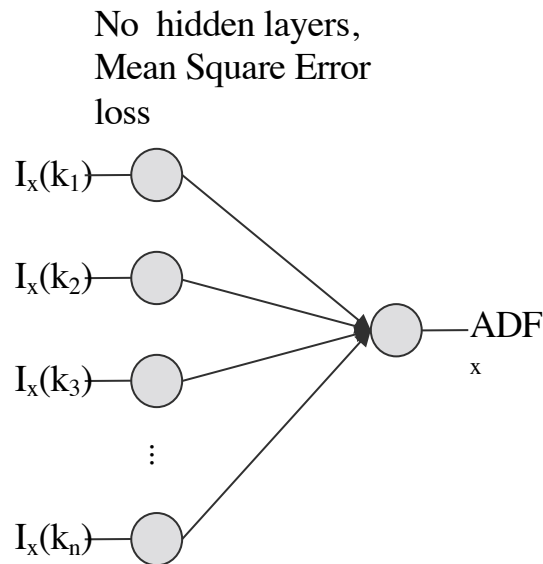
Unsupervised: Cluster Analysis

Li, ...Kalinin, *npj Computational Materials* **5**, (2019) 5

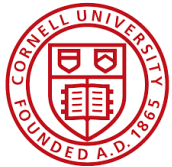




Linear Learning: Model

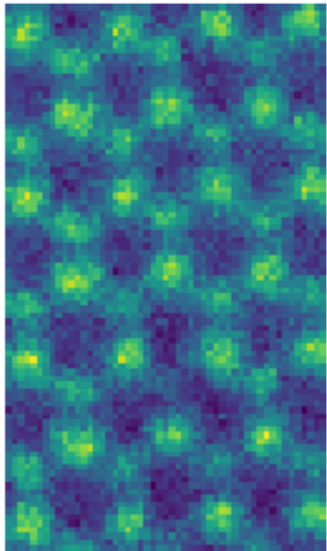
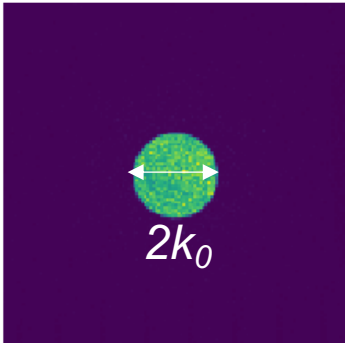


(Sergei says “Beware of overfitting”: Li, ..., Kalinin, *Phys. Rev. E* **100**, (2019), 023308.)



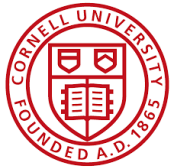
Conventional Imaging: Annular Dark Field (ADF)

ADF(x,y)

A color image showing a noisy, textured pattern with yellow and green spots on a dark blue background, representing an Annular Dark Field (ADF) image.
$$= \int \text{DP}(x,y,k_x,k_y) \times \text{Mask/Filter}(k_x,k_y) d\vec{k}$$
A dark purple square image with a small, bright green circular spot in the center. A white double-headed arrow below the spot is labeled $2k_0$, indicating the diameter of the spot.A black square image with a white annular ring in the center, representing the Mask/Filter function.

Yields a decently interpretable image but resolution is limited to $2k_0$

By training with a target image label that is better than ADF (i.e. ptychography label), can the neural network learn a **transferable** filter that's better than ADF?



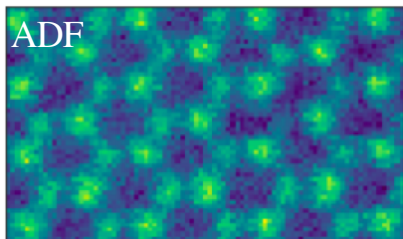
Linear Learning: Transferability?

ADF Mask

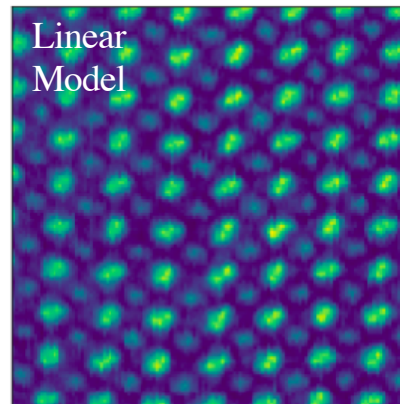
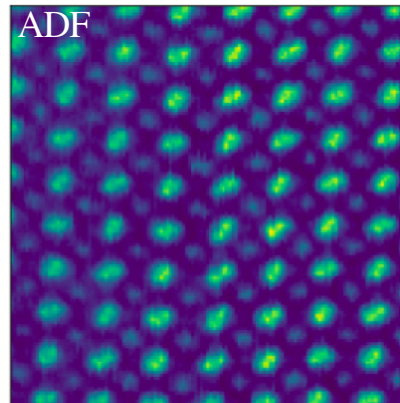
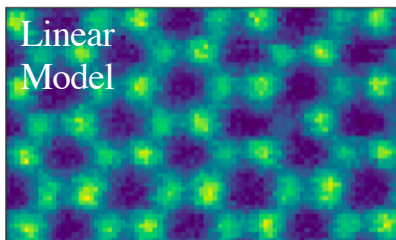
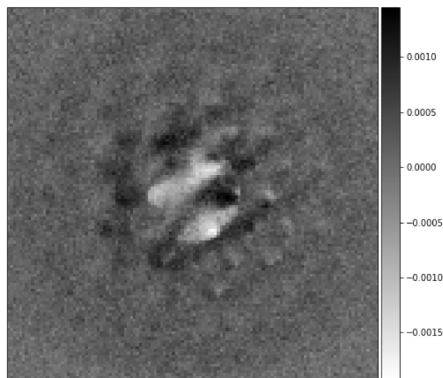


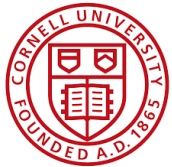
■ = 1

□ = 0

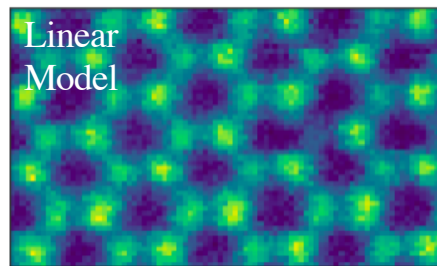
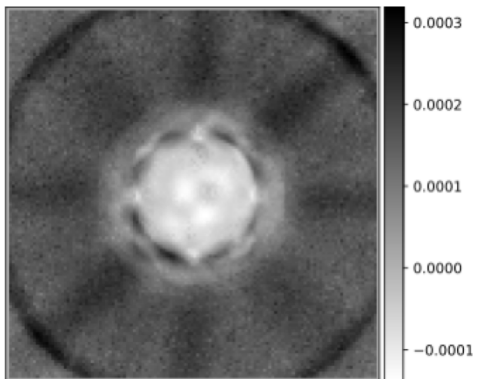


Linear Model

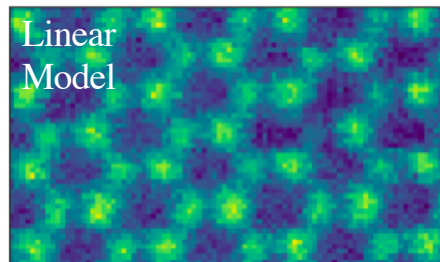
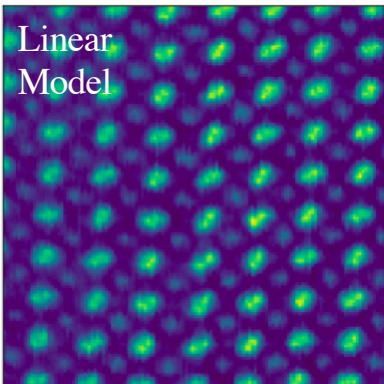




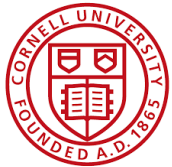
Linear Learning: Strong vs. Weak Scatterer



Training: Mos2 ADF Trained

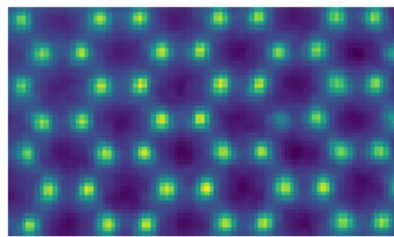
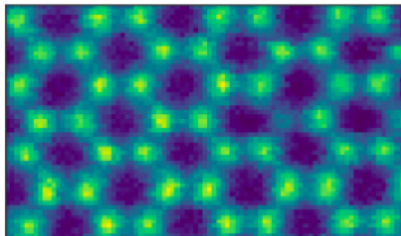
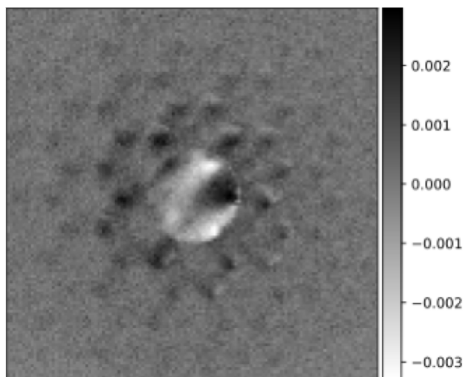


Training: PrScO3 ADF trained



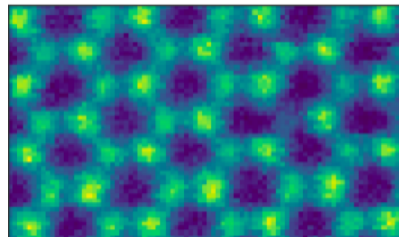
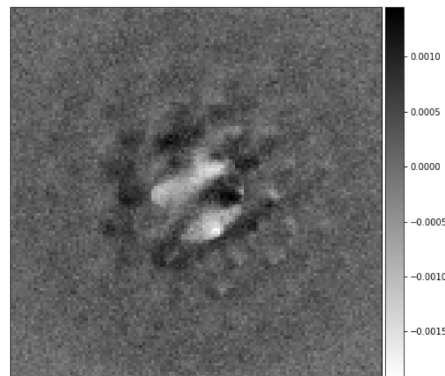
Linear Learning: ADF vs. Ptychography

Linear Model
(MoS₂ Ptych. Trained)

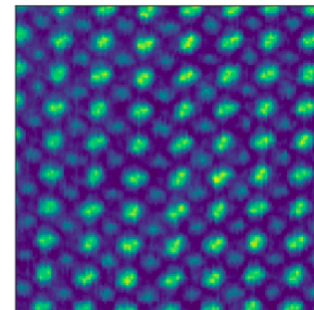


Label:

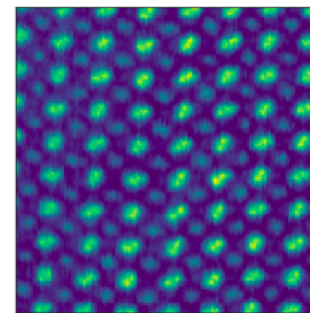
Linear Model
(MoS₂ ADF Trained)



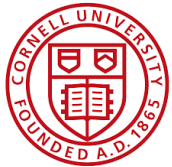
Linear Model
(MoS₂ Ptych. Trained)



Linear Model
(MoS₂ ADF Trained)

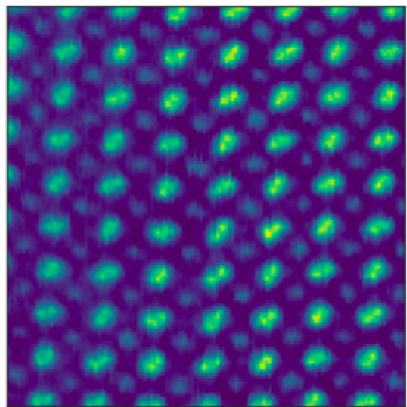


Looks close to ADF

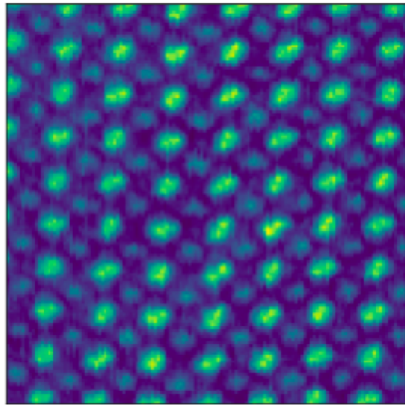


Linear Learning: ADF vs. Ptychography

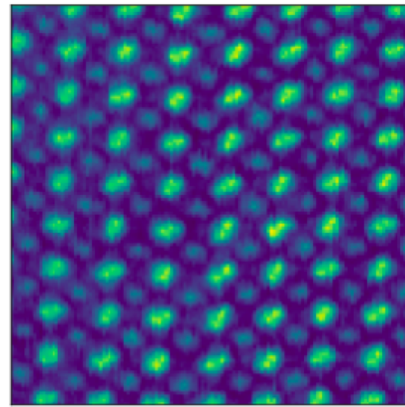
ADF



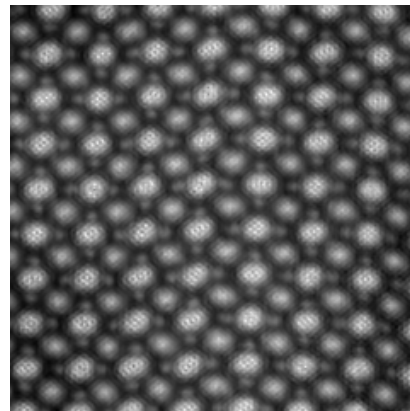
Linear Model
(MoS2 Ptych. Trained)



Linear Model
(MoS2 ADF Trained)

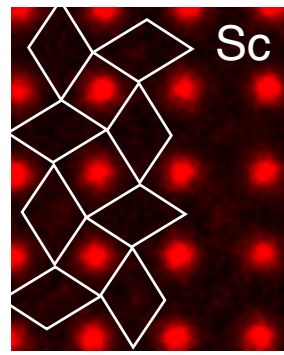
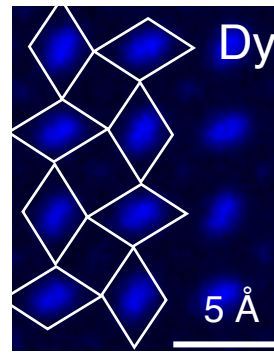
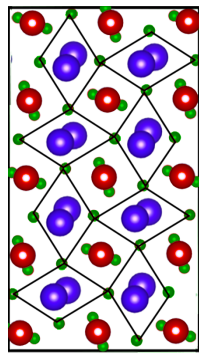


Ptychography

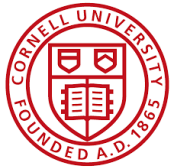


DyScO₃

Looks close to ADF

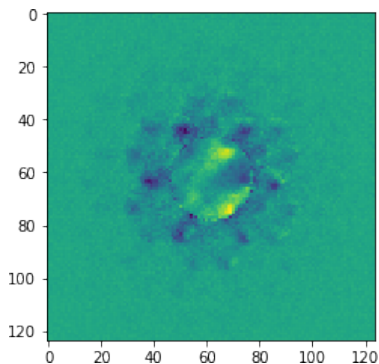


Sc



Addressing Overfitting to Training Lattice

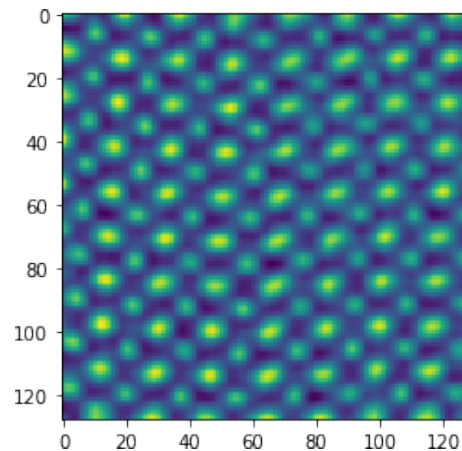
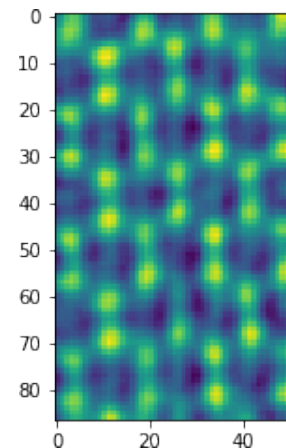
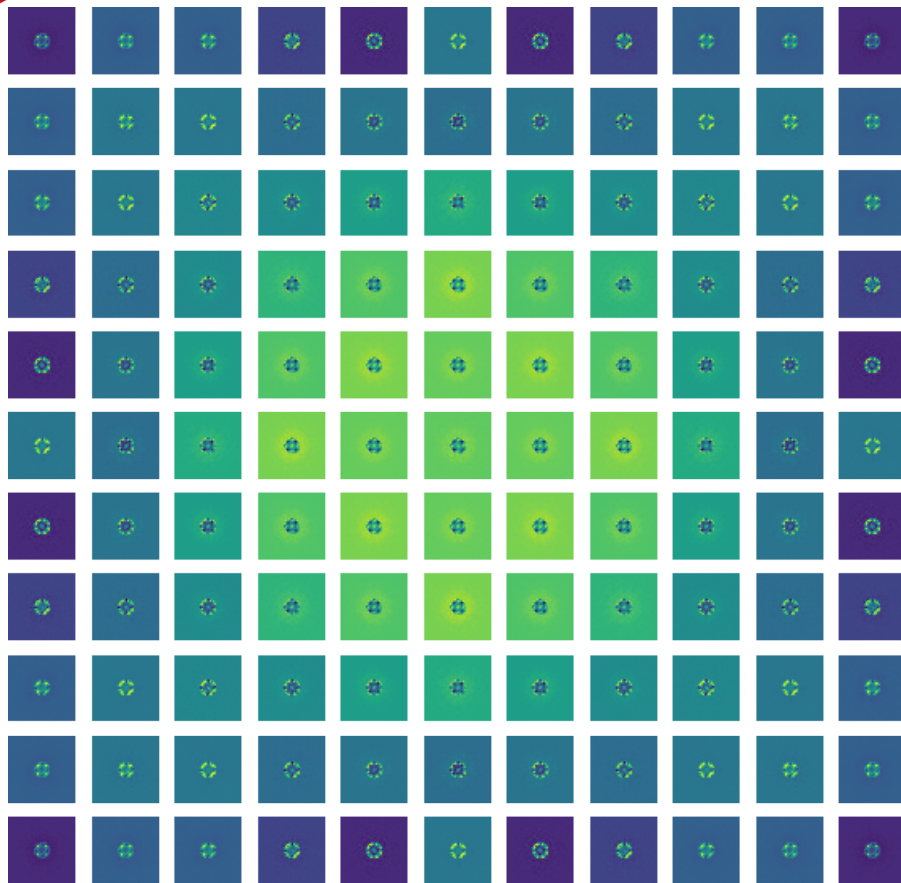
Trained Diffraction Mask

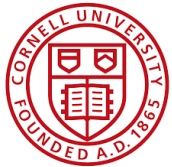


- The trained filter shows a clear six-fold symmetry matching the symmetry of the training lattice image
- Test data has four-fold symmetry
- Overfitting to the training lattice's symmetry negatively affects transfer
- Solution: Train a symmetric filter

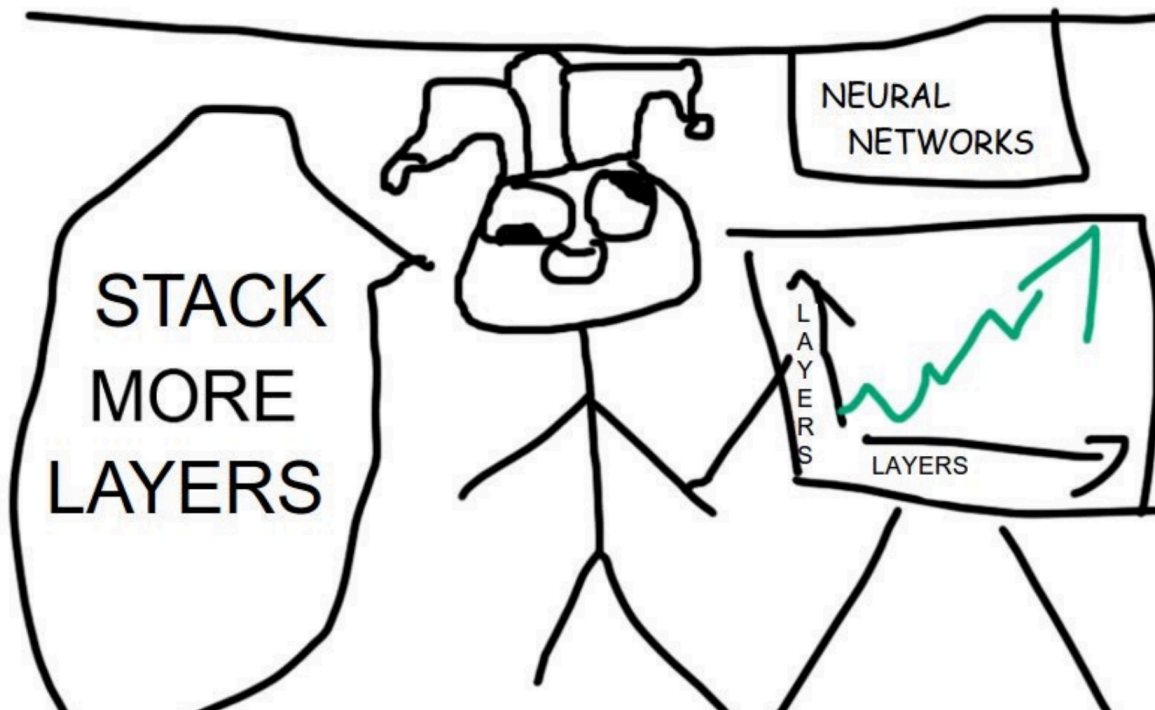


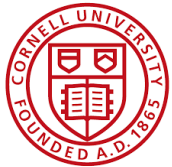
Symmetric Masks and Symmetric 11x11 Neighbors



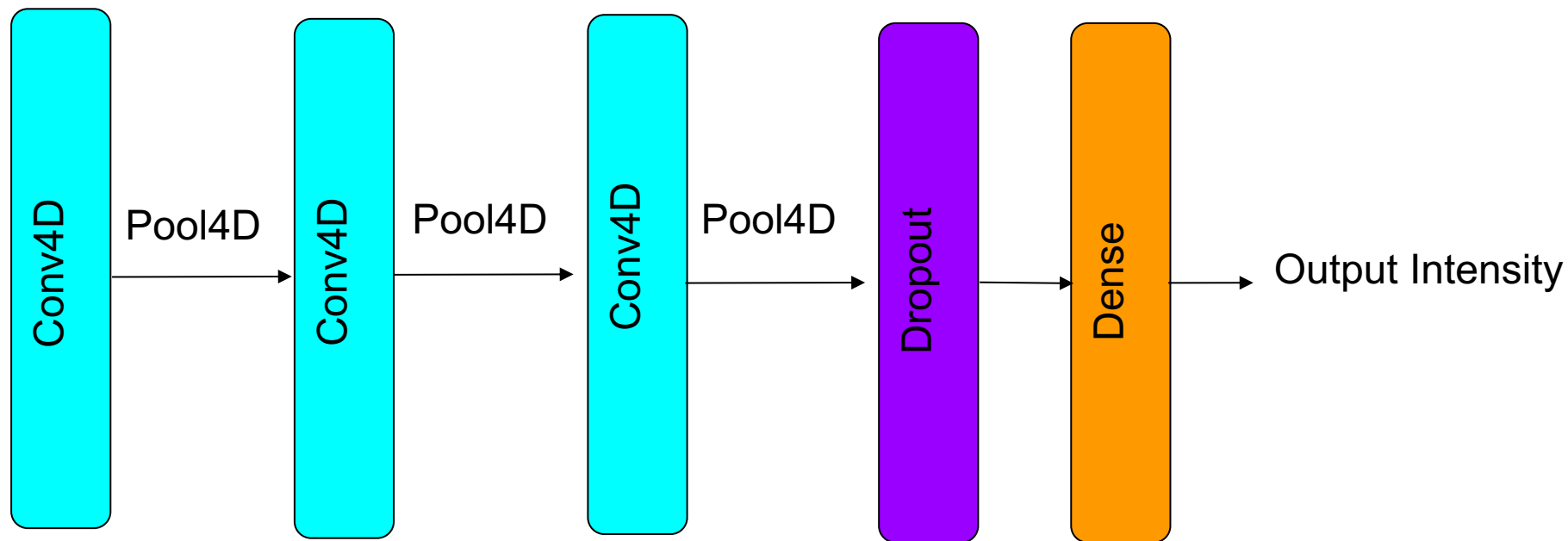


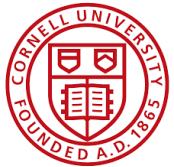
The logical next step





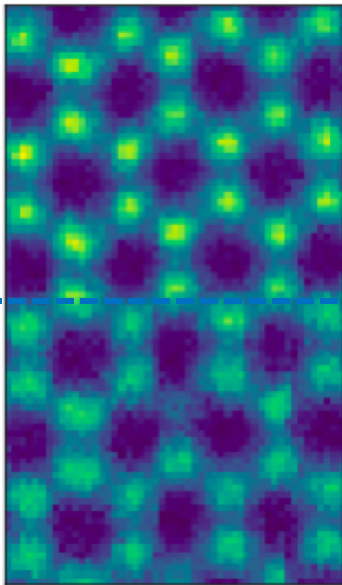
Convolutional Model: Model



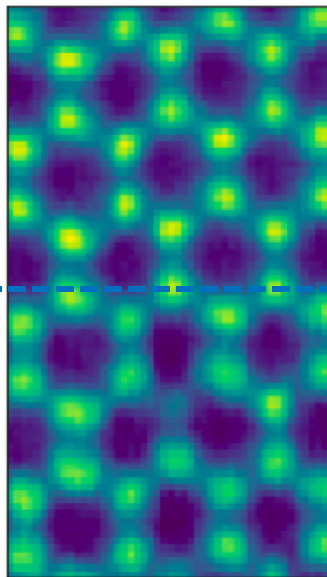


Convolutional Model: Better Ptycho training

Linear

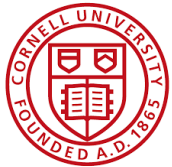


CNN



Train

Fit



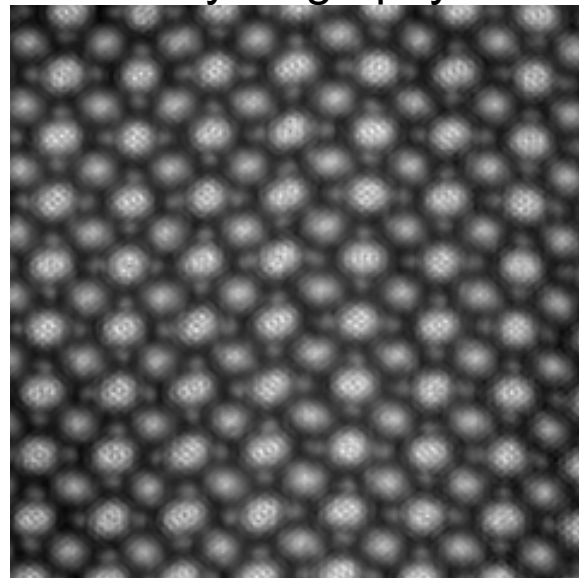
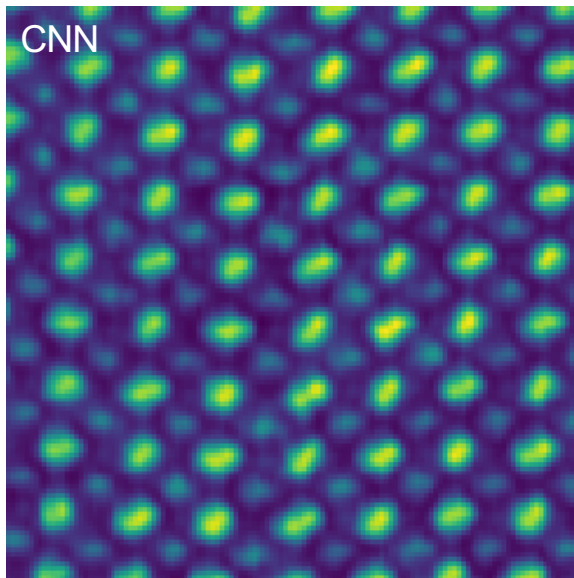
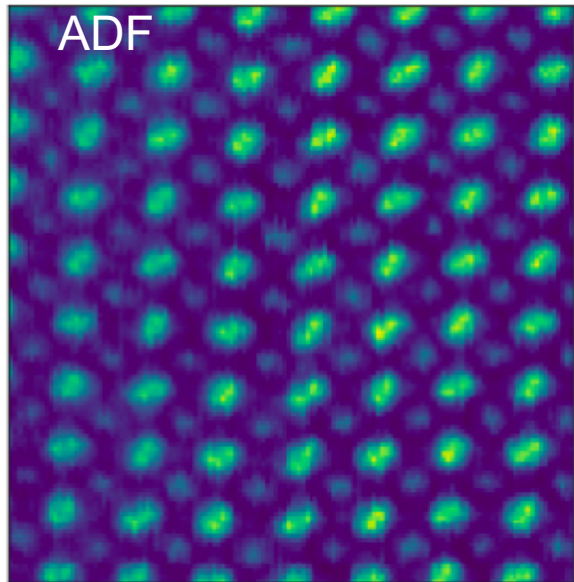
Convolutional Model: Transfer to Thin PrScO_3

CNN trained with MoS₂ Ptych.

Ptychography

ADF

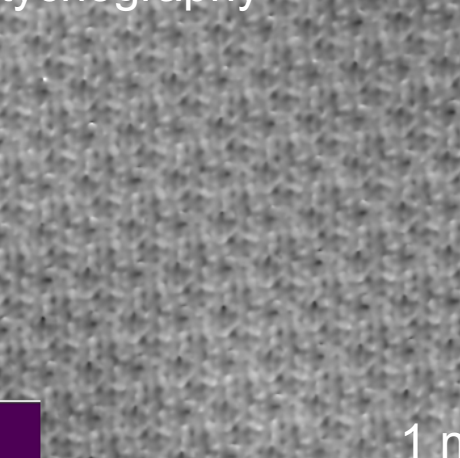
CNN



DyScO_3



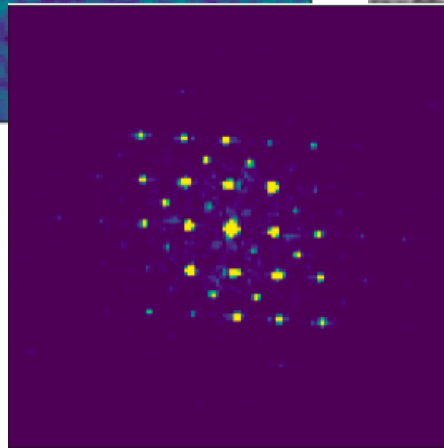
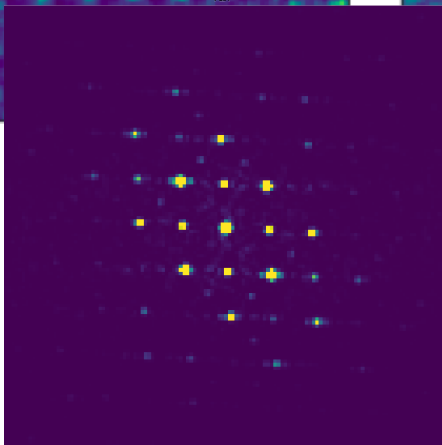
ADF



ptychography

1 nm

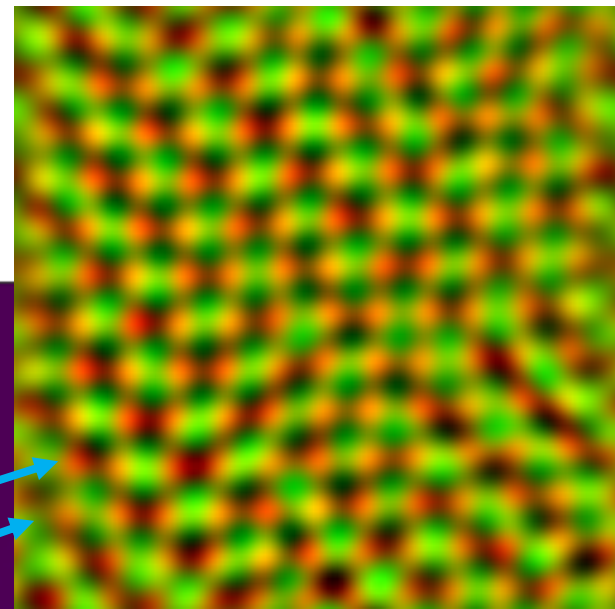
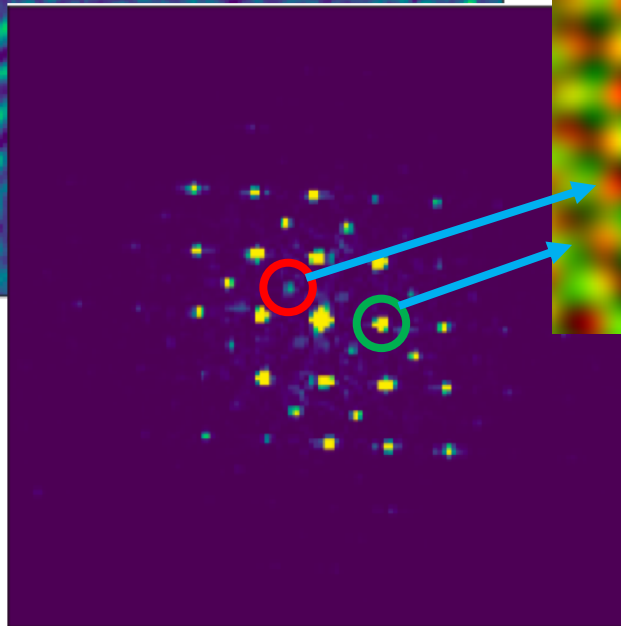
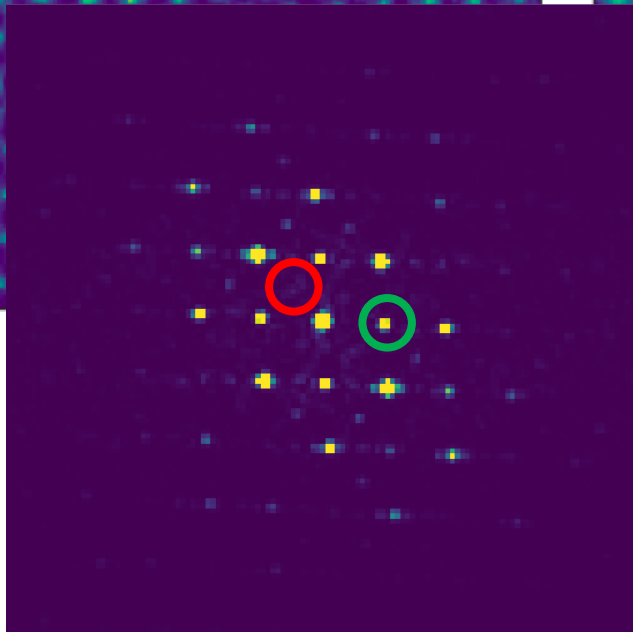
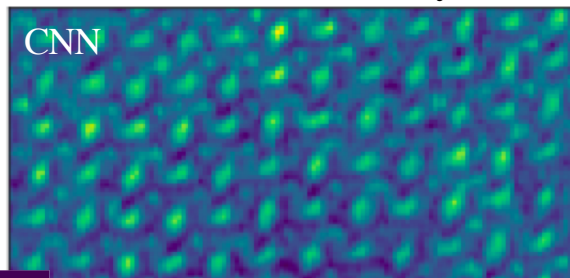
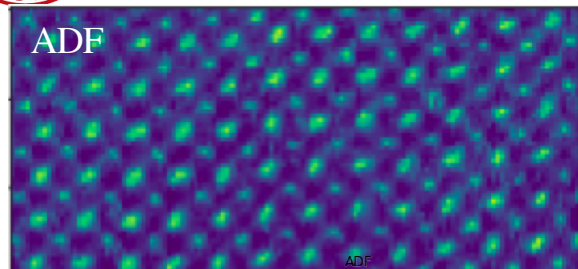
1 nm



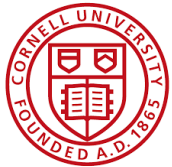


Convolutional Model: Transfer to thick

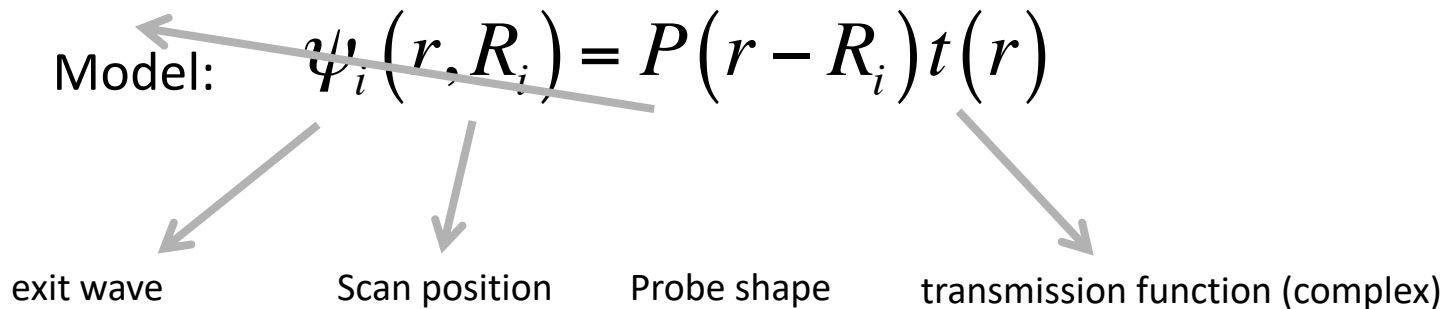
CNN trained with MoS₂ Ptych.



Sensitive to A-site
Splitting



Limits of the Strong Phase Approximation



Object



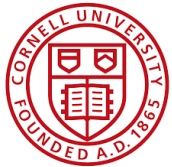
Strong Phase Approximation: $t(r) = \exp[i\sigma V(r)]$



Strong phase approximation is valid if $T \leq 5.2 \frac{(\delta r)^2}{\lambda}$

E. Tsai et al. *Optics Express* **24(25)** 29089-29108 (2016)

Thickness limit for resolving $\delta r = 0.67 \text{ \AA}$ at 120 keV is $\sim 7 \text{ nm}$.



Solving Schrödinger's Equation: Multi-slice Ptychography

$$\psi_{n+1}(x, y) = t_n(x, y) [p(x, y, \Delta z) \otimes \psi_n(x, y)]$$

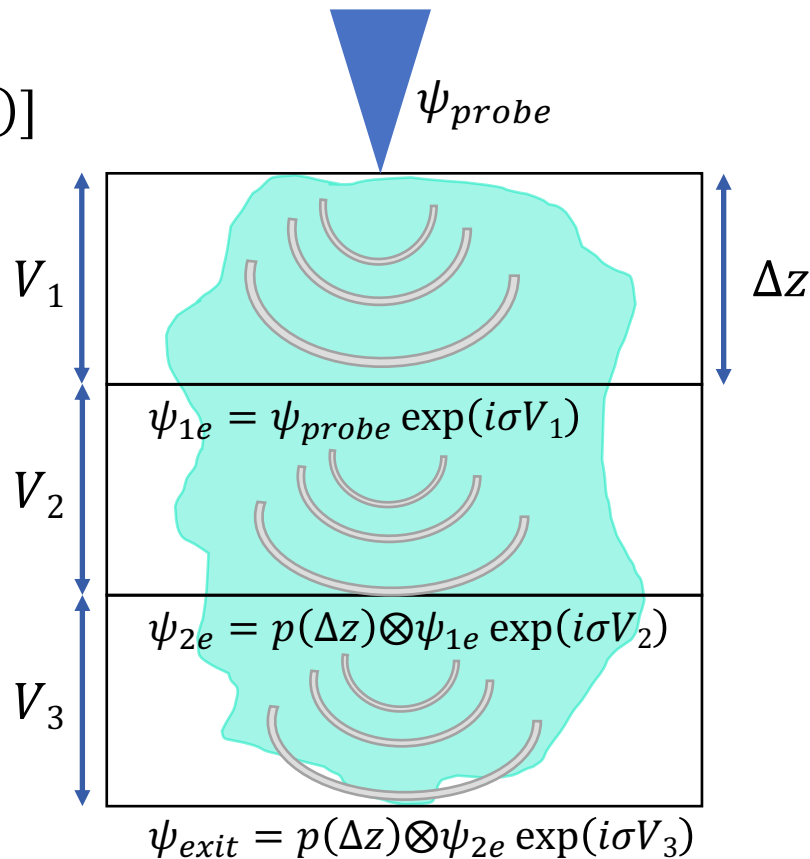
Transmission function
 $t_n(x, y) = \exp(i\sigma V_n)$

Fresnel
propagator

Ptychography:
Reconstruct $V_1 = V_2 = V_3 = \dots$ from scanning
diffraction patterns $(|\psi_{exit}(\vec{k}, \vec{r}_p)|^2)$.

Multi-slice ePIE

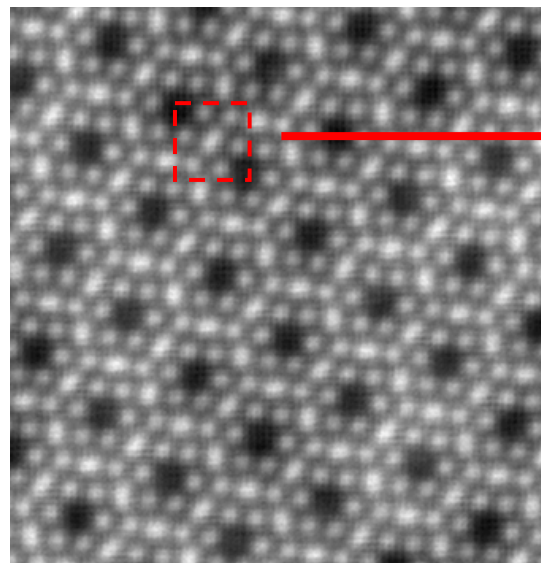
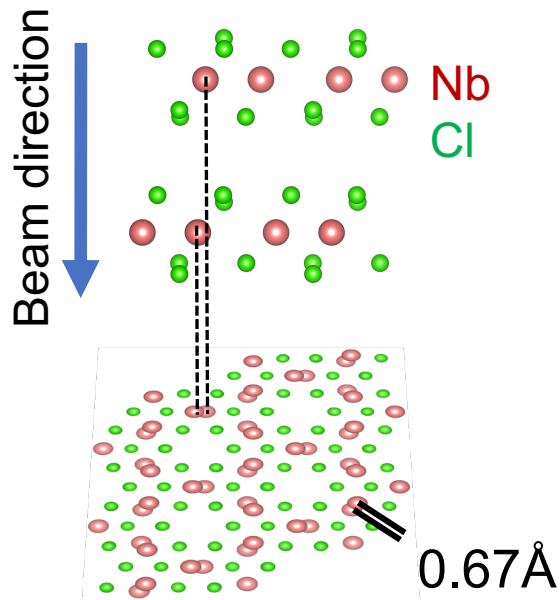
A.M. Maiden et al. J. Opt. Soc. Am. A, 29(8) (2012)



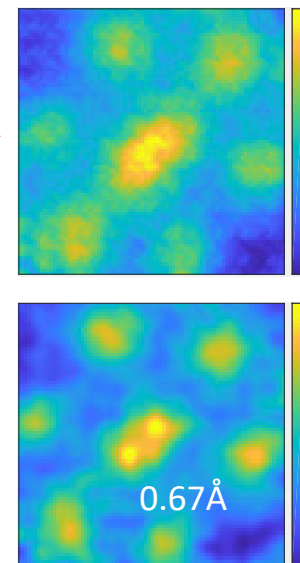
Multislice Ptychography of Thick Nb_3Cl_8

Yi Jiang¹, Zhen Chen², Ismail El Baggari², Lena F. Kourkoutis², Veit Elser², David A. Muller²

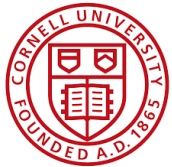
¹Argonne National Laboratory, ²Cornell University



single-slice ptychography

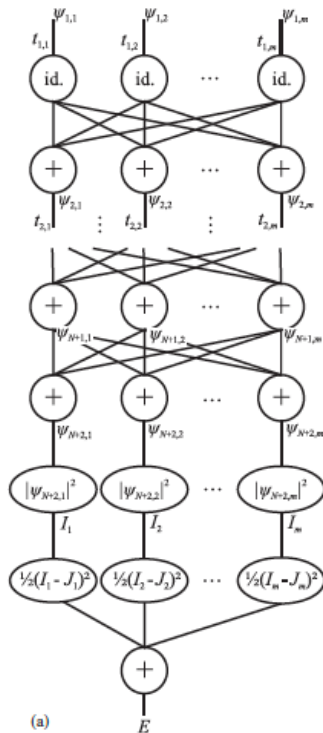


multi-slice ptychography



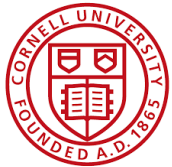
Solving Schrödinger's Equation: Beam Propagation as a NN

GENERAL FRAMEWORK FOR QUANTITATIVE THREE- ...



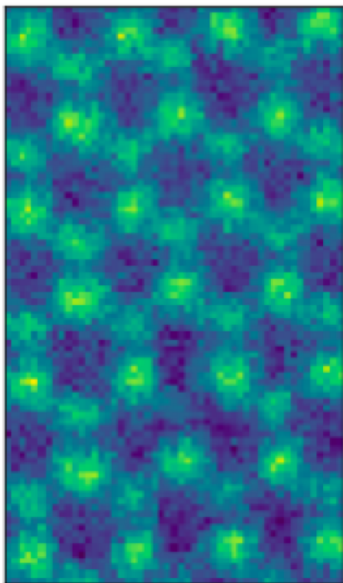
Back propagation to find derivatives of Error wrt node inputs

Also looks ill-conditioned, like MS-ePIE ...

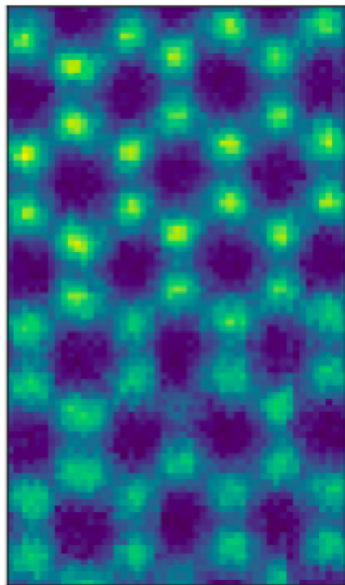


Phase Retrieval for Electron Imaging

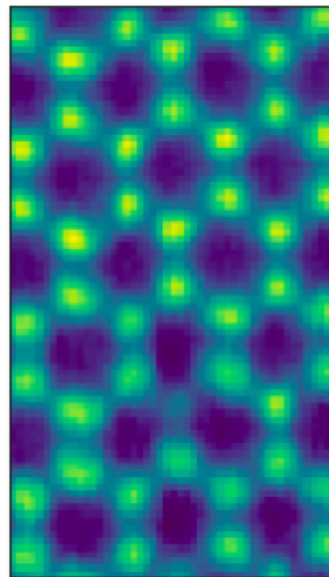
ADF



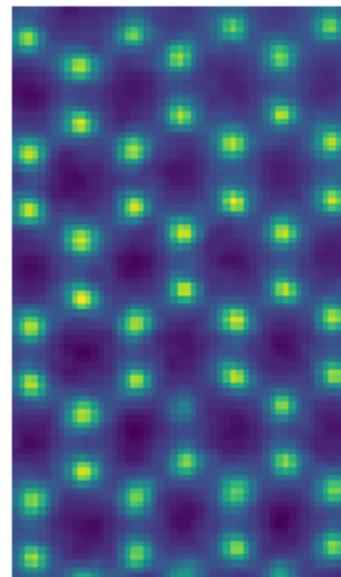
Linear



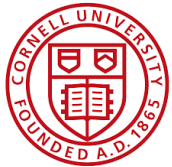
CNN



Ptychography



- An inverse problem for Structure and Physics
- TB/hour of data
- Thick samples: Beam propagation



PARADIM

PLATFORM FOR THE ACCELERATED REALIZATION,
ANALYSIS & DISCOVERY OF INTERFACE MATERIALS



U.S. DEPARTMENT OF
ENERGY

Office of Science



Eun-Ah Kim
Michael Matty
Li Li (GAS)



PAD Hardware
Sol Gruner

Mark Tate
Praful Purohit

Hugh Philips
Kate Shanks

2D Materials
Jiwoong Park

Saien Xie

Kibum Kang
Hui Gao

Kin Fai Mak
Jie Shan

Skyrmions
Greg Fuchs
Song Jin

Emrah Turgut
Nitish Mathur

Bob Burhman