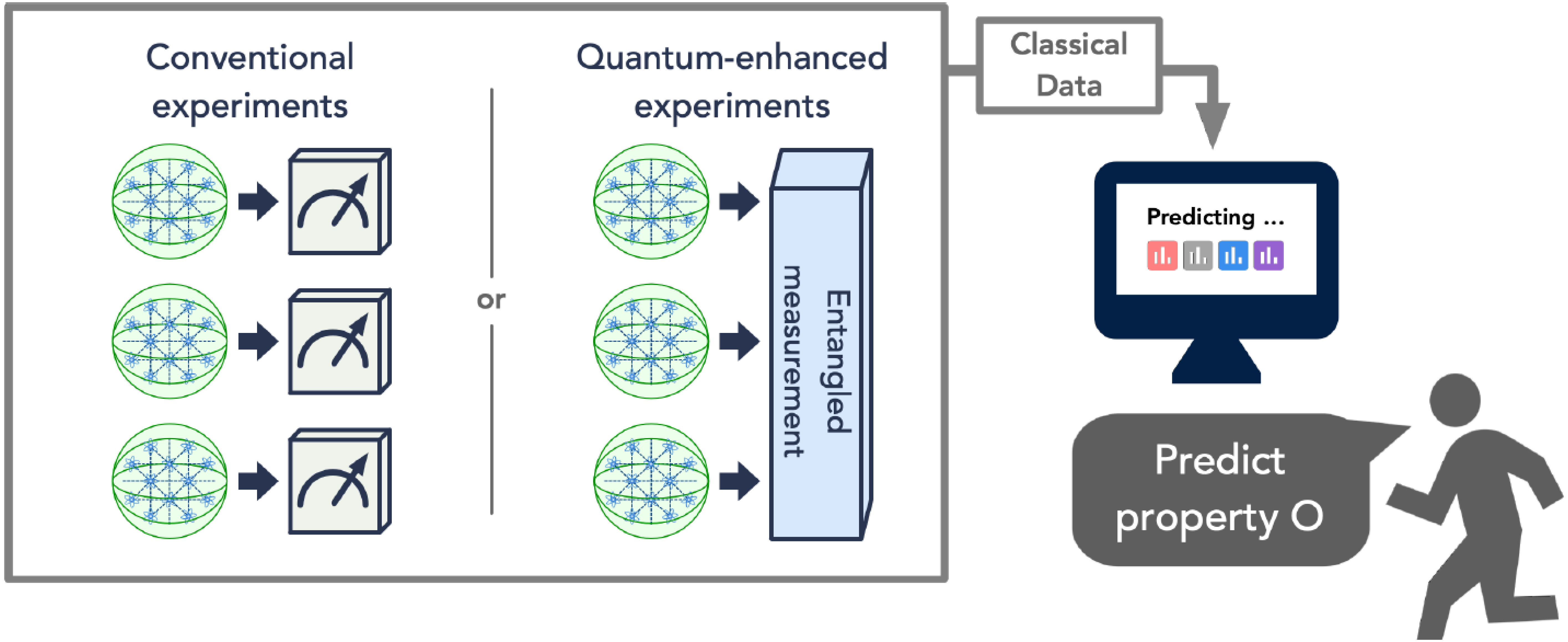
Making Predictions in a Quantum World



aws
IQIM
INSTITUTE FOR QUANTUM INFORMATION AND MATTER

John Preskill
APS March Meeting
6 March 2023

CALIFORNIA INSTITUTE OF TECHNOLOGY 1891
@preskill

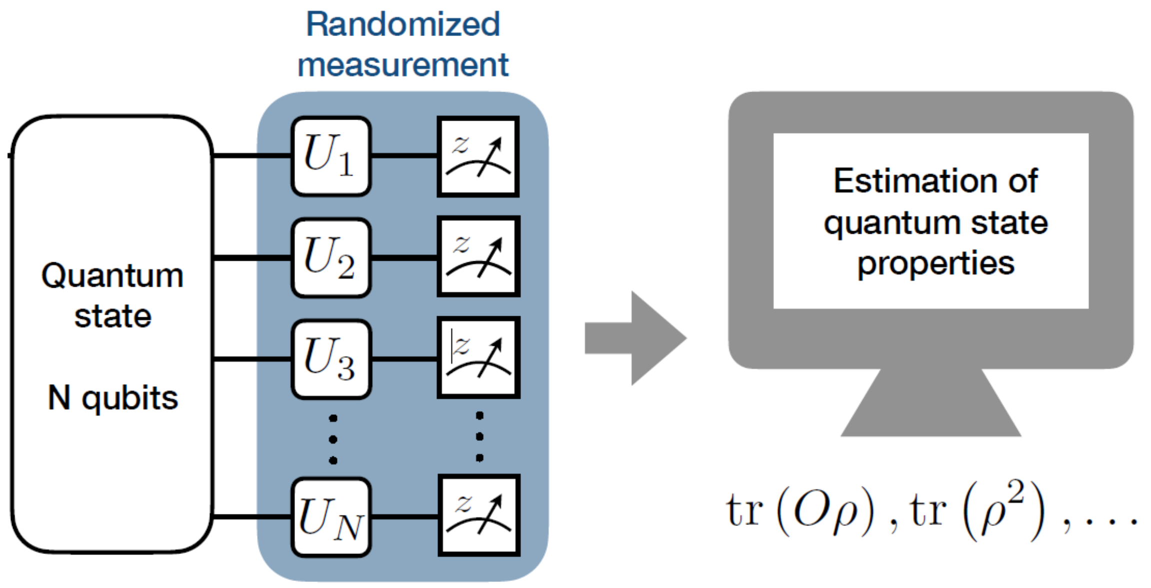
Learning about the quantum world using classical machines

Convert a many-qubit quantum state to a succinct classical description.

Apply classical processing (including machine learning) to the classical description.

Predict properties of exotic quantum systems not previously realized in the lab.

Identify unanticipated quantum phases of matter.



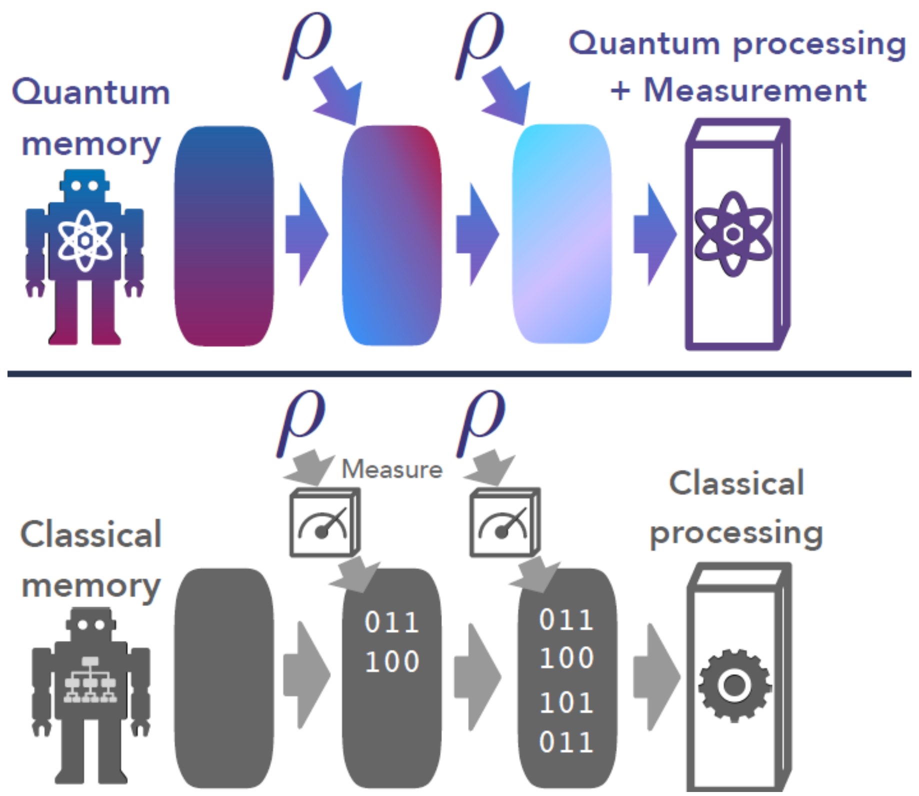
Huang, Kueng, Preskill 2020; Huang, Kueng, Torlai, Albert 2022;
Lewis, Huang, Tran, Lehner, Kueng, Preskill 2023

Learning about the quantum world using quantum machines

Quantum-enhanced measurement strategies: transduce detected quantum data to quantum memory and process it with a quantum computer.

Exponential quantum advantage in learning properties of states and processes.

Unlocking facets of nature that would otherwise remain concealed.



Huang, Kueng, Preskill 2021
Aharonov, Cotler, Qi 2021
Huang, Broughton, Cotler, Chen, Li, Mohseni, Neven, Babbush, Kueng, Preskill, McClean 2022

Quantum tomography

- Quantum state tomography:

  Learn a complete representation of the quantum state.

  $(d \times d \text{ matrix}, \ d = 2^n)$

- Sample optimal protocol (Haah et al. 2017; O'Donnell & Wright 2016):

  Sample complexity: $\mathcal{O}(d^2/\epsilon^2)$

  Quantum resource: Clifford circuits and computational basis measurements

  Classical storage: $\Omega(d^2)$

  Classical post-processing: $\Omega(d^2)$

Shadow tomography

- Shadow tomography:

  Don't learn the quantum state.

  Directly predict M properties: $\mathrm{tr}(O_k \rho), \ \forall 1 \leq k \leq M.$

- Rigorous protocol (Aaronson, Rothblum 2019; Badescu, O'Donnell 2020):

  Sample complexity: $\mathcal{O}(n \log(M)^2 / \epsilon^4)$ ◀ n: system size

  Quantum resource: Quantum memory + exponentially long circuits

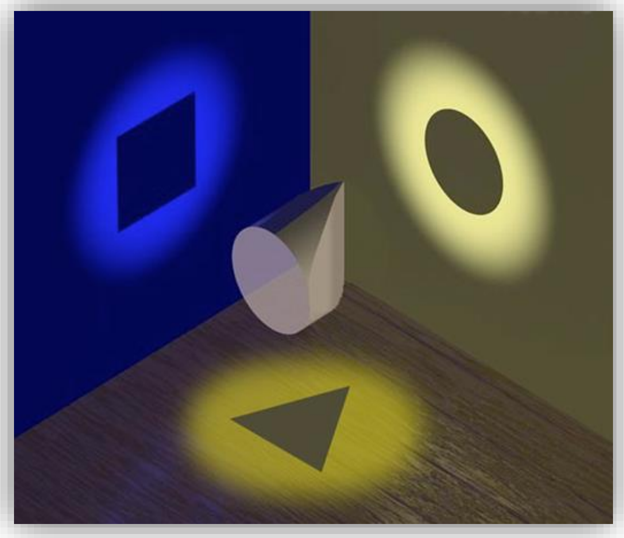
  Storage: $\Omega(d^2)$

  Post-processing: $\Omega(d^2)$

Classical shadows of quantum states

A tractable protocol backed by rigorous theory.

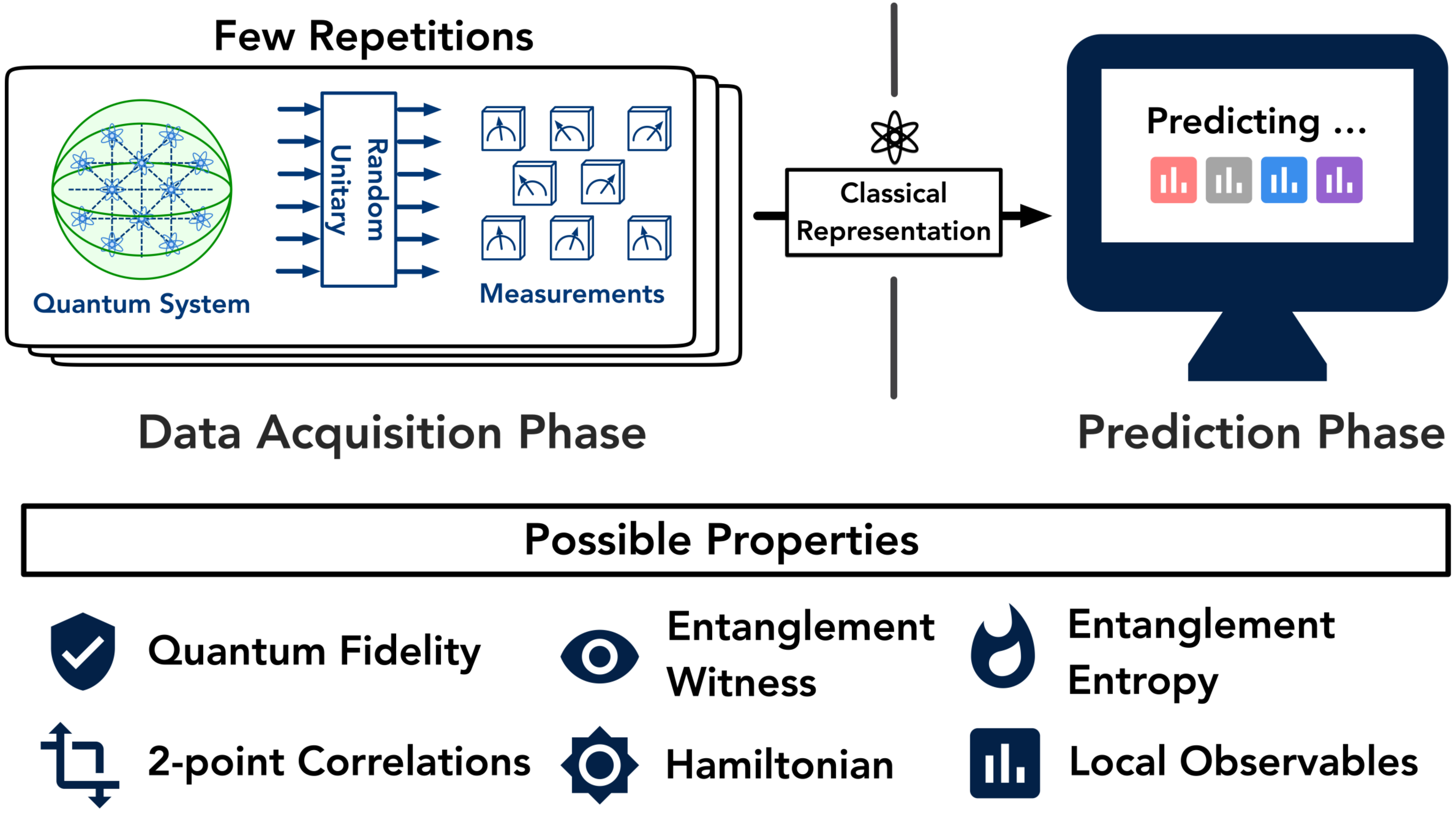
1) A small number of measurements to estimate many properties of a many-qubit quantum state.
2) Succinct classical representations of states, and efficient classical computations for estimates.
3) Rigorous performance guarantees.


classical shadows

Huang, Kueng, Preskill 2020

Classical shadows of quantum states

- Make predictions about a large-scale quantum system from few measurements.

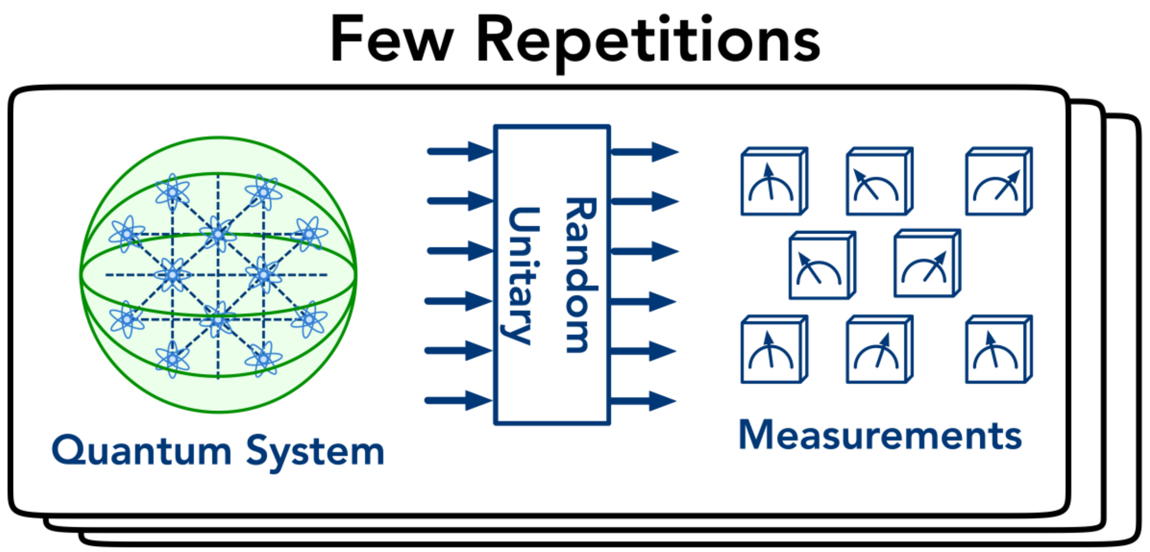


The Procedure:
Data Acquisition Phase

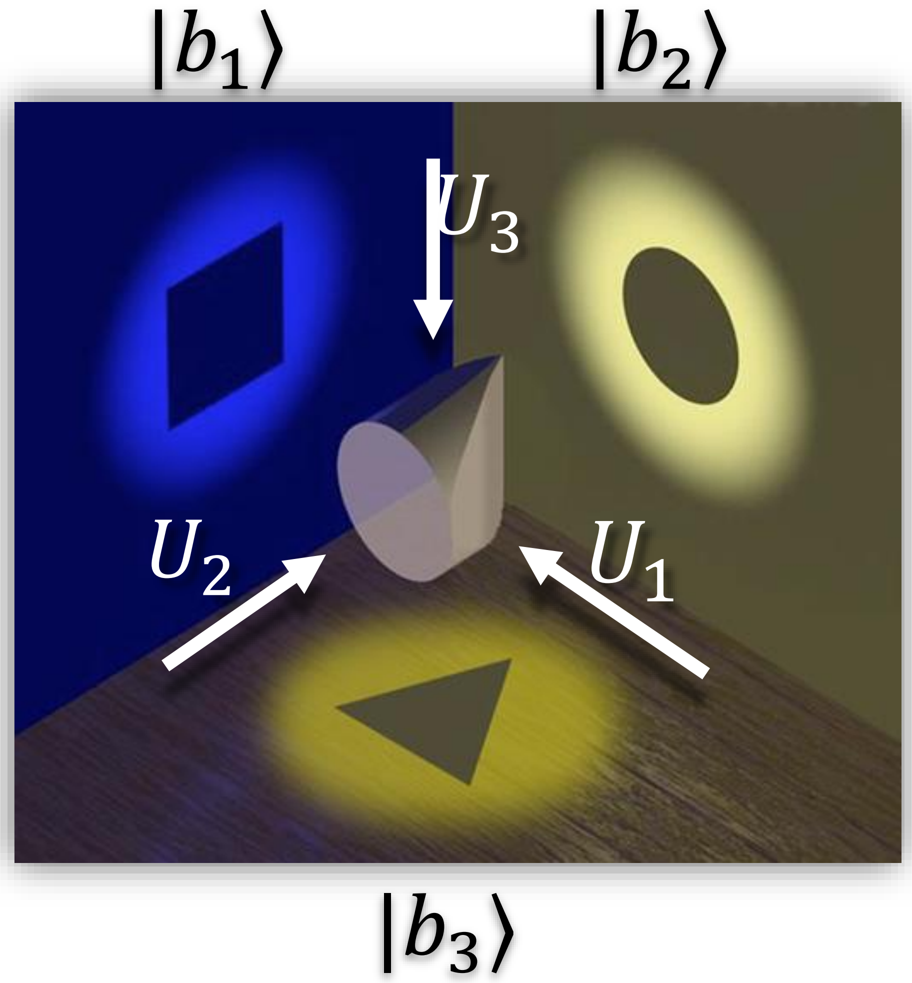
Given multiple copies of n-qubit quantum state ρ and an ensemble of unitary transformations $\{U_i\}$, repeat N times:

- Sample a *random unitary* U_i to rotate the quantum system.
- Measure the system in the computational basis $|b_i\rangle \in \{0,1\}^n$.
- Store the "classical snapshot": $|s_i\rangle = U_i^\dagger |b_i\rangle$.

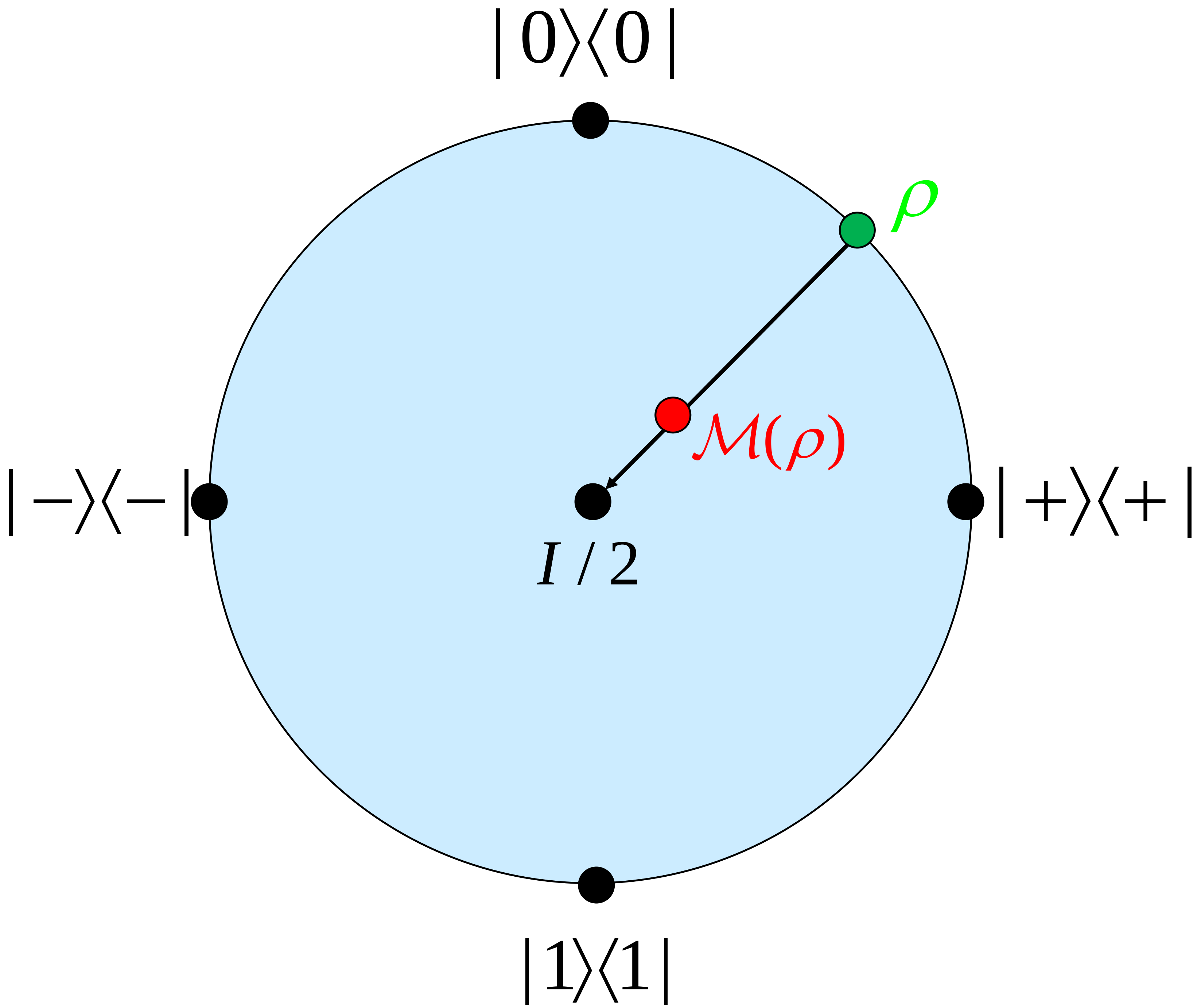
★ $\mathbb{E}[|s_i\rangle\langle s_i|] = \mathcal{M}(\rho)$. ($\mathcal{M}$: some CPTP map)



Data Acquisition Phase



The Procedure:
Data Acquisition Phase



Example: a single rebit

The unitary U is either the identity or a 90 degree rotation of the Bloch sphere.

$$\lim_{N\to\infty} \frac{1}{N} \sum_i |s_i\rangle\langle s_i|$$

$$= \mathbb{E}[\rho]$$

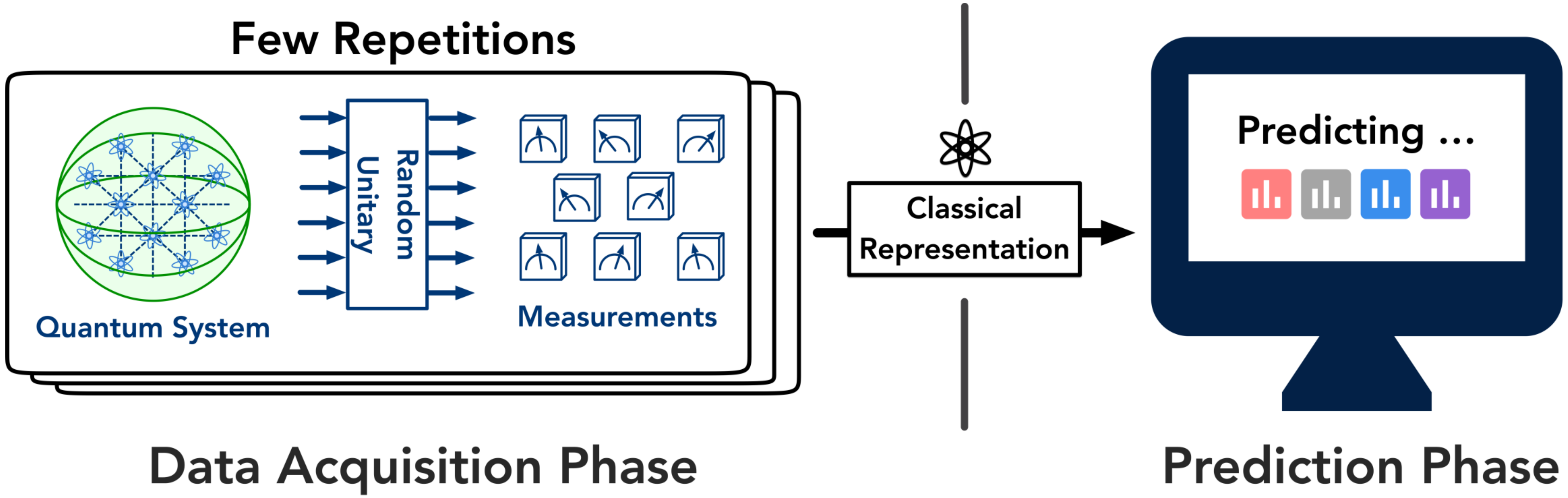
$$= \frac{\rho + I}{3} = \mathcal{M}(\rho)$$

The Procedure: Prediction Phase

Given $S(\rho) = \{|s_1\rangle, \ldots, |s_N\rangle\}$ (the classical shadow),
how to predict properties of the quantum state ρ?

★ $\mathbb{E}[|s_i\rangle\langle s_i|] = \mathcal{M}(\rho)$. ($\mathcal{M}$: some CPTP map)

➡ $\rho = \mathbb{E}[\mathcal{M}^{-1}(|s_i\rangle\langle s_i|)] \Longrightarrow \rho \approx \mathcal{M}^{-1}(|s_i\rangle\langle s_i|)$.



Classical Shadow Theorem

1. Learn a classical representation of an unknown quantum state ρ from $N = \mathcal{O}(B \log(M)/\epsilon^2)$ measurements.
2. Subsequently, given any $O_1, \ldots, O_M$ with $B \geq \max \| O_i \|^2_{\mathrm{shadow}}$, the procedure can use the classical representation to predict $o_1, \ldots, o_M$, where $|o_i - \mathrm{tr}(O_i \rho)| < \epsilon$, for all i.

The shadow norm $\| O \|^2_{\mathrm{shadow}}$ is an upper bound on the variance of our estimator; it depends on the ensemble of unitaries used during the data acquisition phase.

Random Clifford measurement: $\| O \|^2_{\mathrm{shadow}} \leq 3\,\mathrm{tr}(O^2)$ Application: Quantum fidelity $O = |\psi\rangle\langle\psi|$

Random Pauli measurement: $\| O \|^2_{\mathrm{shadow}} \leq 4^w \| O \|^2_{\infty}$ Application: local Hamiltonian $O = H = \sum_a H_a$

Observable O acts on w qubits

Huang, Kueng, Preskill 2020

Local vs. Global

Local (Pauli) measurement:
Low depth, noise resilient, feasible today.
Efficiently predicts local observables.

Global (Clifford) measurement:
Depth scales with system size
Not currently feasible for large systems.
Predicts (some) global observables efficiently.

In between:
Scrambling circuits of intermediate depth.
Does not require local control.
Hu, Choi, You 2021

Noise Robustness

The randomized protocol “twirls” the noise.

It becomes a Pauli channel, which can be efficiently characterized.

Include noise in the channel inversion, yielding unbiased estimators.

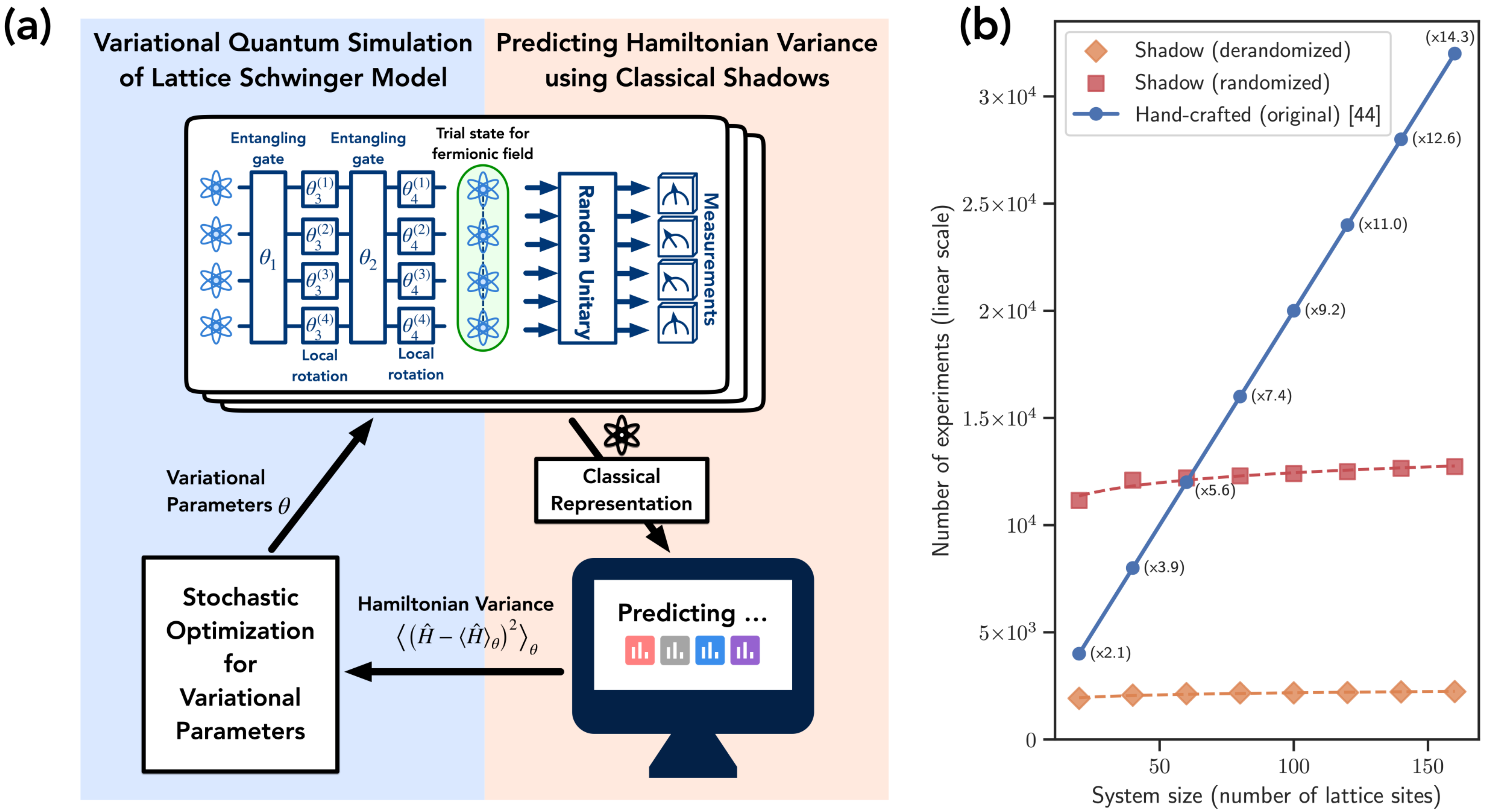
Sampling error in the Pauli channel characterization contributes to variance.

Chen, Yu, Zeng, Flammia 2021

“Measure first, ask questions later.”

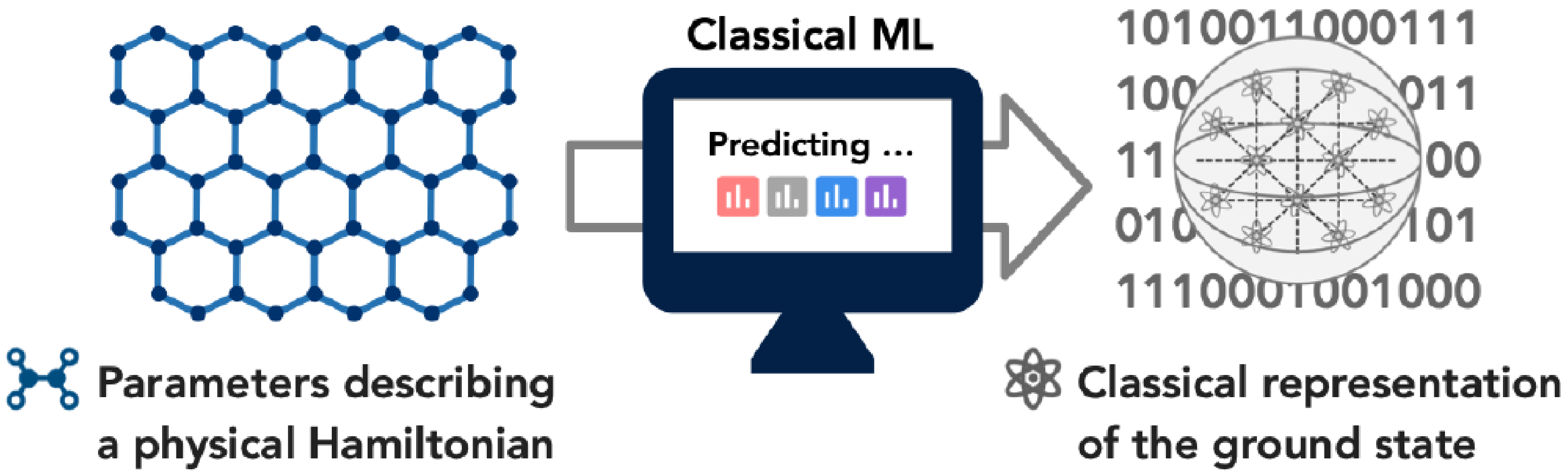
Elben, Flammia, Huang, Kueng, Preskill, Vermersch, Zoller,
The randomized measurement toolbox 2022

Energy variance in 1D quantum electrodynamics



- Innsbruck ion-trap experiment: Kokail, Maier, van Bijnen et al. 2019.
- With classical shadows, # of copies needed to estimate variance of H ~ log(system size)
- Further improvement from derandomization.

Classical machine learning for properties of quantum ground states



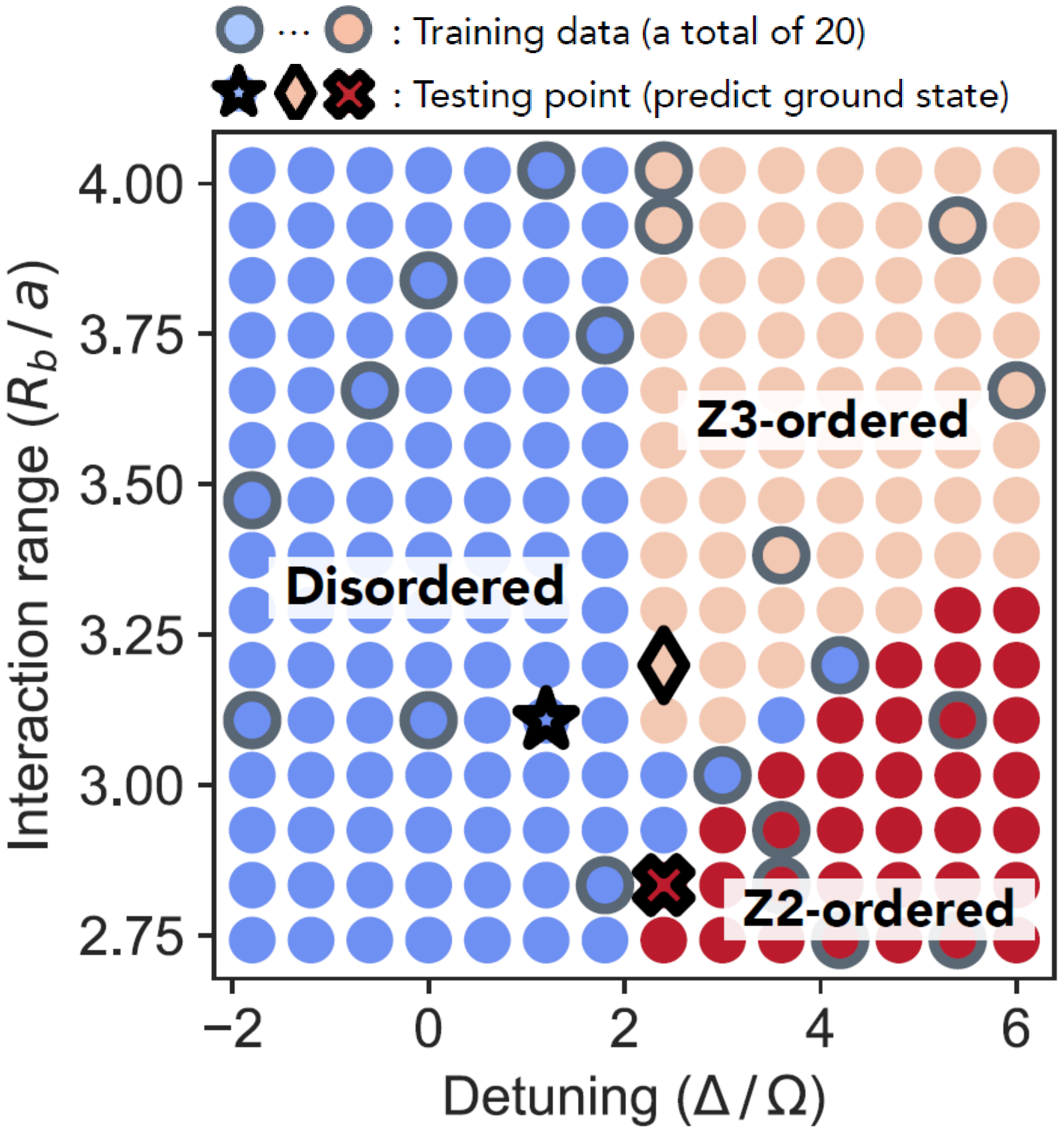
Theorem (*Learning properties of ground states*):
For any smooth family of local Hamiltonians $\{H(x), x \in [-1,1]^m\}$ in a finite spatial dimension with a constant spectral gap, a classical machine learning algorithm can learn to predict an efficient classical representation of the ground state $\rho(x)$ that approximates few-body reduced density matrices up to a constant error. The required amount of training data and computation time are polynomial in m and linear in system size.

Idea: convert training states to their classical shadows. Then use a classical learning algorithm to predict a classical representation for new values of *x*.

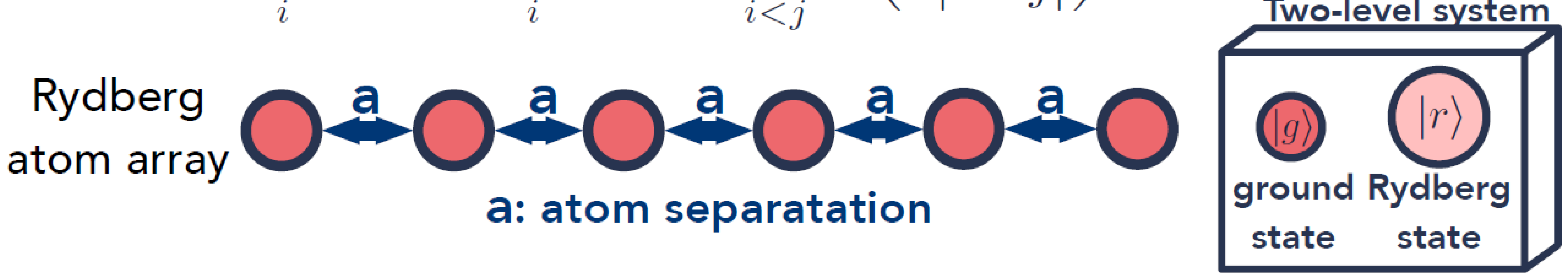
The learning is classical, but we need the quantum platform to prepare and measure the ground state during training. With access to training data, we can solve quantum problems that might be too hard to solve otherwise.

Huang, Kueng, Torlai, Albert, Preskill 2021
Lewis, Huang, Tran, Lehner, Kueng, Preskill 2023

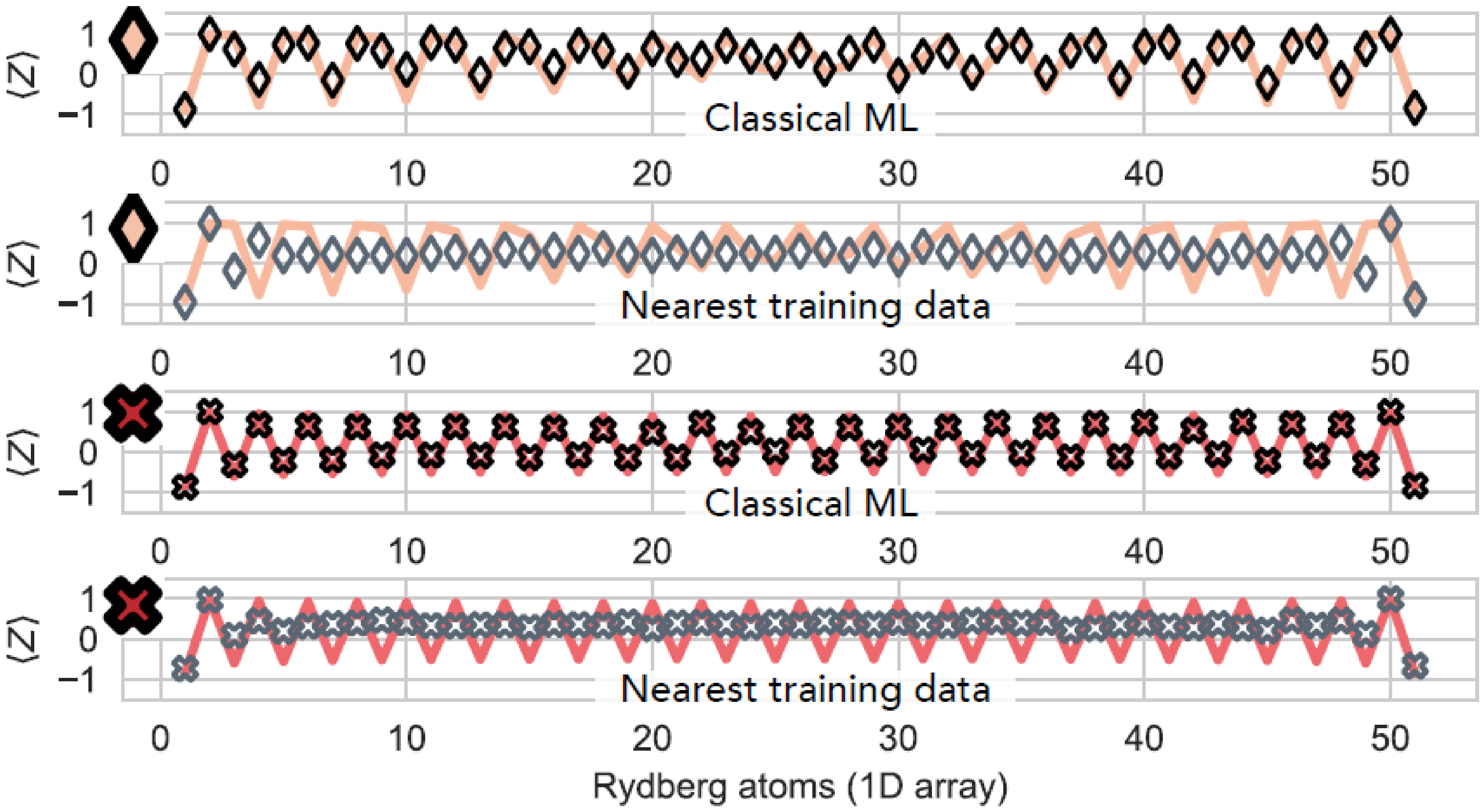
Example: 1D array of Rydberg atoms



$$H = \sum_i \frac{\Omega}{2} X_i - \sum_i \Delta N_i + \sum_{i<j} \Omega \left(\frac{R_b}{a|i-j|} \right)^6 N_i N_j$$



$$N_i = |r_i\rangle\langle r_i|, \quad X_i = |g_i\rangle\langle r_i| + |r_i\rangle\langle g_i|, \quad Z_i = |g_i\rangle\langle g_i| - |r_i\rangle\langle r_i|$$

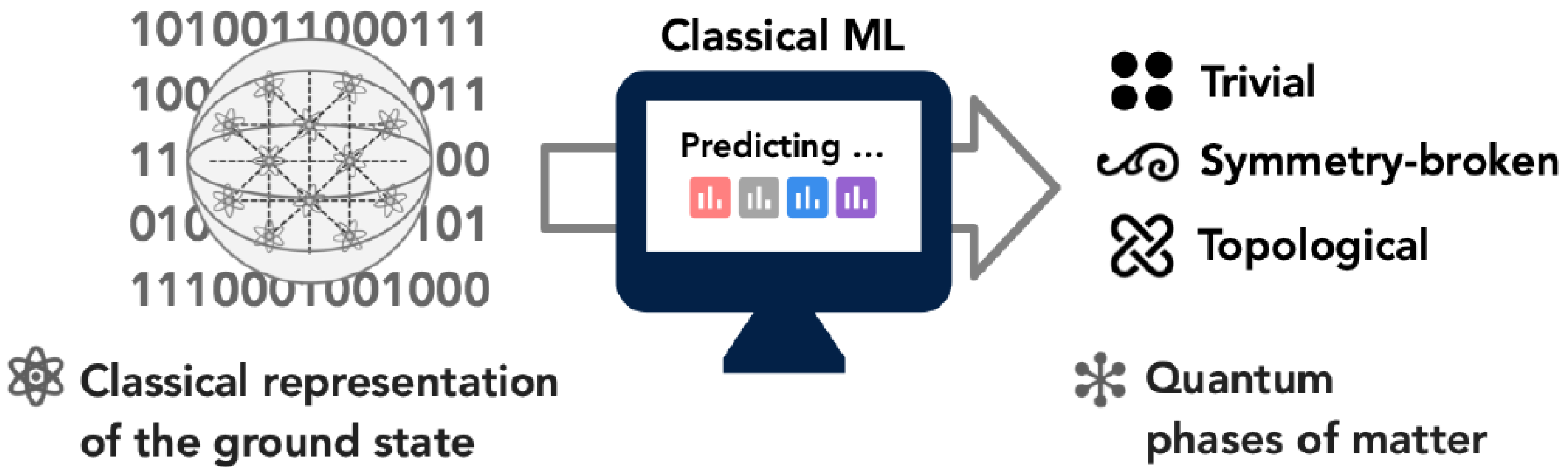


Chain of 51 atoms (as in Bernien *et al.* 2017). We can compute ground state properties using DMRG.

Our rigorous theory does not directly apply, because Hamiltonian is not gapped throughout the parameter regime considered. Yet predictions work well.

500 snapshots taken at each sampled value of x.

Classical machine learning for identifying quantum phases of matter



Theorem (*Identifying quantum phases of matter*):
If there exists a polynomial function of few-body reduced density matrices that classifies phases, then a (supervised) classical machine learning algorithm can learn to classify phases accurately. The required amount of training data and computation time are polynomial in system size.

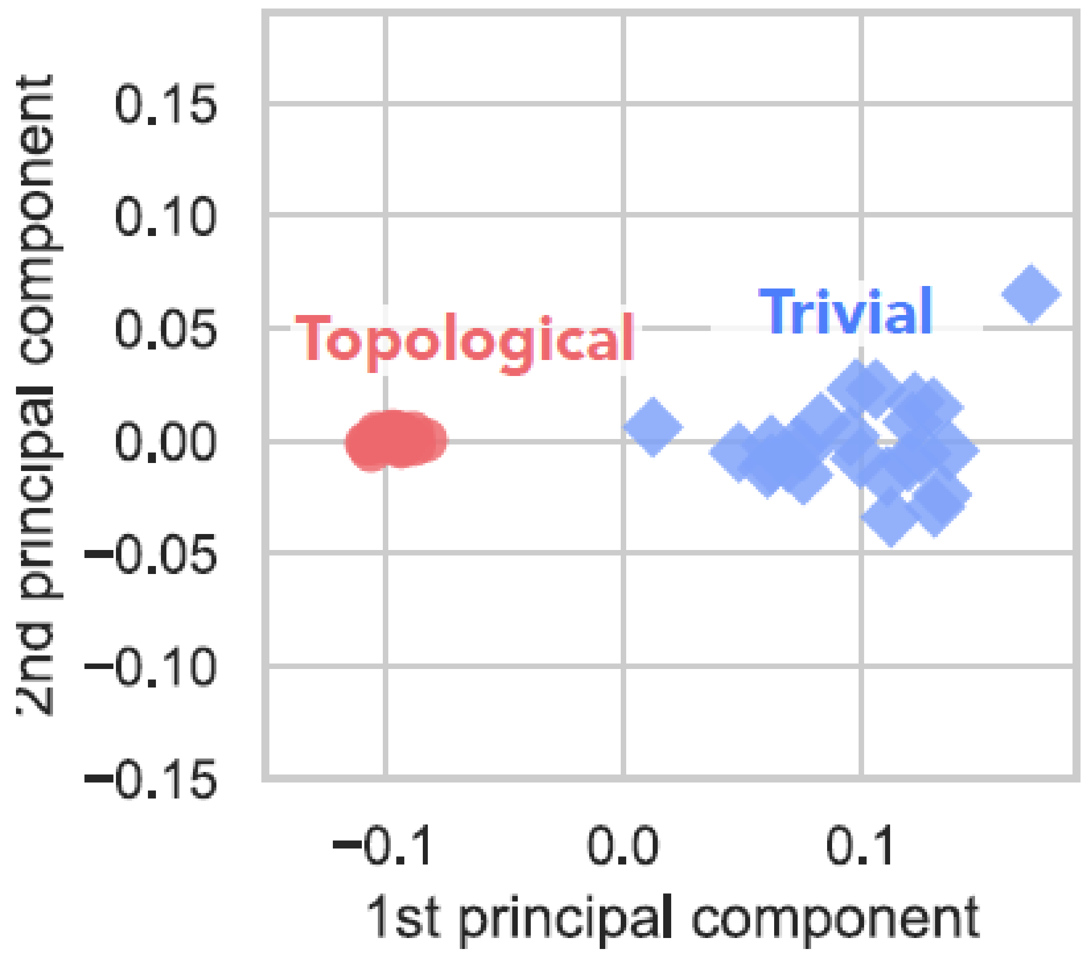
Idea: convert each quantum state to its classical shadow, and learn to classify these shadows.

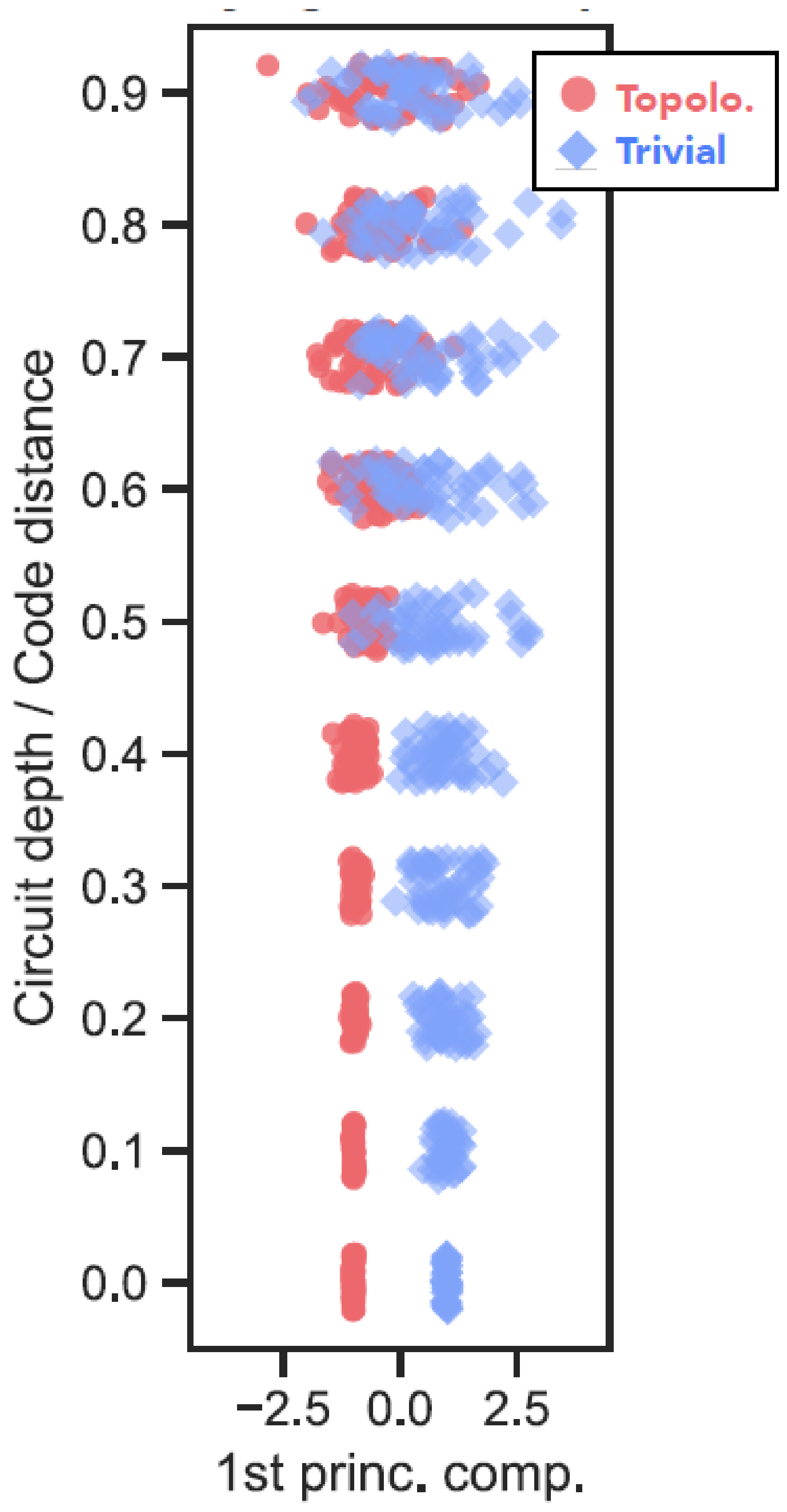
Learning strategy: Map each classical shadow to a feature vector in a high-dimensional space.

The learning algorithm discovers the classifying function, which need not be known in advance.

Huang, Kueng, Torlai, Albert, Preskill 2022

Example: Distinguishing the 2D toric code phase from the trivial phase



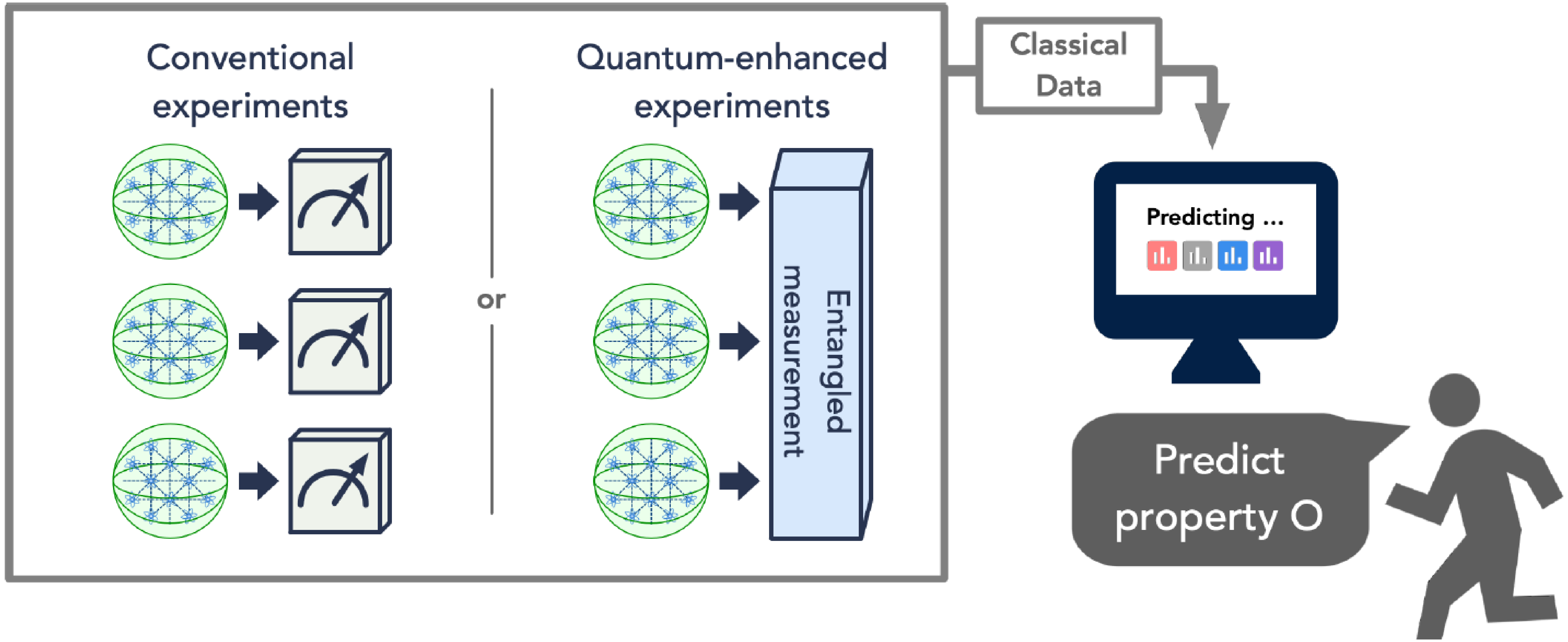


No local circuit of constant depth acting on a product state can reach a topologically ordered state.

Principle components are projections of the data geometry in feature space to a low-dimensional subspace, chosen to maximize the variance of the data.

We consider applying low-depth local quantum circuits to (A) a product state and (B) the toric code state. The resulting classical shadows are cleanly separated in the feature space (and hence a linear classifying function in feature space is easy to learn) until the circuit depth approaches half the code distance.

Conventional experiments vs. quantum-enhanced experiments



How many experiments are needed to learn properties of physical systems, with or without access to quantum memory?

For some tasks, we prove that *exponentially fewer* experiments suffice in the "quantum-enhanced" setting.

And we demonstrate this advantage in experiments using up to 40 qubits on the Sycamore processor.

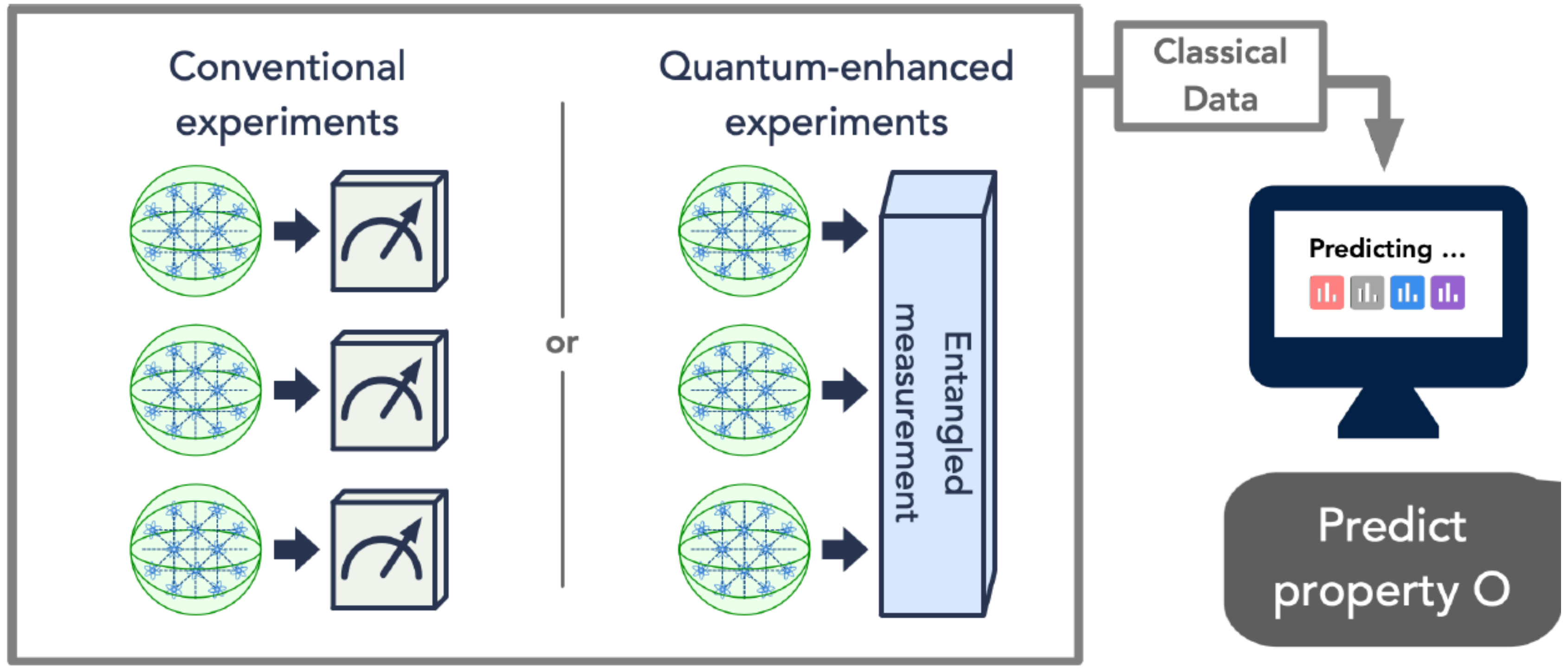
Huang, Broughton, Cotler, Chen, Li, Mohseni, Neven, Babbush, Kueng, Preskill, McClean 2022

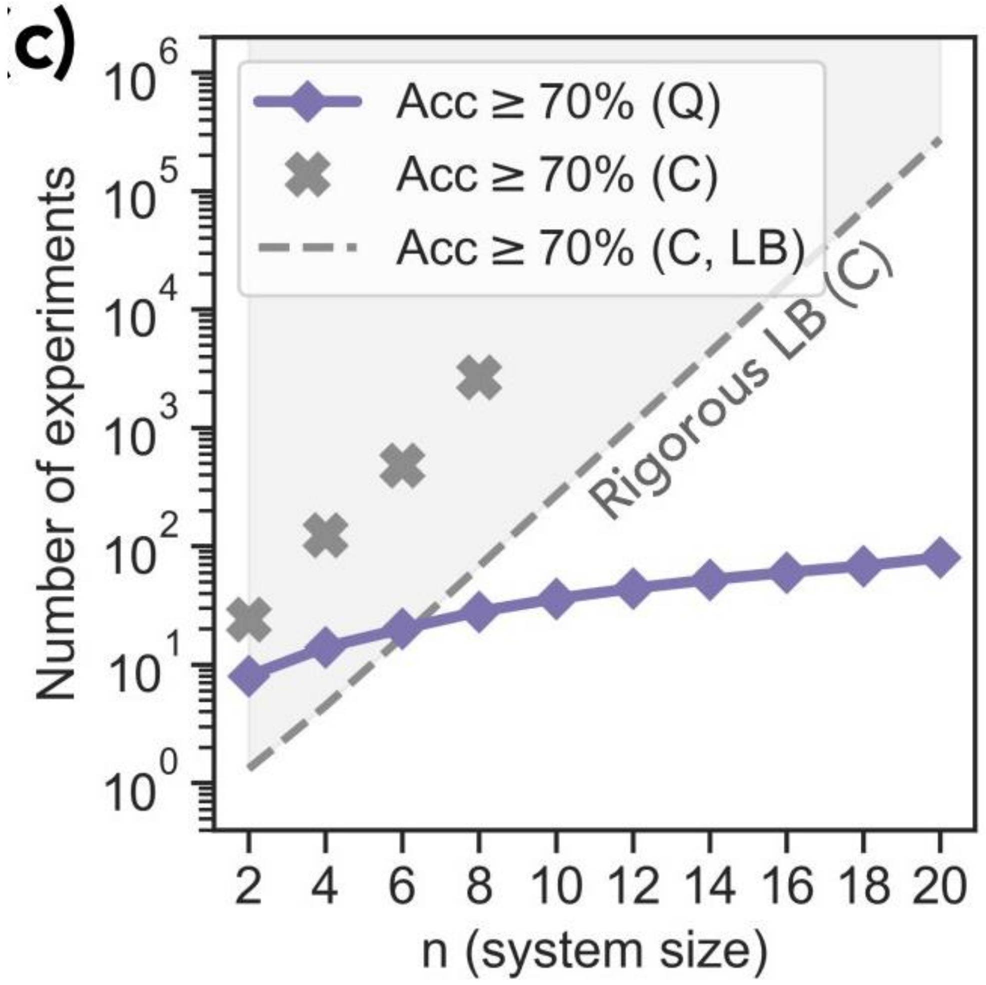
Conventional experiments vs. quantum-enhanced experiments

How many experiments are needed to learn properties of quantum systems, with or without access to quantum memory?

For some tasks, we proved that *exponentially fewer* experiments suffice in the "quantum-enhanced" setting.

And we demonstrated this advantage in experiments using up to 40 qubits on the Sycamore processor.





Exponential quantum advantage in learning expectation values of observables.

Will quantum technology revolutionize how we acquire and process experimental data to learn about the physical world?

Huang, Broughton, Cotler, Chen, Li, Mohseni, Neven, Babbush, Kueng, Preskill, McClean 2022

Making predictions in a quantum world

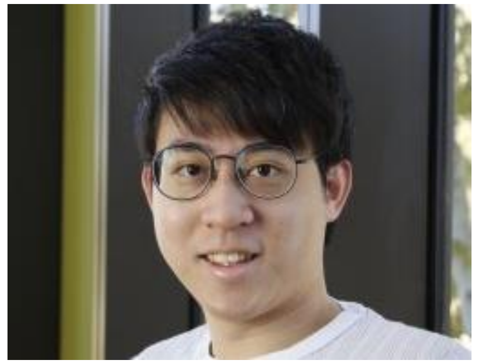
Classical shadows of quantum states: a feasible procedure converting a quantum state to succinct classical data.

O(log M) copies, and efficient classical processing, suffice to predict M properties. “Measure first, ask questions later.”

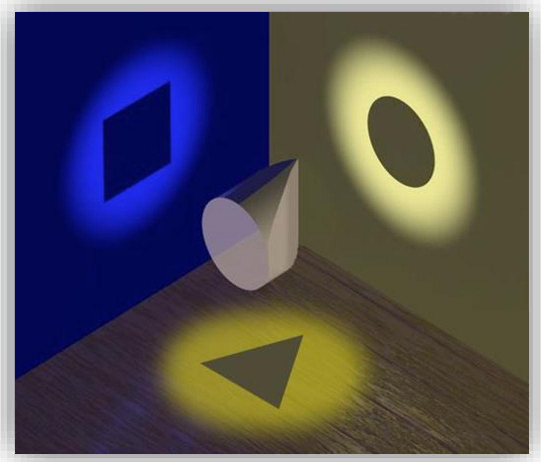
This number of copies is asymptotically optimal for single-copy measurements.

Access to data from quantum experiments may enable classical machine learning to solve quantum problems that would be too hard to solve without access to data.

Quantum-enhanced experiments making use of quantum memory and quantum processing can have an exponential advantage relative to conventional experiments.


H.-Y. (Robert) Huang


Richard Kueng


classical shadows


Giacomo Torlai


Victor Albert


Laura Lewis

Google Quantum AI

Additional Slides

Predicting many properties from very few measurements

We want to estimate expectation values of many w-qubit observables in an n-qubit quantum state ρ, with error at most ε.

We are provided with N copies of ρ. For each copy we perform randomized single-qubit Pauli measurements --- that is, for each qubit in each copy we measure one of the Pauli matrices X, Y, Z, chosen equiprobably, obtaining N succinct "snapshots" of ρ. These snapshots constitute the "classical shadow".

From the snapshots, we *efficiently* compute an approximation (error ε) to expectation values of M w-qubit operators. To succeed with high probability, N copies suffice, where:

$$N = O\left(4^w \log(M) / \varepsilon^2\right)$$

We can also estimate nonlinear properties, such as Rényi entropies. (A polynomial of order k in the density operator can be viewed as the expectation value of an observable acting on k copies of the quantum state.)

Huang, Kueng, Preskill, 2020.

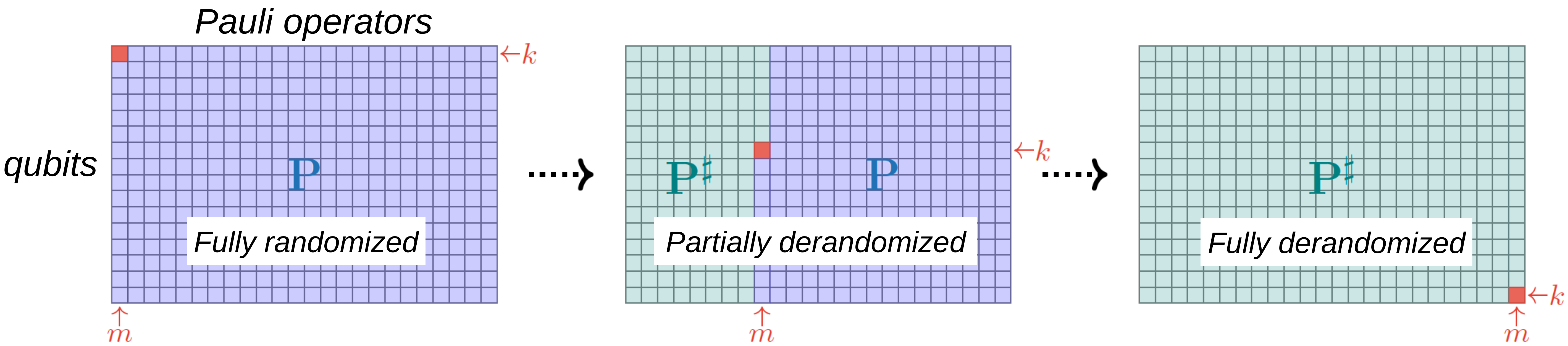
Special case: Predicting Pauli operators

Target (weight w=4):	…	I	I	X	Z	Z	X	I	I	…
Measured Pauli ops:	…	X	Y	X	Z	Z	X	Z	Y	…

Hit probability 3^{-w} ➡ Predict M Pauli ops (weight $\leq w$) with error ε using $O(3^w \log M / \varepsilon^2)$ measurements.

Evans, Harper, Flammia 2019

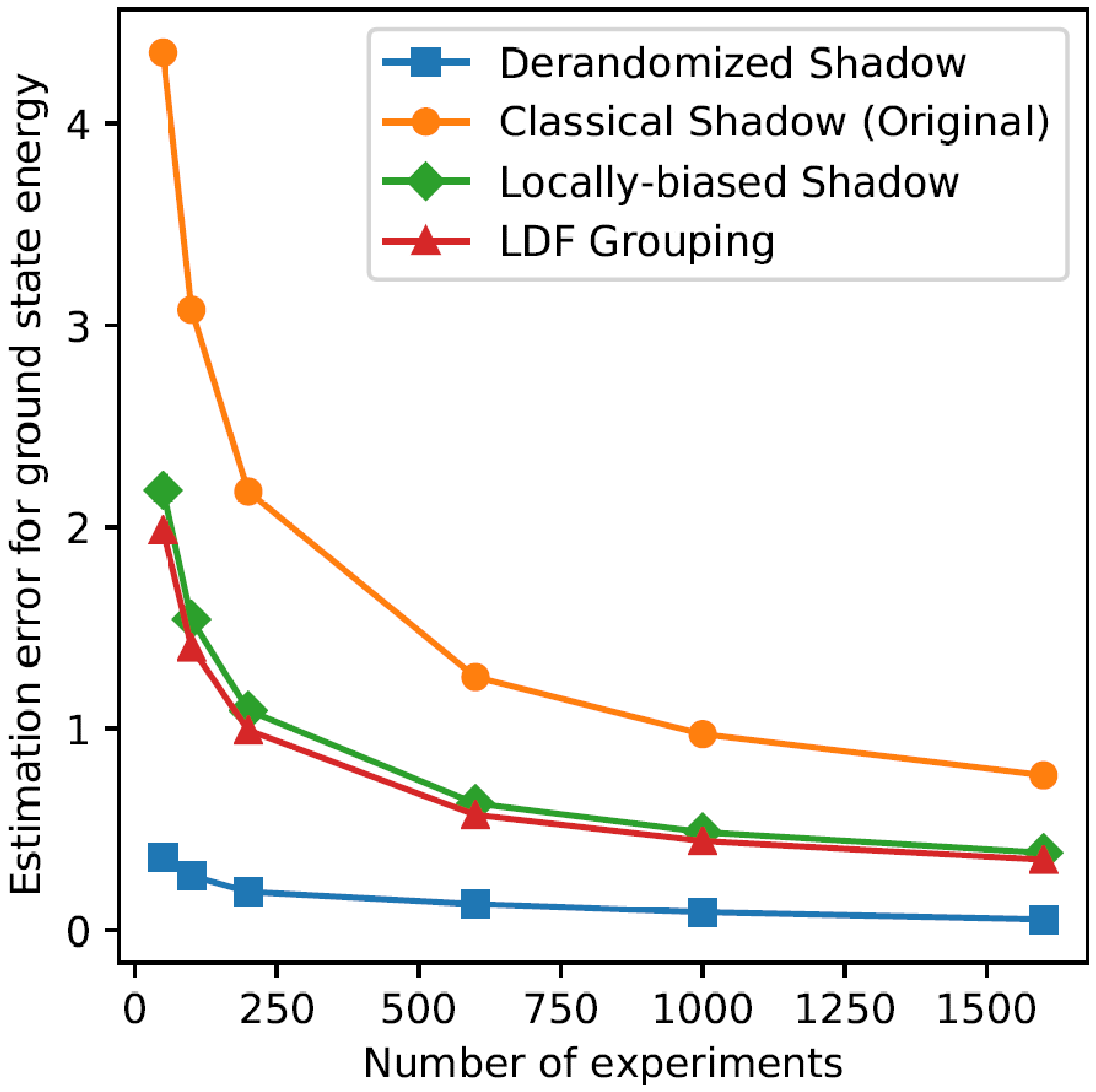
Derandomize:



-- Iteratively replace each randomized single-qubit Pauli measurement by a fixed Pauli.

-- No worse than the randomized protocol, and possibly much better if target observables have structure or include high-weight Pauli ops. *Huang, Kueng, Preskill 2021*

Quantum chemistry



BeH_2 ground state energy estimation error (in Hartree) under Bravyi-Kitaev encoding of fermions, using various measurement schemes.

Locally-biased classical shadows:
Hadfield, Bravyi, Raymond, Mezzacapo 2020.

Largest degree first (LDF) grouping:
Verteletskyi, Yen, Izmaylov 2020.

Derandomized classical shadow
Huang, Kueng, Preskill 2021

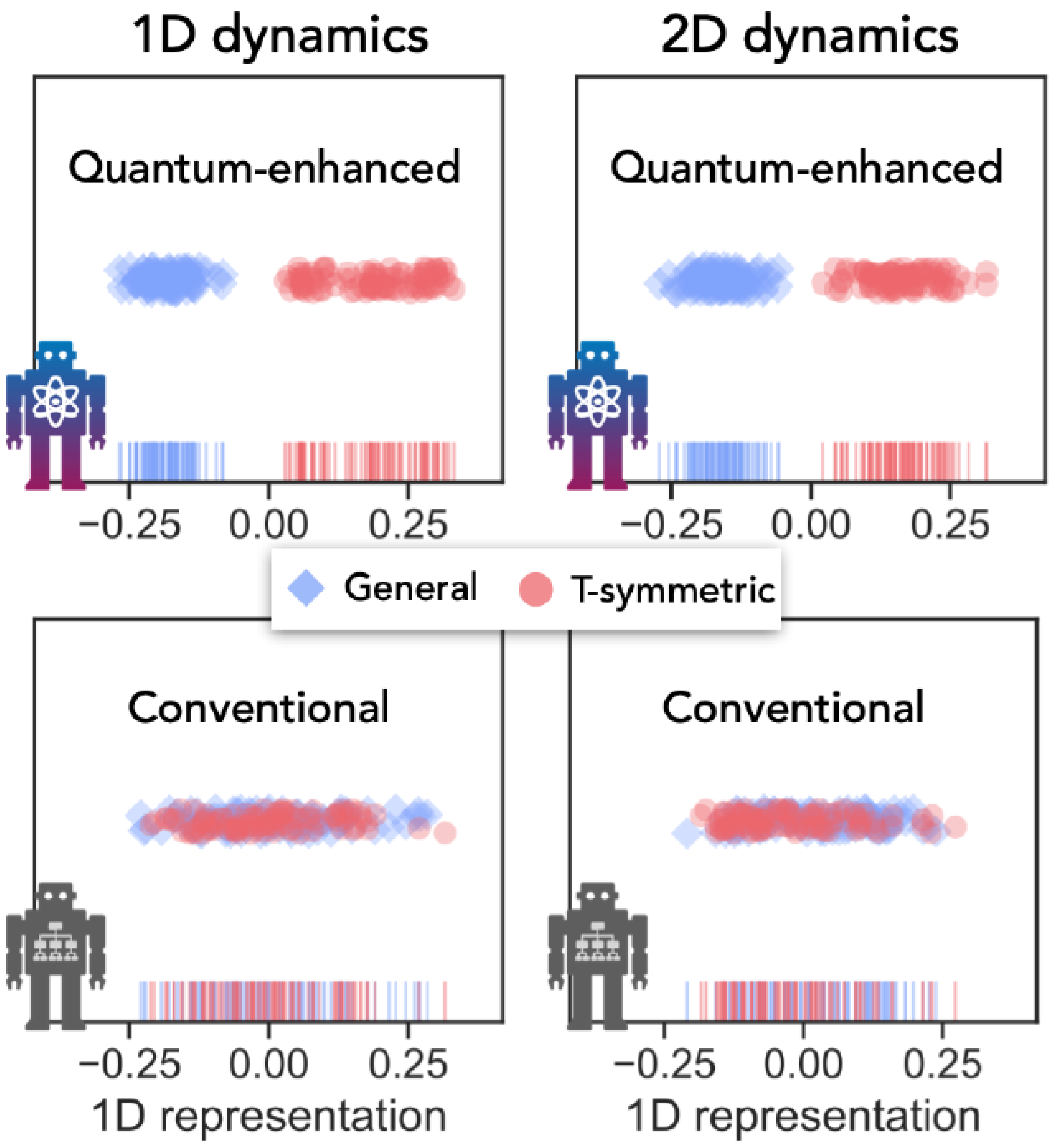
Conventional experiments vs. quantum-enhanced experiments

An unknown unitary evolution operator is drawn from one of two ensembles --- it is either a general unitary matrix or a real orthogonal matrix (time-reversal symmetric).

How well can we distinguish these two symmetry classes?

We generated the transformations as random circuits on Sycamore and applied them to a fixed product input state. In the conventional scenario, we measured all output qubits in the *Y* basis. In the quantum-enhanced scenario we performed Bell measurement across two copies of the output state.

Based on this measurement data, an unsupervised ML could easily distinguish the symmetry classes in the quantum-enhanced scenario but not in the conventional scenario. In both scenarios we ran the quantum circuit 1000 times.



Hard to learn in conventional scenario: Aharonov, Cotler, Qi 2021; Chen, Cotler, Huang, Li 2021