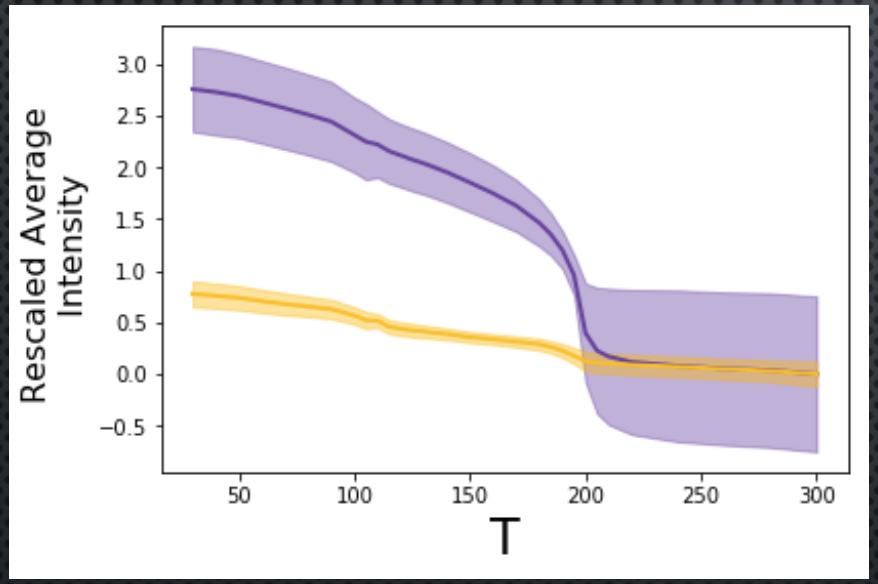
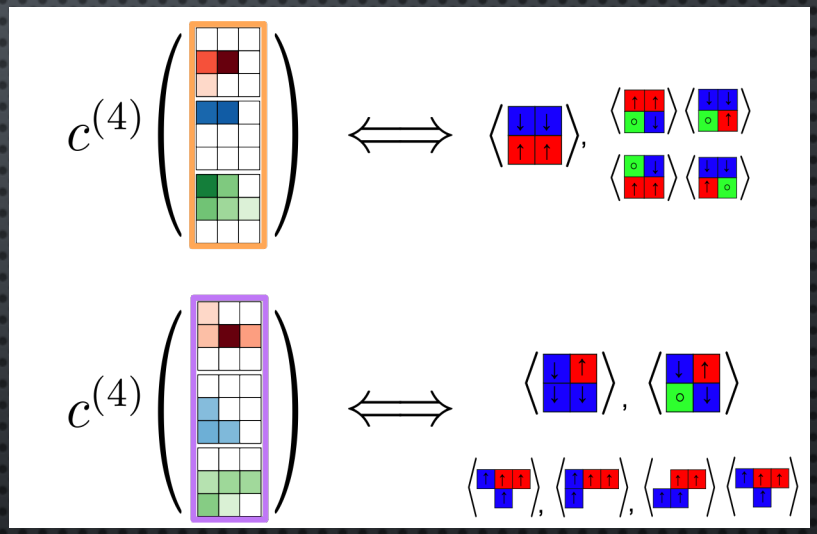
HARNESSING DATA REVOLUTION IN QUANTUM MATTER



Venderley et al, arXiv:2008.03275



Miles et al, arXiv:2011.03474

EUN-AH KIM (CORNELL)

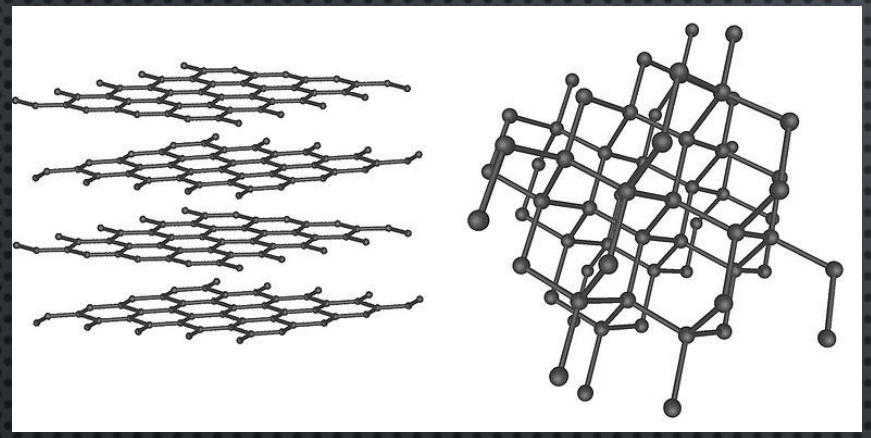
Physics U ML (12/02/2020)

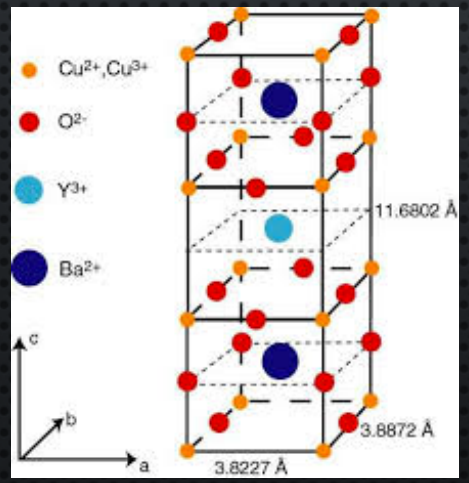
CCMR

PARADIM
AN NSF MATERIALS INNOVATION PLATFORM

U.S. DEPARTMENT OF ENERGY

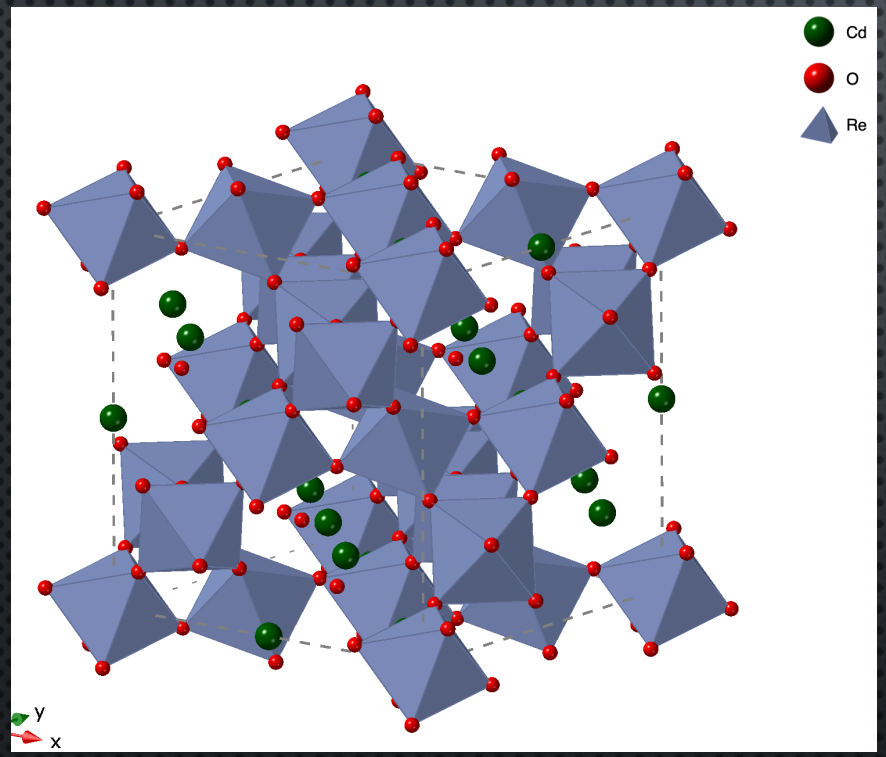
Understand and Control ?





Source: wikipedia

INFORMATION CONTENT



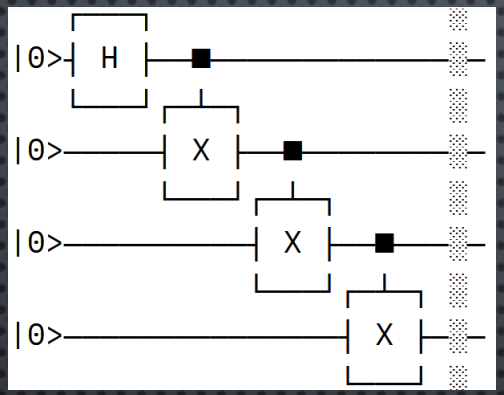
$$\gtrsim \left(\text{Library of Congress} \right)^2$$

Library of Congress ~
20 TB
Source: loc.gov

DATA REVOLUTION

Hilbert Space of Quantum Systems

Source: ibm.com



⋮

\+

Noise

Density matrix

Measurement

$(0_z, 0_z, 0_y, 1_x, \dots)$

$(1_z, 0_x, 0_y, 1_z, \dots)$

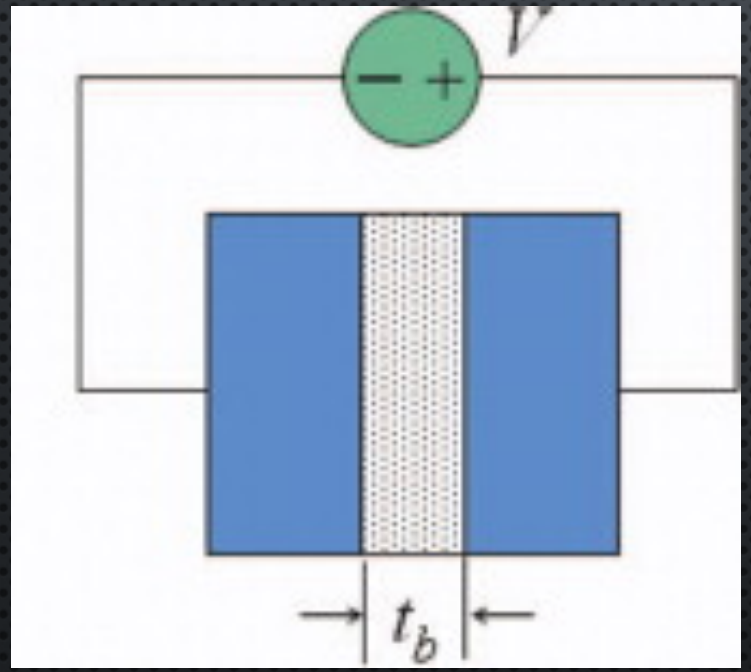
$(0_z, 1_y, 0_z, 1_y, \dots)$

$(0_x, 0_x, 0_y, 0_y, \dots)$

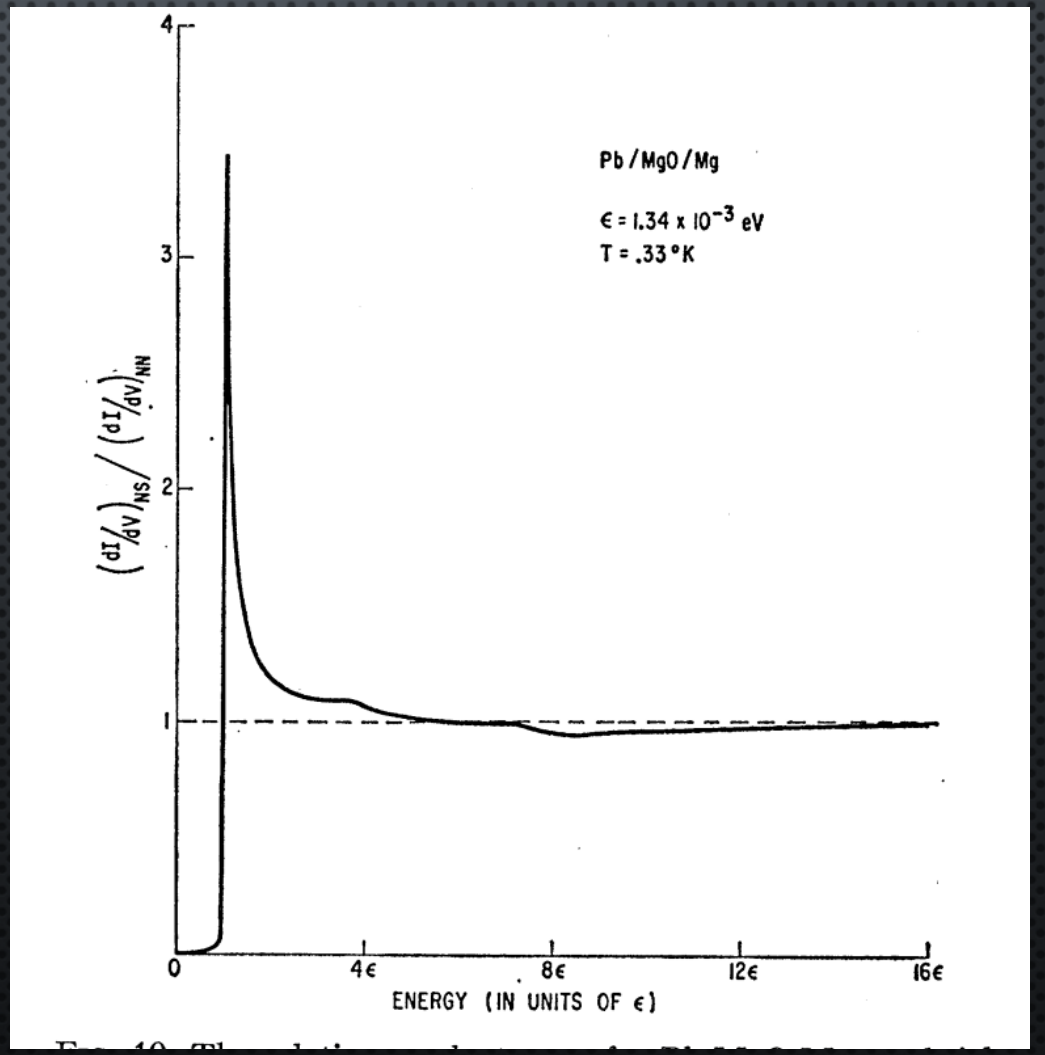
$(1_x, 0_x, 1_z, 0_y, \dots)$

⋮

TUNNELING DENSITY OF STATES, IN 1962

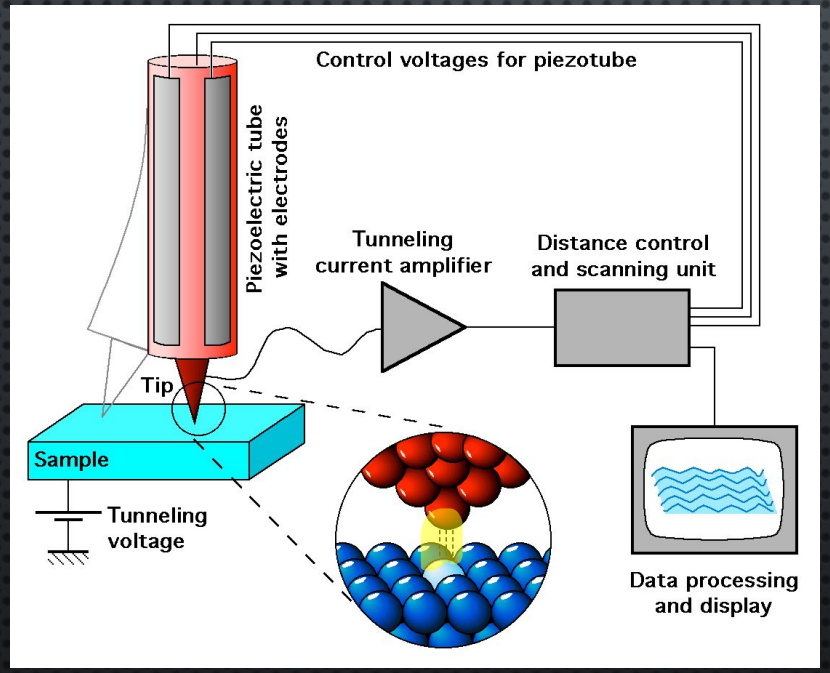


Differential conductance dI/dV @ V proportional to N(E=eV)

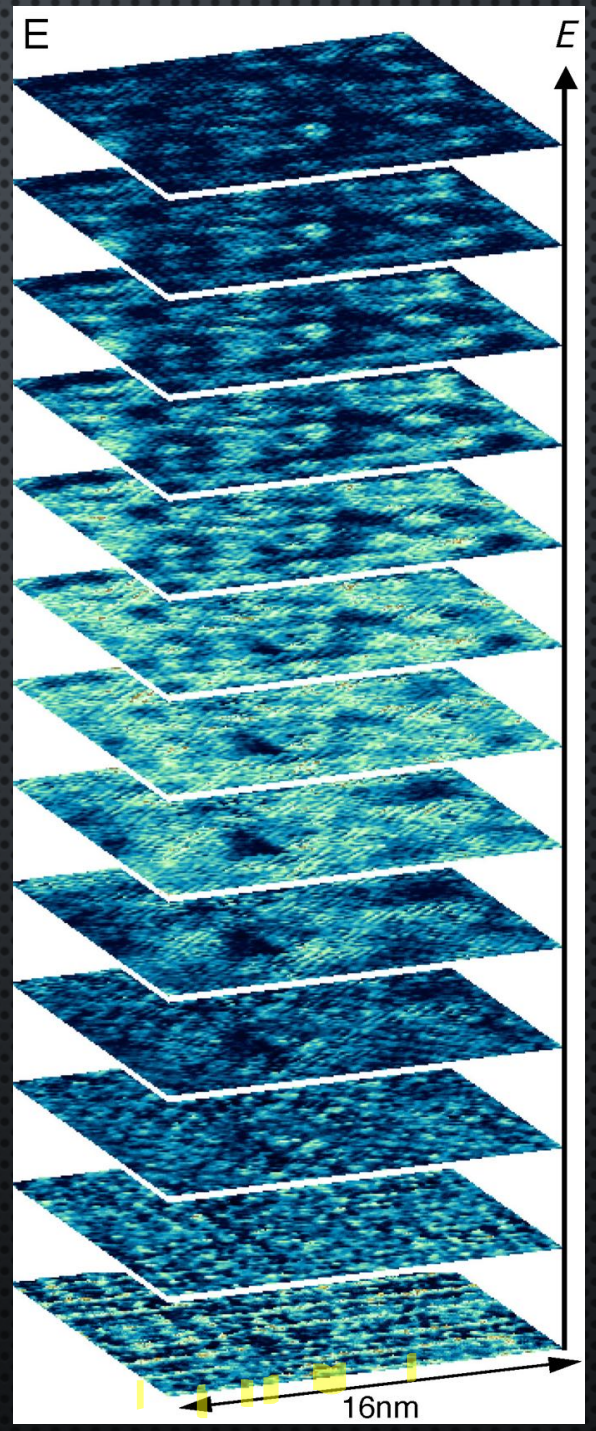


Giaever et al, Phys. Rev. 126, 941 (1962)

Tunneling Density of States, in 2000's



Imaging N(r,E):
Scanning Tunneling Spectroscopy



X-ray diffraction in 1913

The Reflection of X-rays by Crystals.

By W. H. Bragg, M.A., F.R.S., Cavendish Professor of Physics in the University of Leeds; and W. L. Bragg, B.A., Trinity College, Cambridge.

(Received April 7,—Read April 17, 1913.)

Proceedings of Royal Society A, 01, July 1913

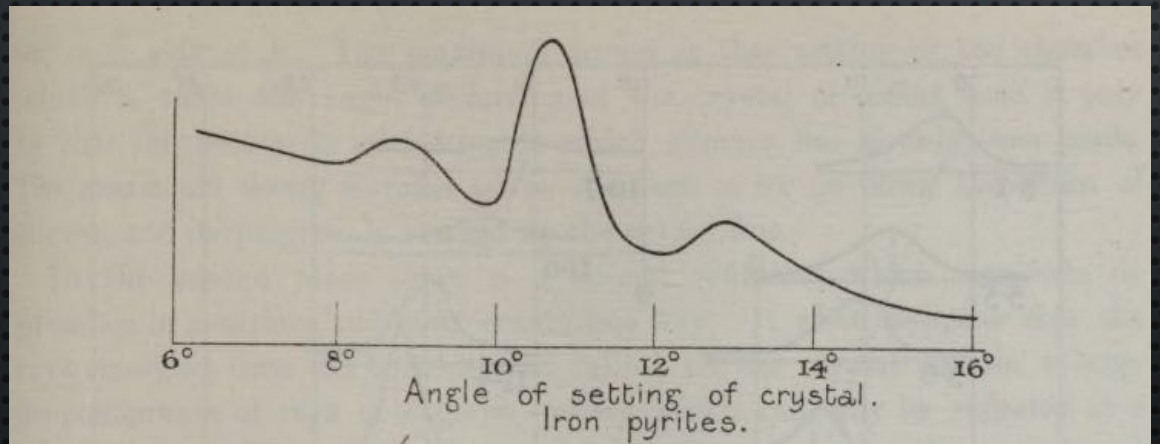
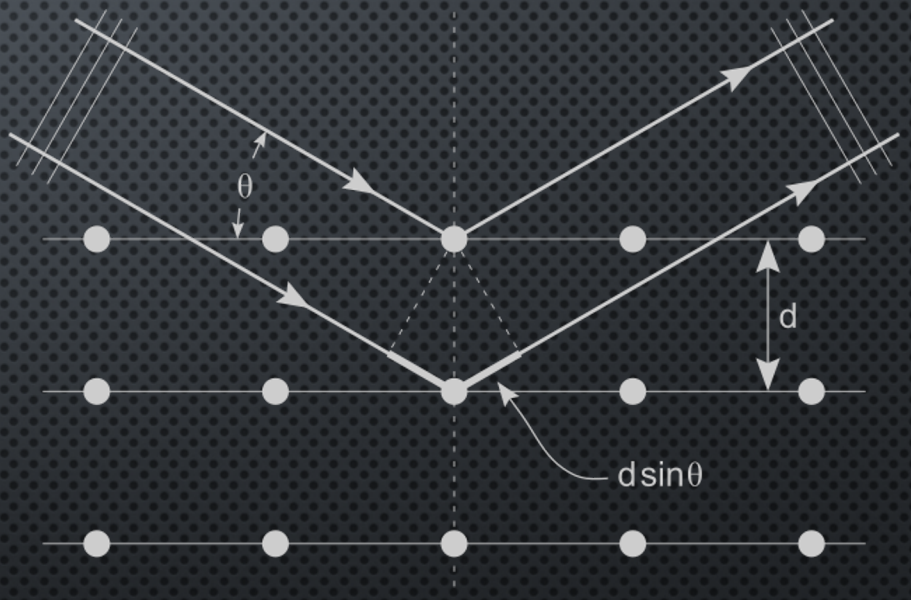


Fig. 2.—Reflection from face (100) of iron pyrites, at varying angles of incidence. Abscissa—Angle of incidence of rays on crystal face; Ordinate—Strength of reflected beam, arbitrary scale.

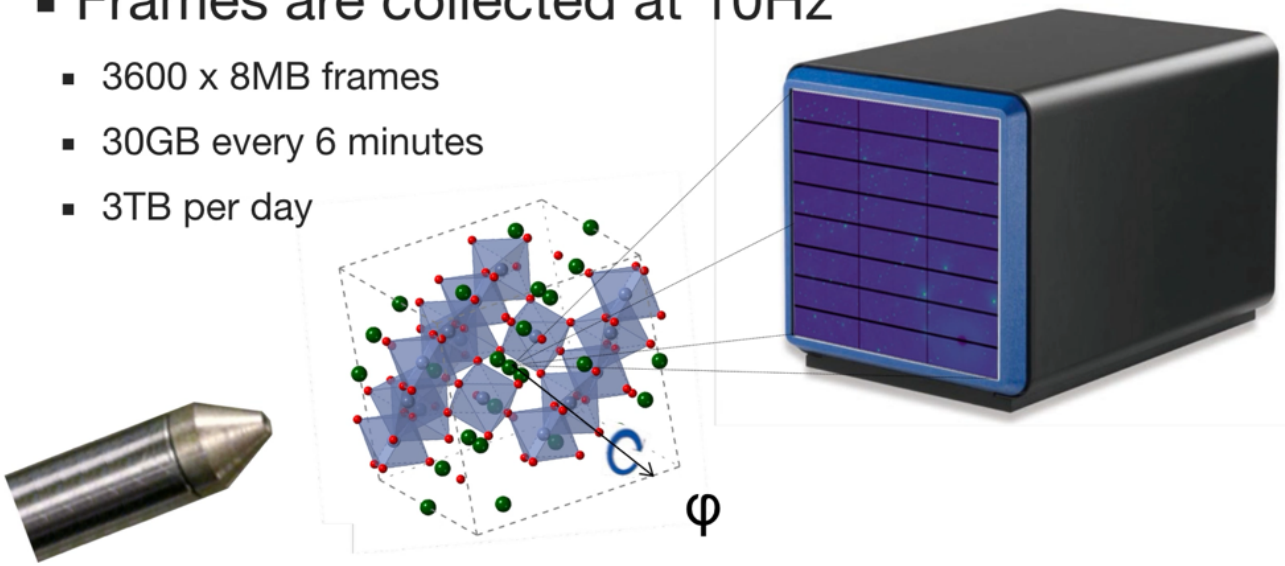


$$n\lambda = 2d\sin\theta$$

X-RAY SCATTERING GEOMETRY

Continuous Rotation Method

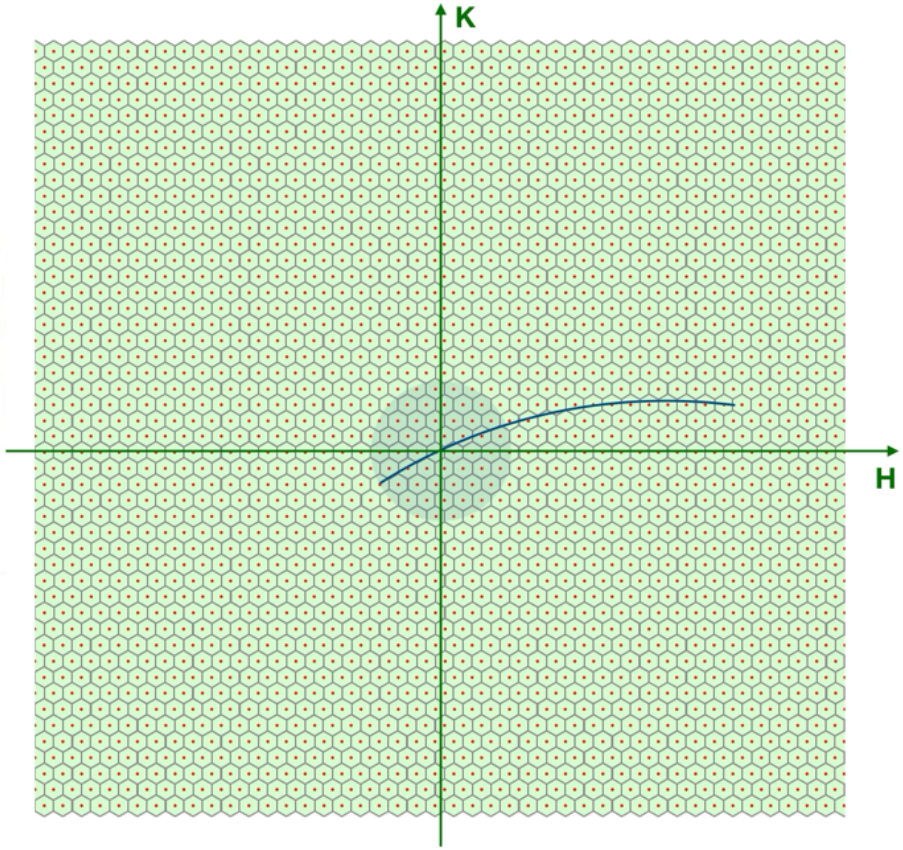
- The sample is continuously rotated at 1°s^{-1}
- Frames are collected at 10Hz
  - 3600 x 8MB frames
  - 30GB every 6 minutes
  - 3TB per day



Incident Beam

Sample

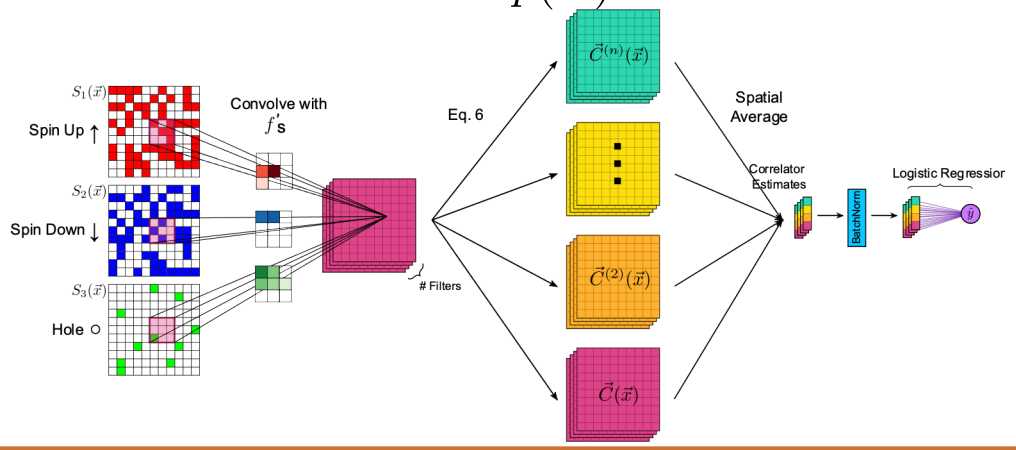
Pilatus 2M CdTe Detector



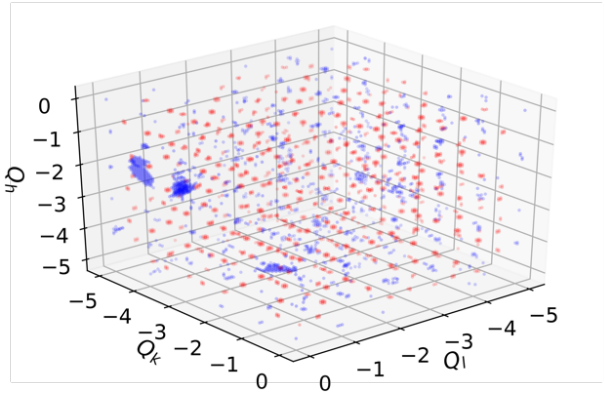
Reciprocal Space

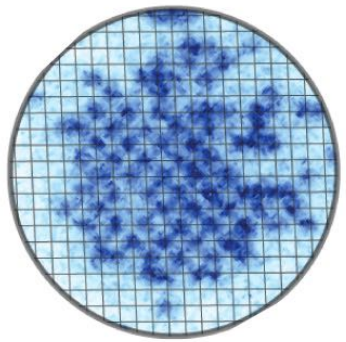
Data Science

$$p(A|B) = \frac{p(B|A)p(A)}{p(B)}$$

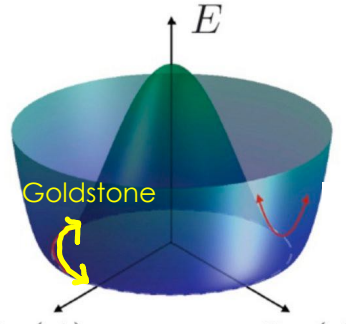


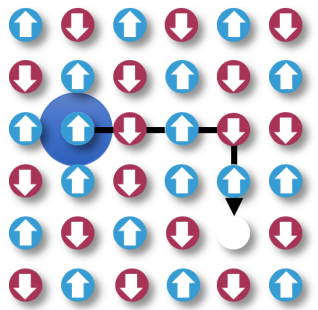
Experimental Data



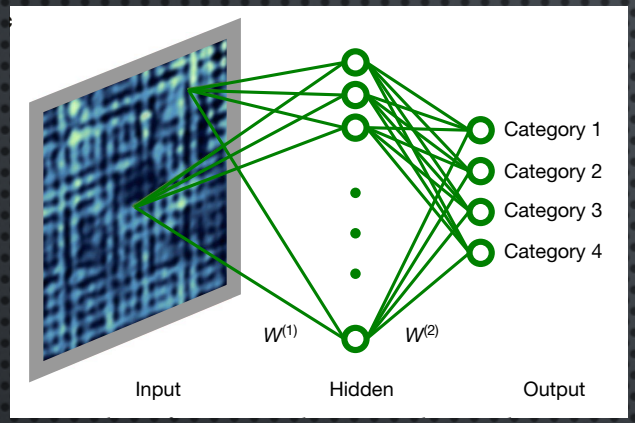


Theoretical Understanding

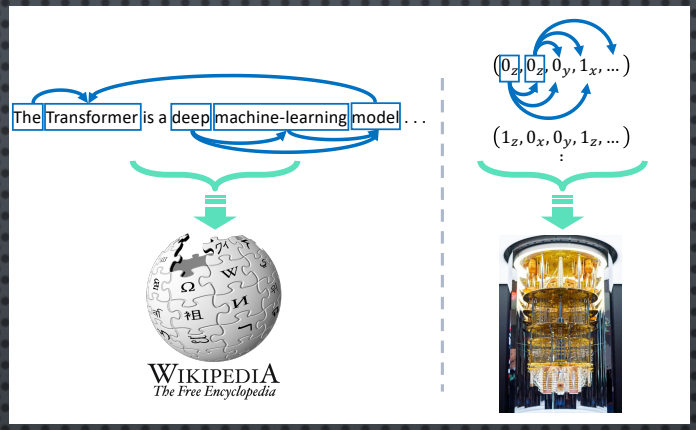




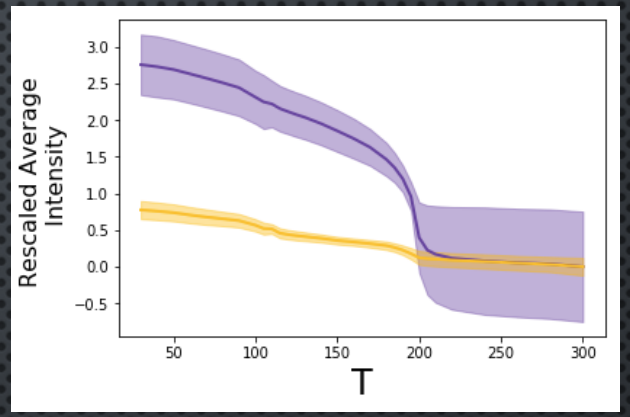
Harnessing Data Revolution



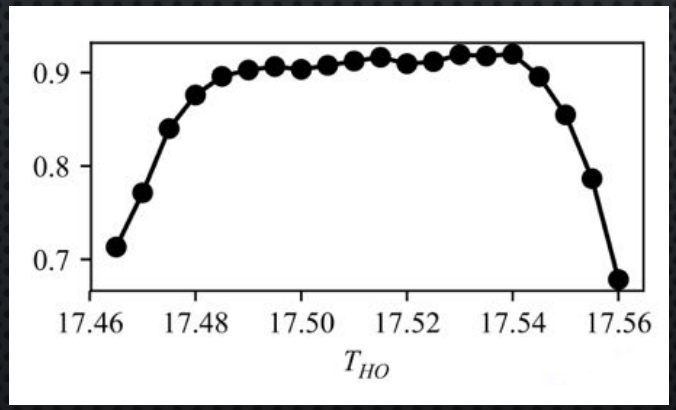
STS on high Tc cuprate BSSCO
Testing Charge Order Hypothesis
Zhang et al, Nature 570, 484 (2019)



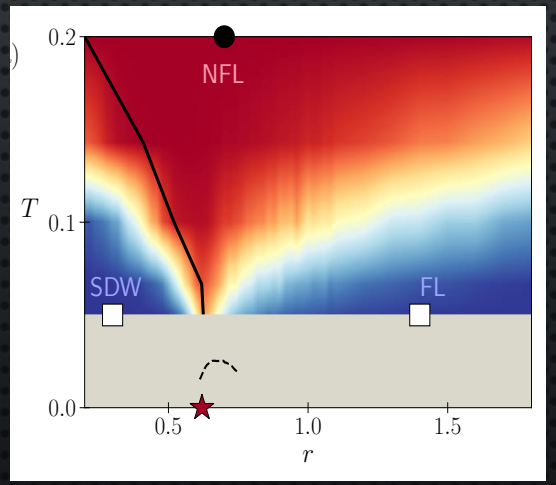
Attention-based Quantum Tomogrphy
Cha et al, arXiv:2006.12469



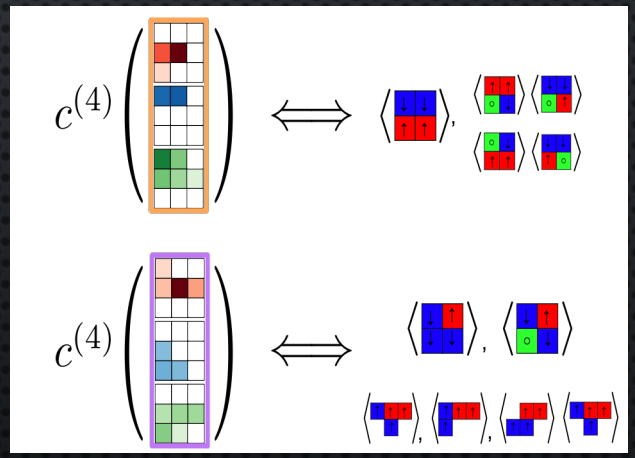
X-TEC
Venderley et al, arXiv:2008.03275



RUS on URu_2Si_2
Constraining the hidden order OP
Ghosh, Matty et al. Sci. Adv.(2020)



Detecting Non-Fermi Liquid Transport
Driskell et al, arXiv: 2007.07898



QGM on Fermi Hubbard model
Miles et al, arXiv:2011.03474

Use minimalistic approaches that integrates key physics principles.

Unsupervised ML for Voluminous XRD data: X-TEC (XRD Temperature Clustering)


J. Venderley


M. Matty


G. Pleiss
(Cornell, CS)


K. Winberger
(Cornell, CS)

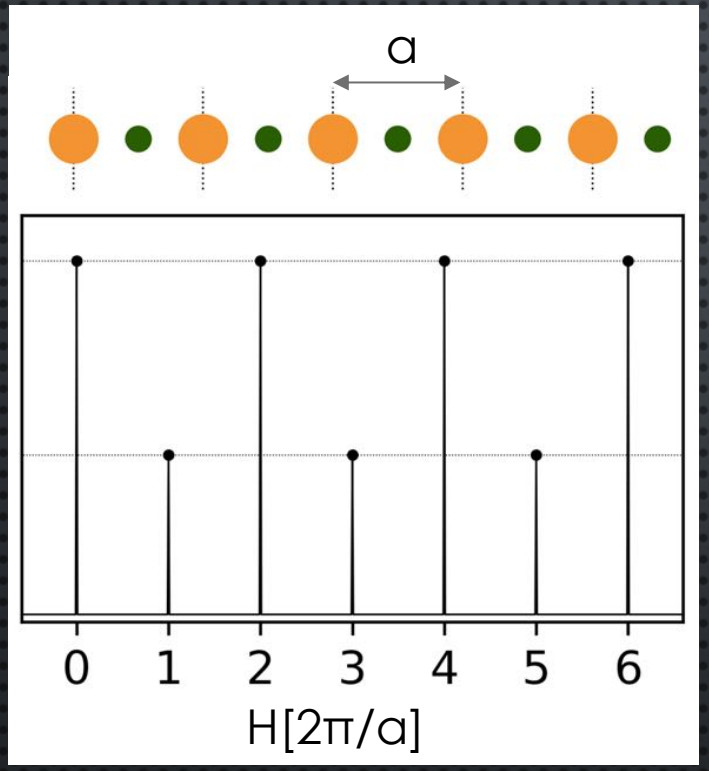

J. Ruff
(CHESS)


M. Krogstad
(ANL)

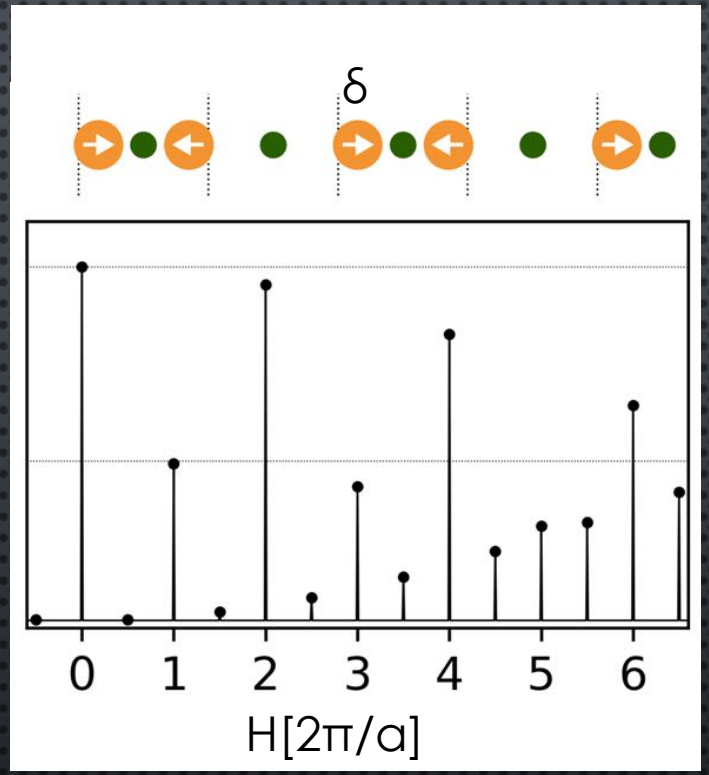

R. Osborn
(ANL)

J. Venderley, EAK et al,
arXiv:2008.03275

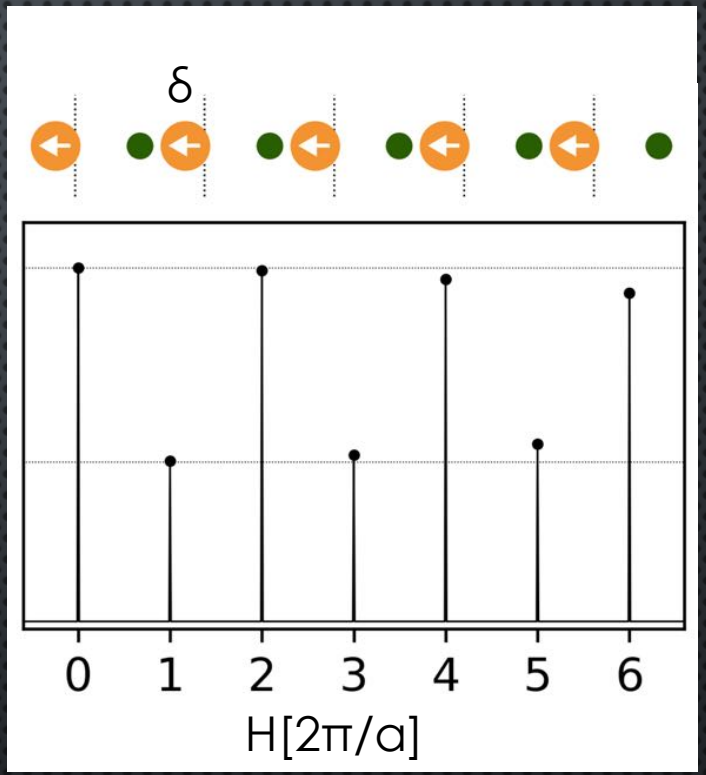
CDW v.s. IUC as seen from XRD I(q)



Fully Symmetric (undistorted)

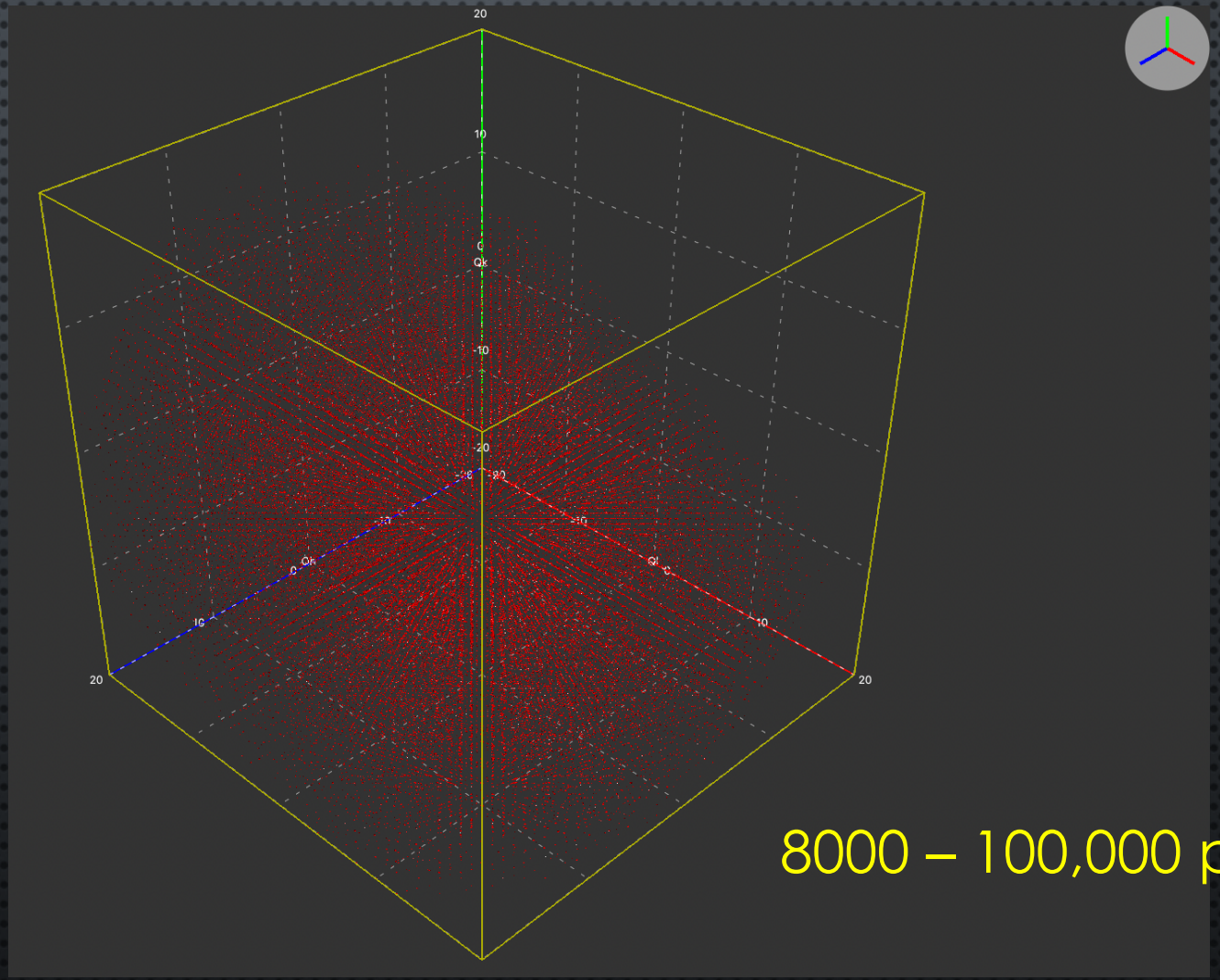


CDW with period 2a: superlattice peaks require multiple BZ's due to form factors



Intra Unitcell (IUC) order: Only reflected in the form factors of Bragg peaks, no superlattice peaks

How to Detect Subtle Ordering?



8000 – 100,000 peaks

Searching for special $\vec{Q}$'s

Searching for that special piece?

Searching for that special piece?

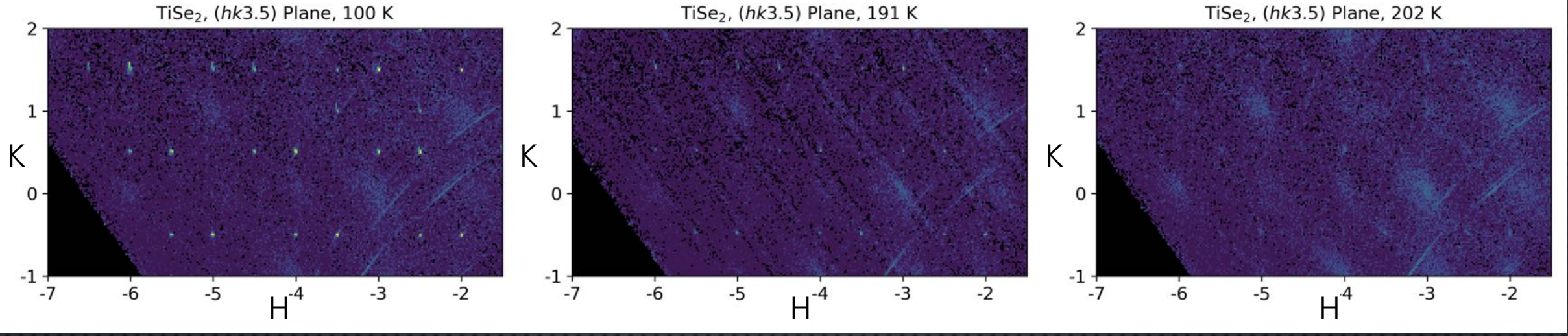
What if we don't know the sorting criteria?

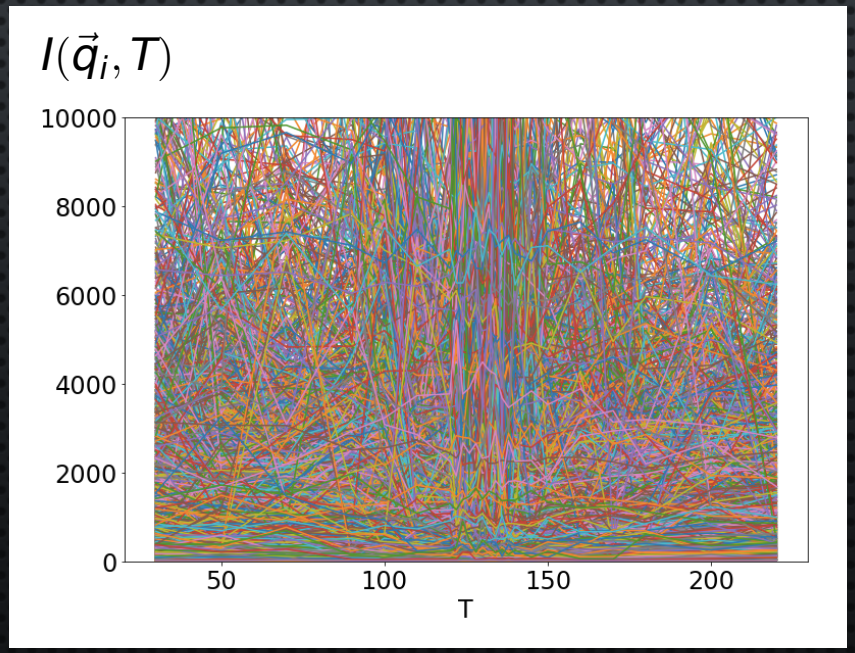
- $I(\vec{q};T) \propto |\widetilde{\rho}(\vec{q};T)|^2$: Fourier amplitude of density

$$\delta F = \delta E - T\delta S$$

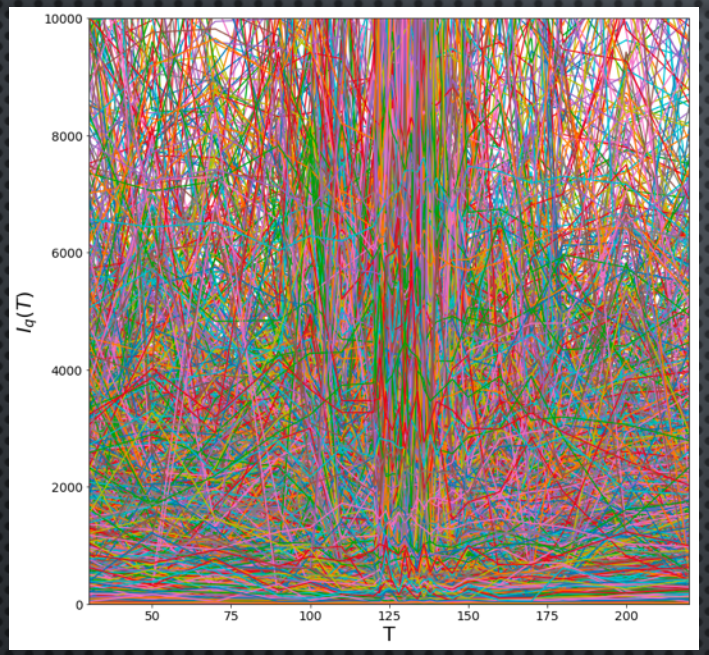
- Expect different clusters of *T*- series among "population" of $\vec{q}$'s

Temperature Evolution of Intensities $I(\vec{q};T)$

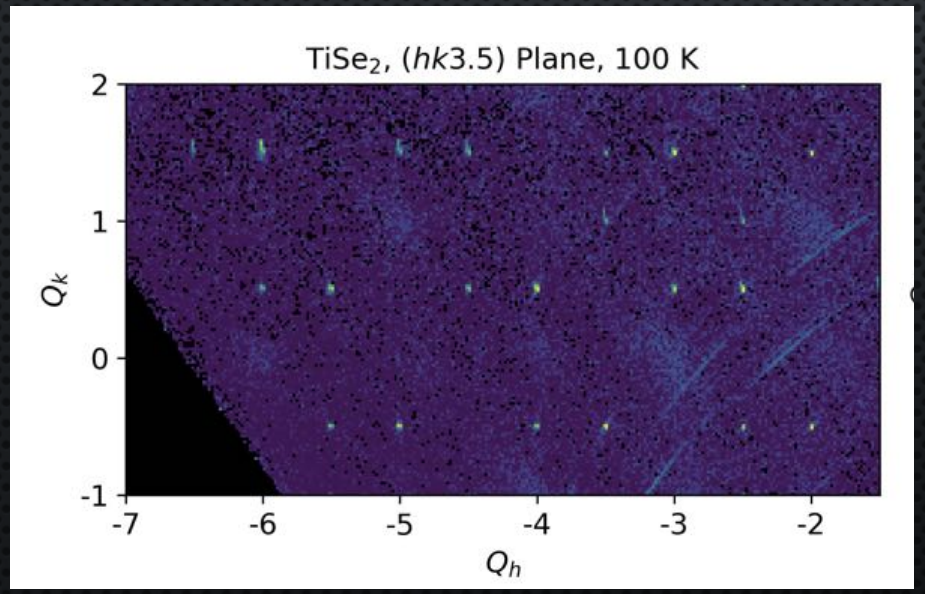




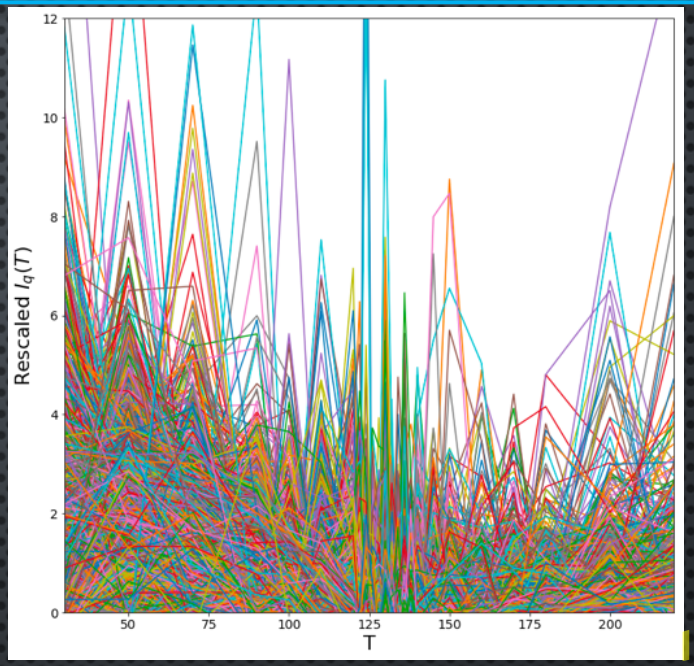
Preprocessing

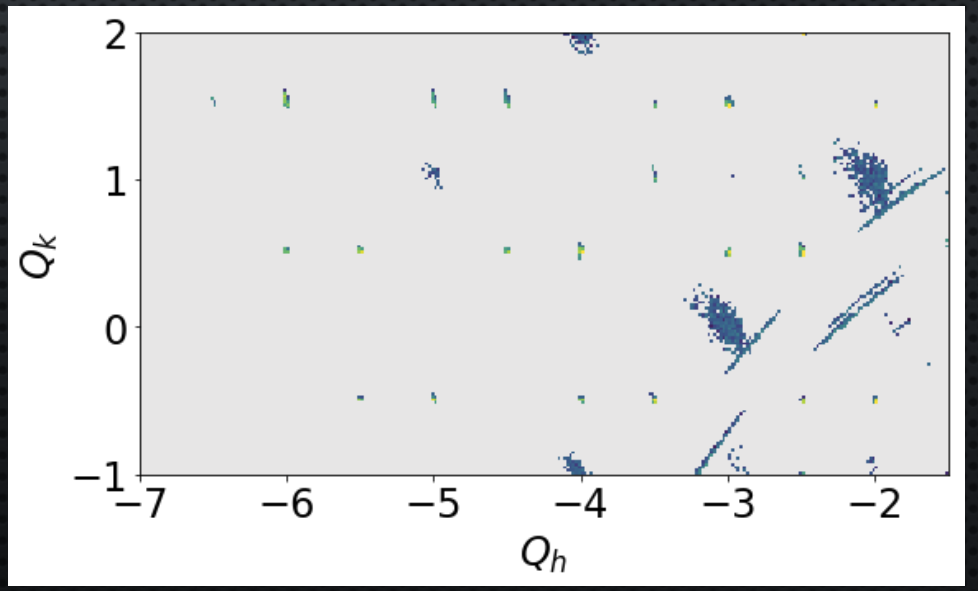


Threshold

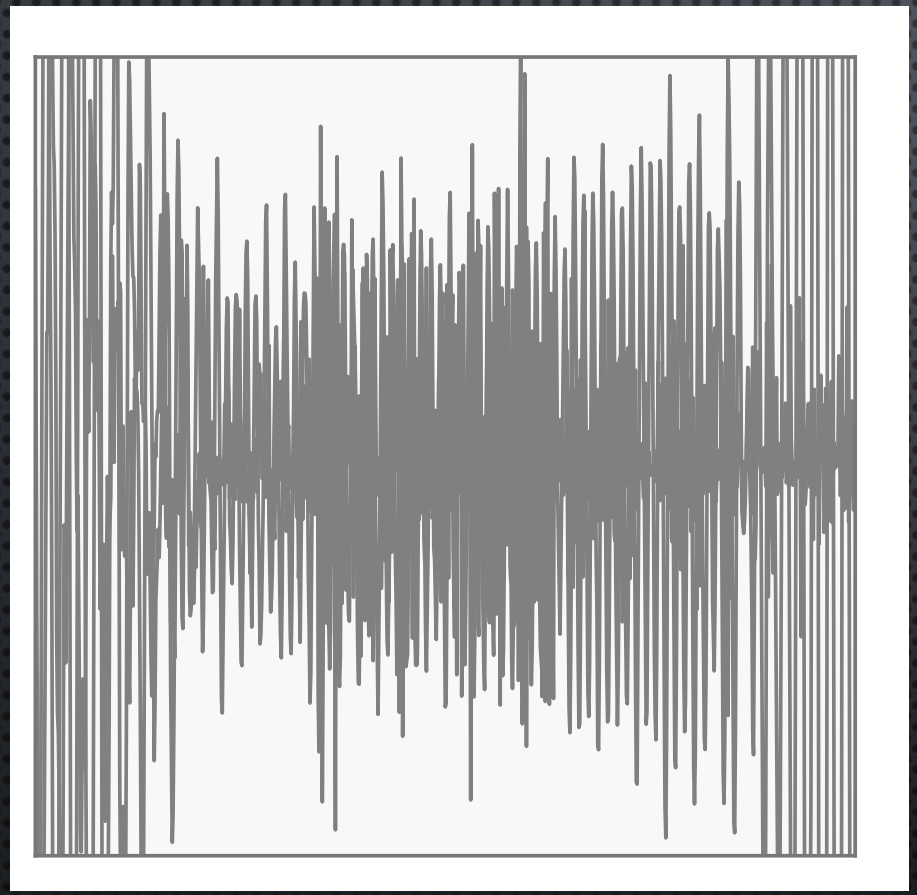


and Rescale

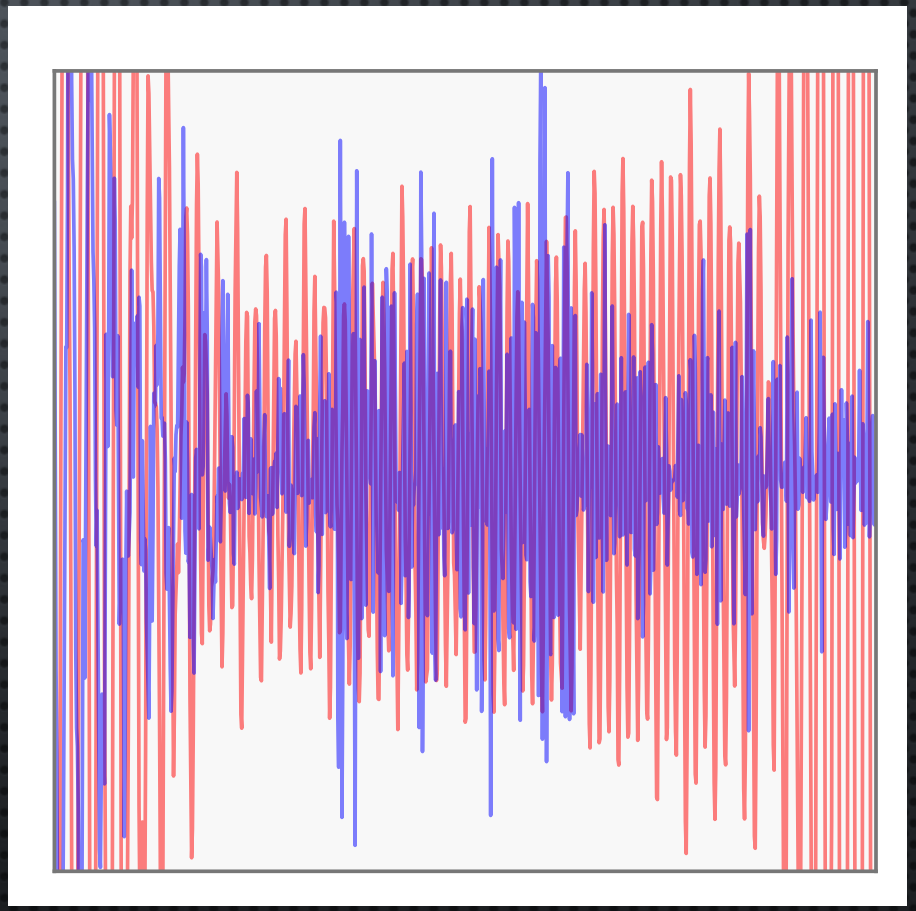




Speaker Verification



time



time

Gaussian Mixture Model

- Preprocessed temperature series for each $\vec{q}_i$:

$$\tilde{\mathbf{I}}(\vec{q}_i) \equiv \{\tilde{I}(\vec{q}_i, T_j); j = 1, \cdots, d^T\}$$

- Conjecture K distinct multivariate normal dist.

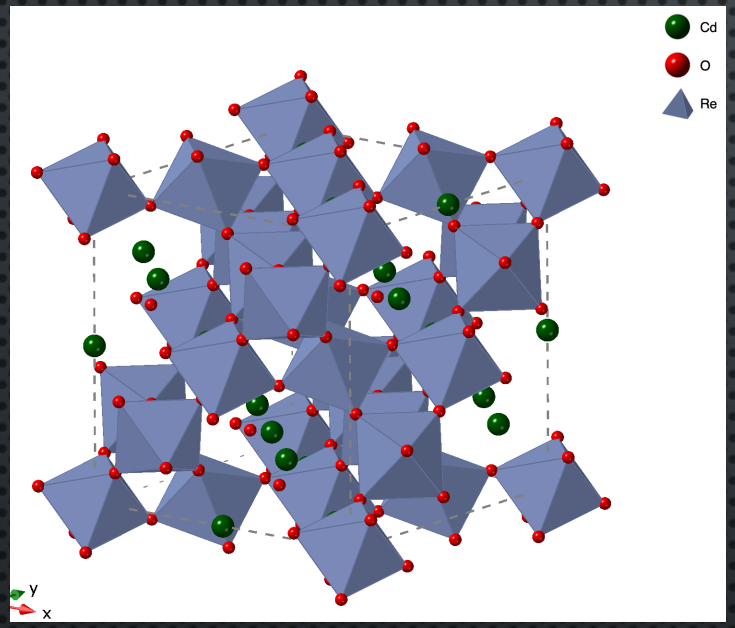
$$\mathcal{N}\left(\tilde{\mathbf{I}}(\vec{q}_i)|\mathbf{m}_k, \mathbf{s}_k\right) \equiv \frac{1}{(2\pi)^{d_T/2}} \frac{1}{\sqrt{\det \mathbf{s}_k}} e^{-\frac{1}{2}\left[\tilde{\mathbf{I}}(\vec{q}_i - \mathbf{m}_k)^\dagger \mathbf{s}_k^{-1} (\tilde{\mathbf{I}}(\vec{q}_i) - \mu_k)\right]}.$$

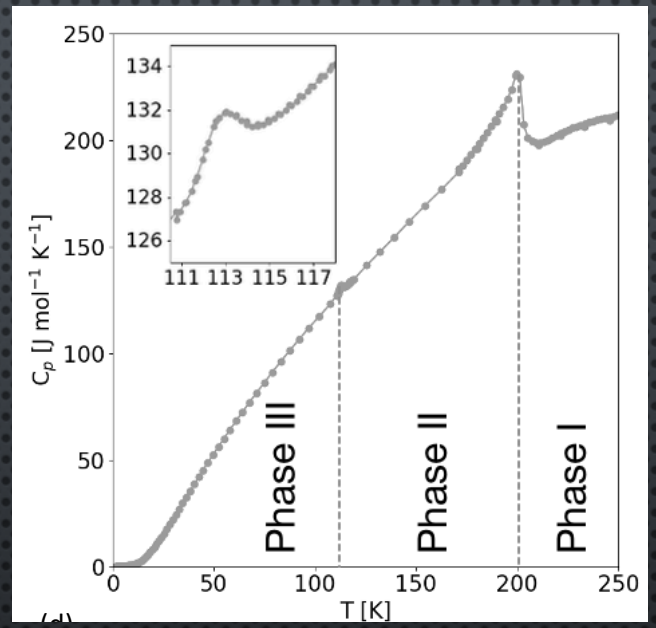
$$\log p\left(\{\tilde{\mathbf{I}}(\vec{q}_i)\}|\pi, \mathbf{m}, \mathbf{s}\right) = \sum_{\vec{q}_i} \log \left[\sum_{k=1}^{K} \pi_k \mathcal{N}\left(\tilde{\mathbf{I}}(\vec{q}_i)|\mathbf{m}_k, \mathbf{s}_k\right)\right]$$

- The cluster assignement for the seires $\tilde{\mathbf{I}}(\vec{q}_i)$

$$w_i^k = \frac{\pi_k \mathcal{N}\left(\tilde{\mathbf{I}}(\vec{q}_i)|\mathbf{m}_k, \mathbf{s}_k\right)}{\sum_k \pi_k \mathcal{N}\left(\tilde{\mathbf{I}}(\vec{q}_i)|\mathbf{m}_k, \mathbf{s}_k\right)}$$

Order Parameter of Phase II in Pyrochlore $Cd_2Re_2O_7$





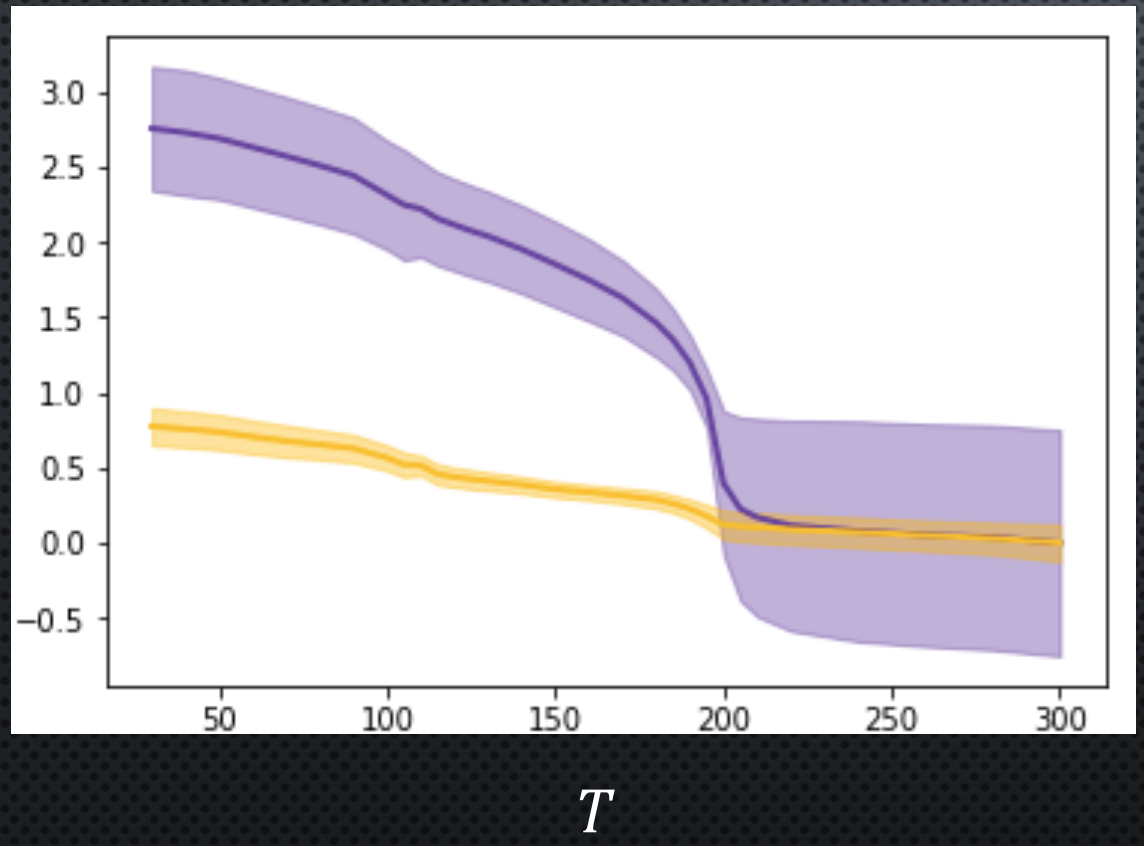
A parity-breaking electronic nematic phase transition in the spin-orbit coupled metal $Cd_2Re_2O_7$

J. W. Harter,[1,2] Z. Y. Zhao,[3,4] J.-Q. Yan,[3,5] D. G. Mandrus,[3,5] D. Hsieh[1,2]*

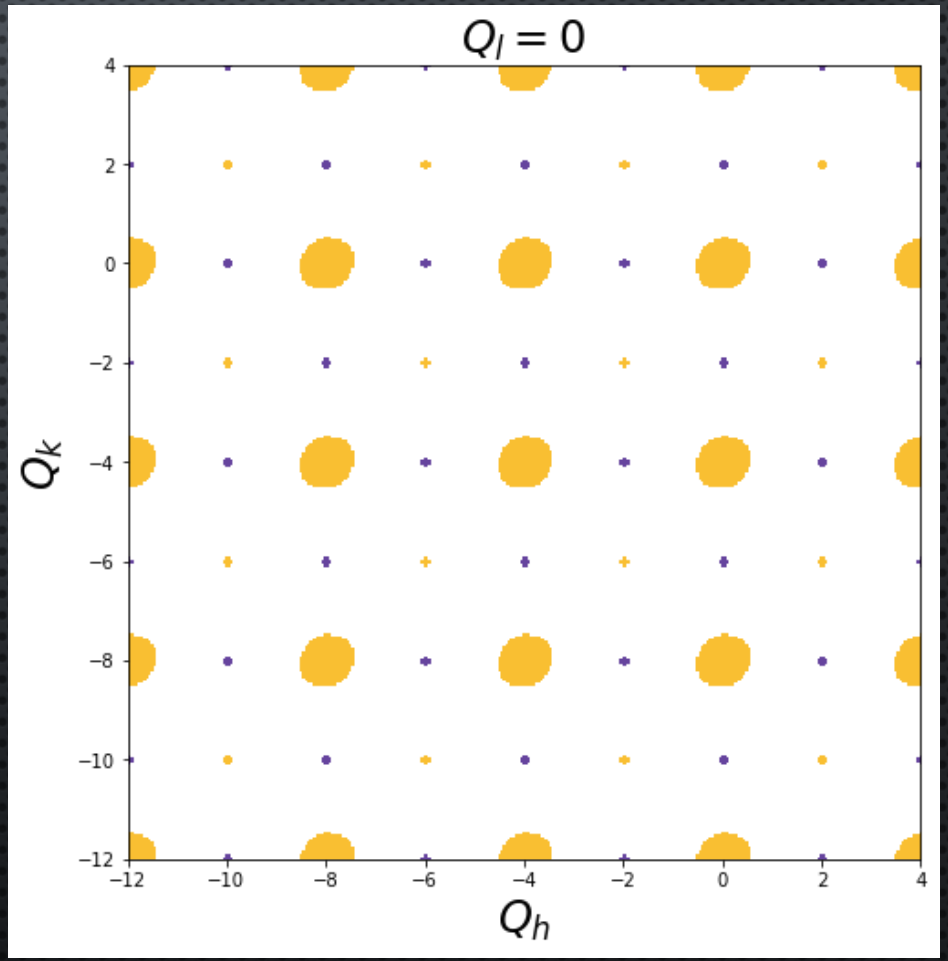
Harter *et al.*, *Science* **356**, 295–299 (2017) 21 April 2017

- First known Pyrochlore SC.
- Primary Order Parameter of Phase II :
  2 dim E_U vs 1 dim $_{T2U}$

Clustering and Interpretation

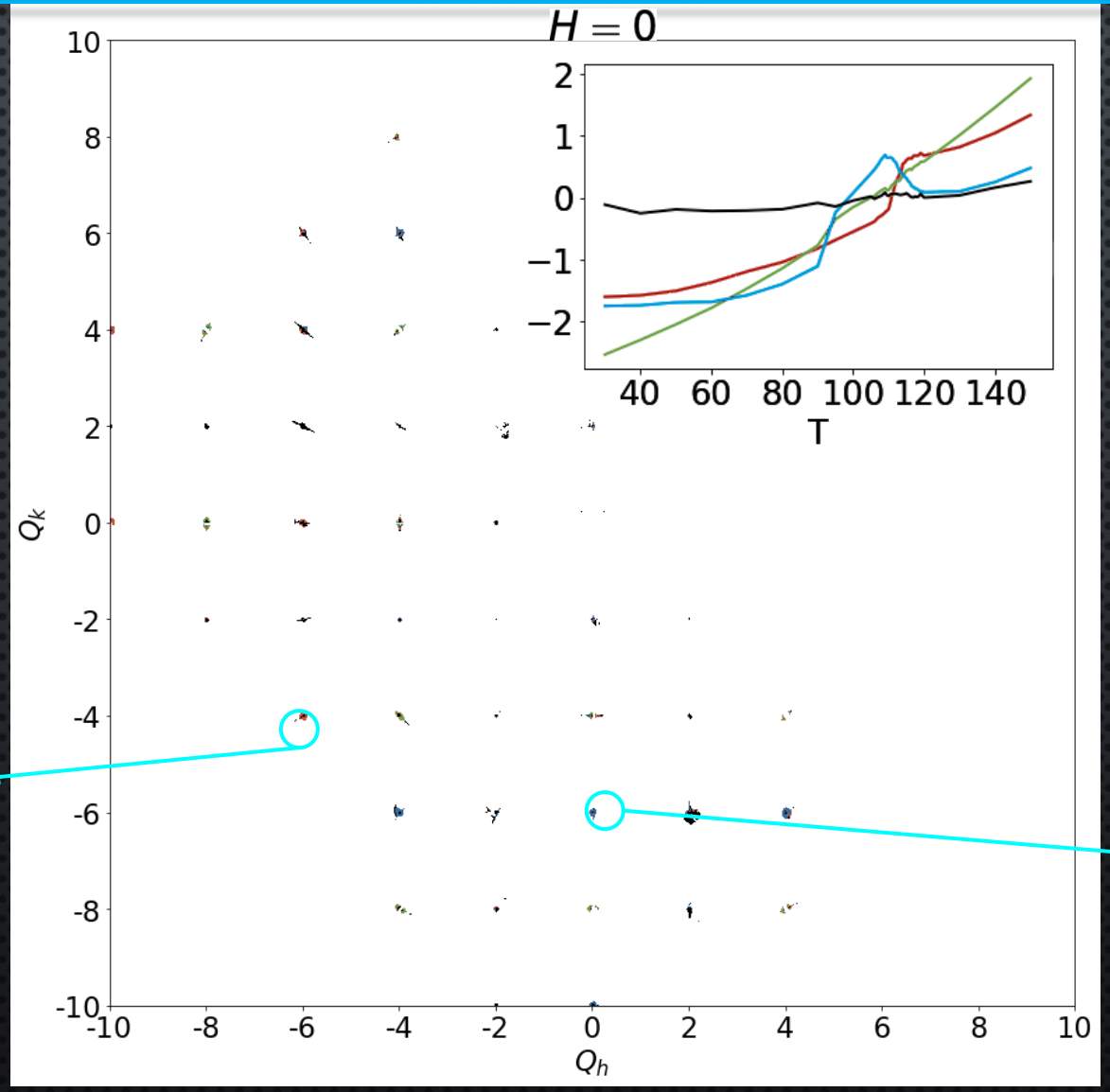


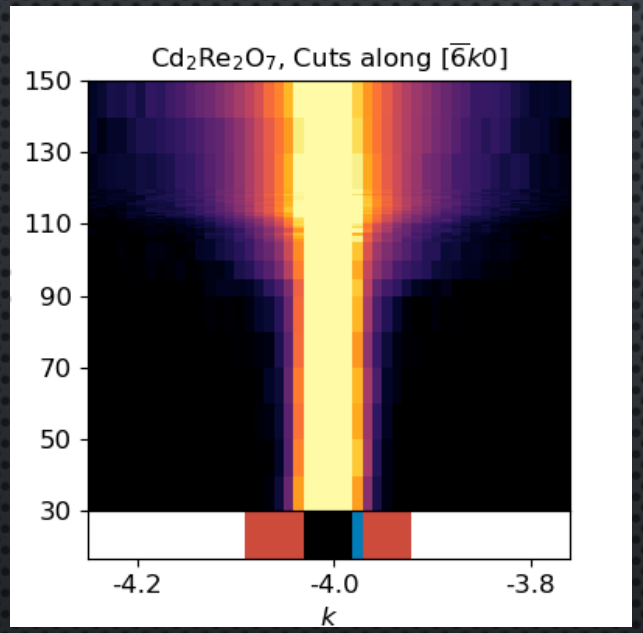
15,000 BZ's (8TB) in 15 min

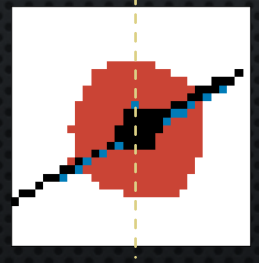


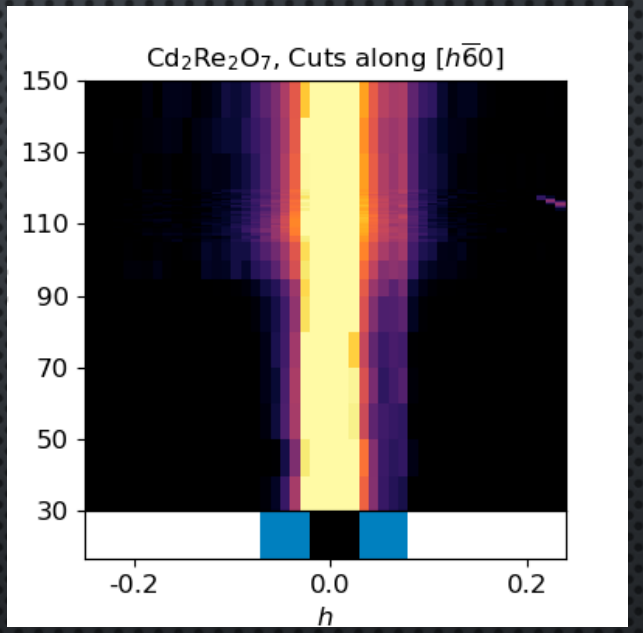
Purple: one of (H,K,L) 4n+2 , indicate I-4m2 or I-4

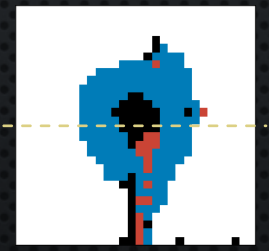
Looking inside Bragg peaks











CCNN (Correlation CNN) for Quanum Images

CORRELATOR CONVOLUTIONAL NEURAL NETWORKS: AN INTERPRETABLE ARCHITECTURE FOR IMAGE-LIKE QUANTUM MATTER DATA

(arXiv:2011.03474)

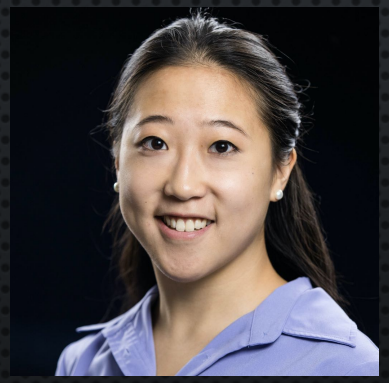
Cole Miles

R. Wu
(Cornell, CS)

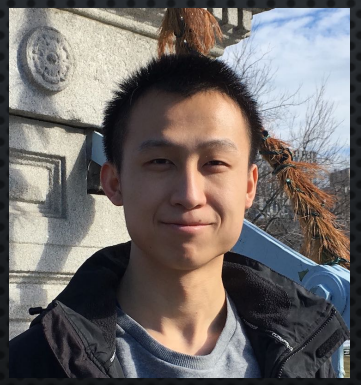
K. Winberger
(Cornell, CS)

A. Bohrdt
(Harvard)

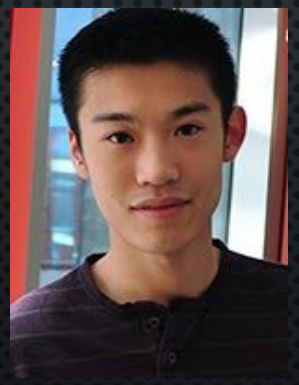
E. Demler
(Harvard)



C. Chiu
(Princeton)



M. Xu
(Harvard)



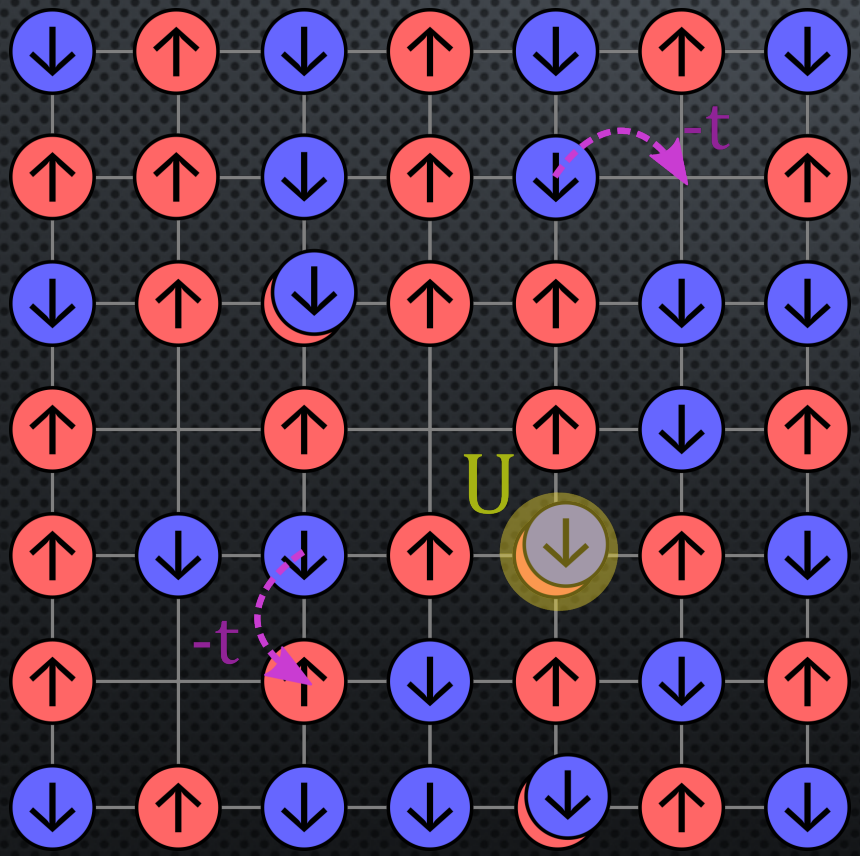
G. Ji
(Harvard)

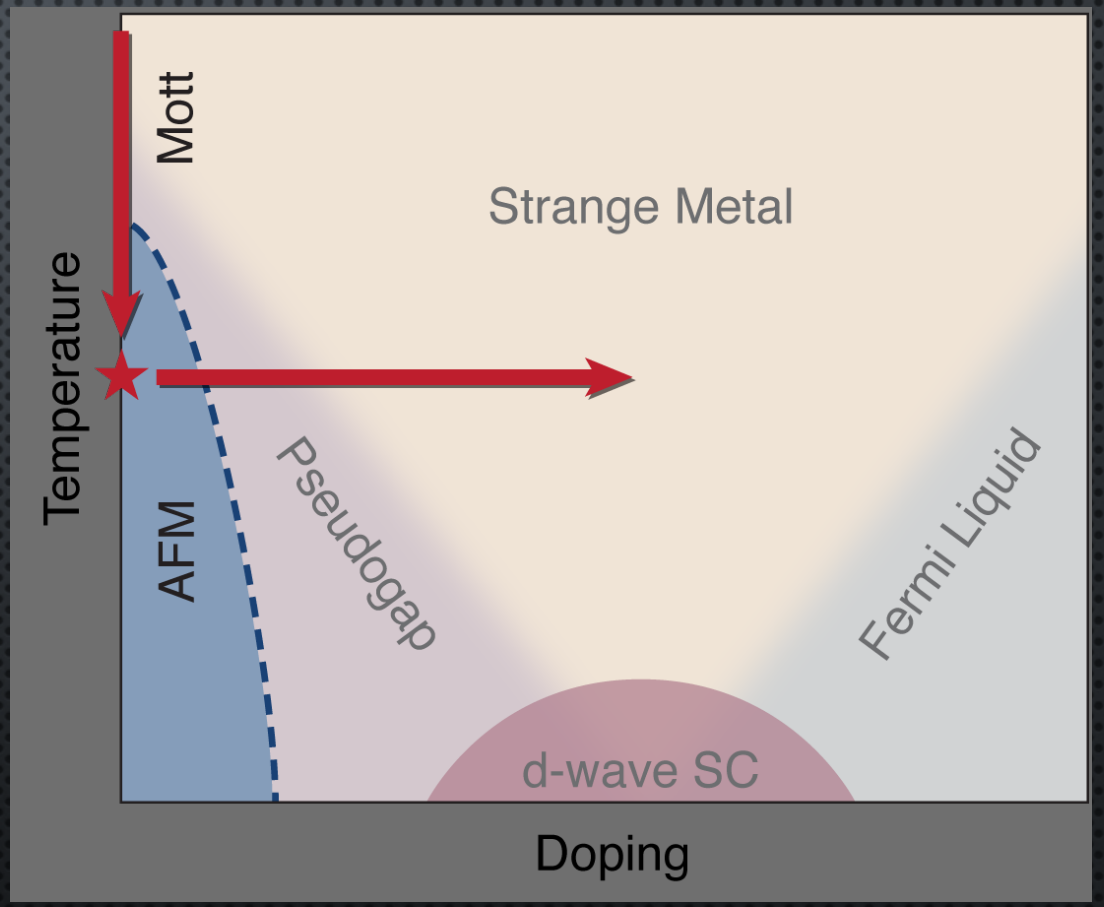
M. Greiner
(Harvard)

DOE CSGF

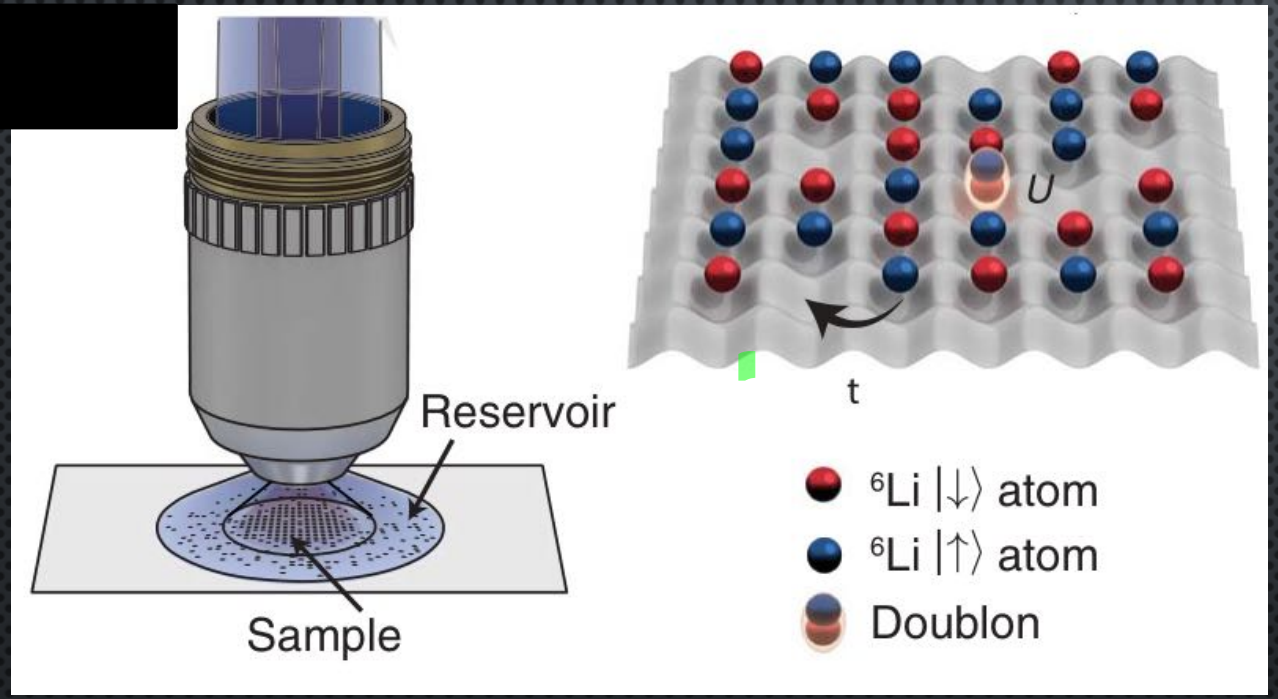
FERMI-HUBBARD MODEL

$$H = -t \sum_{\langle i,j \rangle,\sigma} c^{\dagger}_{i,\sigma} c_{j,\sigma} + U \sum_{i} n_{i\uparrow} n_{i\downarrow}$$





QUANTUM GAS MICROSCOPY

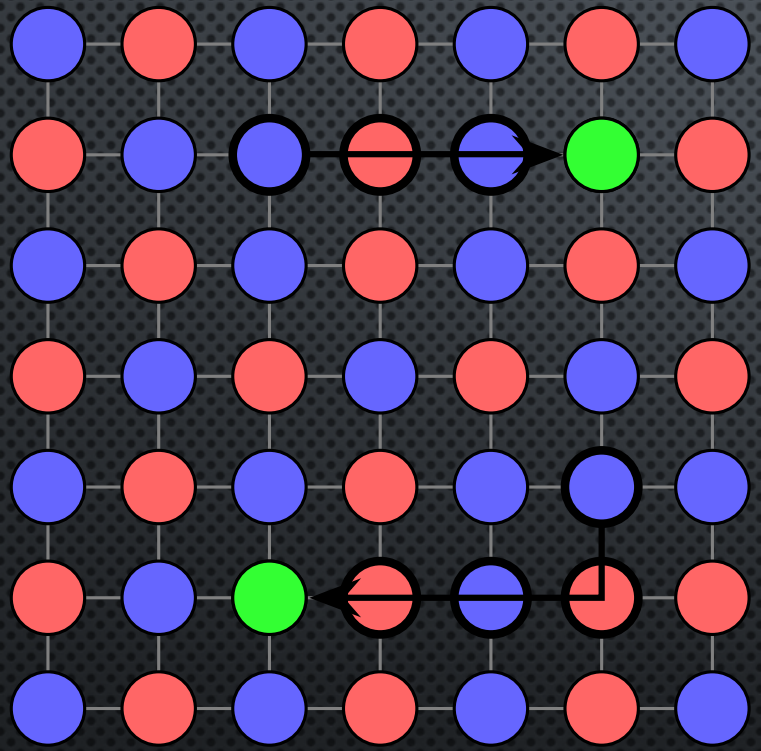


Sample

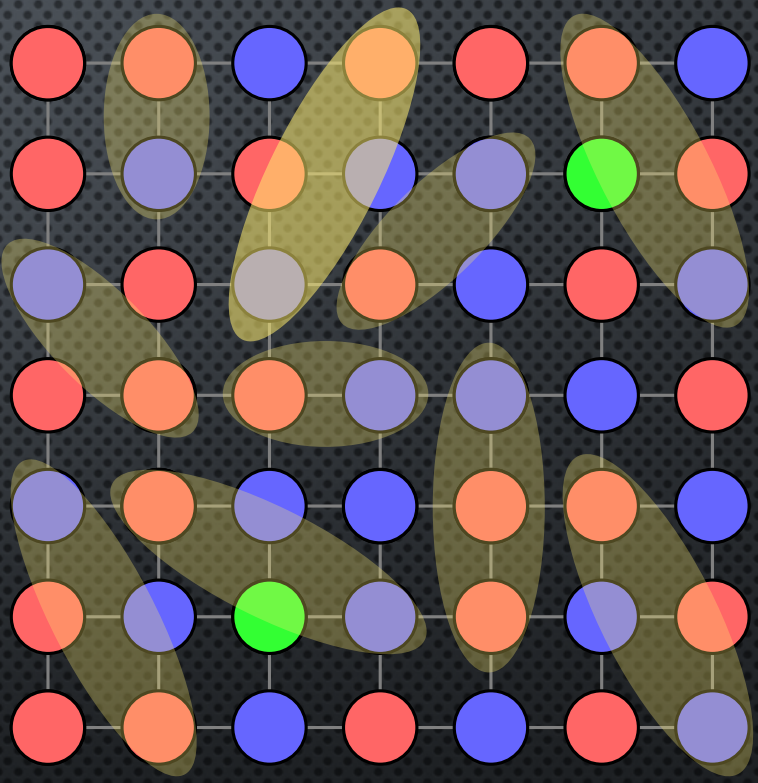
Real-Space Occupation Snapshots

Taken from Chiu, C. S. et al., *Science* **365** (6450), 2019

Which ansatz better describes data?



Geometric String

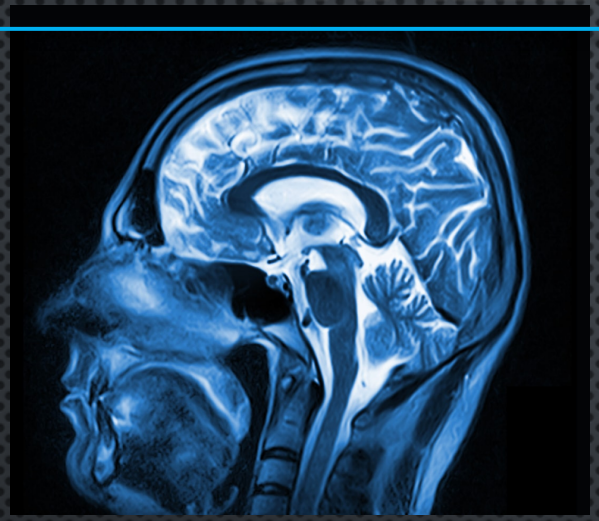


π-flux theory with singlet bonds

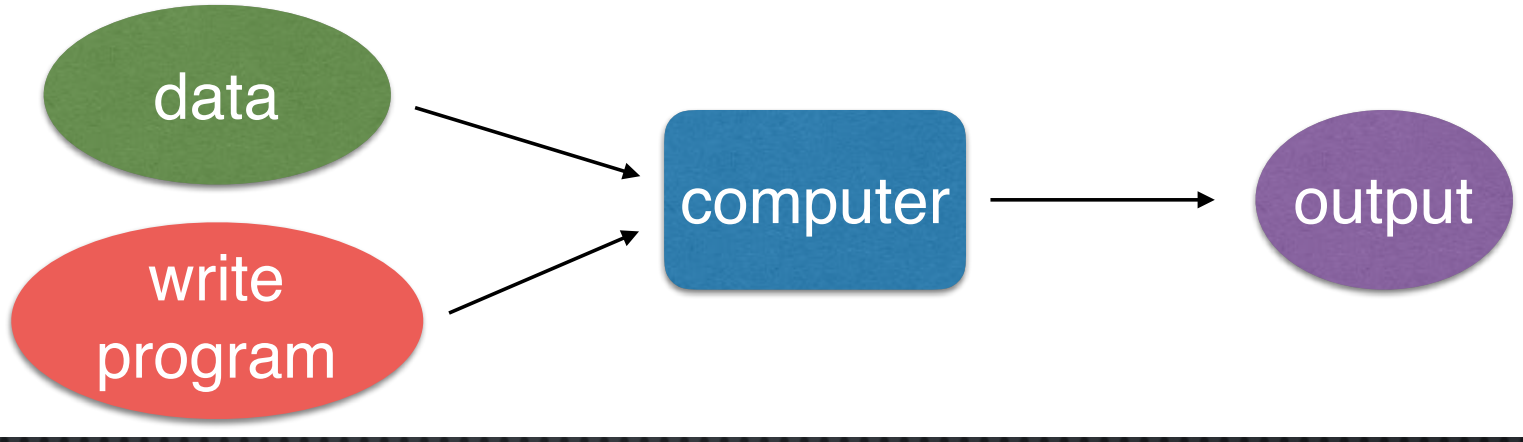
Conventional Approach

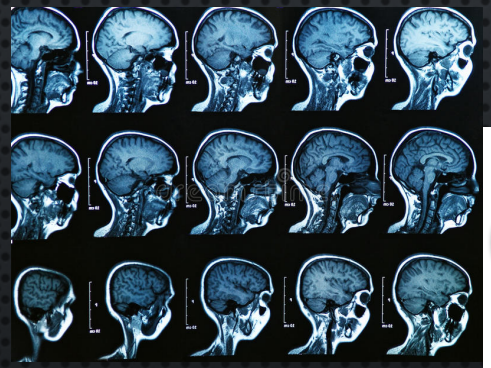
1. Magnetization
2. Spin-spin (2site) correlation

CHALLENGES OF COMPLEXITY: ALTZHEIMER

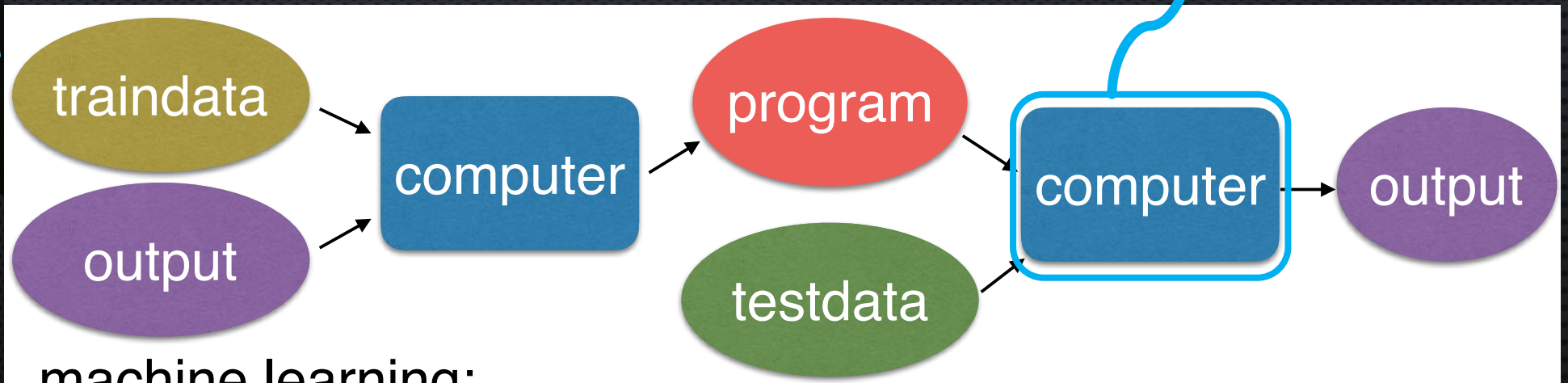


traditional cs:





Learned Machine



machine learning:

Off-the-shelf CNN

1. Over-parametrization
2. A black-box

CONVOLUTIONAL NEURAL NETWORKS (CNNS)

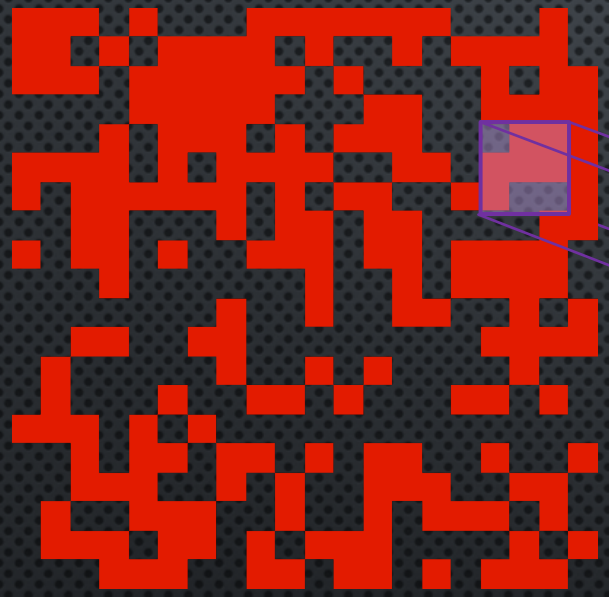
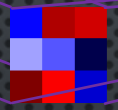


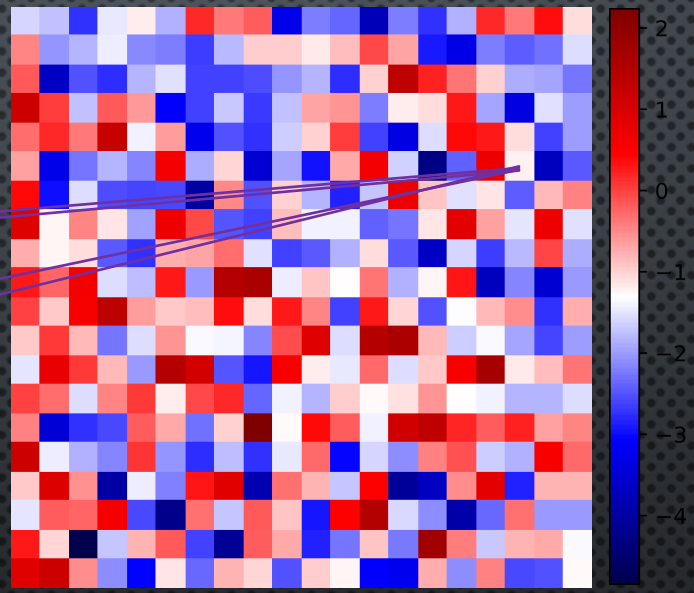
Image data

$S(x)$



Convolution al filter *(learned)*

$f(a)$



Convolutional map

$$C(x) = \sum_a f(a)S(x+a)$$

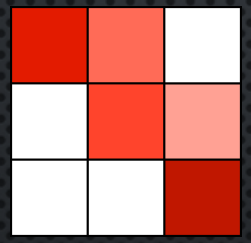
Nonlinearity

$\hat{y}$

Class Prediction

POWERS OF A CONVOLUTION AS NON-LINEARITY?

Example Filter

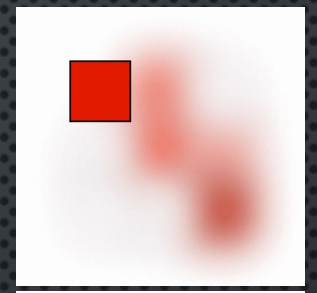


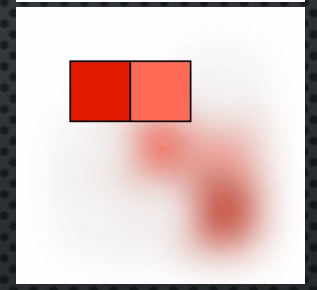
1ST ORDER $C_x = f_a S_{x+a}$

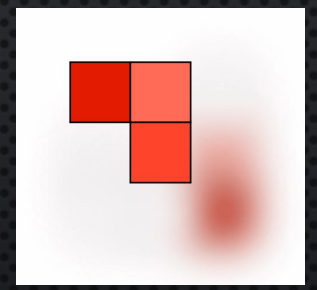
2nd order $C_x^2 = (f_a S_{x+a})^2 = f_a f_b S_{x+a} S_{x+b}$

3rd order $C_x^3 = (f_a S_{x+a})^3 = f_a f_b f_c S_{x+a} S_{x+b} S_{x+c}$

(sums over a, b, c implied)

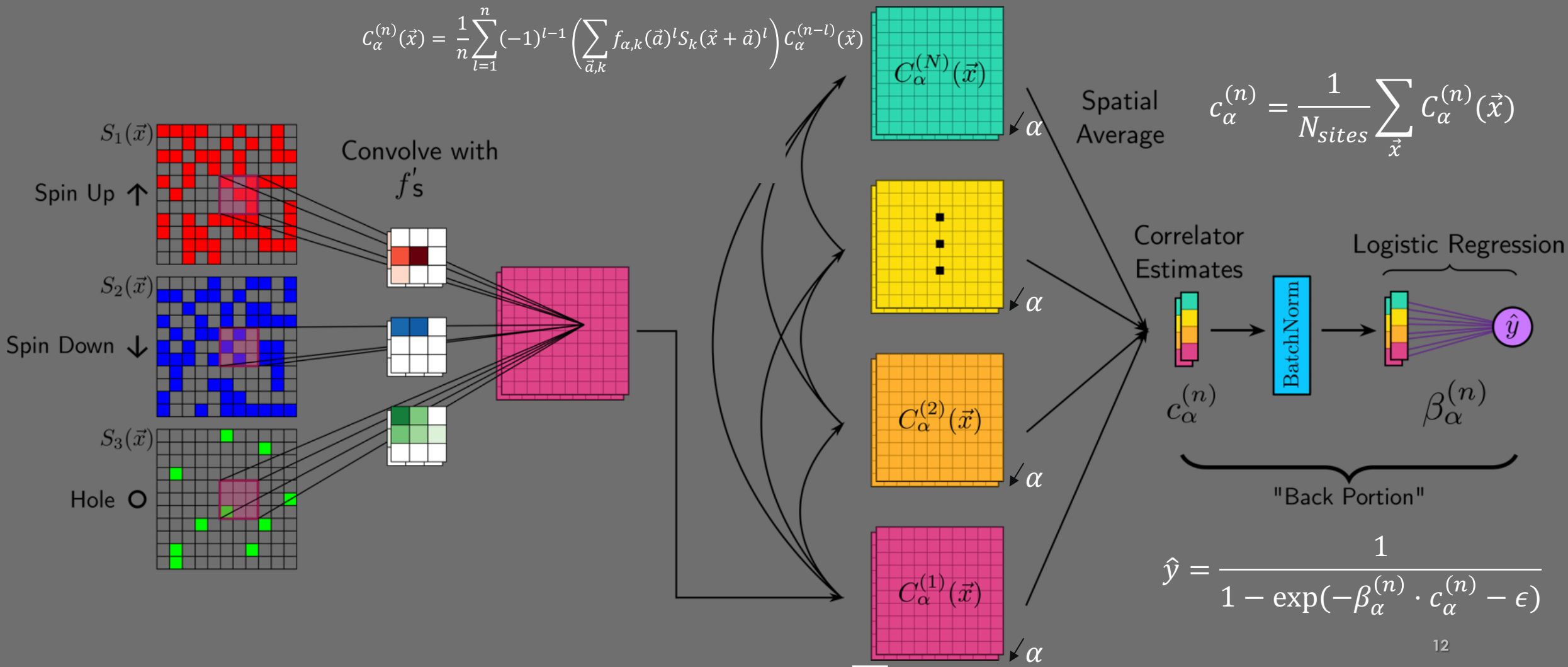






CORRELATION CNN(CCNN)

$$C_\alpha^{(n)}(\vec{x}) = \frac{1}{n}\sum_{l=1}^{n}(-1)^{l-1}\left(\sum_{\vec{a},k} f_{\alpha,k}(\vec{a})^l S_k(\vec{x}+\vec{a})^l\right) C_\alpha^{(n-l)}(\vec{x})$$



$$C_\alpha^{(1)}(\vec{x}) = \sum_{\vec{a},k} f_{\alpha,k}(\vec{a})\, S_k(\vec{x}+\vec{a})$$

Cascade Construction of Correlation Filters

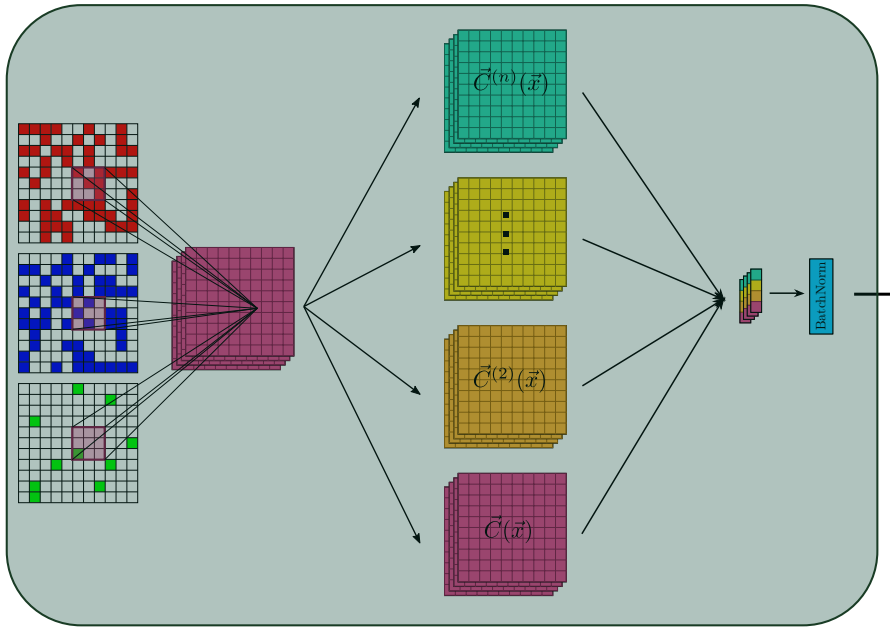
$$\vec{C} \equiv \sum_{\vec{a}} \mathbf{f}(\vec{a}) \cdot \vec{S}(\vec{x} + \vec{a})$$

$$\vec{C}^{(n)} = \frac{1}{n!} \sum_{\vec{a}_1 \neq \ldots \neq \vec{a}_n} \prod_{j=1}^{n} \mathbf{f}(\vec{a}_j) \cdot \vec{S}(\vec{x} + \vec{a}_j)$$

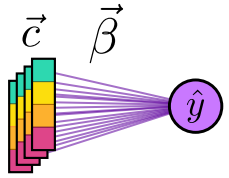
$$\vec{C}(\vec{x})^{(n)} = \frac{1}{n} \sum_{m=1}^{n} (-1)^{m-1} \left(\sum_{\alpha} f_{\alpha}^{m} S(\vec{x} + \vec{a})^{m} \right) \vec{C}^{(n-m)}$$

Regularization Path Analysis

Hold Fixed



Retrain

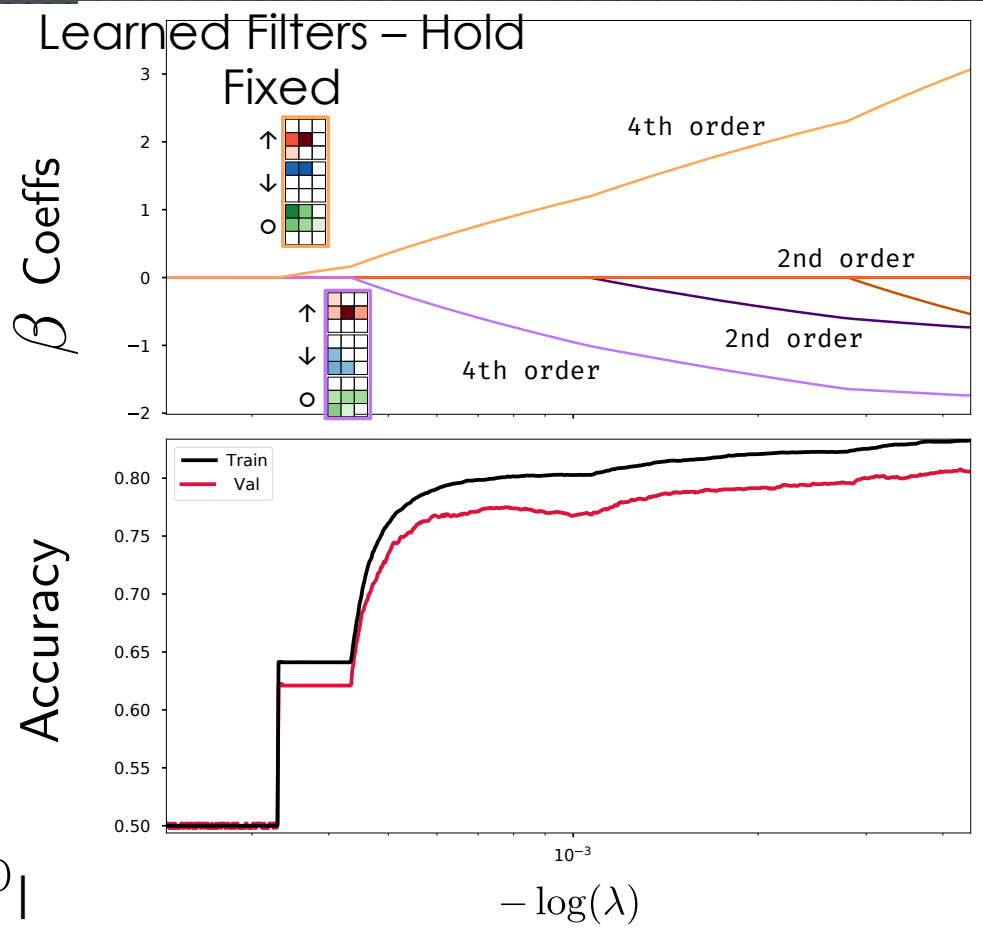


$$\hat{y} = \frac{1}{1+e^{-(\vec{\beta}\cdot\vec{c}+\epsilon)}}$$

Repeatedly re-train logistic regression under new loss, ramping down λ:

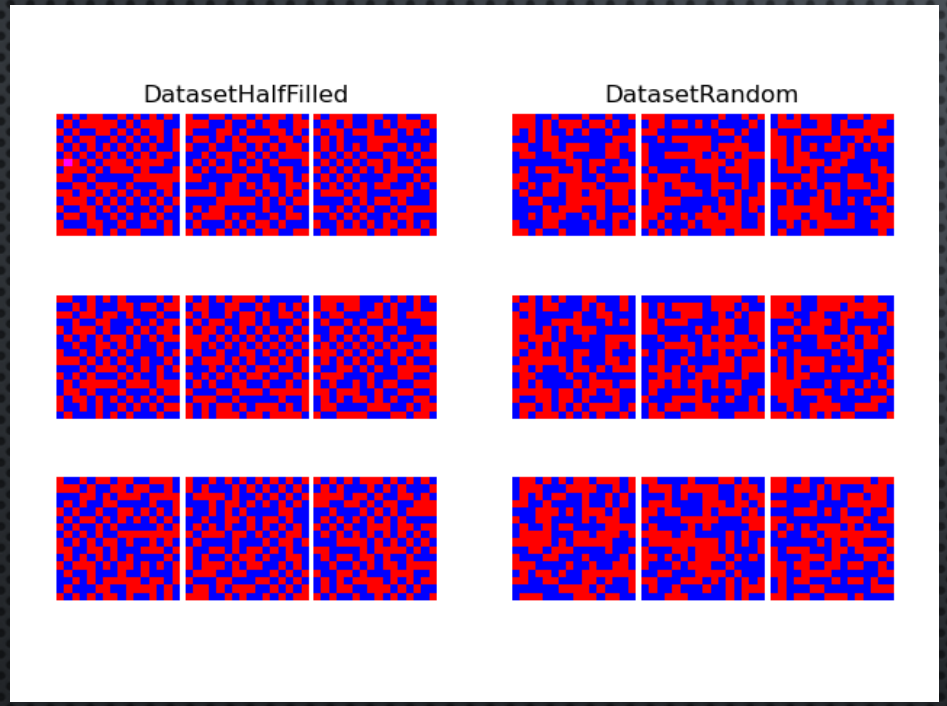
$$L(y,\hat{y}) = -y\ \log\hat{y} - (1-y)\log(1-\hat{y}) + \lambda\sum_{\alpha,n}|\beta_\alpha^{(n)}|$$

Learned Filters – Hold Fixed

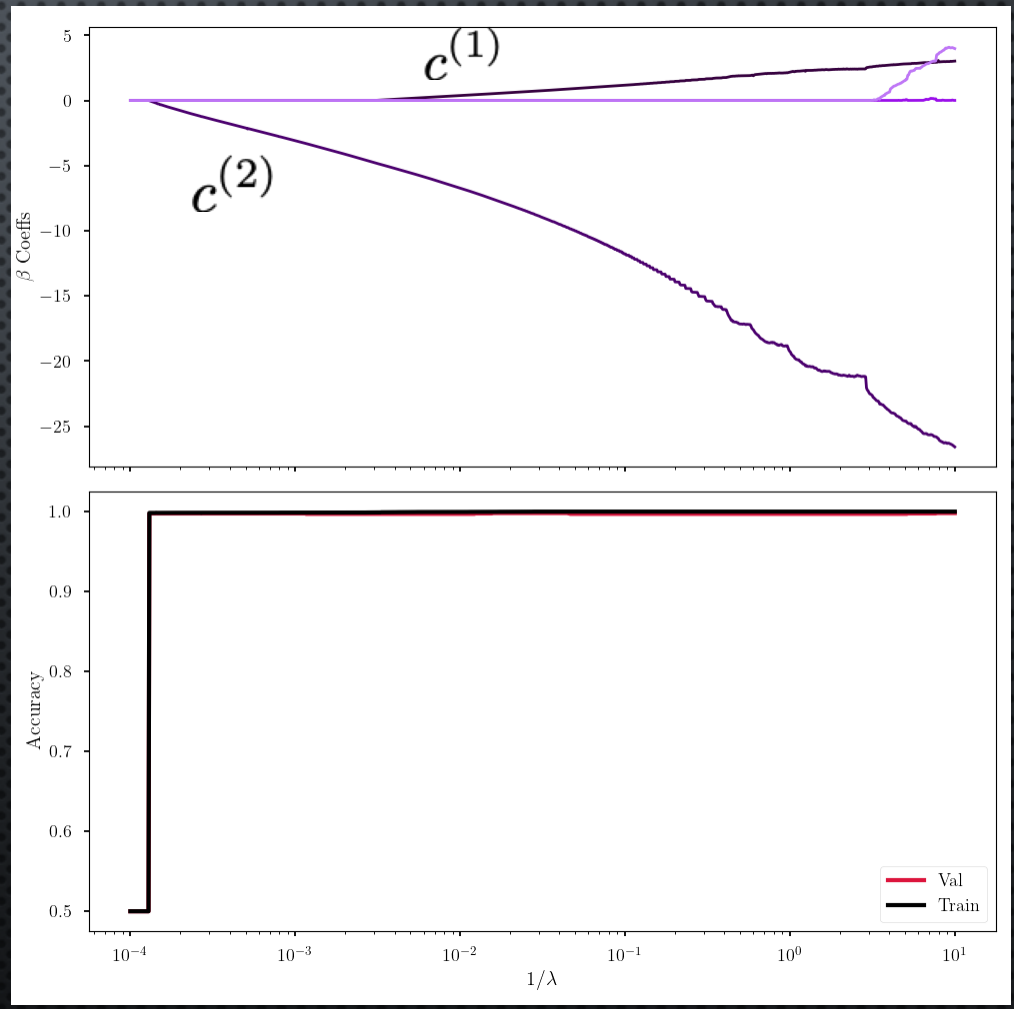


Fluctuating AFM vs Random Image

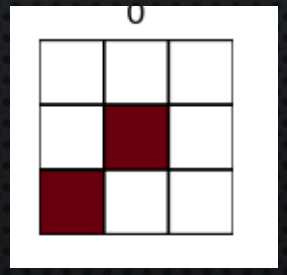
Labeled Training Data

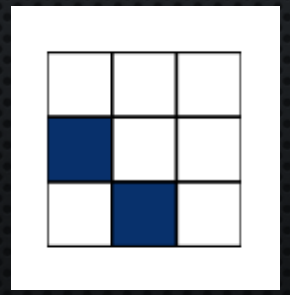


Regularization Path Analysis

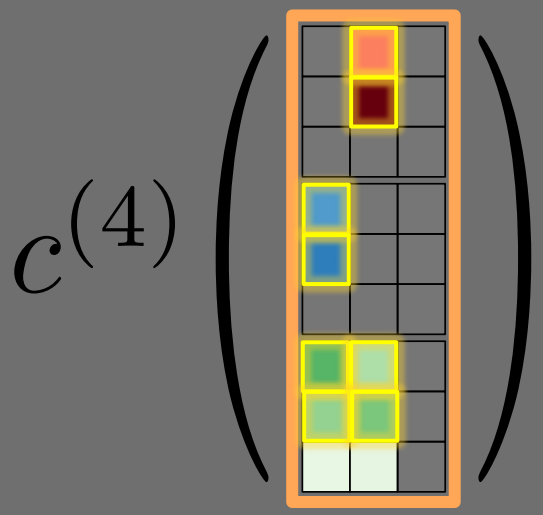


Learned Filters



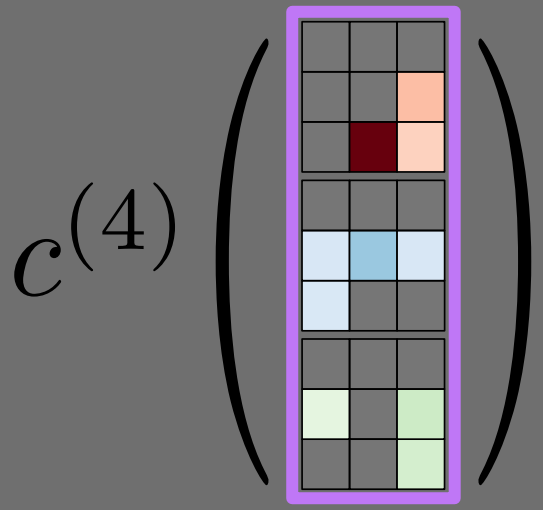


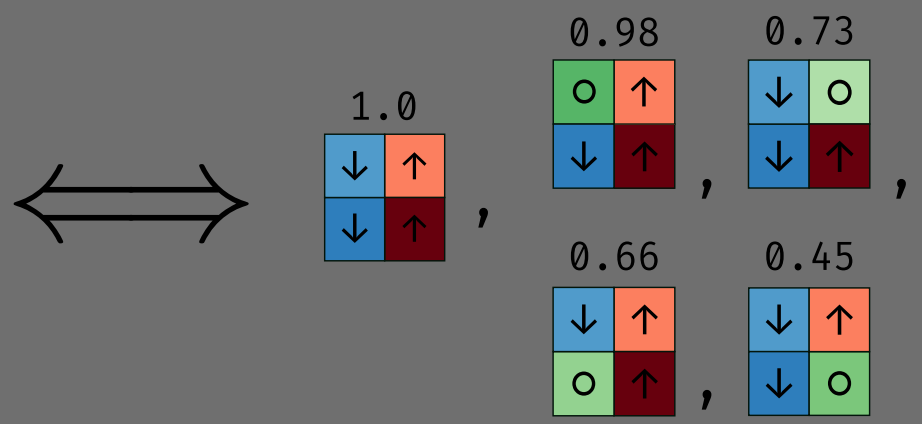
Learned Correlations

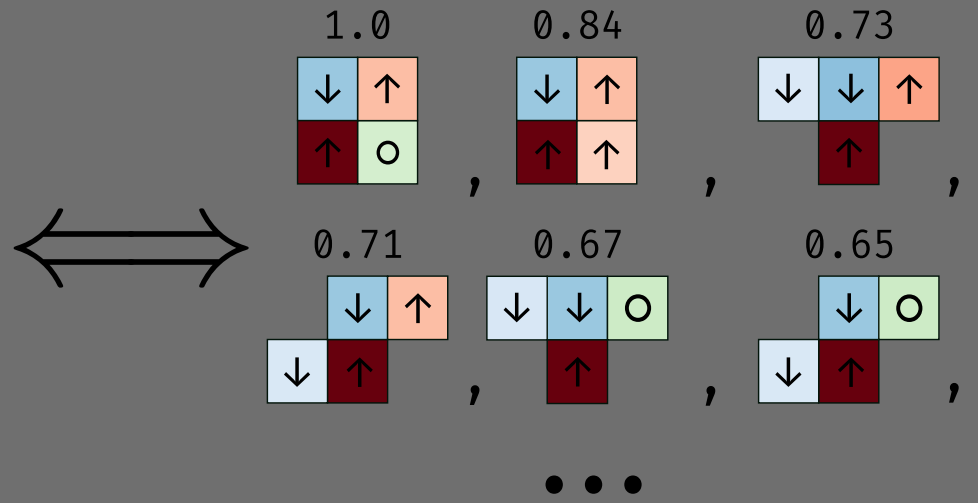


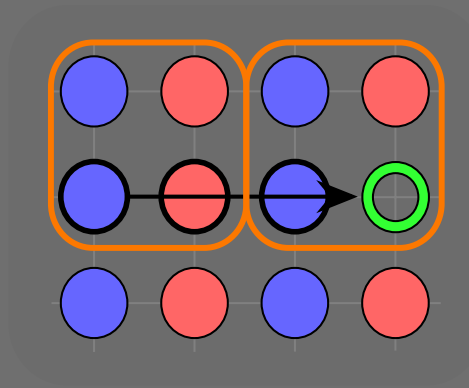
Geometric String

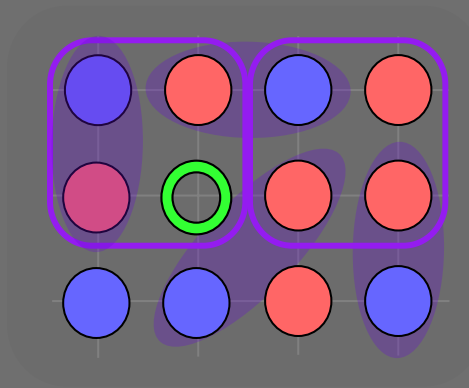
π-flux











Harnessing Data Revolution in Quantum Matter

- X-TEC (XRD TEmperature series Clustering): Use the fundamental role of temperature to discover orders and fluctuations in reciprocal space. arXiv:2008.03275

- CCNN (Correlation Convolution Neural Network): Cascaded correlation kernel detects crucial correlation functions in quantum matter imaging. arXiv:2011.03474