Title: Data-centric learning of Quantum Many-body States with Classical Machines

Speakers: Eun-Ah Kim

Collection: Machine Learning for Quantum Many-Body Systems

Date: June 13, 2023 - 1:30 PM

URL: https://pirsa.org/23060049

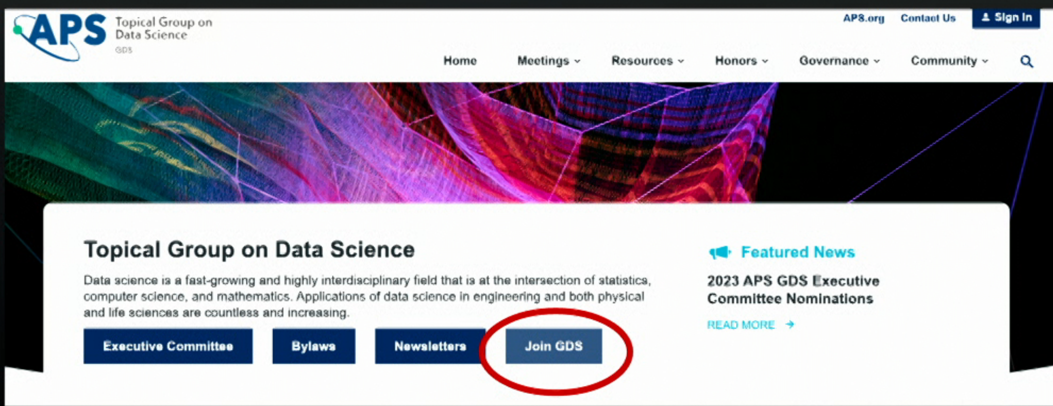
DATA-CENTRIC LEARNING OF QUANTUM MANY-BODY STATES WITH CLASSICAL MACHINES

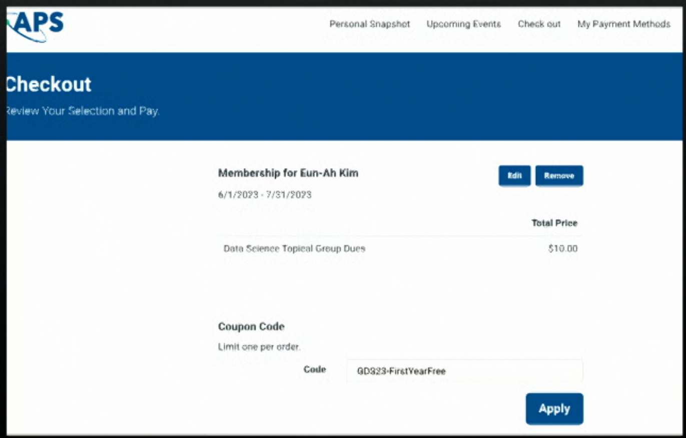
EUN-AH KIM

CORNELL UNIVERSITY, HARVARD RADCLIFFE INSTITUTE, EWHA WOMANS UNIVERSITY

PERIMETER, 6/13/2023

APS Topical Group in Data Science

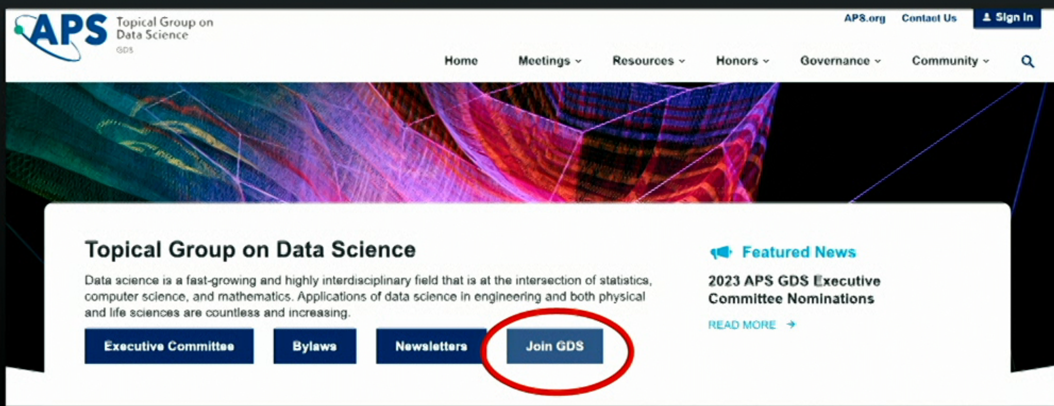


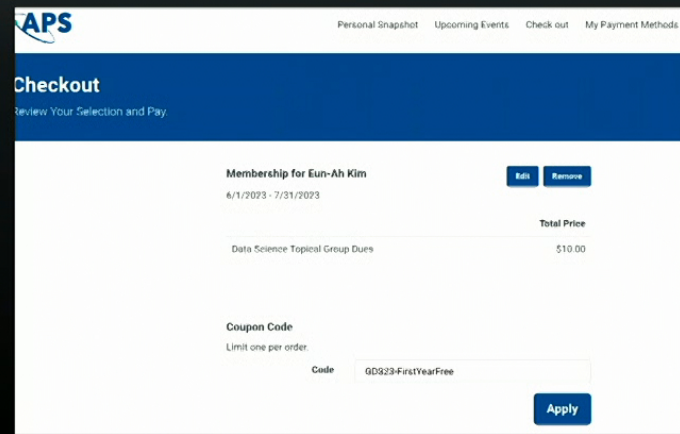


https://engage.aps.org/gds/home

GDS23-FirstYearFree

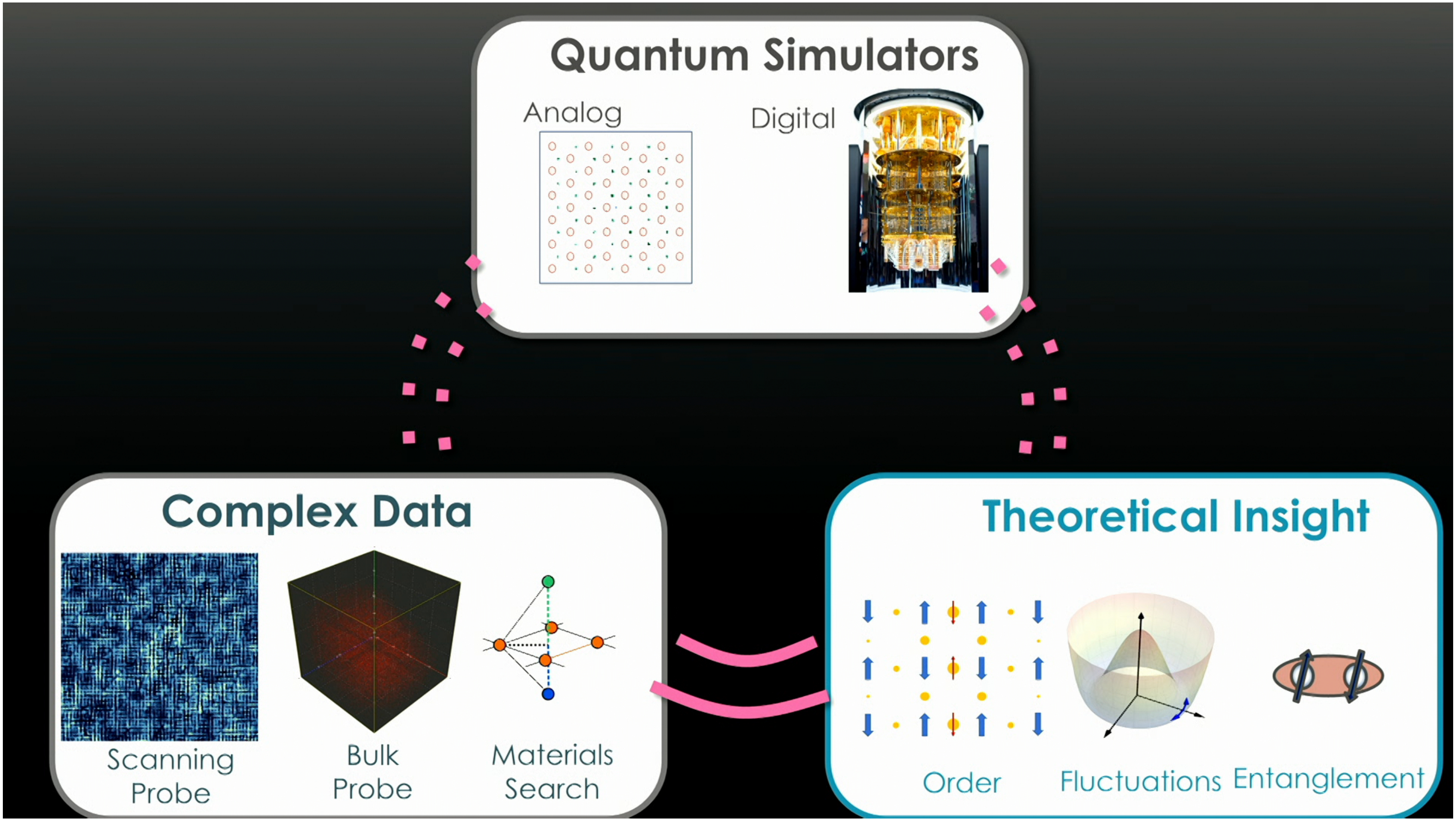
APS Topical Group in Data Science





https://engage.aps.org/gds/home

GDS23-FirstYearFree


Quantum Simulators
Analog
Digital
Complex Data
Scanning Probe
Bulk Probe
Materials Search
Theoretical Insight
Order
Fluctuations
Entanglement

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

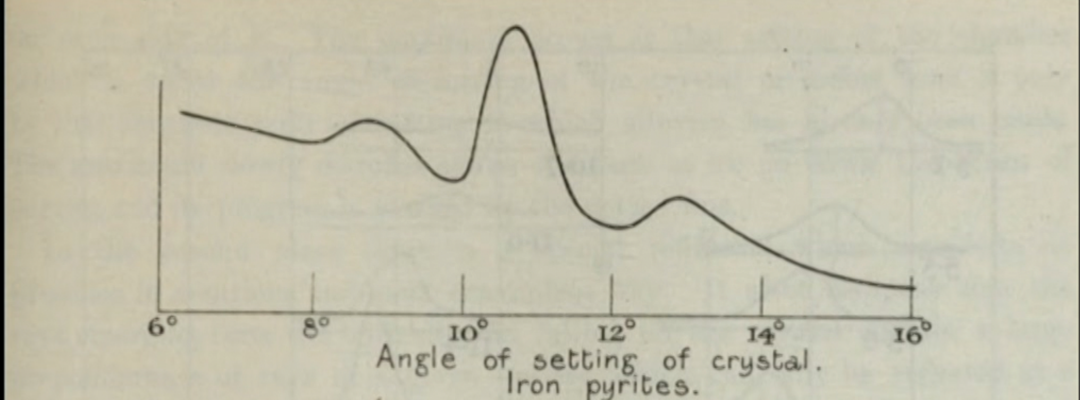
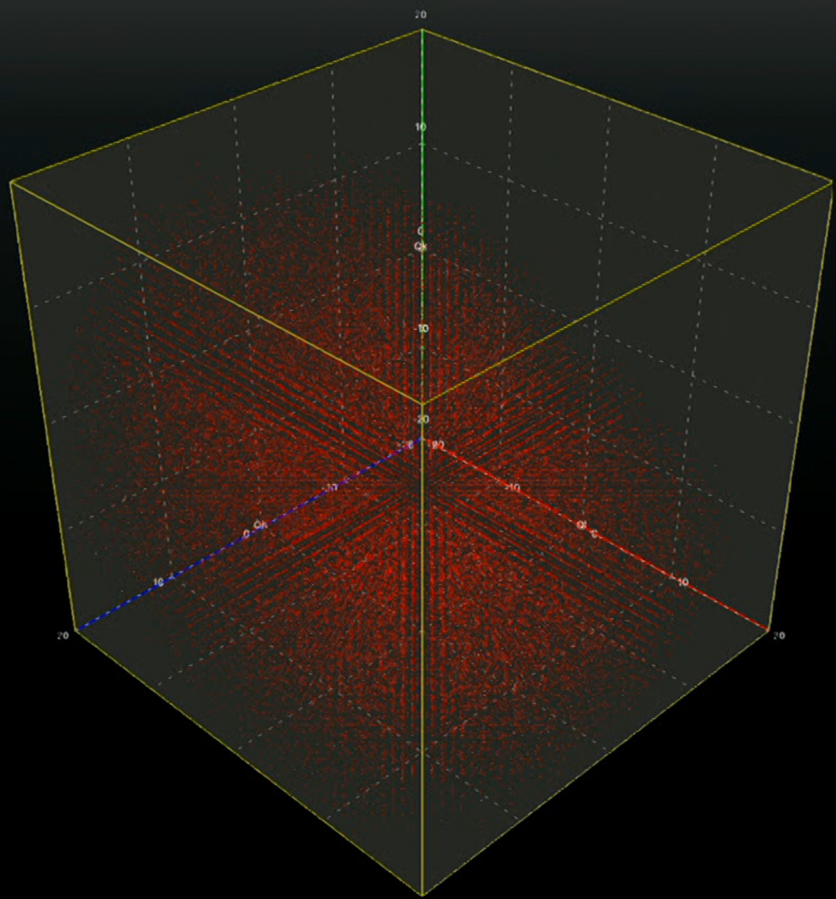
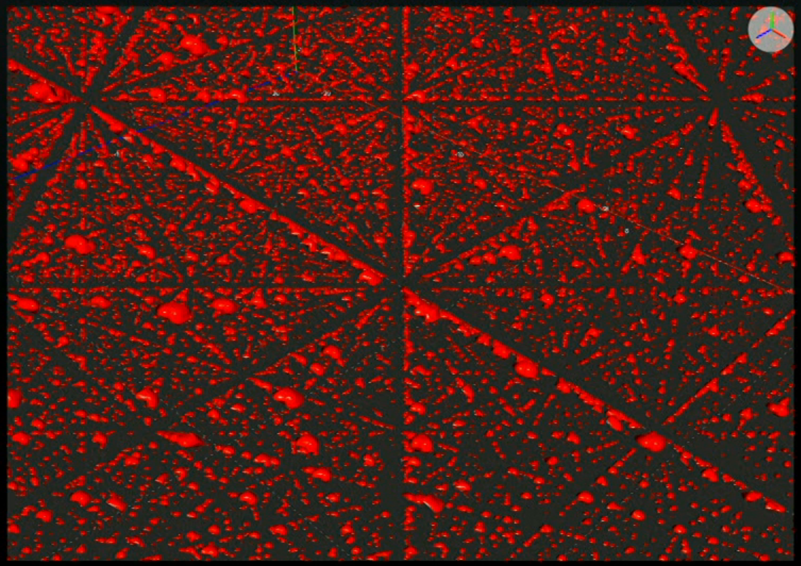


Fig. 2.—Reflection from face (100) of iron pyrites, at varying angles of incidence. Abscissa—Angle of incidence of rays on crystal face; Ordinate—Strength of reflected beam, arbitrary scale.

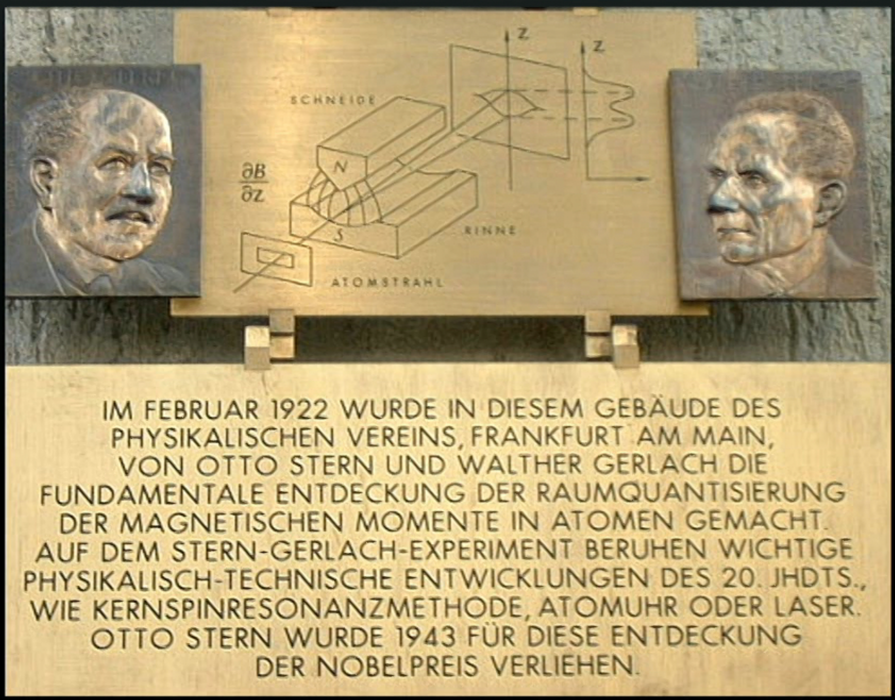
$$n\lambda = 2d \sin\theta$$

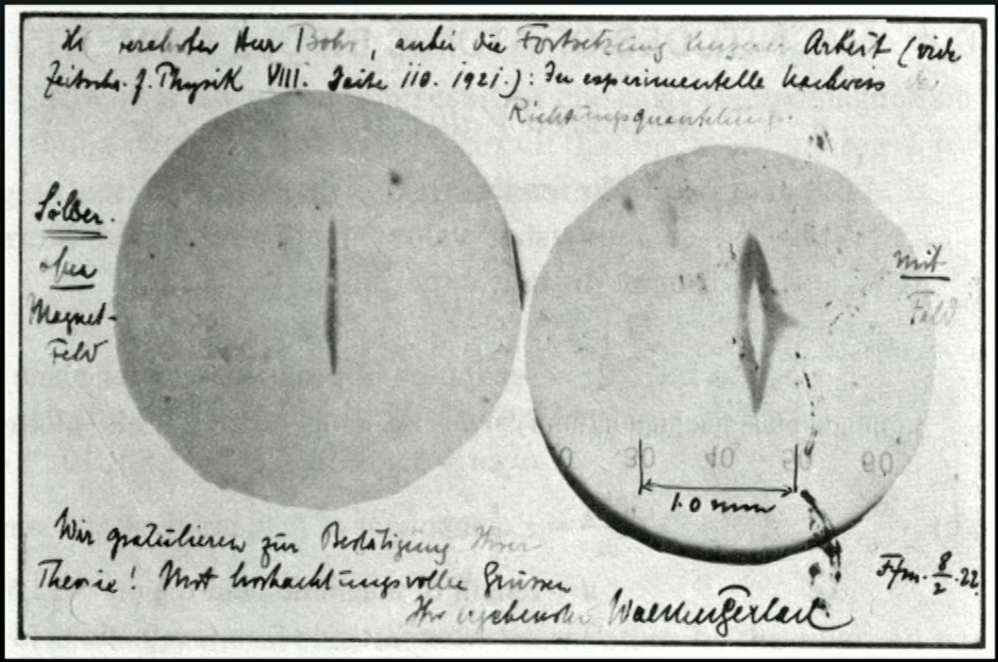
X-ray diffraction Today





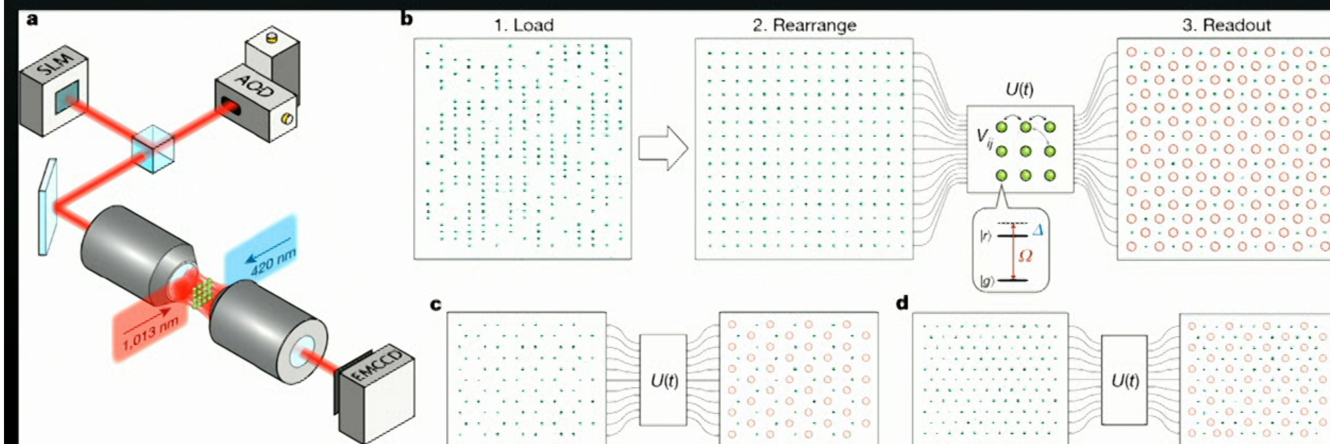
Projective Measurements in 1922



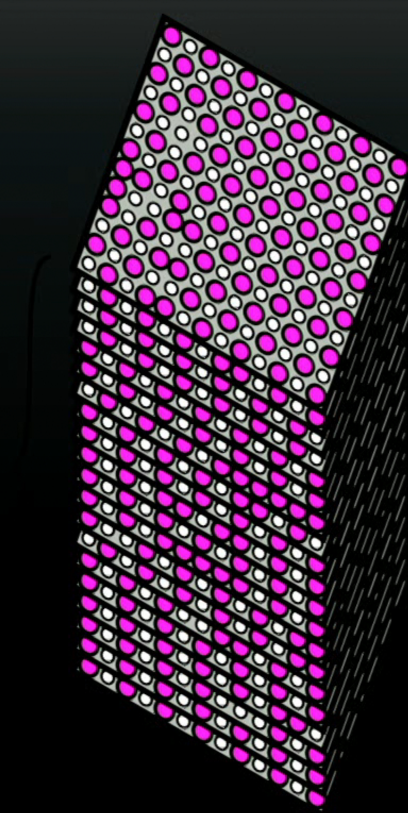


Projective Measurements Today

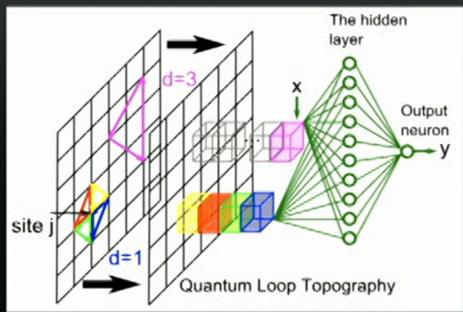
Quantum Simulators



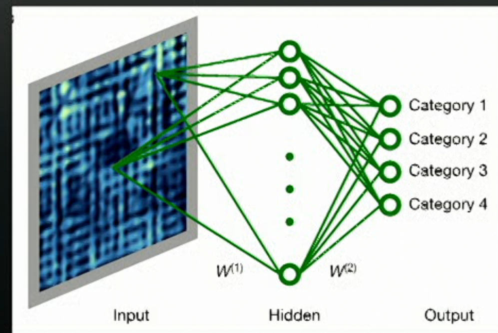
Taken from Ebadi. et al., *Nature* **595** (227), 2021



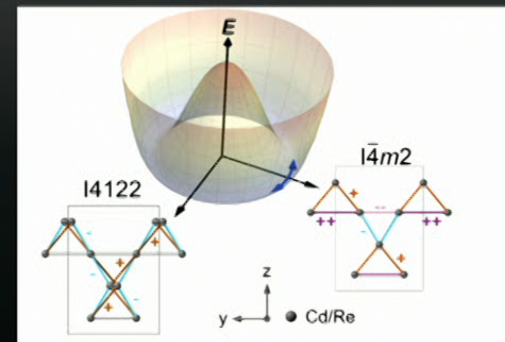
Learning Quantum many-body States



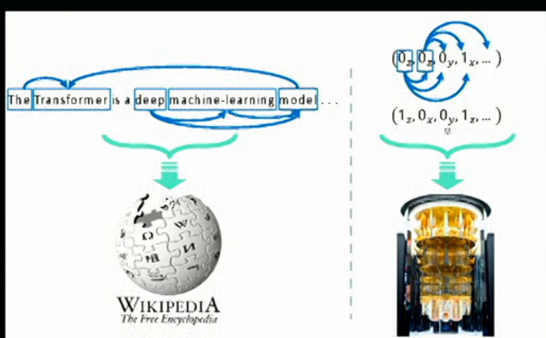
Quantum loop topography
Zhang & Kim, PRL 118, 216401 (2017)



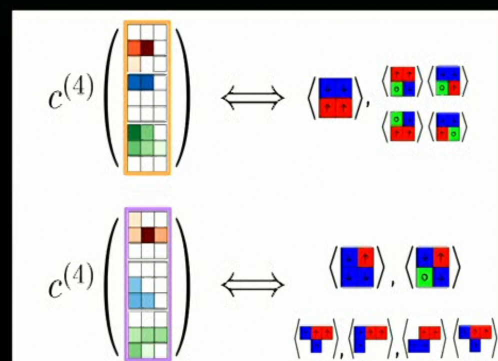
STS on high Tc cuprate BSSCO
Testing Charge Order Hypothesis
Zhang et al, Nature 570, 484 (2019)



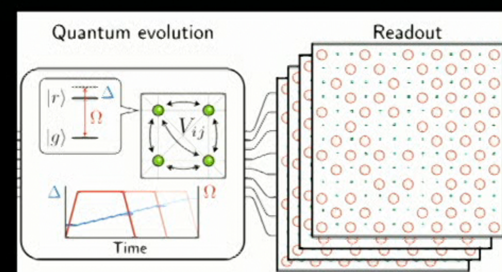
Xray-Temperature Clustering
Venderley et al, PNAS (2022)



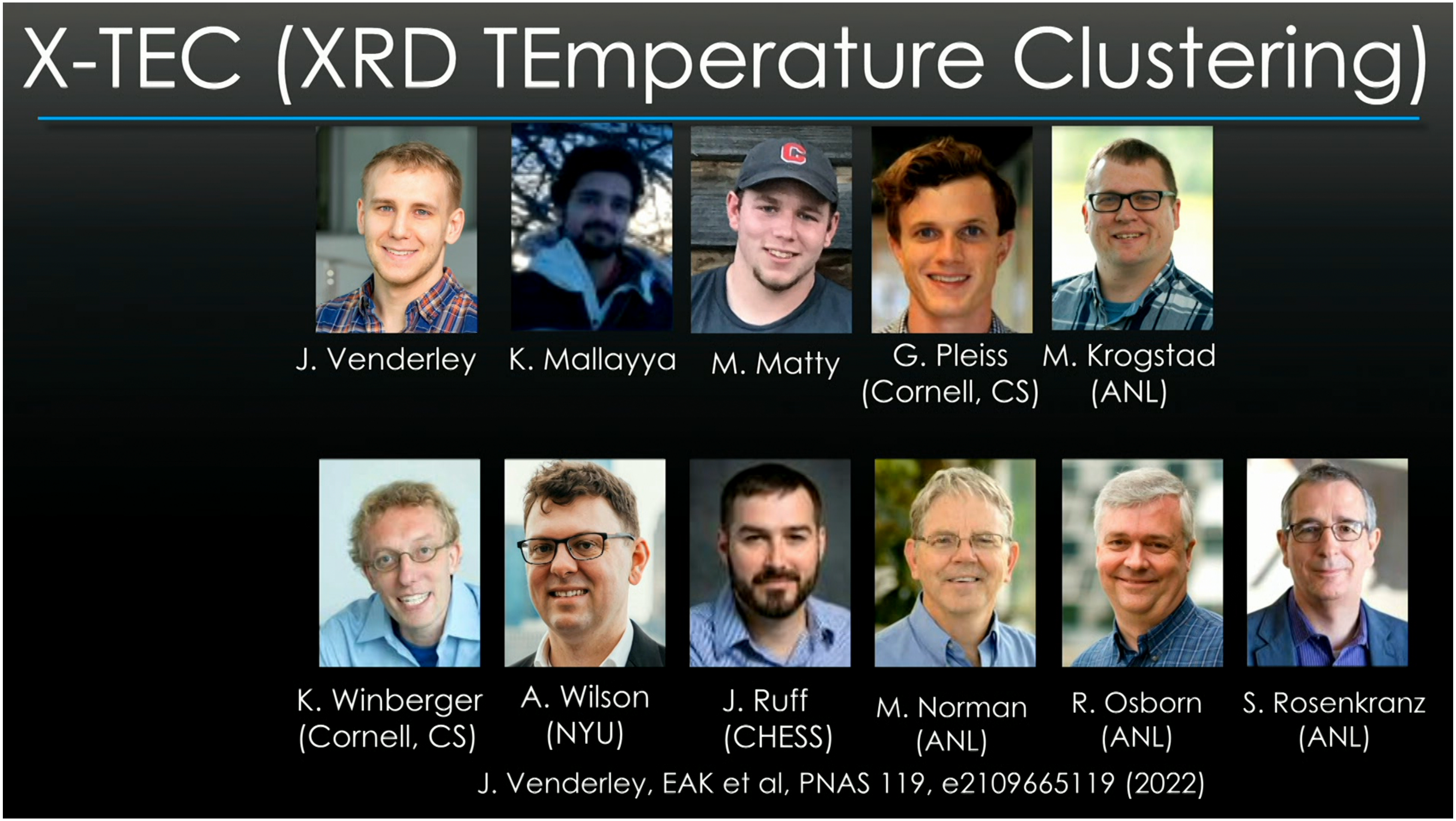
Attention-based Quantum Tomogrphy
Cha et al, MLST (2021)



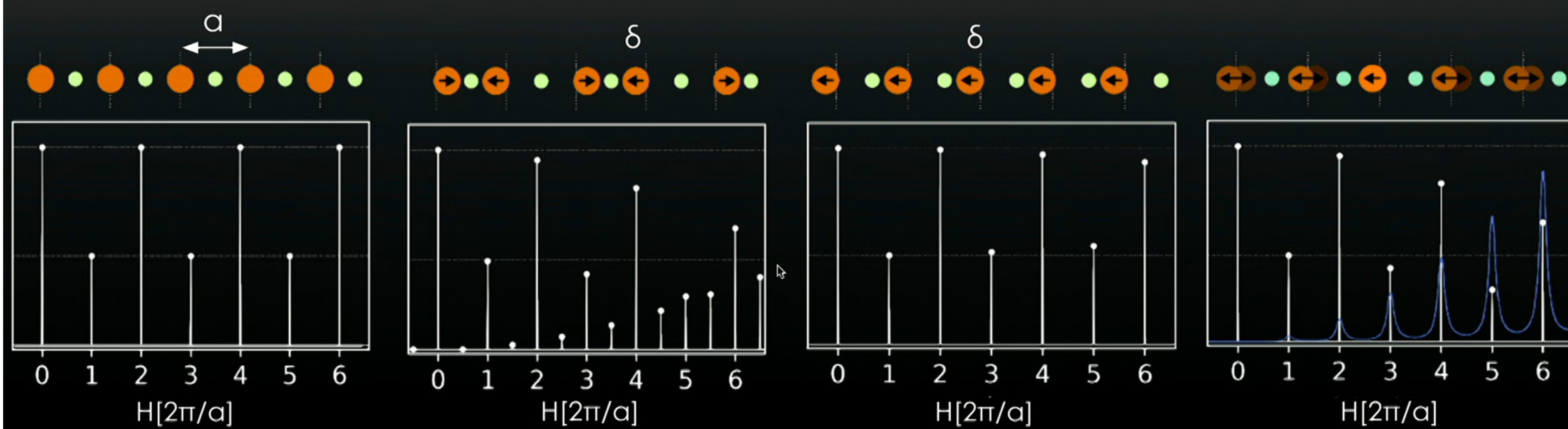
QGM on Fermi Hubbard model
Miles et al, Nat. Comm. 12, 3905 (2021)



Quantum Simulator Snapshots
Miles et al, PRR 5, 013026 (2023)


X-TEC (XRD TEmperature Clustering)
J. Venderley
K. Mallayya
M. Matty
G. Pleiss (Cornell, CS)
M. Krogstad (ANL)
K. Winberger (Cornell, CS)
A. Wilson (NYU)
J. Ruff (CHESS)
M. Norman (ANL)
R. Osborn (ANL)
S. Rosenkranz (ANL)
J. Venderley, EAK et al, PNAS 119, e2109665119 (2022)

Orders and fluctuarions as seen in XRD



Fully Symmetric (undistorted)

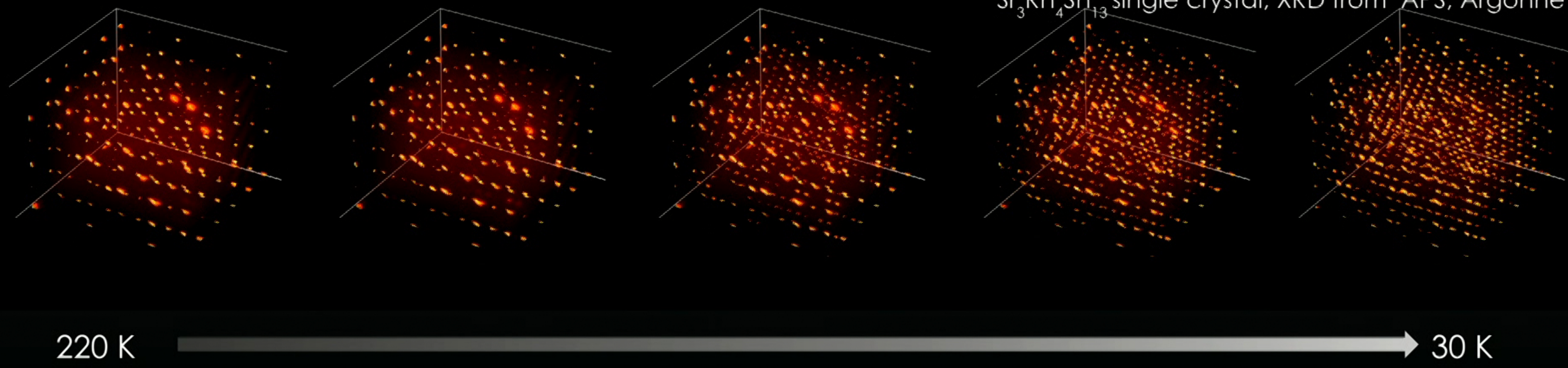
CDW with period 2a: superlattice peaks

IUC order: Form factors of Bragg peaks.

Fluctuations

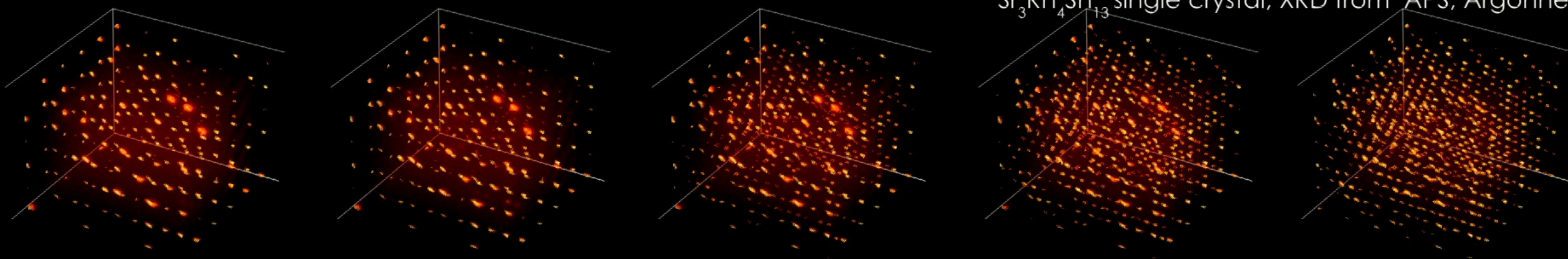
Temperature evolution of XRD

$Sr_3Rh_4Sn_{13}$ single crystal, XRD from APS, Argonne



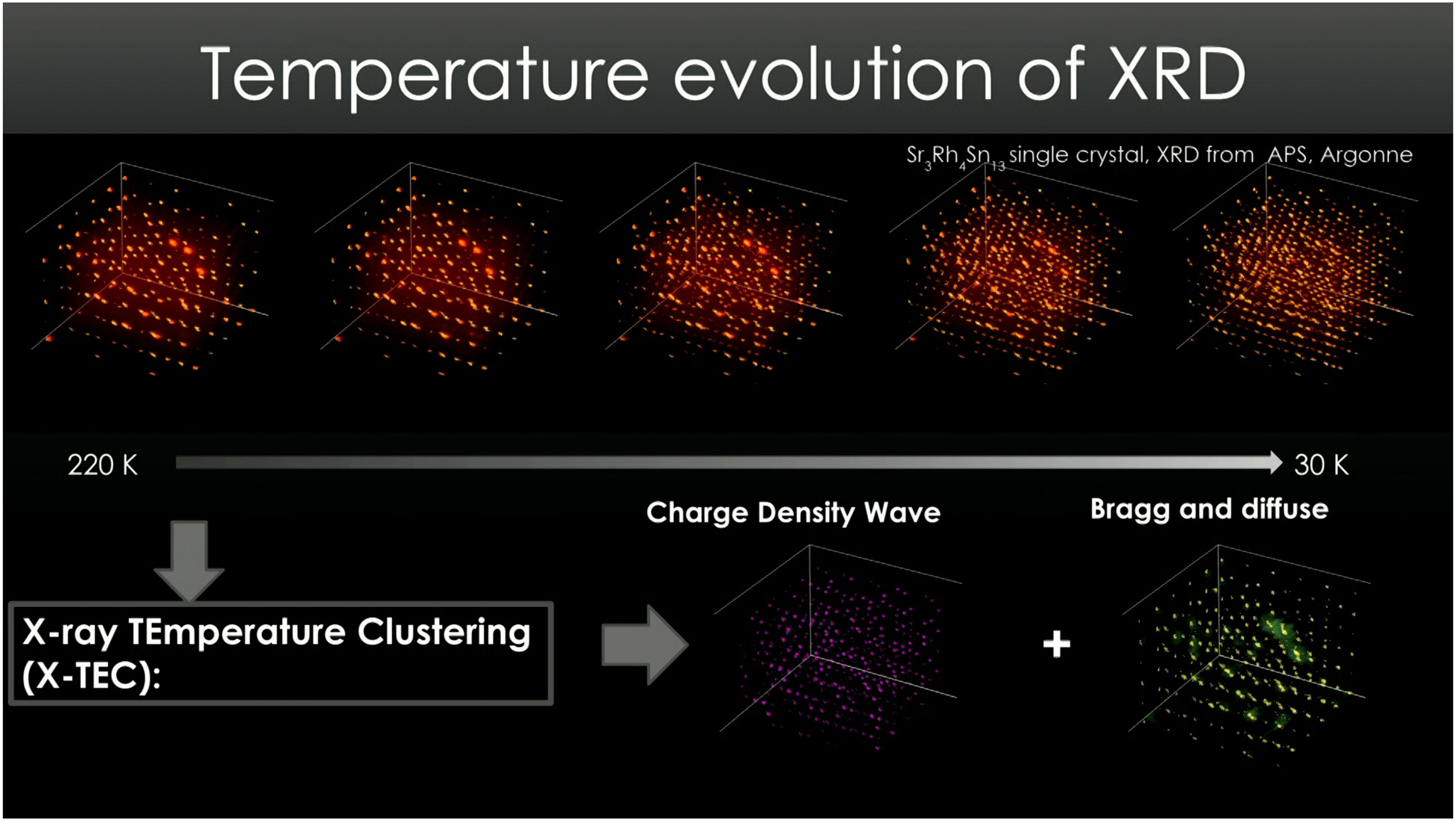
Temperature evolution of XRD

$Sr_3Rh_4Sn_{13}$ single crystal, XRD from APS, Argonne



220 K → 30 K

X-ray TEmperature Clustering (X-TEC):


Temperature evolution of XRD
Sr3Rh4Sn13 single crystal, XRD from APS, Argonne
220 K
30 K
Charge Density Wave
Bragg and diffuse
X-ray TEmperature Clustering (X-TEC):
+

Learn the sorting criteria for emergence?

$$\boxed{F = E - TS}$$

Learn the sorting criteria for emergence?

$$\boxed{F = E - TS}$$

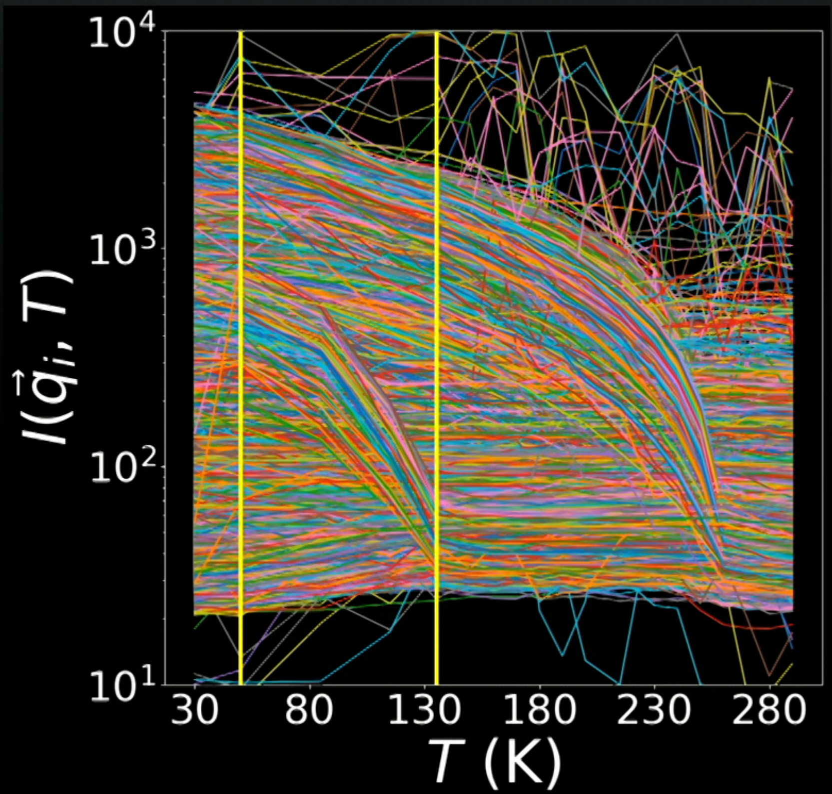
- Temperature series (T_d - points) for each $\vec{q}_i$

Learn the sorting criteria for emergence?

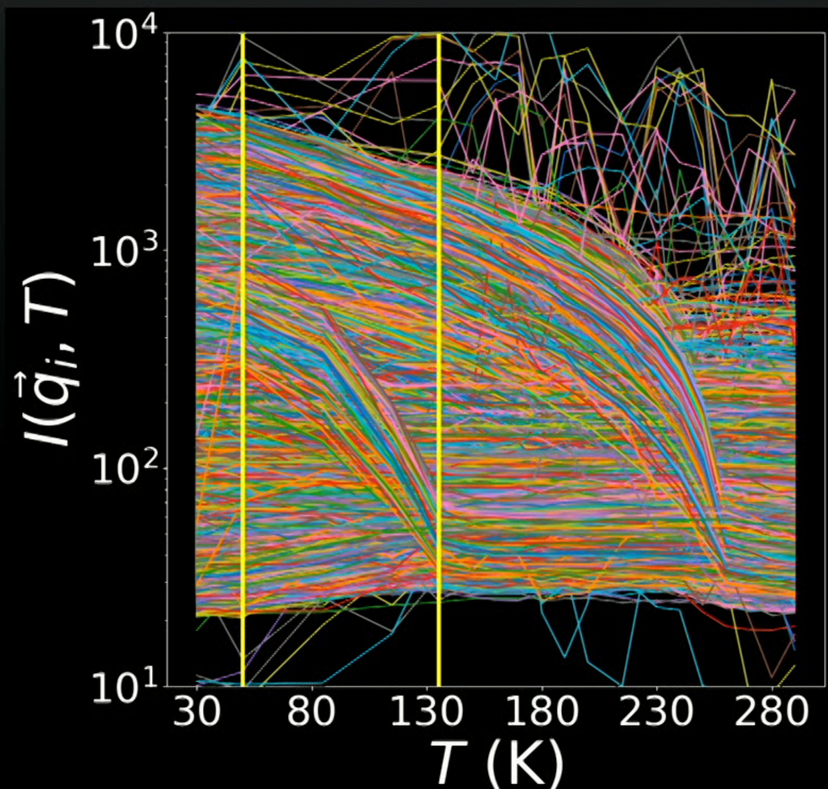
$$\boxed{F = E - TS}$$

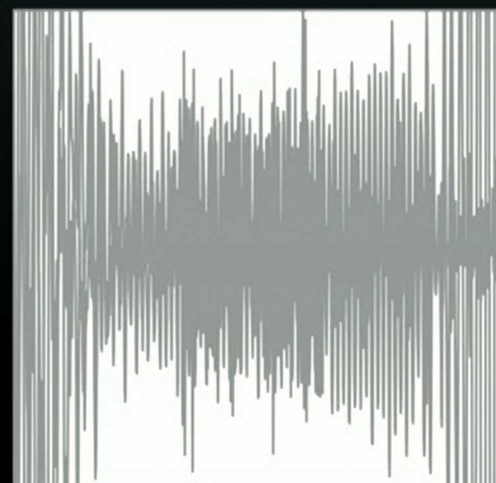
- Temperature series (T_d - points) for each $\vec{q}_i$
- A point in T_d -dim space
- Expect different clusters of T-series among the population of $\vec{q}_i$

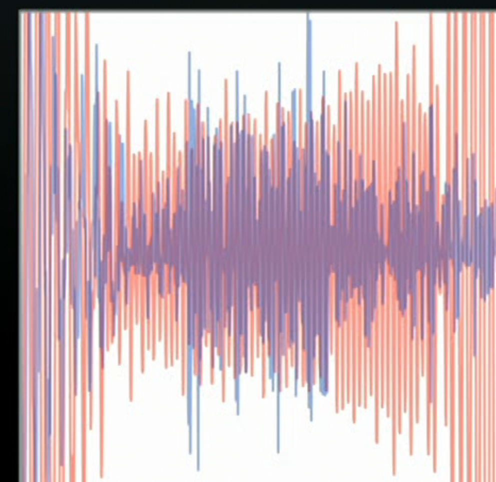
Raw T-Series and Speaker Verification



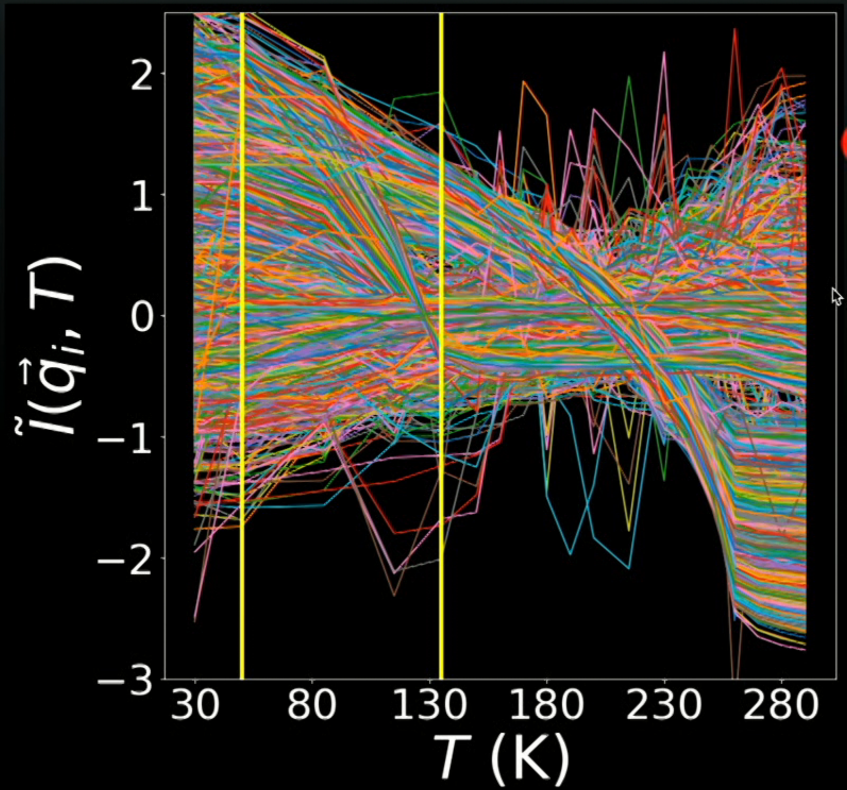
Raw T-Series and Speaker Verification

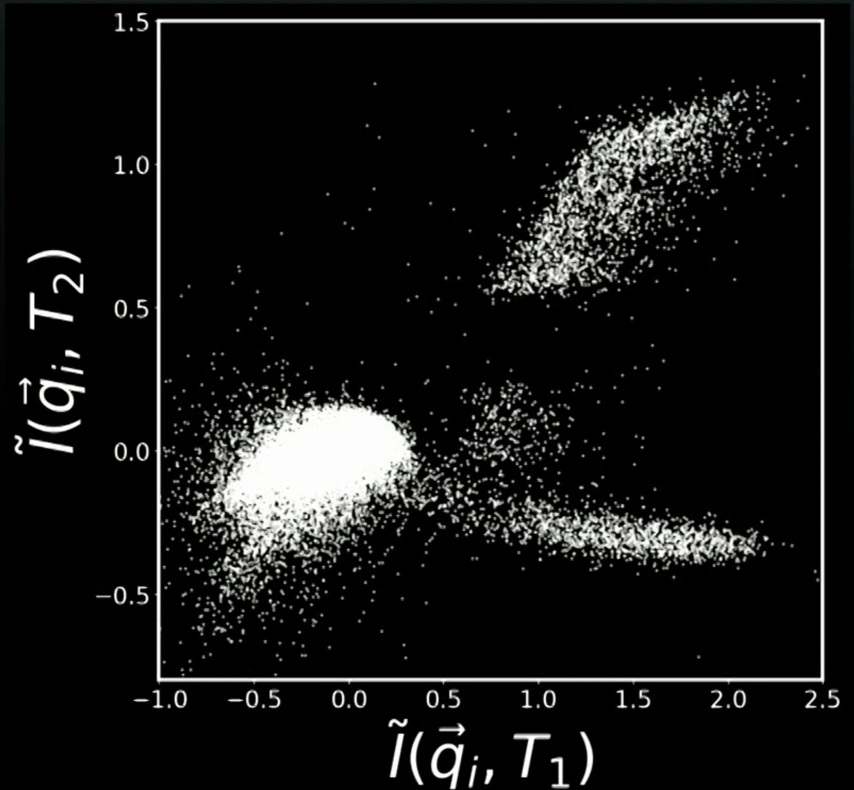






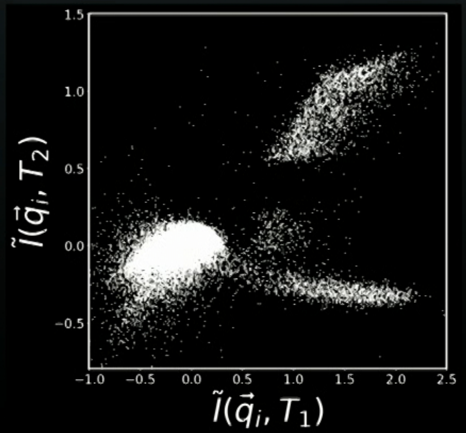
T-series as a point in high-dim space





Clustering based on Probabilistic Modeling

d-temperature series ⇔ A point in d-dimension

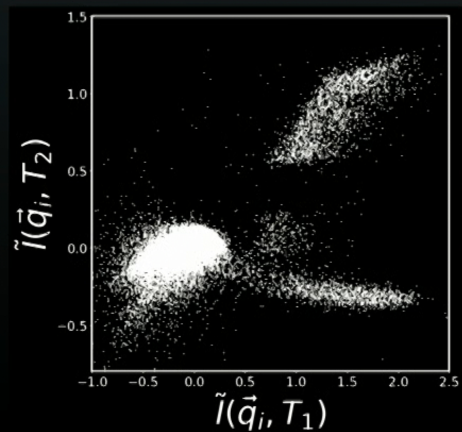


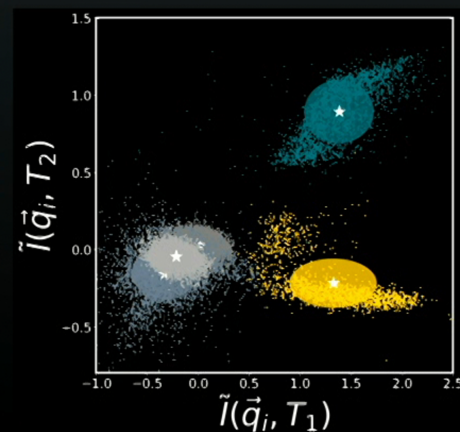
- Gaussian Mixture Model

$$p(I) = \sum_{k=1}^{K} \pi_k \mathcal{N}(I|m_k, s_k)$$

Clustering based on Probabilistic Modeling

d-temperature series ⟺ A point in d-dimension



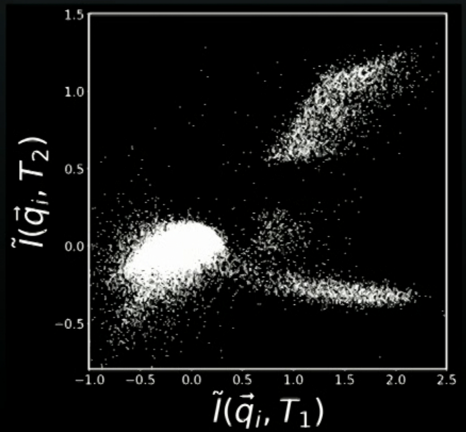


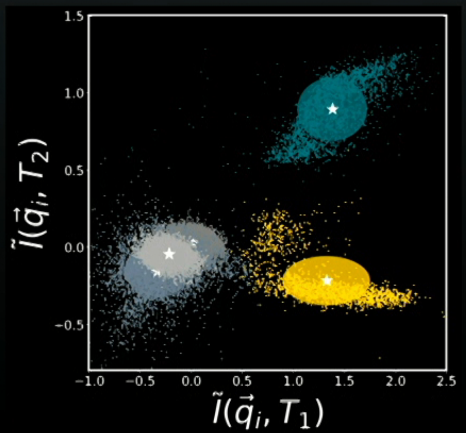
- Gaussian Mixture Model

$$p(I) = \sum_{k=1}^{K} \pi_k \mathcal{N}(I | m_k, s_k)$$

Clustering based on Probabilistic Modeling

d-temperature series ⬌ A point in d-dimension





- Gaussian Mixture Model

$$p(I) = \sum_{k=1}^{K} \pi_k \mathcal{N}(I|m_k, s_k)$$

Learning the model and clustering

- Find the maximum likelihood solution for $\{I\} = \{I(\vec{q}_i)\}$

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

- Cluster assignment for each data point $\vec{q}_i$

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

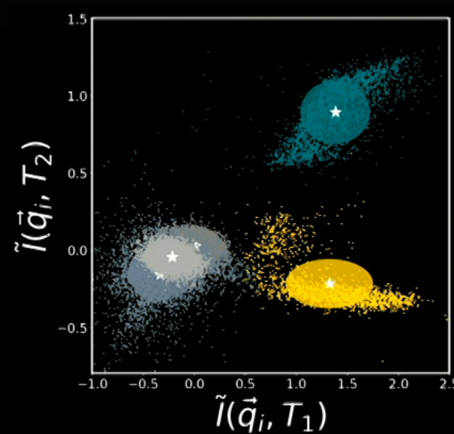
Learning the model and clustering

- Find the maximum likelihood solution for $\{I\} = \{I(\vec{q}_i)\}$

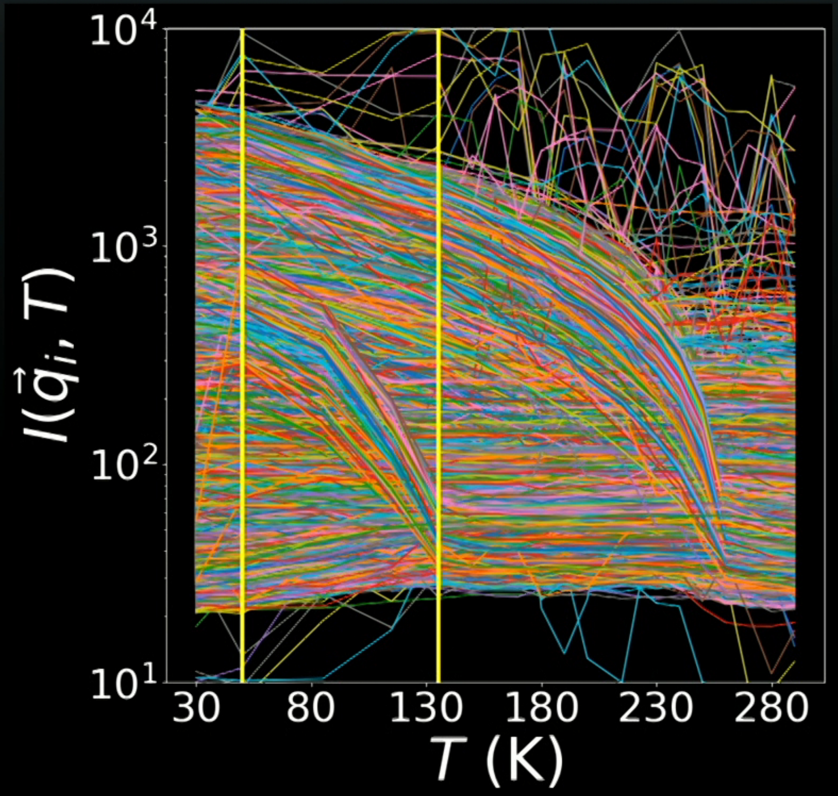
$$\log p\Big(\{\tilde{\mathbf{I}}(\vec{q}_i)\}|\pi, \mathbf{m}, \mathbf{s}\Big) = \sum_{\vec{q}_i} \log\left[\sum_{k=1}^{K} \pi_k \mathcal{N}\Big(\tilde{\mathbf{I}}(\vec{q}_i)|\mathbf{m}_k, \mathbf{s}_k\Big)\right]$$

- Cluster assignment for each data point $\vec{q}_i$

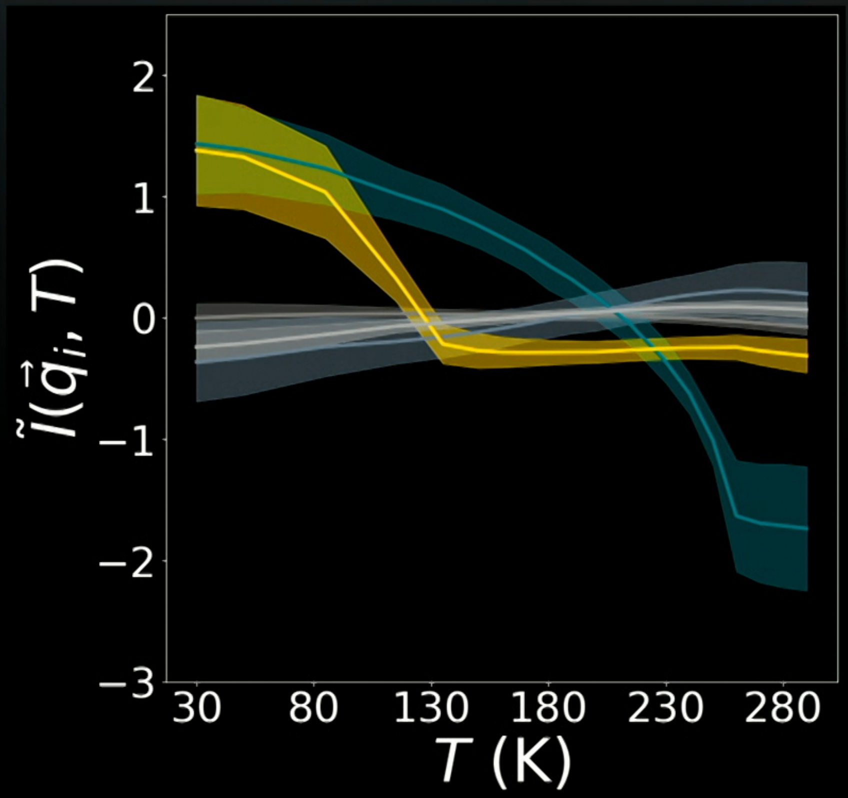
$$w_i^k = \frac{\pi_k \mathcal{N}\Big(\tilde{\mathbf{I}}(\vec{q}_i)|\mathbf{m}_k, \mathbf{s}_k\Big)}{\sum_k \pi_k \mathcal{N}\Big(\tilde{\mathbf{I}}(\vec{q}_i)|\mathbf{m}_k, \mathbf{s}_k\Big)}$$



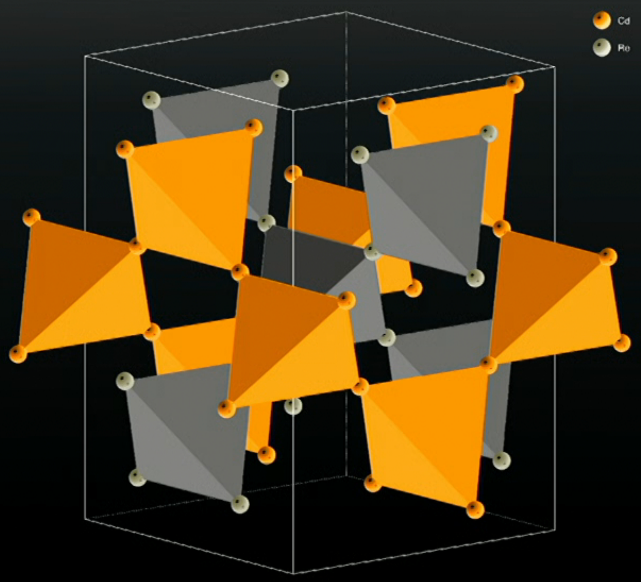
Discovery of order parameters

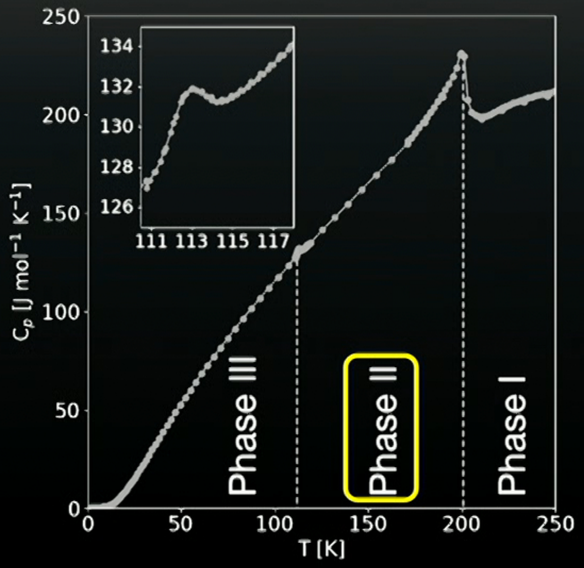


Discovery of order parameters



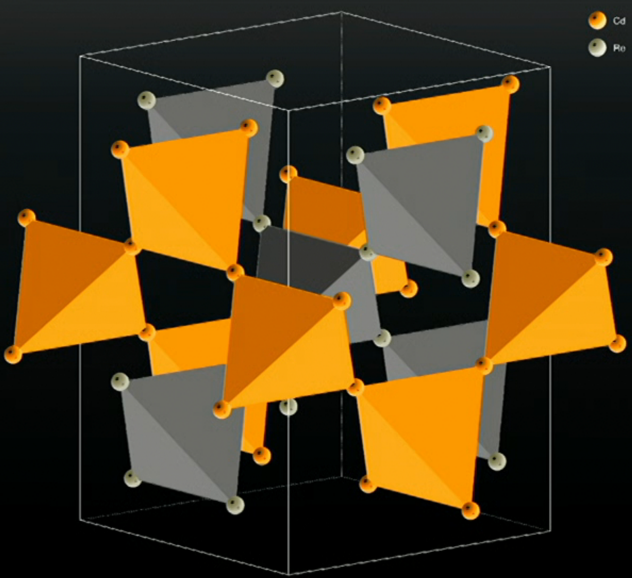
Pyrochlore $Cd_2Re_2O_7$

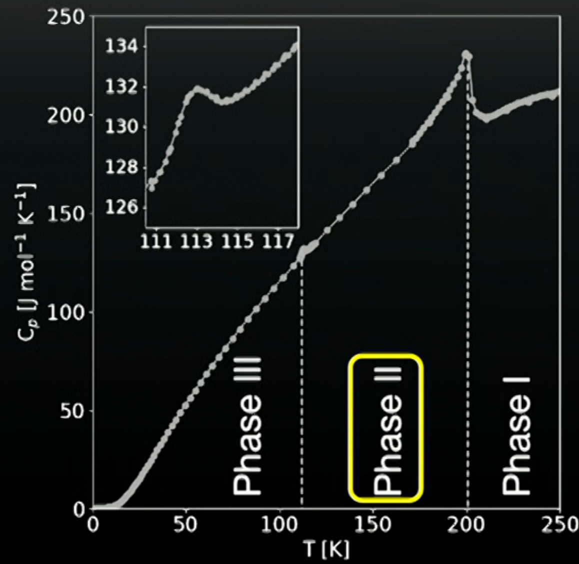


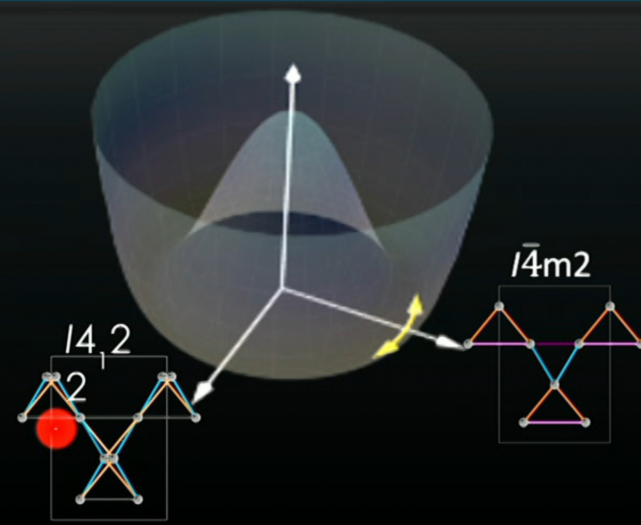


- Large signal in C_p vs small displacements, 5d oxide.
- Ordering of Phase II : 2 dim E_u vs 1 dim T_{2u}

Pyrochlore $Cd_2Re_2O_7$



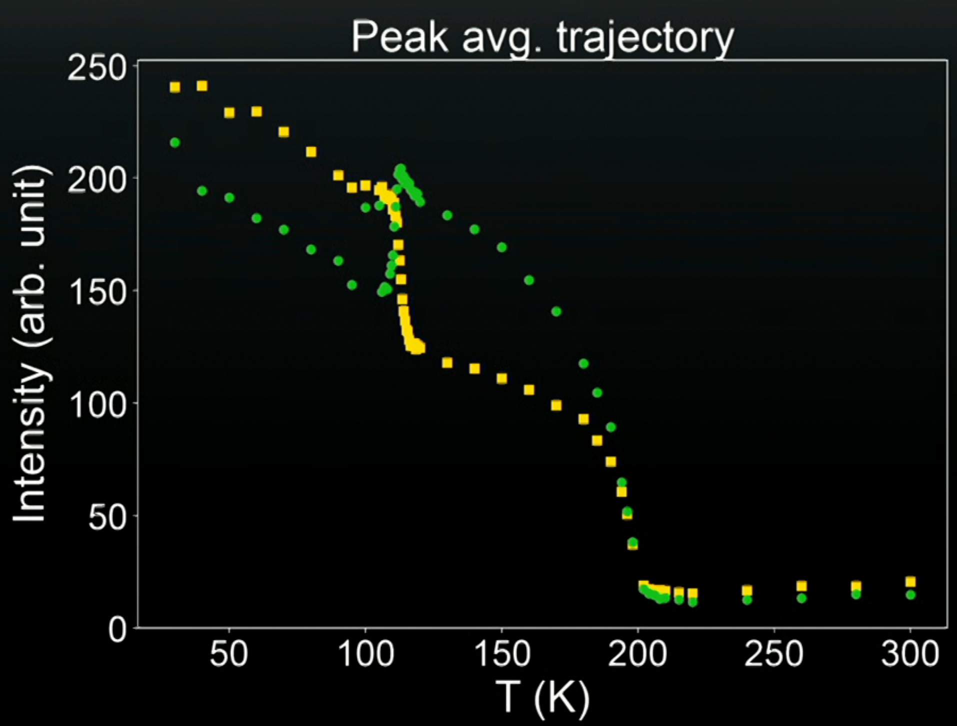




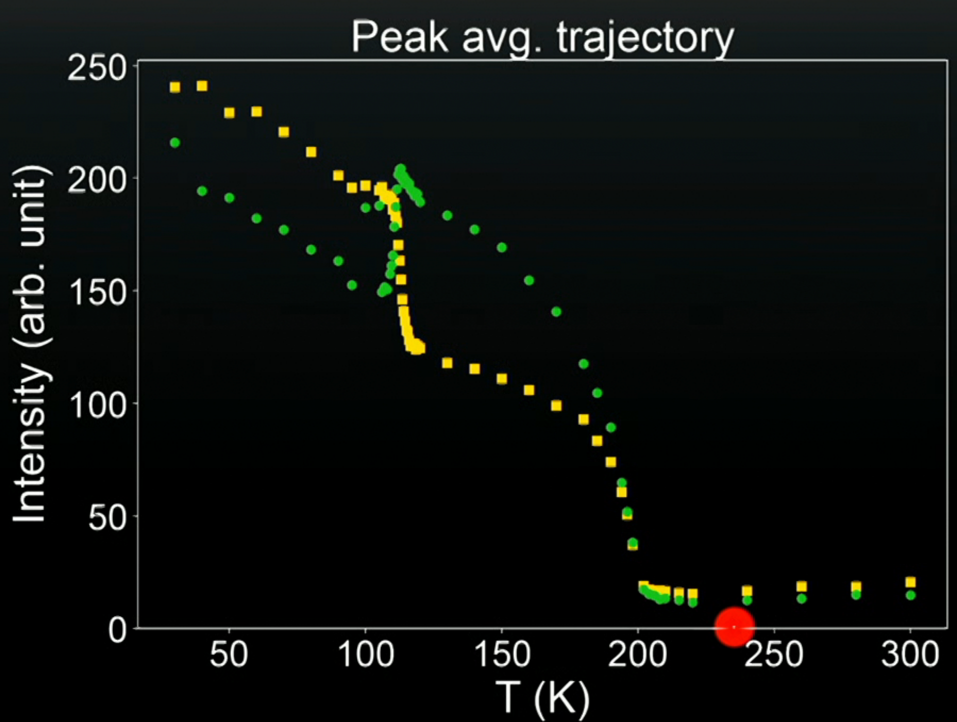
- Large signal in C_p vs small displacements, 5d oxide.
- Ordering of Phase II : 2 dim E_u vs 1 dim T_{2u}

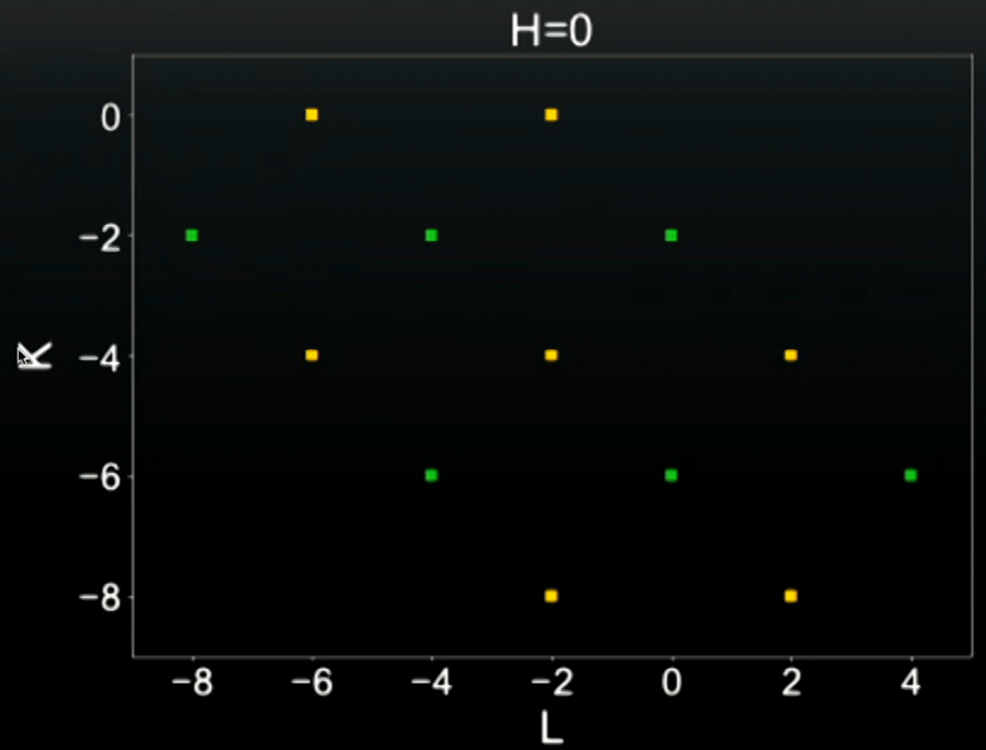
8 TB in ~10 min

Discovery I: Re displacement



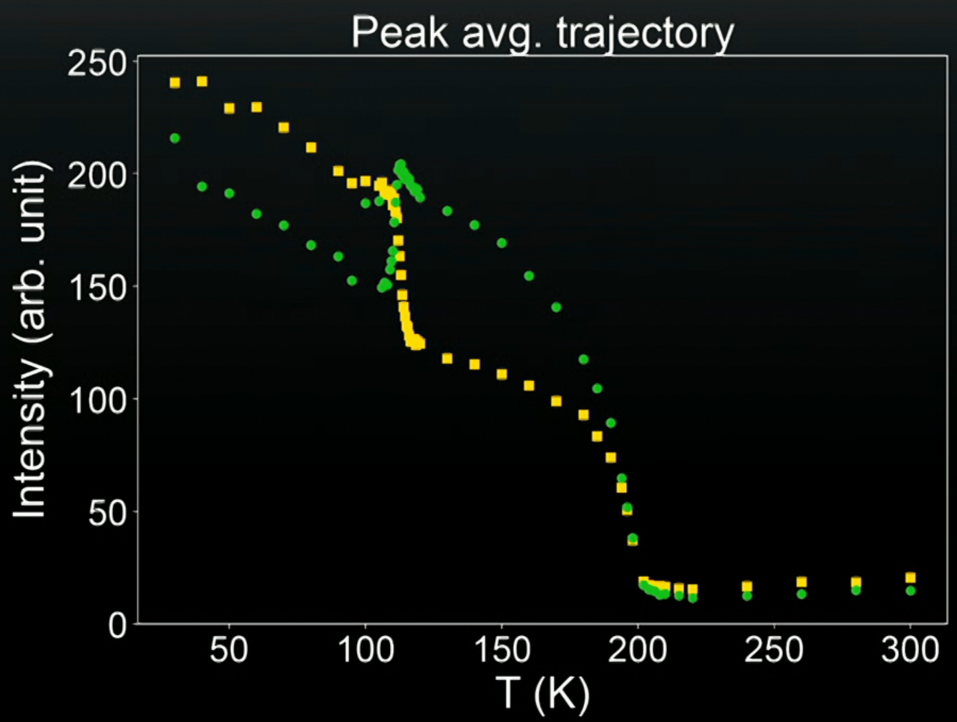
Discovery I: Re displacement

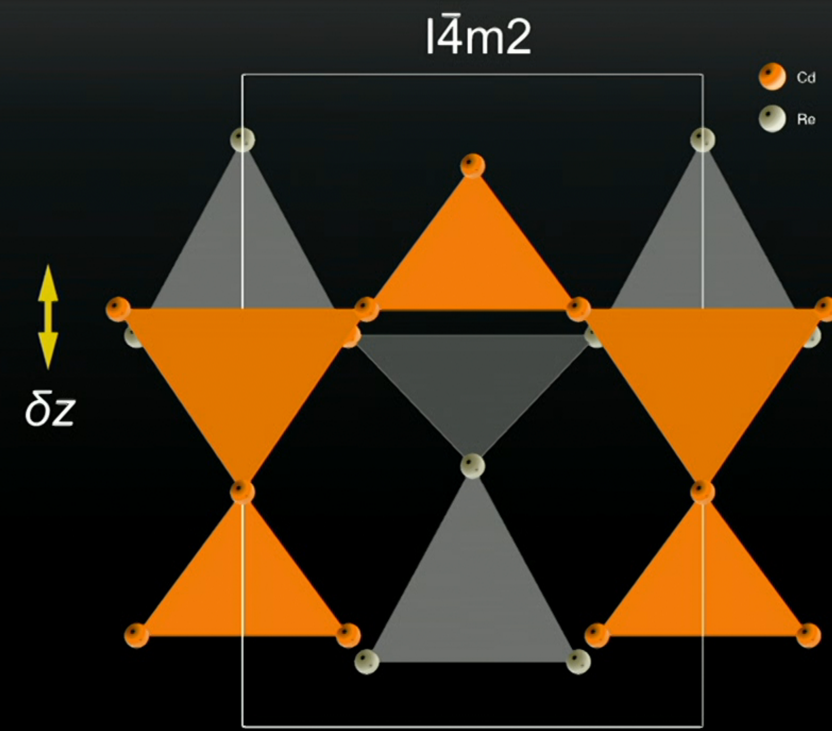




$(4n_1, 4n_2, 4n_3 + 2)$ vs $(4n_1 + 2, 4n_2, 4n_3)$ or $(4n_1, 4n_2 + 2, 4n_3)$

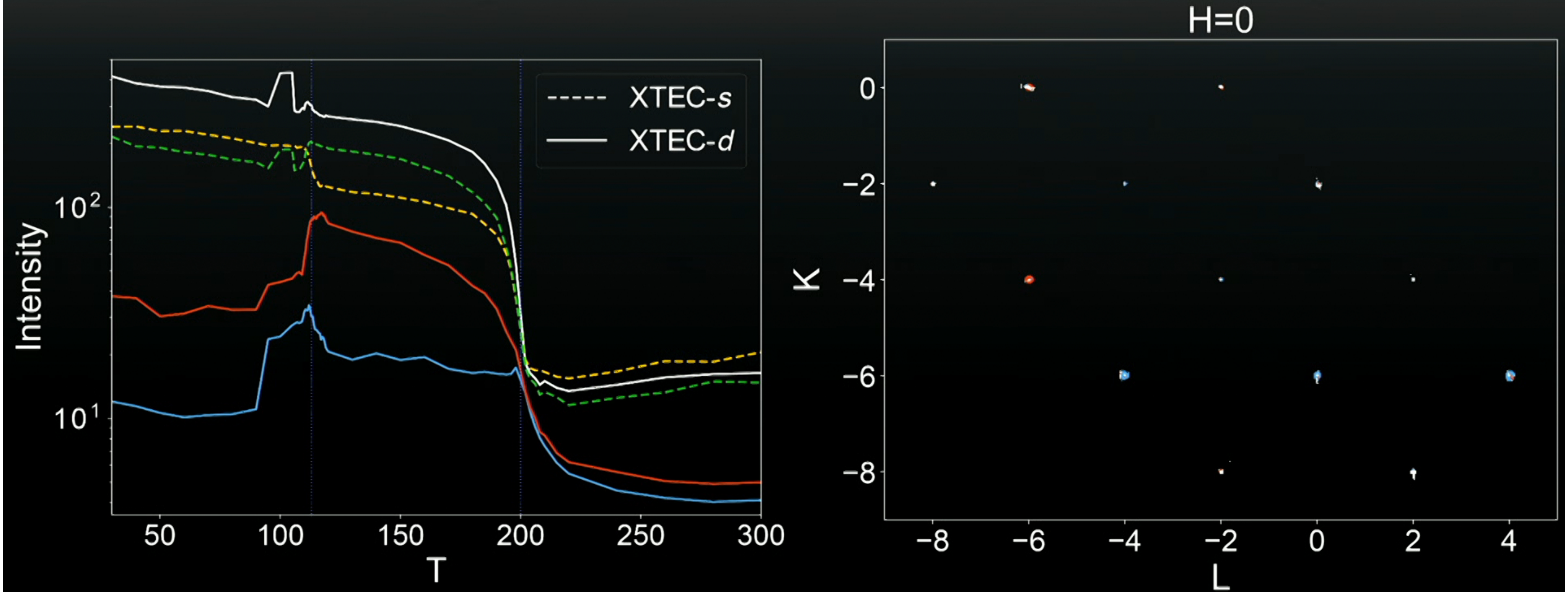
Discovery I: Re displacement





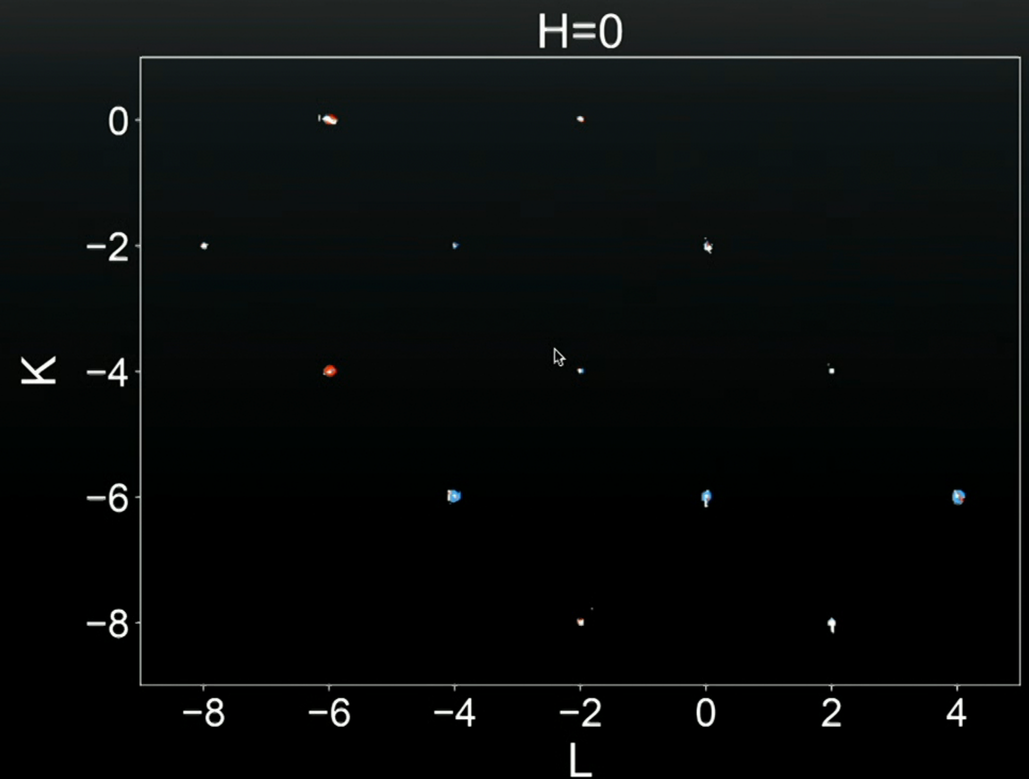
$(4n_1, 4n_2, 4n_3 + 2)$ vs $(4n_1 + 2, 4n_2, 4n_3)$ or $(4n_1, 4n_2 + 2, 4n_3)$

Discovery II: Goldstone mode



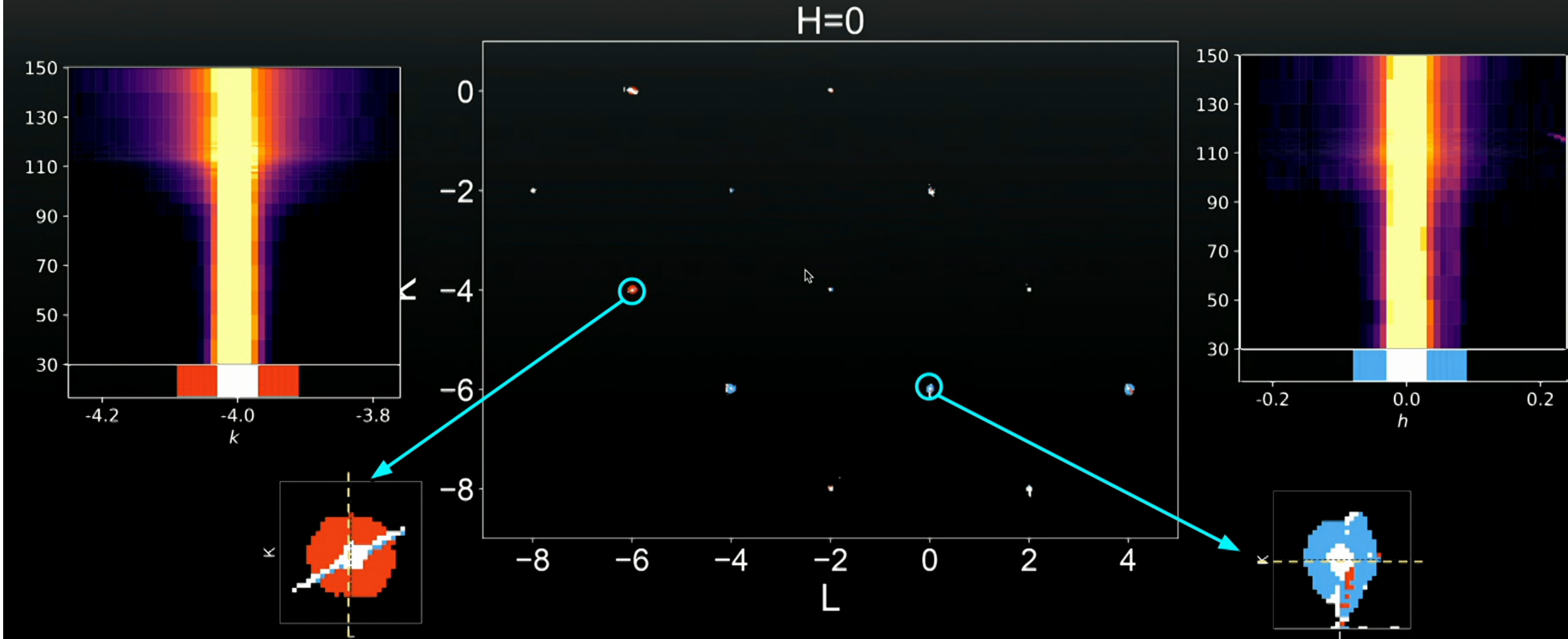
$(4n_1, 4n_2, 4n_3 + 2)$ vs $(4n_1 + 2, 4n_2, 4n_3)$ or $(4n_1, 4n_2 + 2, 4n_3)$

Discovery II: Goldstone mode



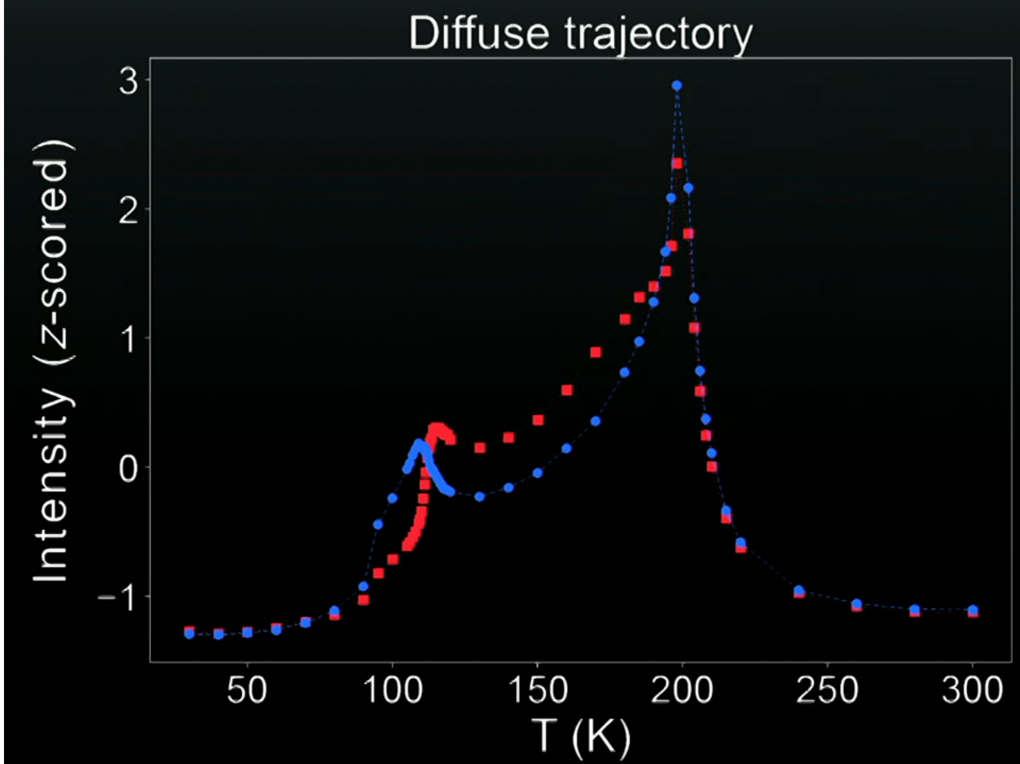
$(4n_1, 4n_2, 4n_3 + 2)$ vs $(4n_1 + 2, 4n_2, 4n_3)$ or $(4n_1, 4n_2 + 2, 4n_3)$

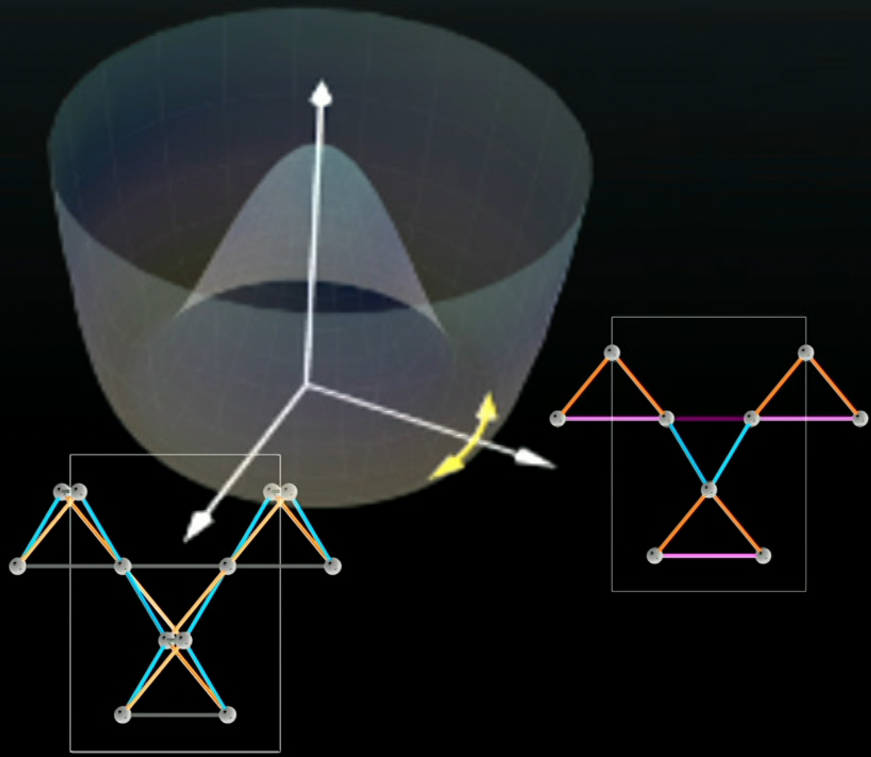
Discovery II: Goldstone mode



$(4n_1, 4n_2, 4n_3 + 2)$ vs $(4n_1 + 2, 4n_2, 4n_3)$ or $(4n_1, 4n_2 + 2, 4n_3)$

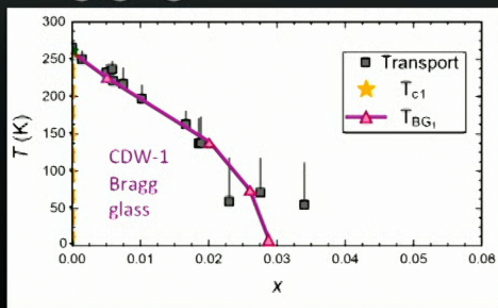
Discovery II: Goldstone mode





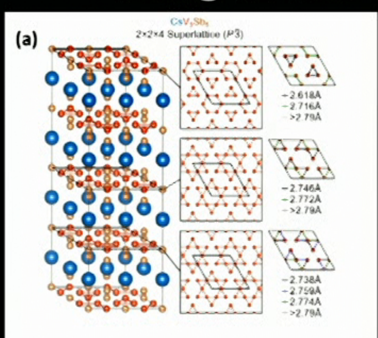
Versatility

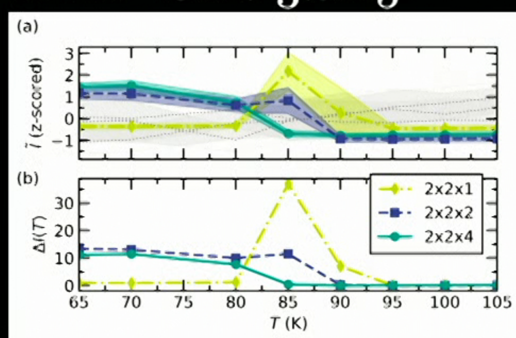
- Bragg glass in Pd-ETe3



Mallayya, EAK et al (arXiv:2207.14795)

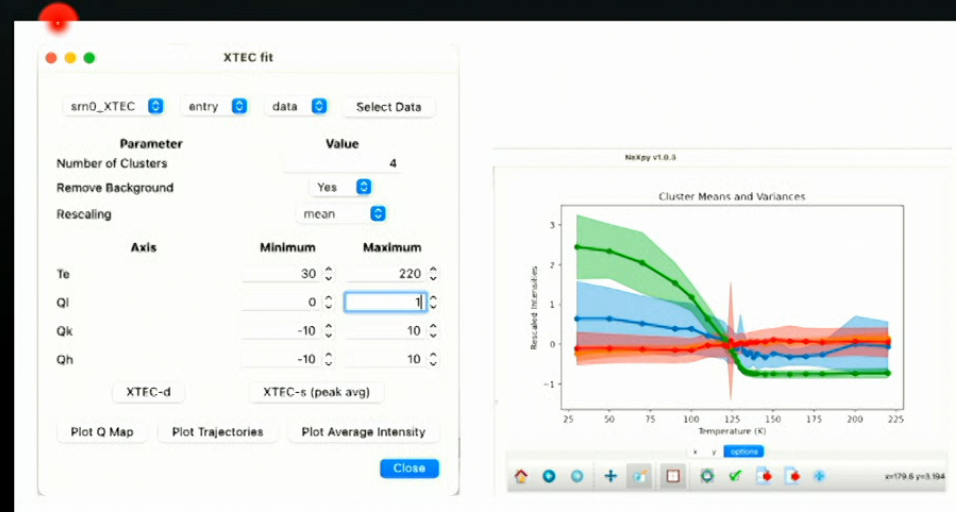
- Kagome metal CsV_3Sb_5





Kautzsch, EAK, Wilson et al, PRM (2023)

- Plug-in for NeXPy for Advanced Photon Source and CHESS

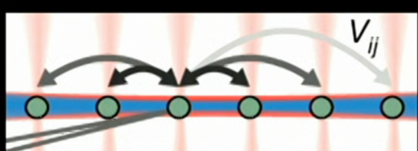


- Time-resolved XRD
- 4D STEM data

“Decoding”
Quantum Simulator Data

Projective measurements

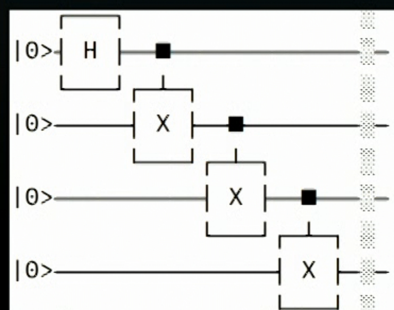
ibm.com



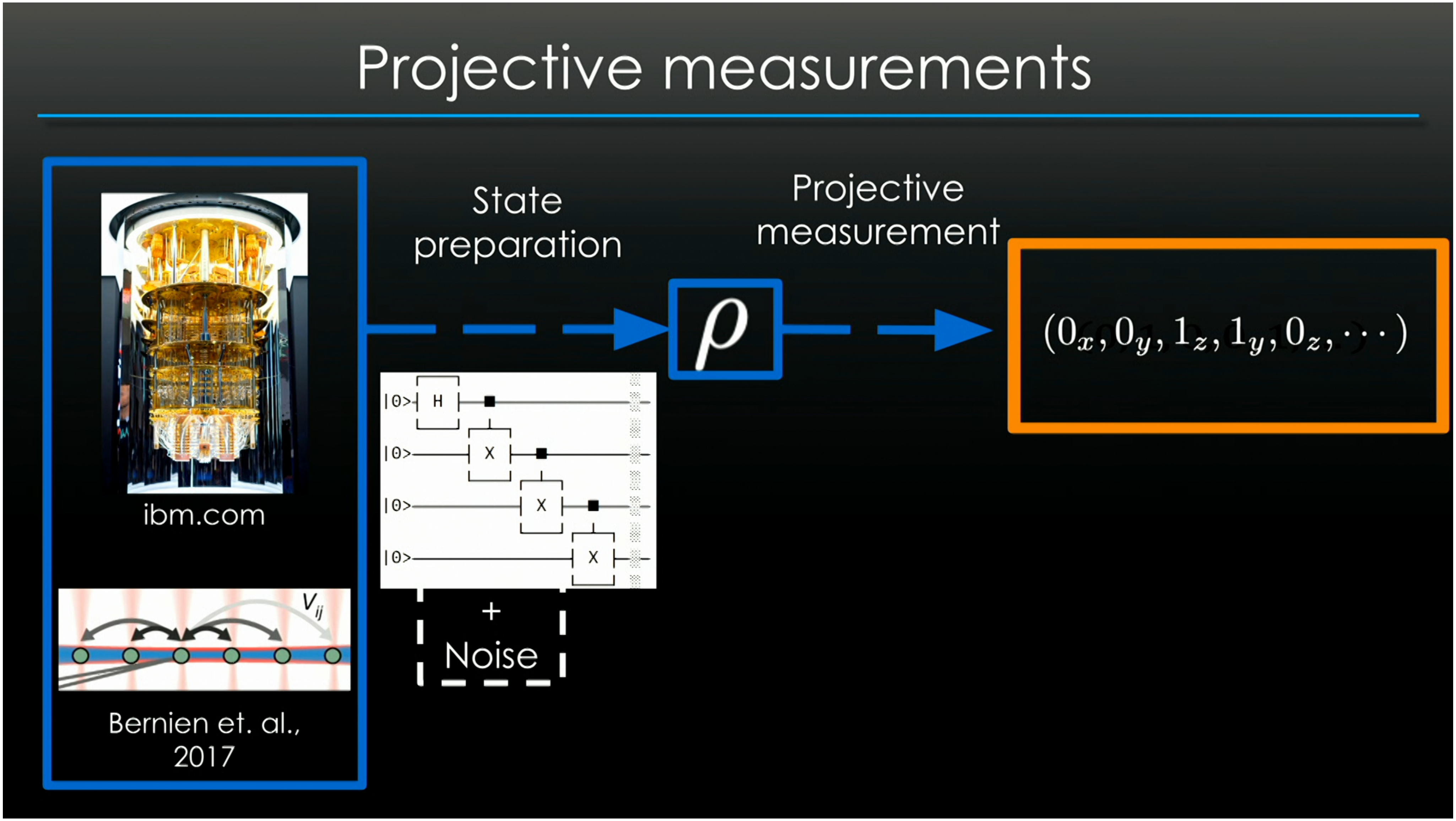
Bernien et. al., 2017

State preparation

ρ



+ Noise



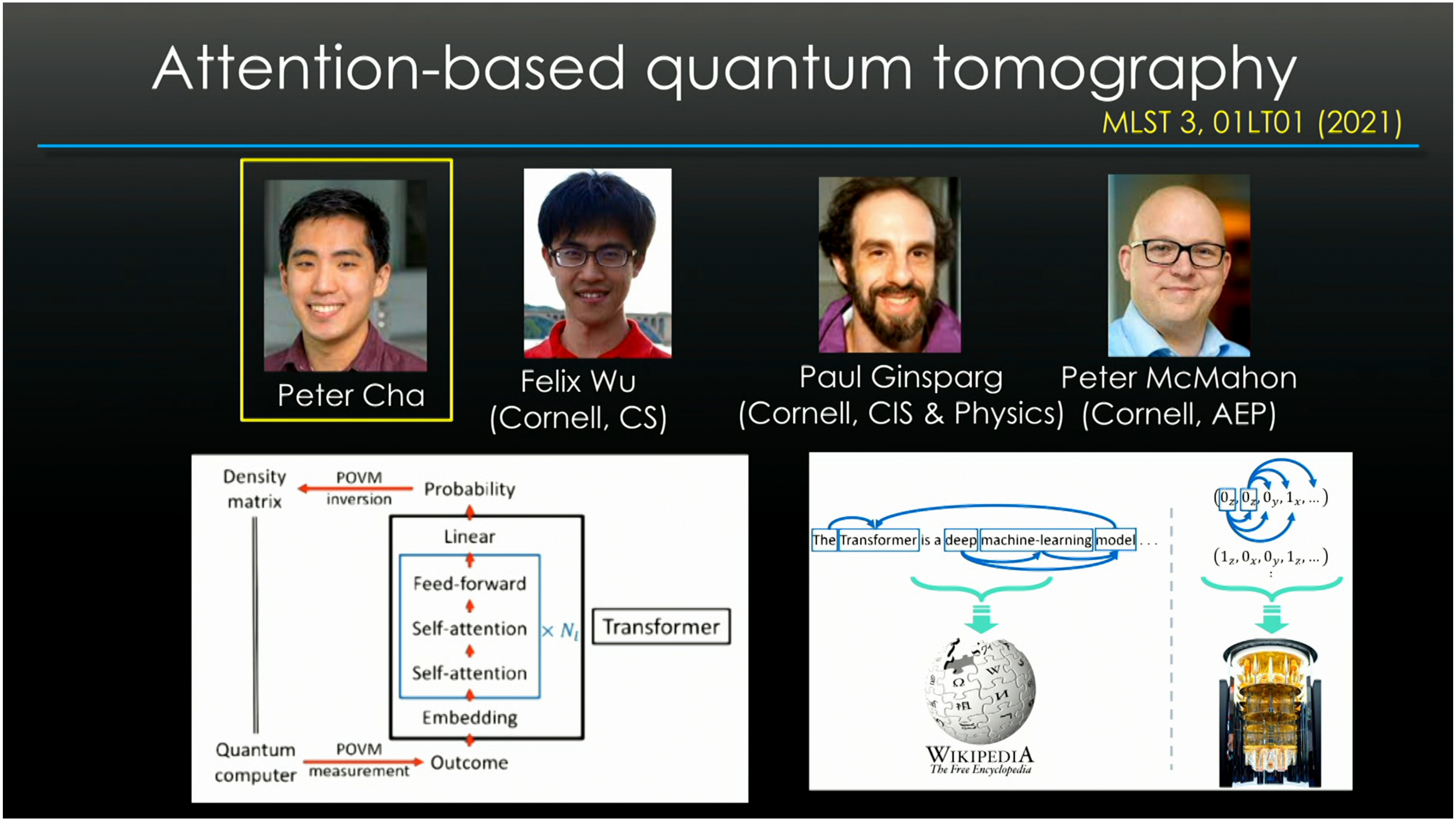
Projective measurements
ibm.com
Bernien et. al., 2017
State preparation
Projective measurement
ρ
$(0_x, 0_y, 1_z, 1_y, 0_z, \cdots)$
|0>
H
X
+ Noise

Challenge of Tomography

- N_q-QUBIT SYSTEM $\leftrightarrow \sim 4^{N_q}$ COMPLEX #'S
- $N_q = 10 \sim 10^6$ #'S ~ 1 MB
- $N_q = 20 \sim 10^{12}$ #'S ~ 1 TB
- $N_q = 53 \sim 10^{30}$ #'S ~ 1 QUINTILLION TB

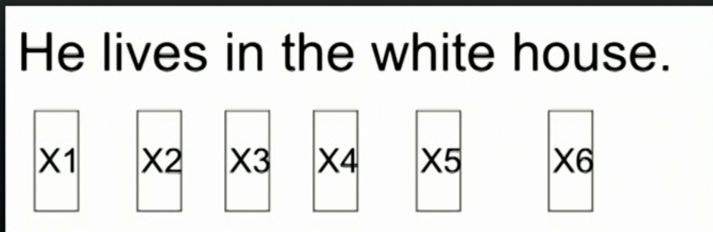
Library of Congress ~ 20 TB
Source: loc.gov

- MAXIMUM LIKELIHOOD ESTIMATION (MLE): $N_q < 10$



Attention-based quantum tomography
MLST 3, 01LT01 (2021)
Peter Cha
Felix Wu
(Cornell, CS)
Paul Ginsparg
(Cornell, CIS & Physics)
Peter McMahon
(Cornell, AEP)
Density matrix
POVM inversion
Probability
Linear
Feed-forward
Self-attention
Self-attention
$\times N_l$
Transformer
Embedding
Quantum computer
POVM measurement
Outcome
The Transformer is a deep machine-learning model . . .
$(0_z, 0_z, 0_y, 1_x, \dots)$
$(1_z, 0_x, 0_y, 1_z, \dots)$
WIKIPEDIA
The Free Encyclopedia

Self-attention



Self-attention

He lives in the white house.

X1 X2 X3 X4 X5 X6

$$S_{ij} = \langle QX_i, KX_j \rangle$$

$$A_{ij} = \frac{\exp(S_{ij})}{\sum_{ik} \exp(S_{ik})}$$

$$\hat{X}_i = \sum_j A_{ij} V_j$$

- Q,K,V are to be learned.

Self-attention

He lives in the white house.

X1 X2 X3 X4 X5 X6

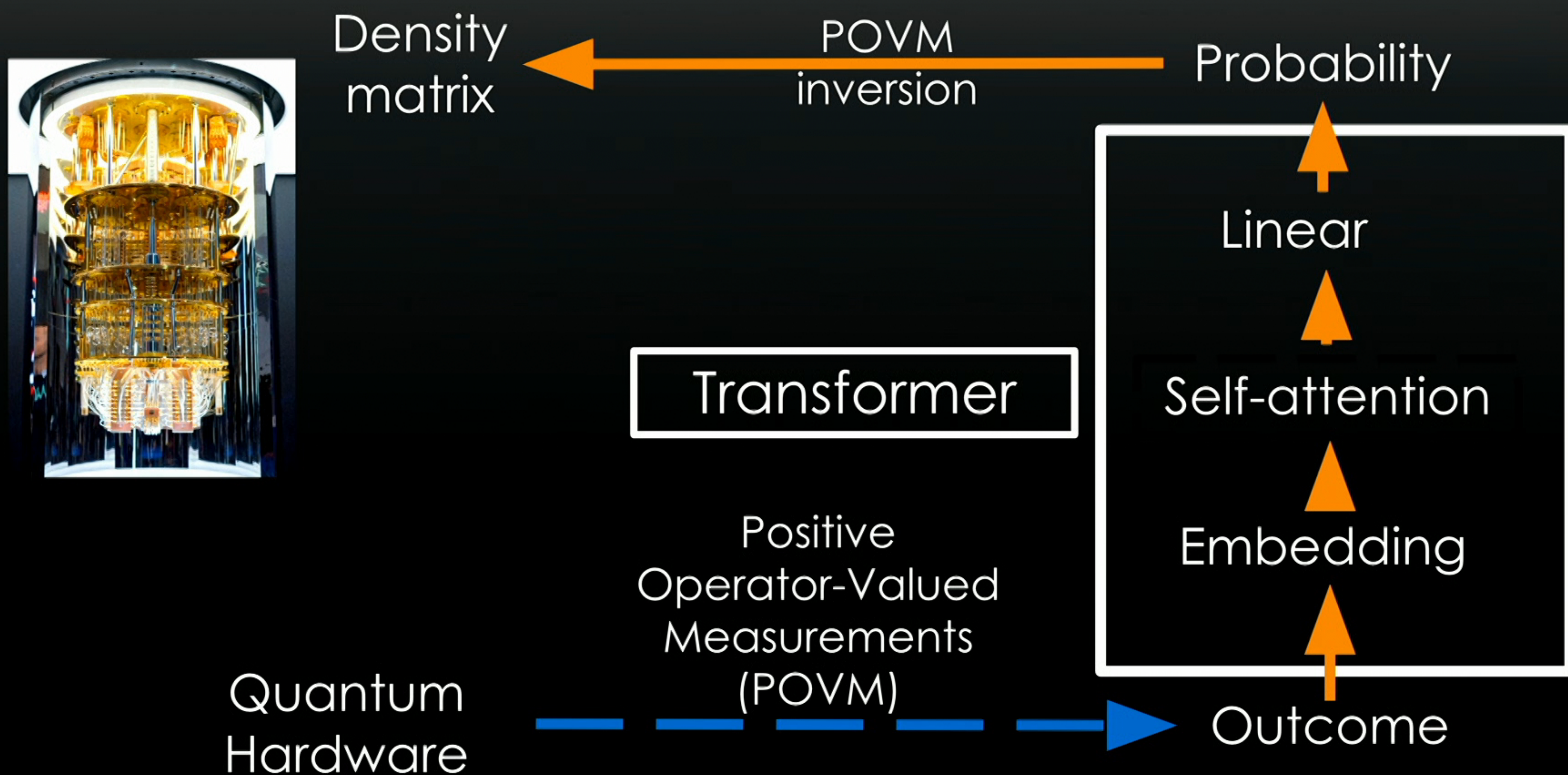
$$S_{ij} = \langle QX_i, KX_j \rangle$$

$$A_{ij} = \frac{\exp(S_{ij})}{\sum_{ik} \exp(S_{ik})}$$

$$\hat{X}_i = \sum_j A_{ij} V_j$$

- Q,K,V are to be learned.
- High, non-local correlations are used.

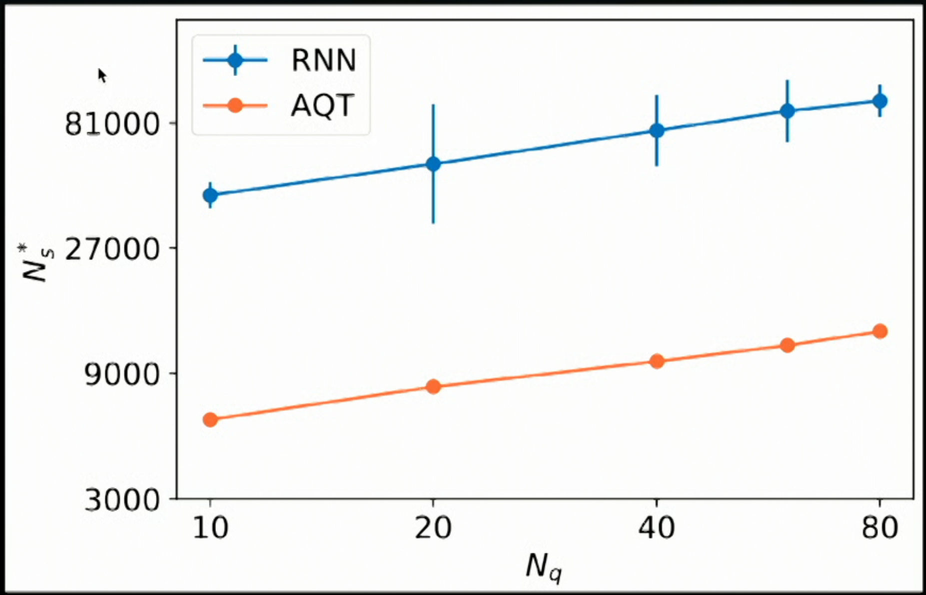
Attention-based Quantum Tomogrphy (AQT)



Best among ML based tomographies

$$|\psi_{GHZ}\rangle = \frac{1}{\sqrt{2}}(|0\rangle^{\otimes N} + |1\rangle^{\otimes N})$$

- Classical Fidelity $F_c(p_0, p_1) = \sum_{\vec{a}} \sqrt{p_0(\vec{a})p_1(\vec{a})}$
- Minimum sample size for Fc=0.99



Comparable to MLE tomography

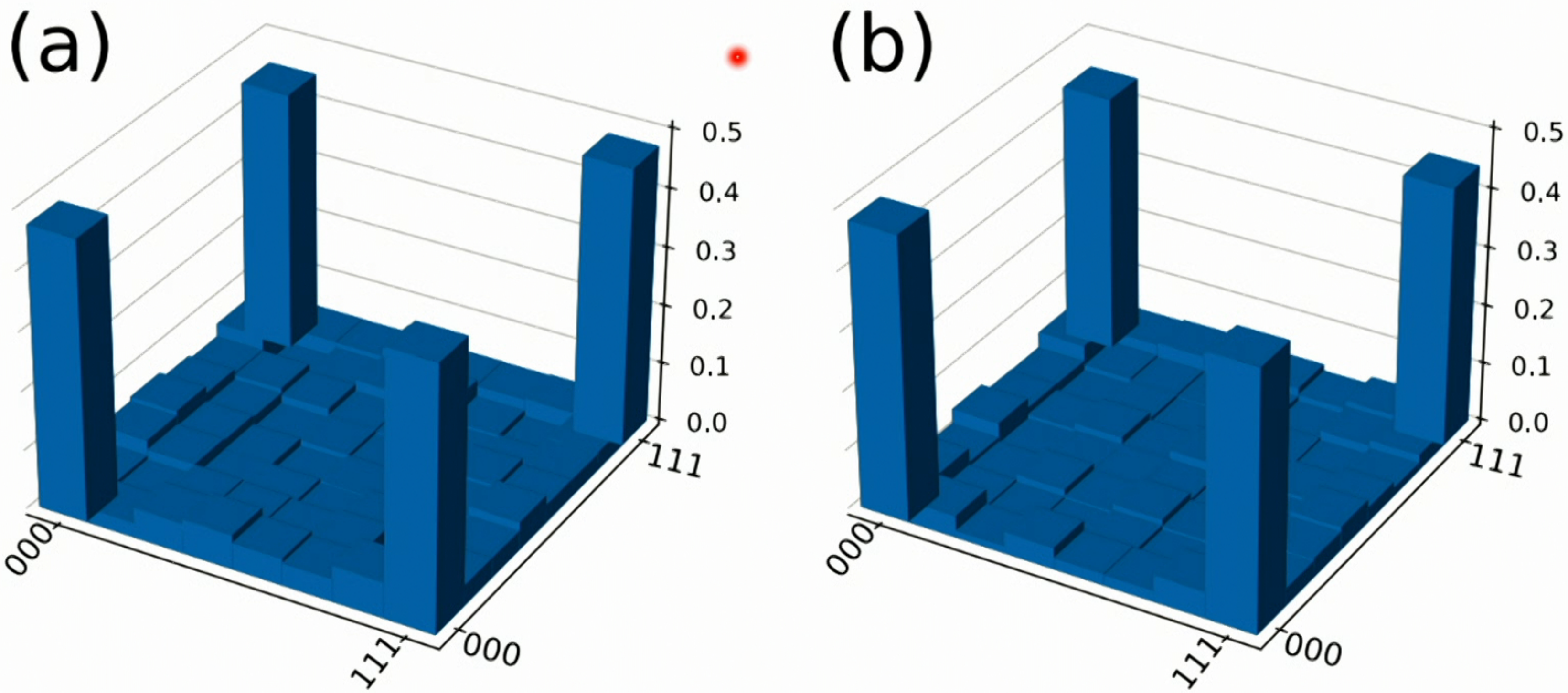
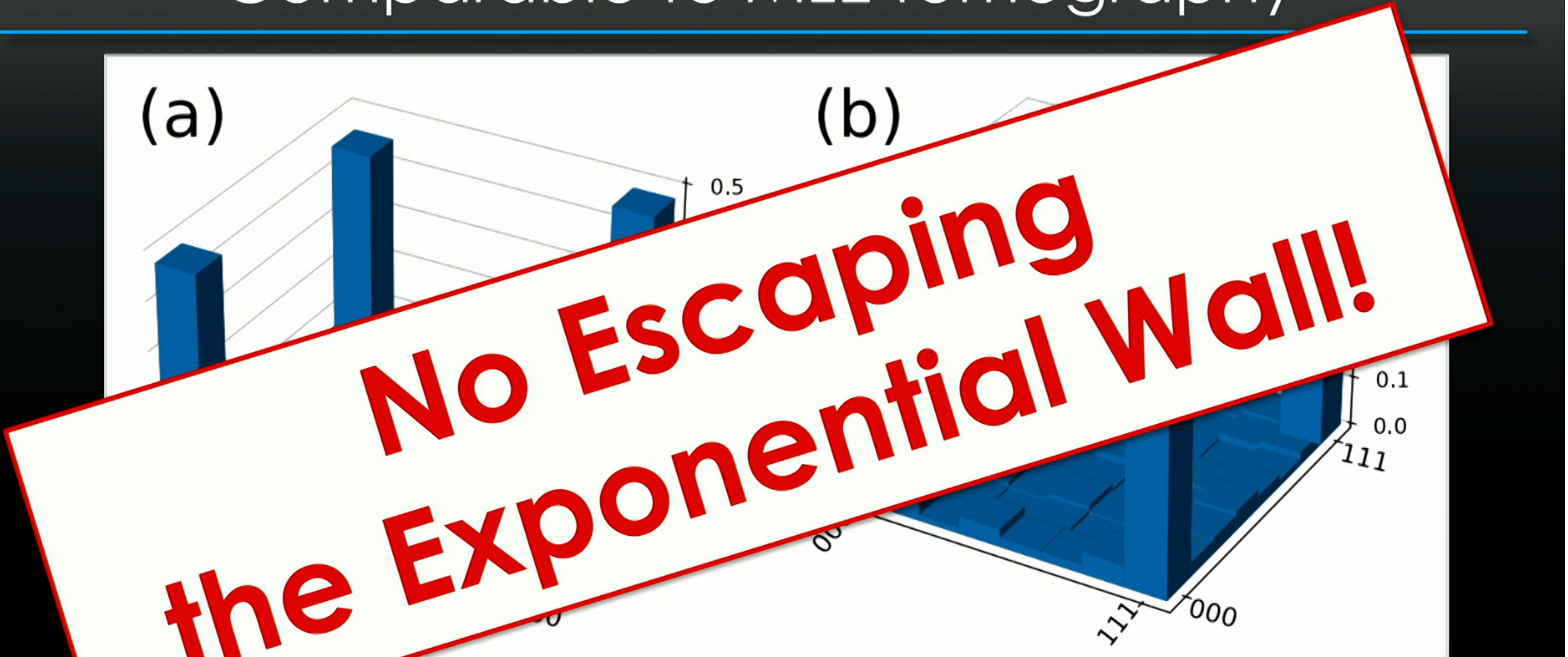


Figure 3. Benchmarking AQT (a) to MLE tomography offered by IBM's Qiskit library (b) for a noisy 3-qubit GHZ state data generated on the IBMQ_OURENSE quantum computer. Each bar represents the absolute value of a density matrix (DM) element.

Comparable to MLE tomography



…rking AQT (a) to MLE tomography offered by IBM's Qiskit library (b) for a noisy 3-qubit GHZ state data …on the IBMQ_OURENSE quantum computer. Each bar represents the absolute value of a density matrix (DM) element.

Take 2. Phase recognition from single instances: Ordered Phases (classical)

MACHINE LEARNING DISCOVERY OF NEW PHASES IN PROGRAMMABLE RYDBERG QUANTUM SIMULATOR SNAPSHOTS

PRR 5, 013026 (2023)

Cole Miles

K. Winberger
(Cornell, CS)

R. Samajdar
(Harvard)

S. Sachdev
(Harvard)

S. Ebadi
(Harvard)

T. Wang
(Harvard)

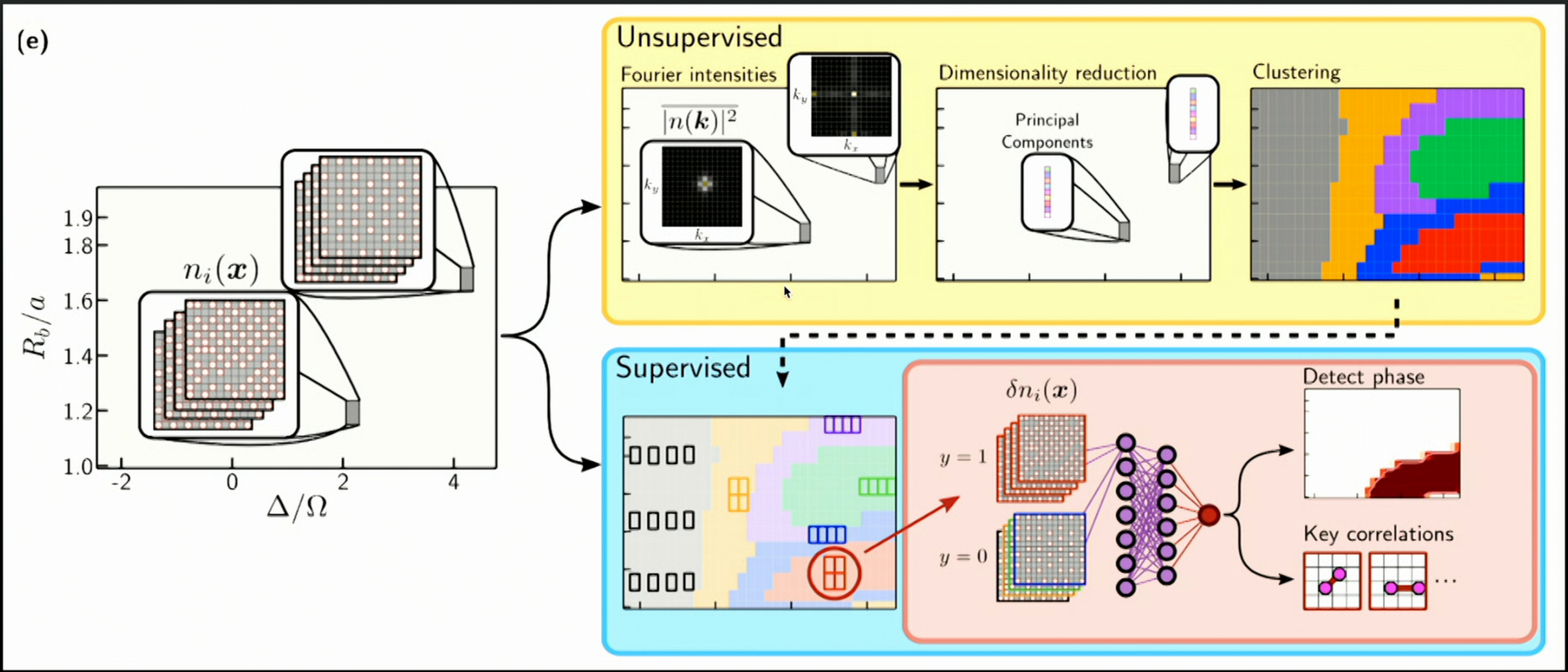
H. Pichler
(Innsbruck)

M. Greiner
(Harvard)

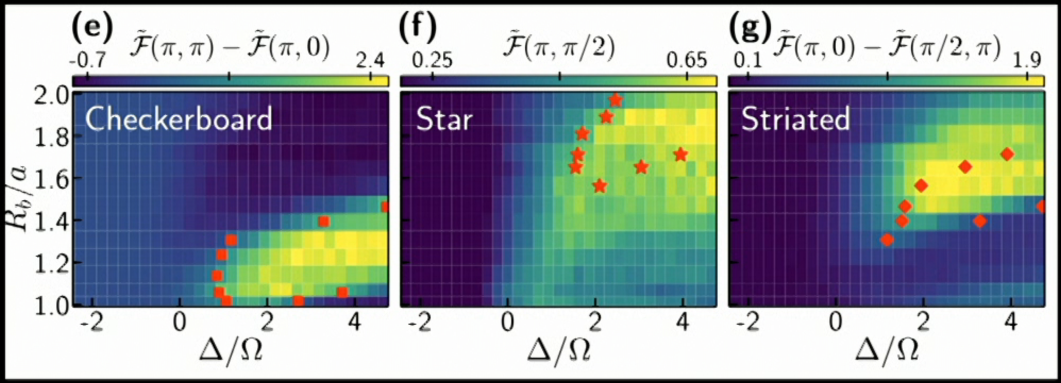
M. Lukin
(Harvard)

DOE CSGF

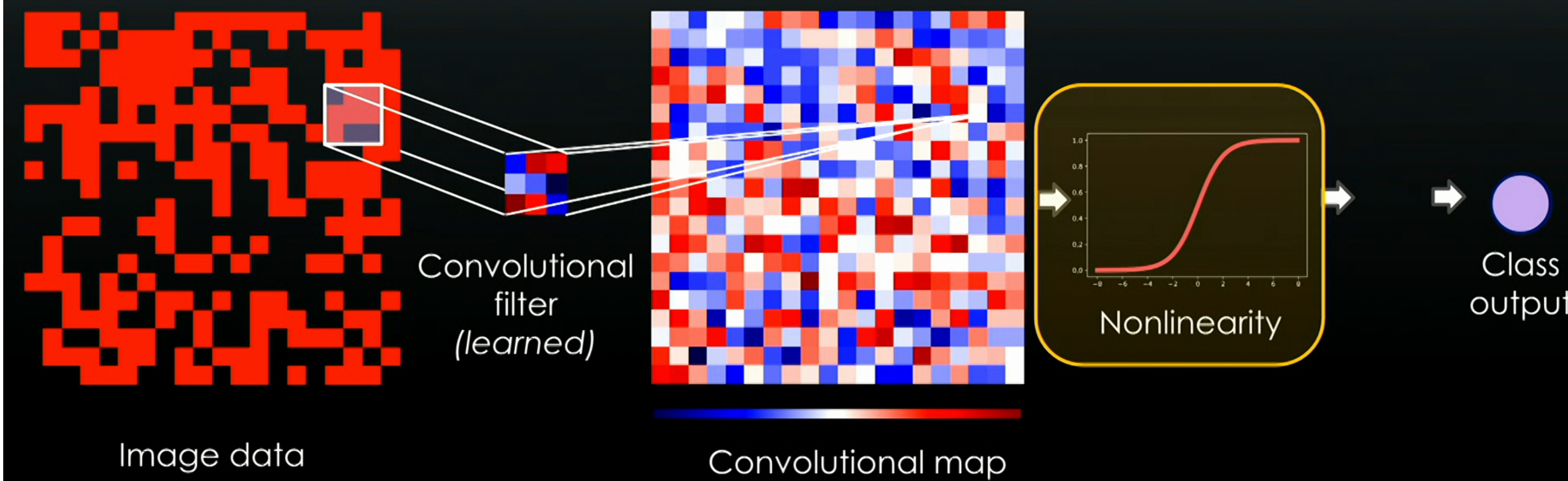
Hybrid-CCNN


(e)
Unsupervised
Fourier intensities
Dimensionality reduction
Principal Components
Clustering
Supervised
Detect phase
Key correlations
$n_i(\boldsymbol{x})$
$\overline{|n(\boldsymbol{k})|^2}$
$\delta n_i(\boldsymbol{x})$
$y = 1$
$y = 0$
R_b/a
Δ/Ω
1.9
1.8
1.6
1.4
1.2
1.0
-2
0
2
4

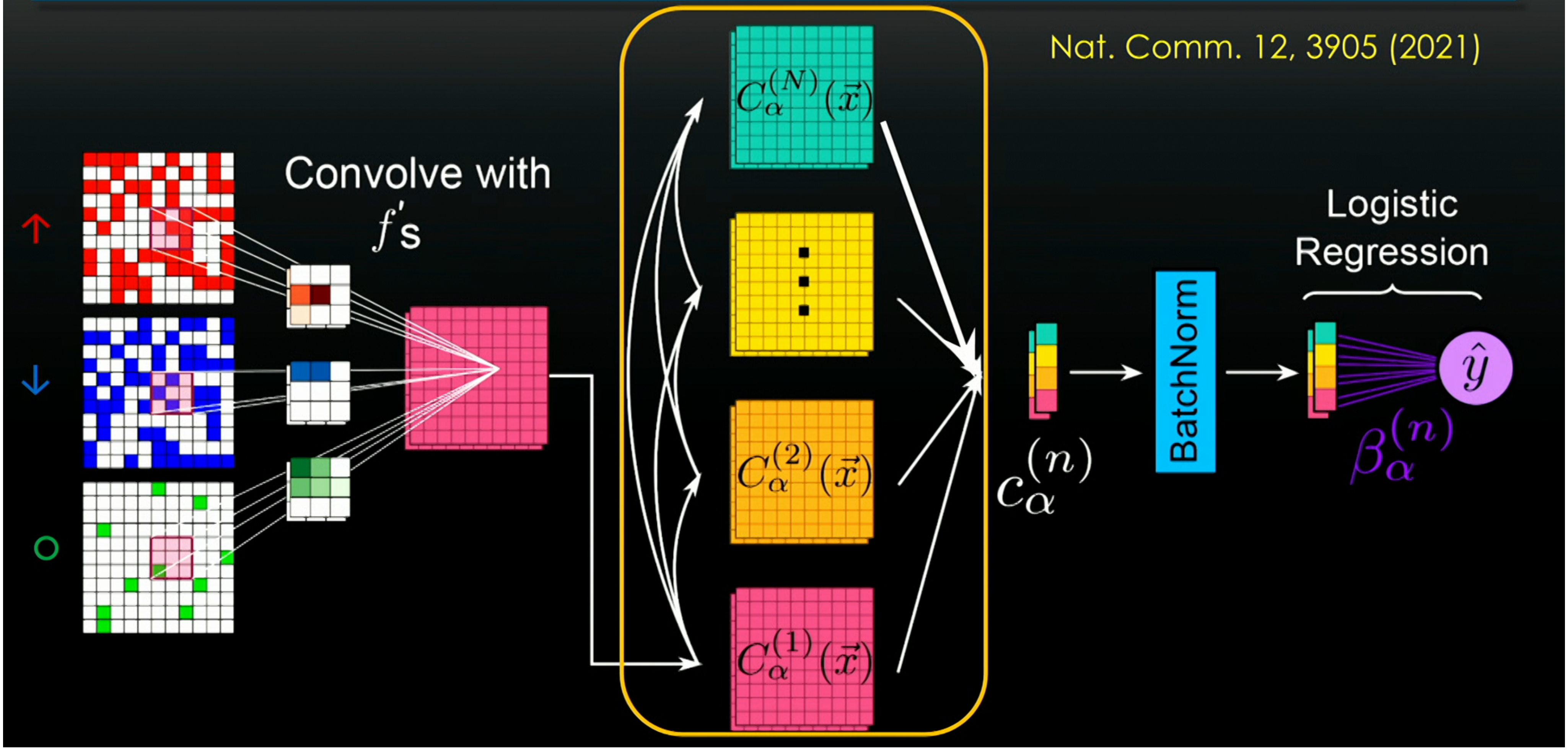
Identification of Phases



Convolutional Neural Networks (CNN)

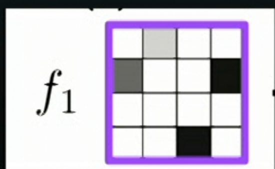


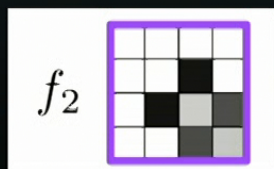
Correlation CNN (CCNN)



Regularization Path Analysis (RPA)

1.Fix learned filters

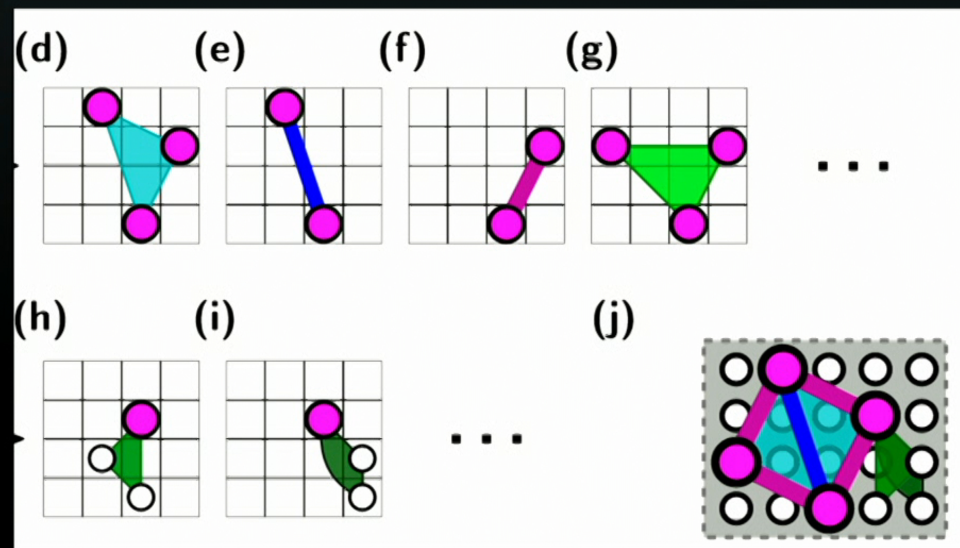




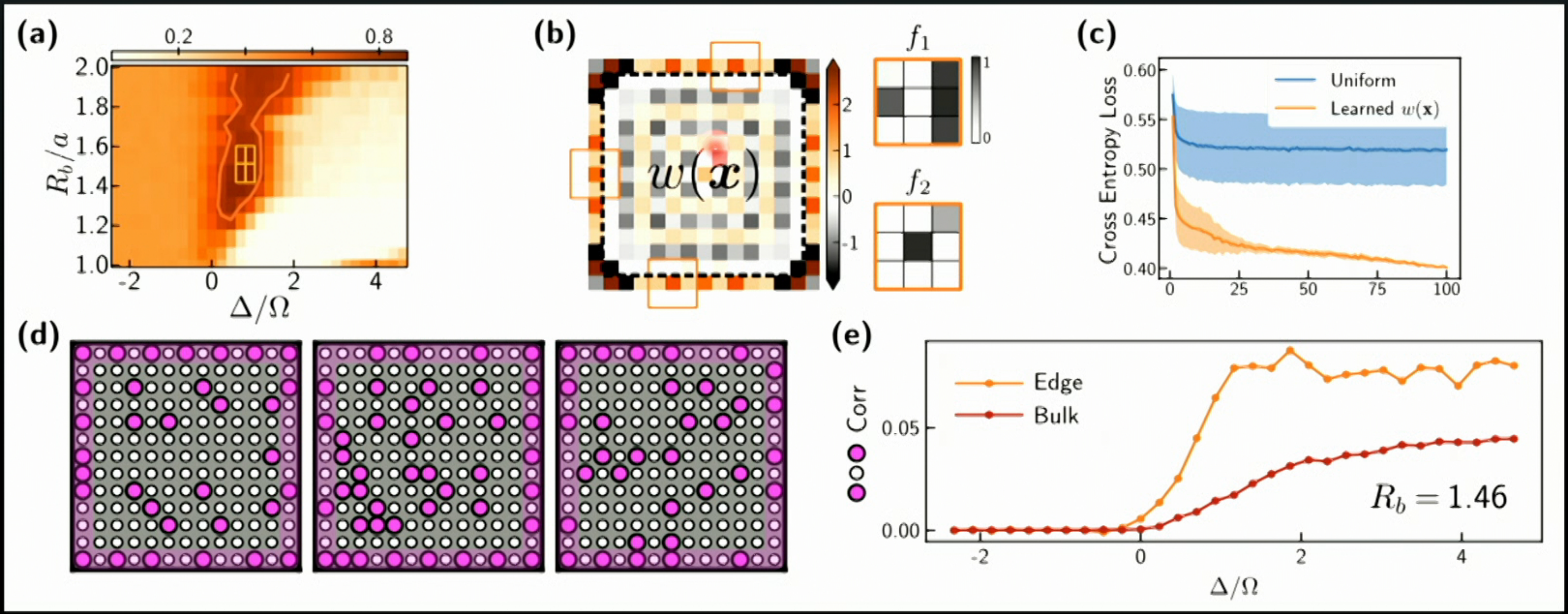
2. Retrain the final layer under new loss

$$\mathcal{L} = \frac{1}{N}\sum_i (-y_i \log \hat{y}_i - (1 - y_i)\log(1 - \hat{y}_i)) + \gamma \sum_{\alpha,\boldsymbol{a}} |f_\alpha(\boldsymbol{a})|,$$

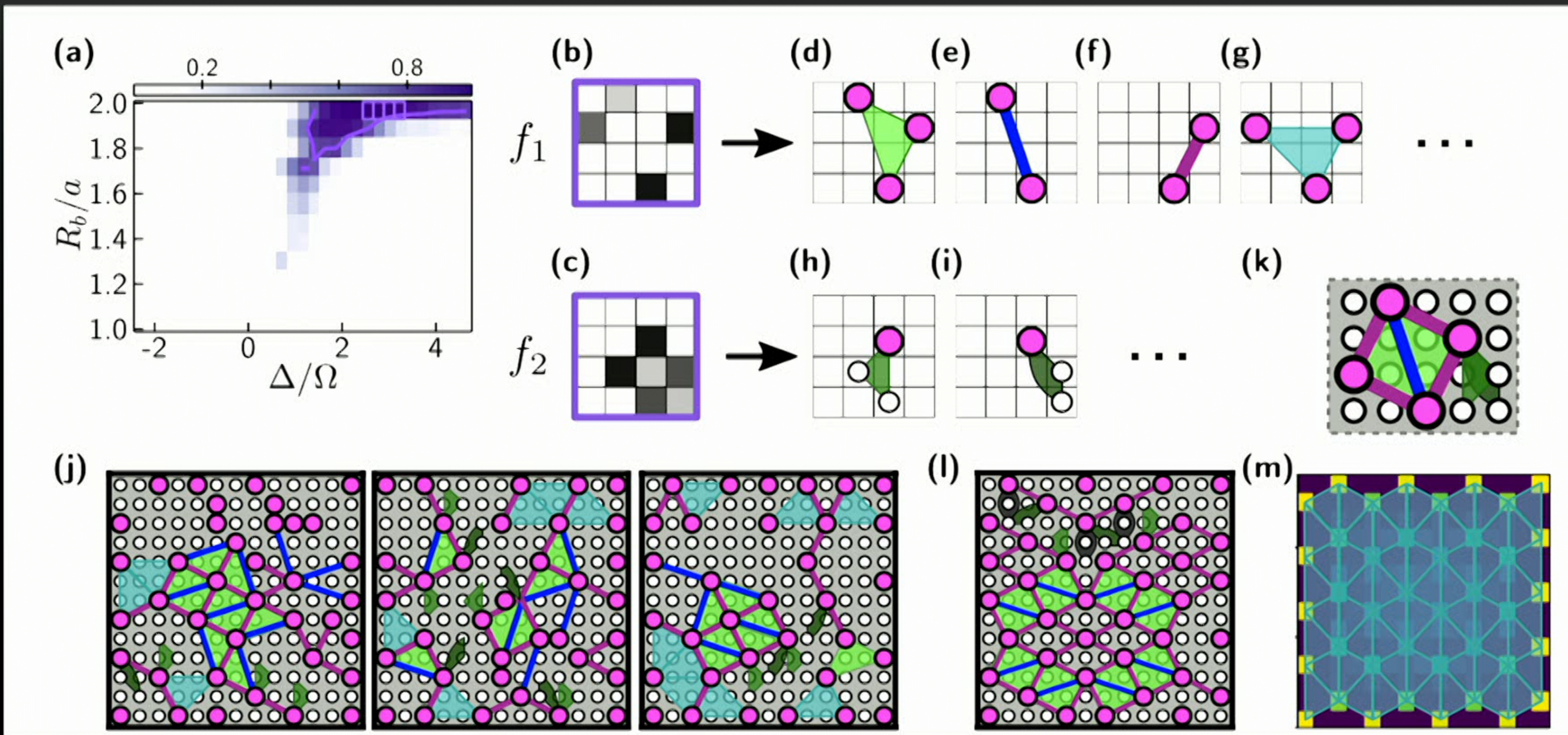
3. Uncover signature motifs



I: Edge-ordering



II. Staggered/Rhombic + Nematic



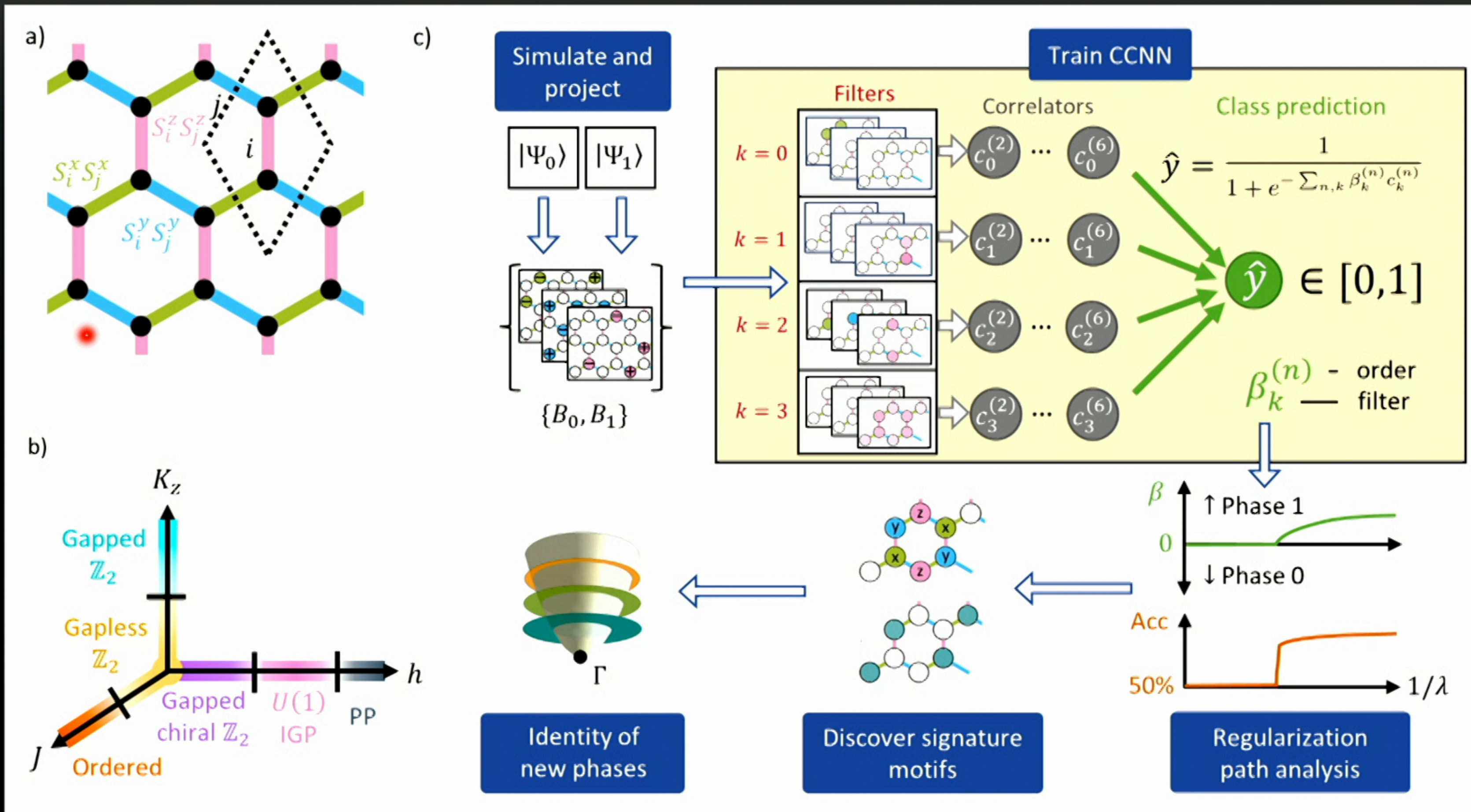
Machine learning feature discovery of spinon Fermi surface

Kevin Zhang

https://engage.aps.org/gds/home

Kevin Zhang, Feng Shi, Yuri Lensky, Nandini Trivedi, Eun-Ah Kim (arXiv:2306.03143)

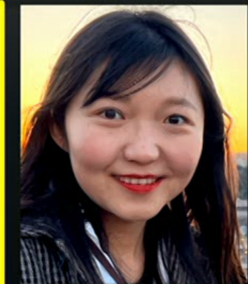
Discovering features of quantum state



Take 3. Learn quantum complexity from a collection of bitstrings

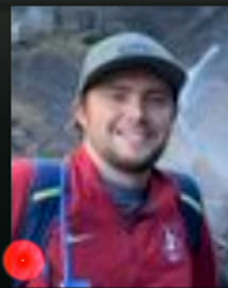
Attention to Quantum Complexity

H. Kim



C. Wan
(Cornell, CS)

K. Varma

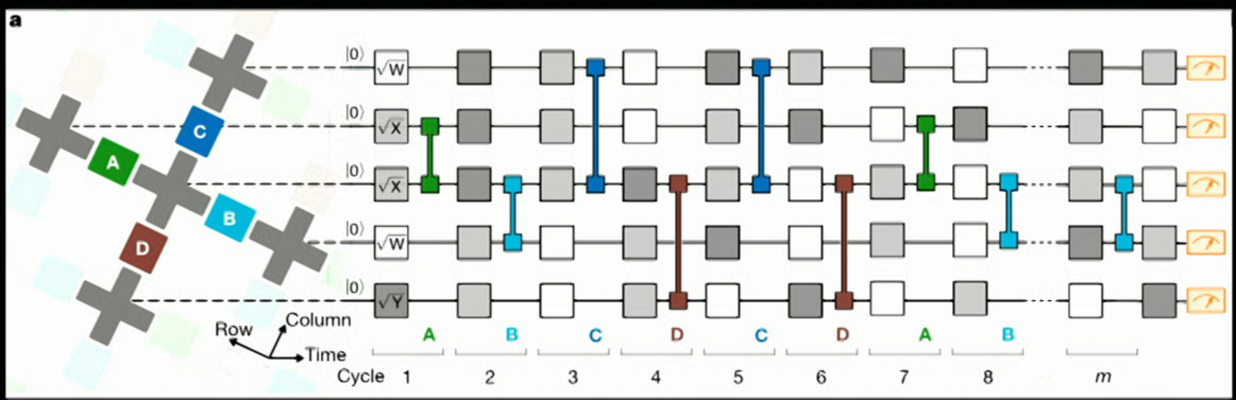


J. Hoke
(Stanford)

Yuri Lensky
(Cornell/Google)

P. Roushan
(Google)

K. Winberger
(Cornell, CS)



H.Kim, EAK et al, in preparation (2023)

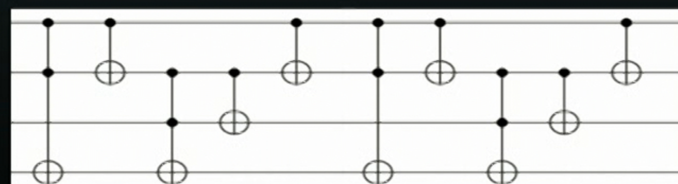
Quantum supremacy using a programmable superconducting processor

Arute et al, Nature 574, 505 (2019)

Complexity of Random Circuits

Minimum no. of operations to construct $|\Psi\rangle$

$|000...\rangle$

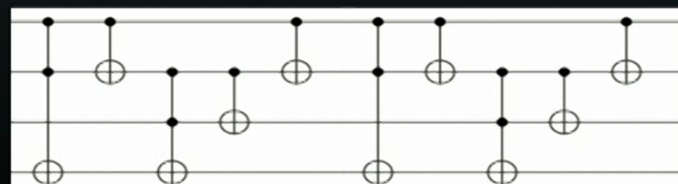


$|\Psi\rangle$

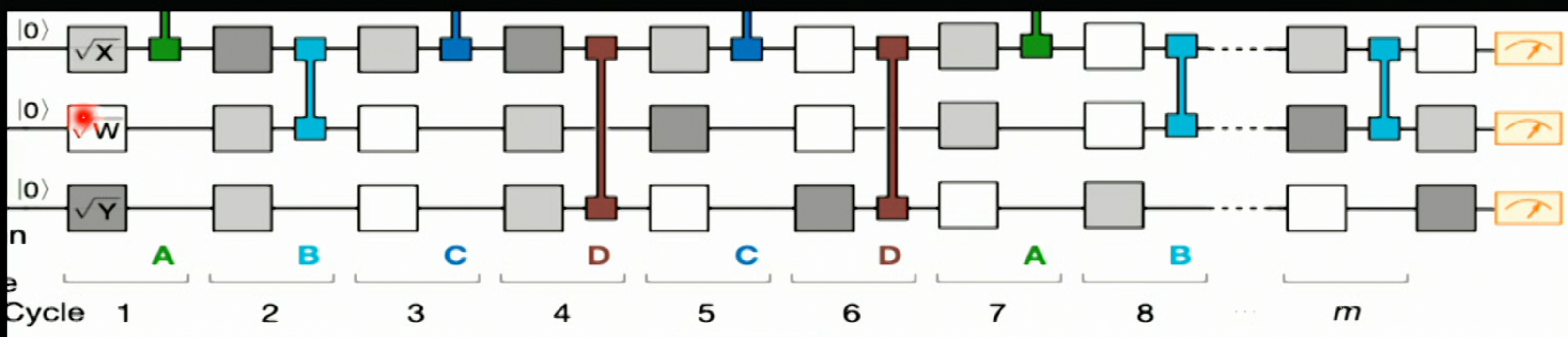
Complexity of Random Circuits

Minimum no. of operations to construct $|\Psi\rangle$

$|000...\rangle$



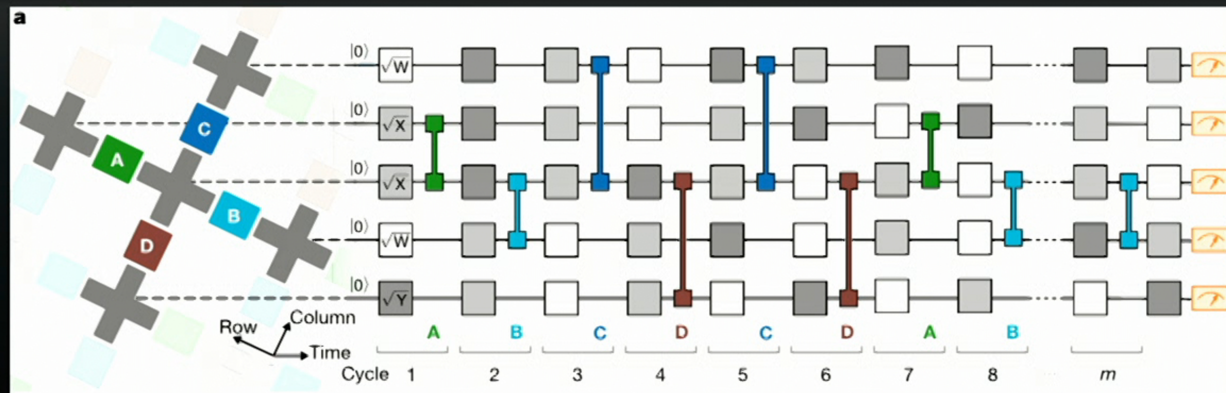
$|\Psi\rangle$



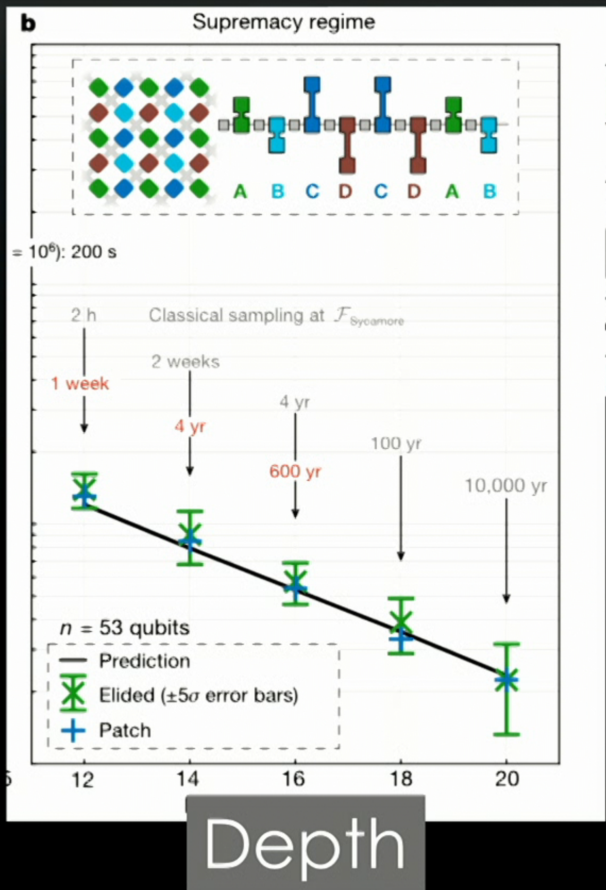
Increasing depth $\Longrightarrow$ **Increasing Complexity**

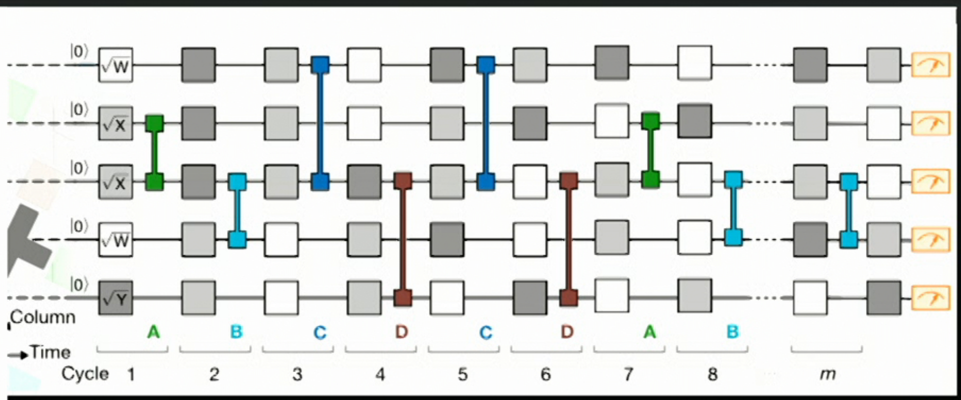
Arute, F., Arya, K., Babbush, R. *et al.* *Nature* 574, 505–510 (2019)

Quantum Supremacy?



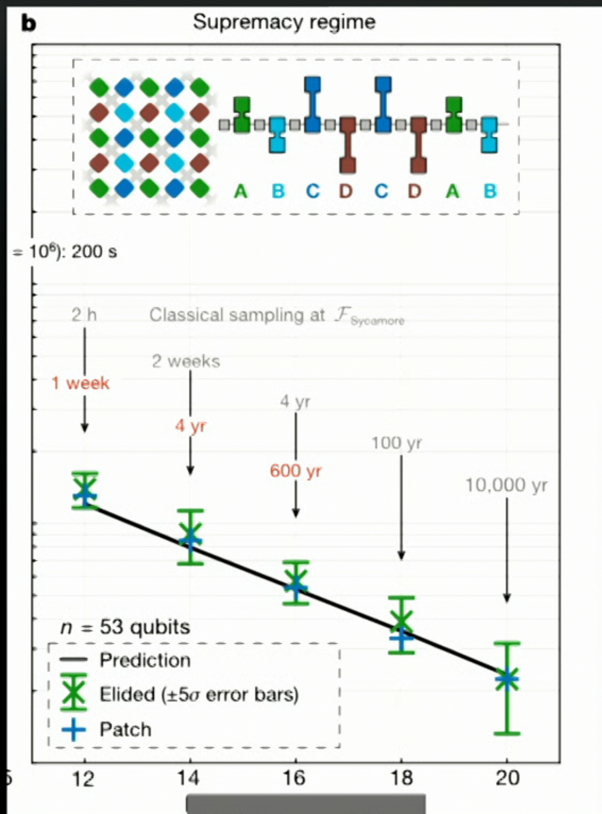
Quantum Supremacy?



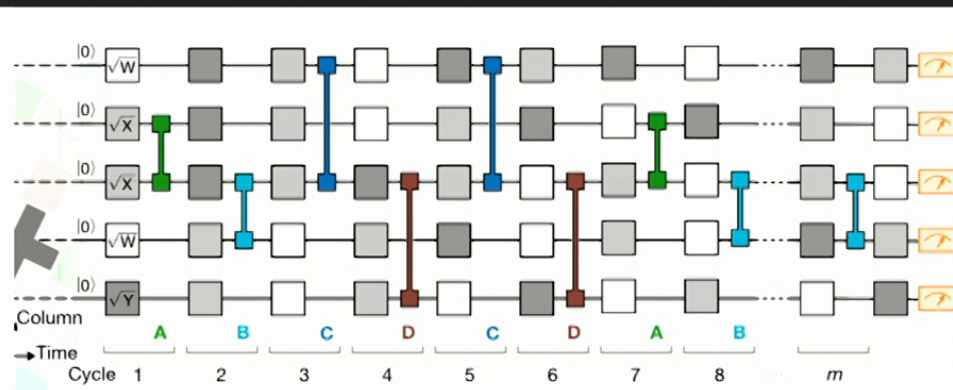


Quantum Supremacy?

$\mathcal{F}_{\mathrm{XEB}}$



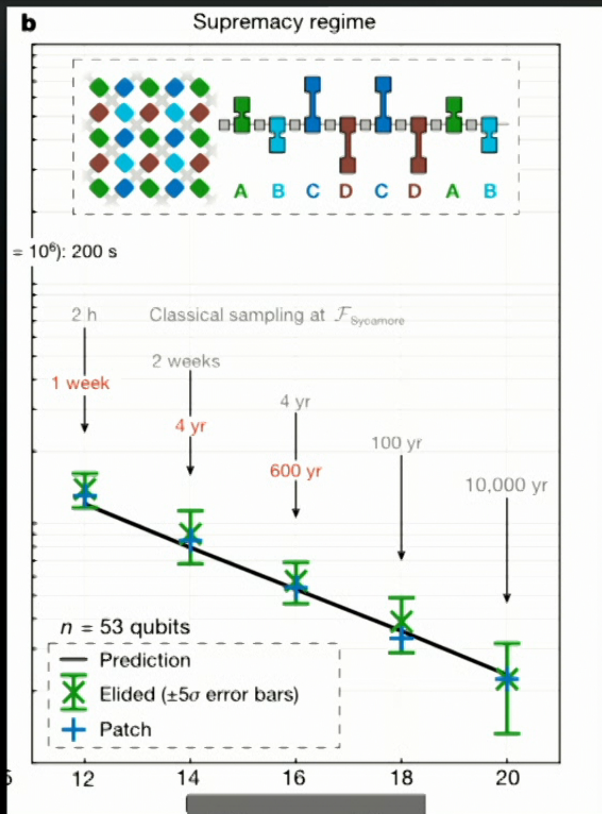
Depth



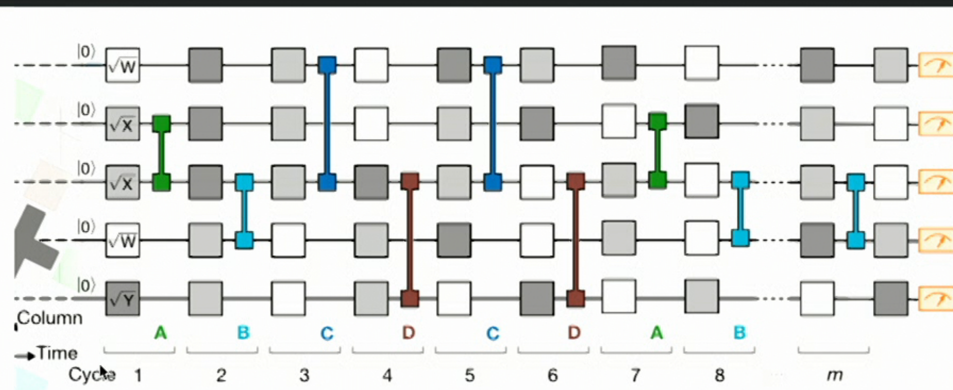
$$\mathcal{F}_{\mathrm{XEB}} = 2^n \langle P(x_i) \rangle_i - 1$$

Quantum Supremacy?

$\mathcal{F}_{\mathrm{XEB}}$



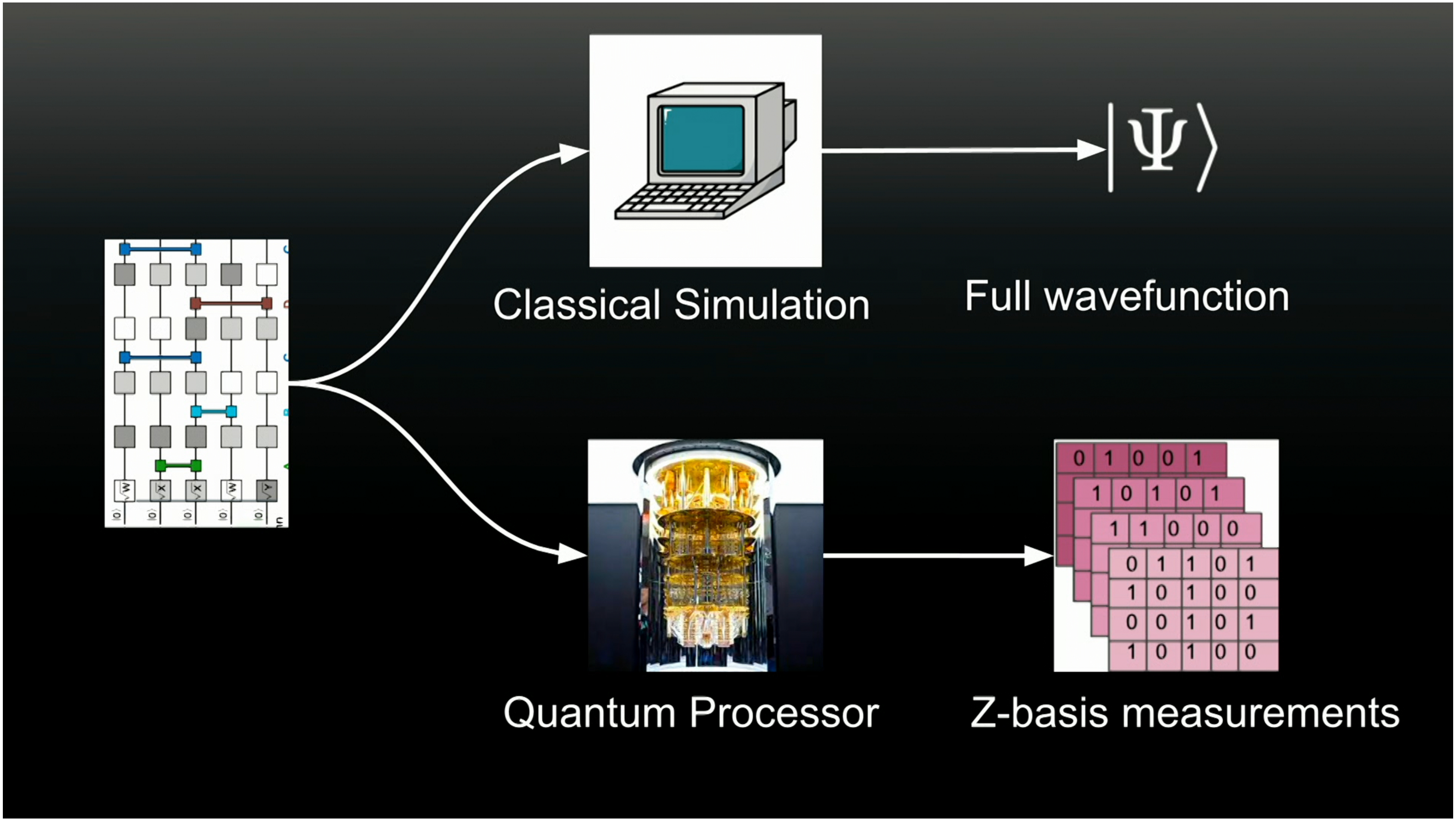
Depth



$$\mathcal{F}_{\mathrm{XEB}} = 2^n \langle P(x_i) \rangle_i - 1$$

Complex :

$$\mathcal{F}_{\mathrm{XEB}} \to 1$$



Classical Simulation
$|\Psi\rangle$
Full wavefunction
Quantum Processor
Z-basis measurements

k-design

- K-design random circuit: match perfect infinite-depth up to k-moments

$$U^{\otimes k} \otimes (U^\dagger)^{\otimes k}$$

k-design

- K-design random circuit: match perfect infinite-depth up to k-moments

$$U^{\otimes k} \otimes (U^\dagger)^{\otimes k}$$

- K is expected to grow with depth and system size.

Self-attention

He lives in the white house.

X1 X2 X3 X4 X5 X6

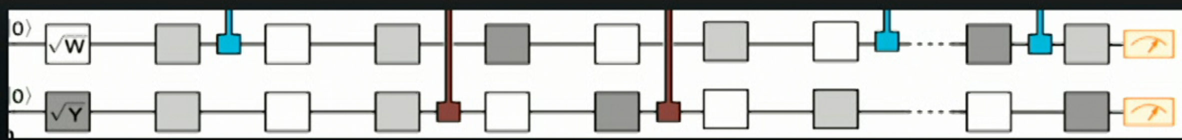
$$S_{ij} = \langle QX_i, KX_j \rangle$$

$$A_{ij} = \frac{\exp(S_{ij})}{\sum_{ik} \exp(S_{ik})}$$

$$\hat{X}_i = \sum_j A_{ij} V_j$$

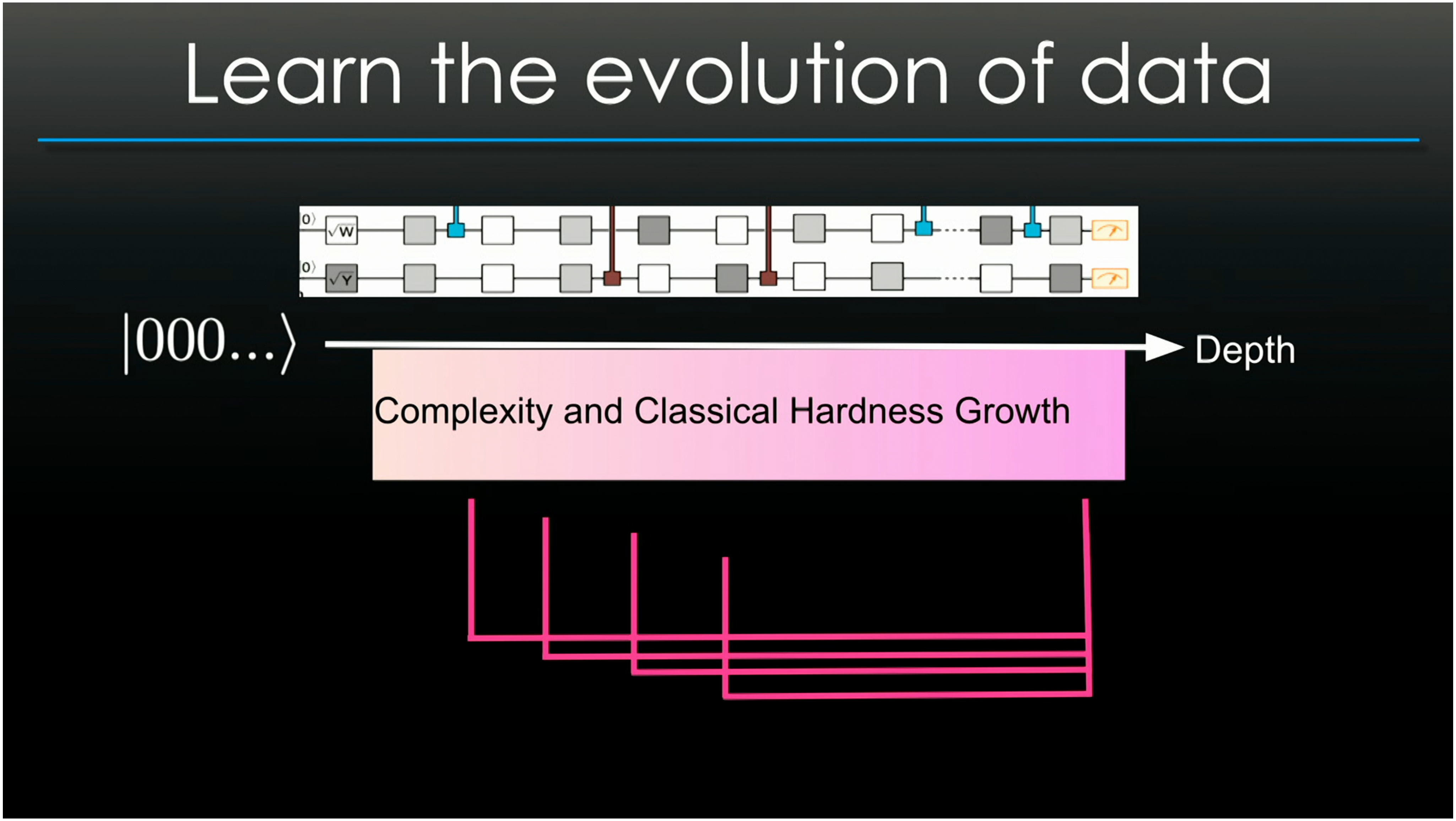
- Q,K,V are to be learned.
- High, non-local correlation are used.

Learn the evolution of data

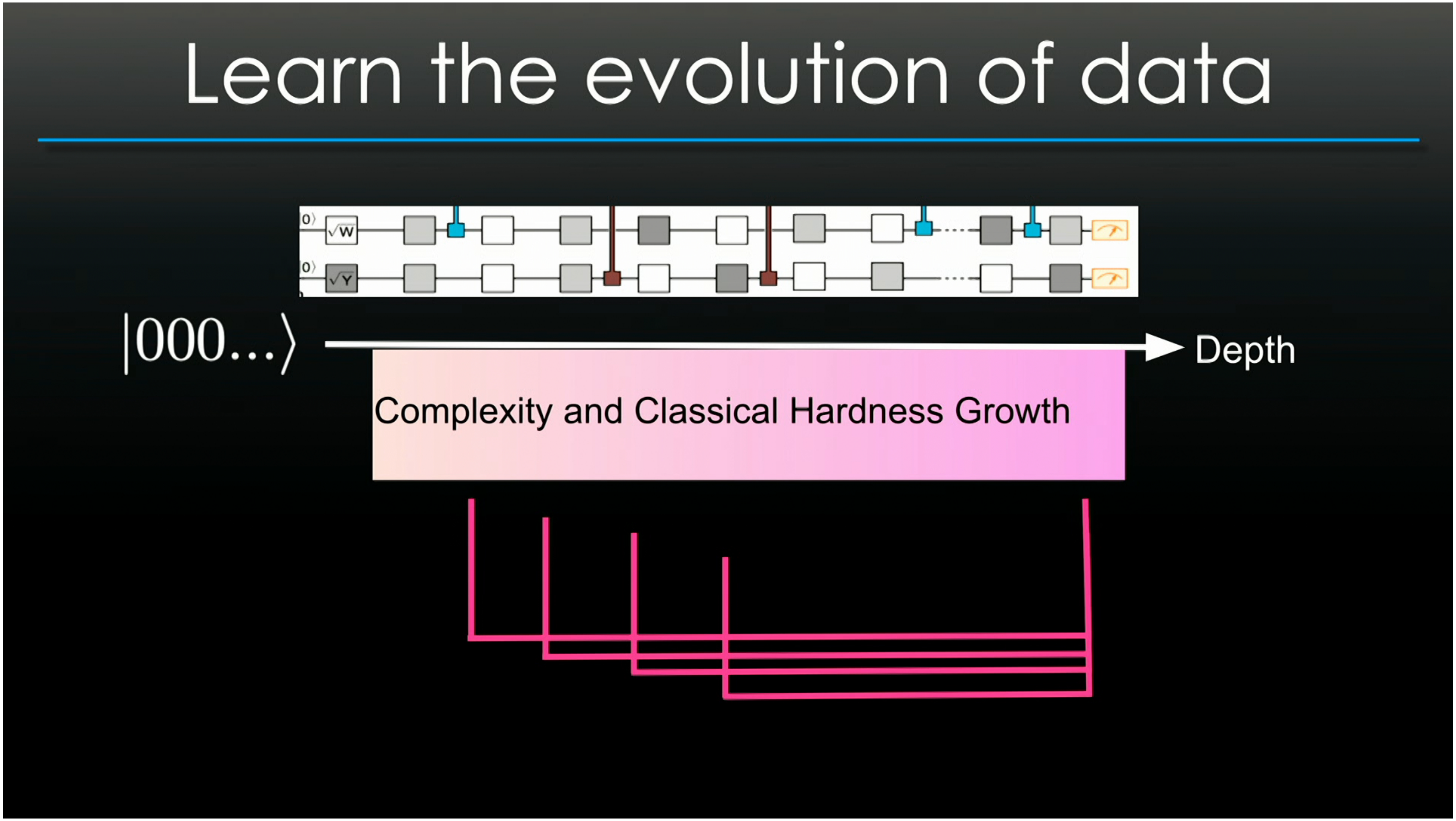


$|000...\rangle$ → Depth

Complexity and Classical Hardness Growth



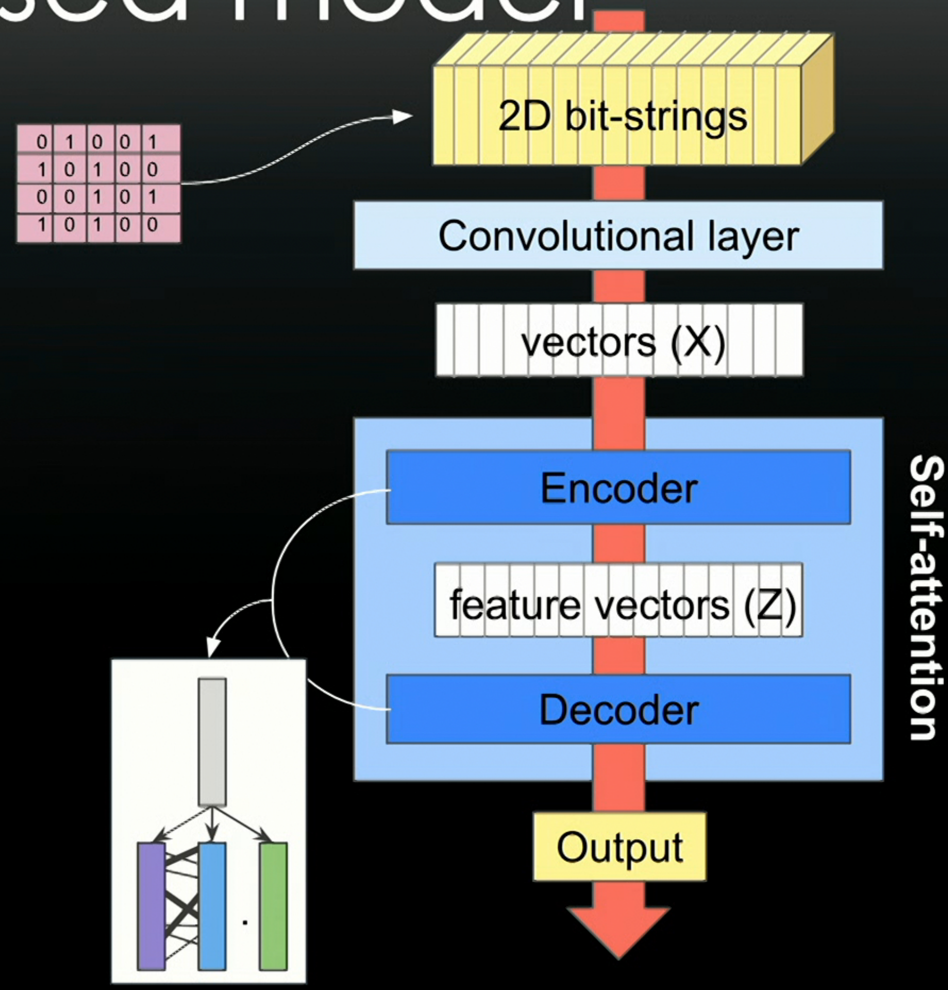
Learn the evolution of data
√W
√Y
|000...⟩
Depth
Complexity and Classical Hardness Growth


Learn the evolution of data
$|000...\rangle$
Depth
Complexity and Classical Hardness Growth

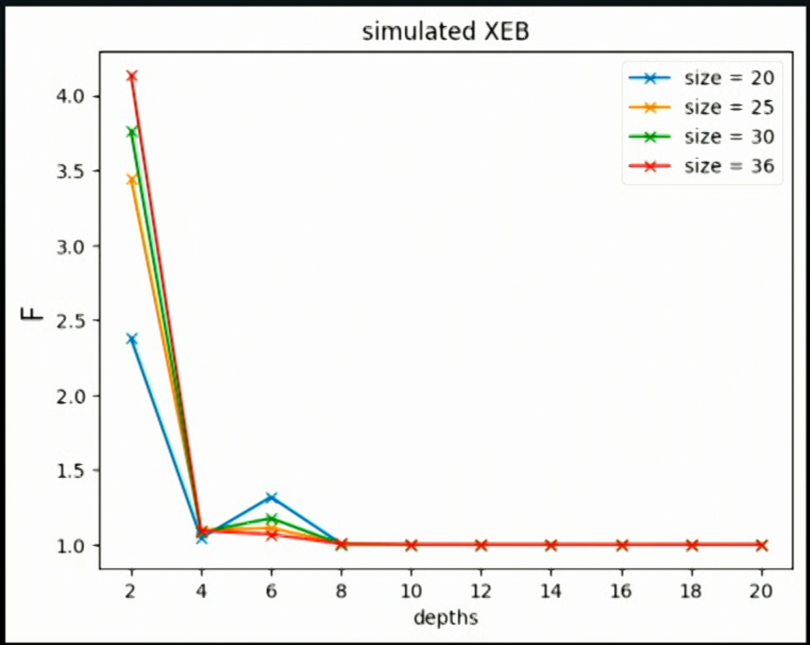
Attention based model

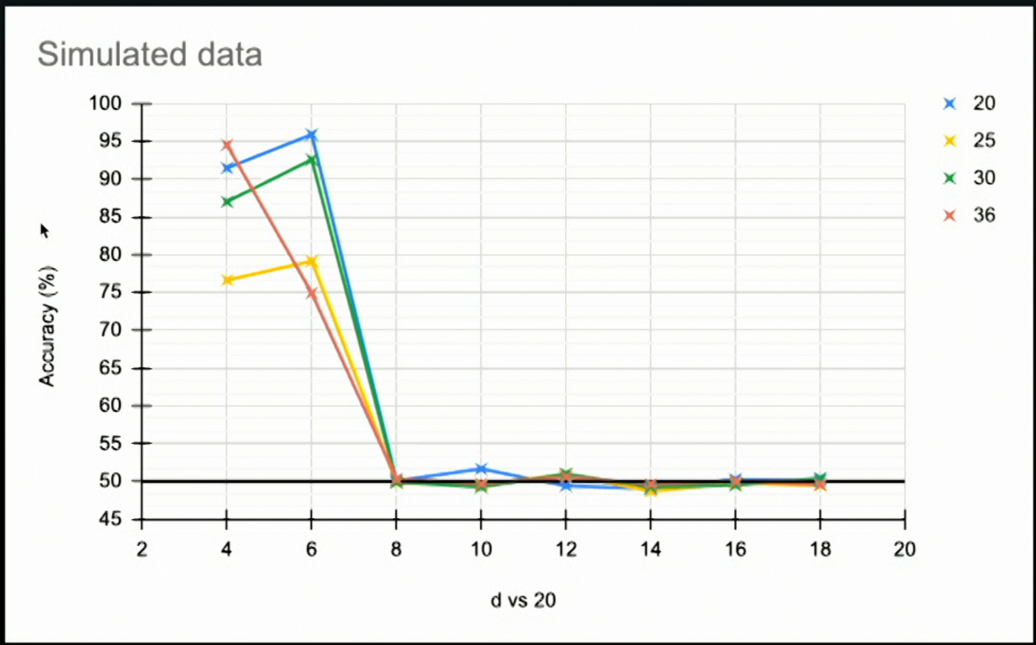
Input: 2D bit-strings

- One convolutional layer
- Attention blocks
- Output: Binary depth classification

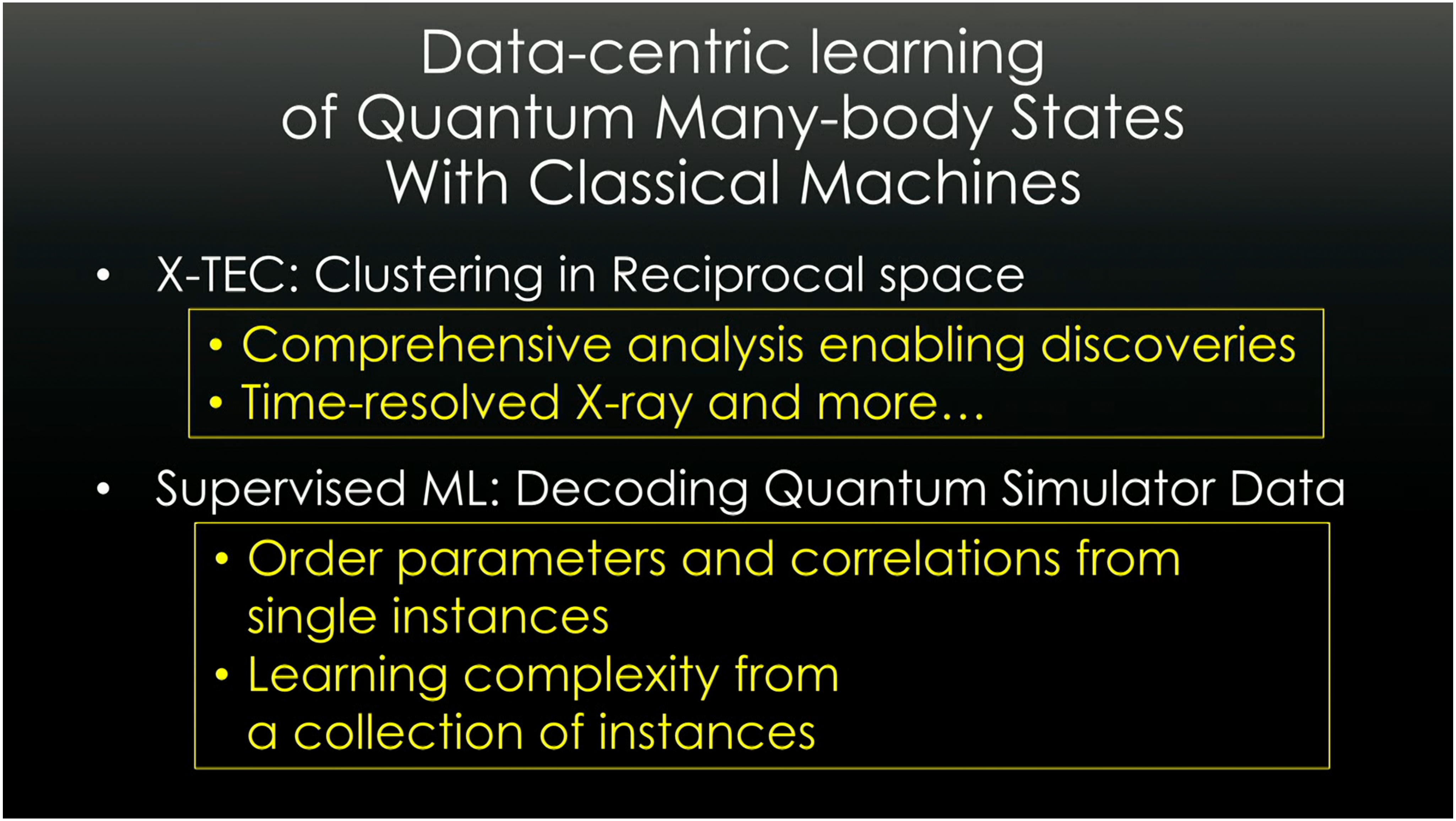


Preliminary results





H.Kim, EAK et al, in preparation (2023)


Data-centric learning
of Quantum Many-body States
With Classical Machines
• X-TEC: Clustering in Reciprocal space
• Comprehensive analysis enabling discoveries
• Time-resolved X-ray and more…
• Supervised ML: Decoding Quantum Simulator Data
• Order parameters and correlations from single instances
• Learning complexity from a collection of instances

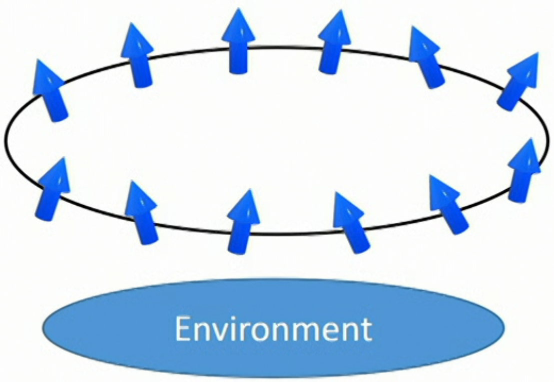
Dorothy Vaughan and West Area Computers becoming first programmers of IBM 7090 in 1961,
From `Hidden Figures', 20th Century Fox (2016)

Dynamics of Open Quantum Systems

Why simulate Open Quantum Systems?

..to discover new physics:

A lack of computational tools prevents the exploration of new interesting physics:

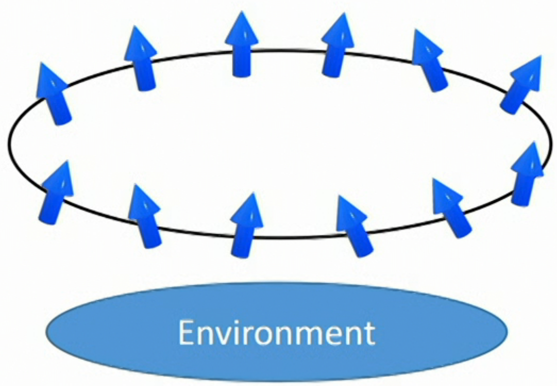


Dynamics of Open Quantum Systems

Why simulate Open Quantum Systems?

..to discover new physics:

A lack of computational tools prevents the exploration of new interesting physics:



..as a benchmarking tool:

Quantum simulators are sensitive to outside noise – require tools to benchmark these devices: