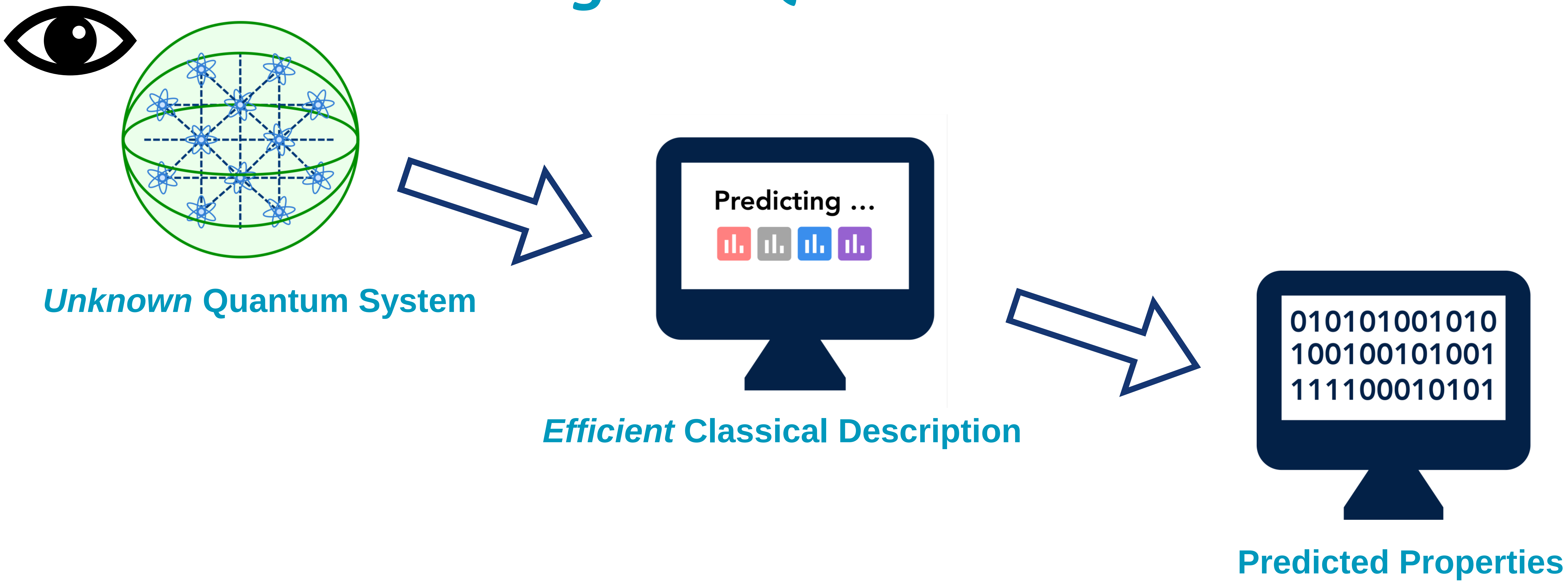
Learning in a Quantum World



INSTITUTE FOR QUANTUM INFORMATION AND MATTER

John Preskill

Strings 2023

27 July 2023

@preskill

From Sabrina:
This edition of the Strings conference will feature four "Challenge" talks designed to alert our community to new and promising ideas in adjacent subjects. We would love for you to give one of these talks.

From Rob:
We were hoping that you could give us a future looking talk on new prospects/directions for quantum information and quantum gravity/strings.

Strings 1998

You start with the brane
And the brane is BPS.

Then you go near the brane
And the space is AdS.

Who knows what it means?
I don’t, I confess.

Ehhh! Maldacena!

Jeff Harvey

Entanglement theory

Computational complexity

Error correction

Spacetime emerges from entanglement

The geometrization of quantum information in quantum gravity has been a fruitful and mysteriously successful research direction …

Netta Engelhardt

Boundary = physical qubits

Low-energy bulk = logical qubits

Entanglement = wormhole

⬇

Wormhole teleportation

Effective field theory:

Low energy

Low curvature

Low *complexity*

The nonisometric nature of the code is hidden by a veil of computational complexity.

Daniel Harlow

Quantum information

More qubits

Higher gate fidelity

New platforms

Programmable quantum simulators

Quantum error correction meets experiment

Applications?

Learning about the quantum world

Curse of dimensionality

Capture essential features with simple classical models

Generalize from data, predict what we’ll see in new situations

Useful tasks

Predict properties of chemical compounds and materials we haven't encountered in the lab before.

Recognize when a qualitatively new phase of matter (in equilibrium or out of equilibrium) has been created.

Preparing target quantum states using available experimental tools.

Extracting a usable signal from very noisy data.

Learning states, observables, and processes

-- Collect training data

-- Create a (classical) model

-- Predict properties

How much training data do we need, and how hard is it to construct and use the model to predict accurately?

What is experimentally feasible, now or in the future?

Huang, Kueng, Preskill 2020; Huang, Kueng, Torlai, Albert, Preskill 2022;
Huang, Kueng, Preskill 2021; Huang, Chen, Preskill 2022
Huang, Broughton, Cotler, Chen, Li, Mohseni, Neven, Babbush, Kueng, Preskill, McClean 2022
Lewis, Huang,Tran, Lehner, Kueng, Preskill 2023

Learning quantum *states*

ρ is an *arbitrary* unknown *n*-qubit quantum state.

We have access to *N* identically prepared copies of ρ.

We measure the copies (e.g., randomized, nonadaptive on each copy).

Our goal is to predict $\mathrm{tr}(\rho O_\alpha)$ with error ≤ ε for *all* the observables $\{O_\alpha\}$ in a restricted family, with guaranteed success probability ≥ 1- δ.

We might also wish to predict nonlinear functions of ρ, such as Rényi entropies (expectation of $\mathrm{tr}(\rho^{\otimes k} O)$).

Classical shadows of quantum states

Ensemble of unitary transformations $\{U_i\}$.

For each copy of ρ, sample random U_i, apply to ρ, measure in standard basis, obtain bit string $|b_i\rangle$.

Snapshot: $|s_i\rangle = U_i^\dagger |b_i\rangle$.

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

$\mathcal{M}$ is an (invertible) CPTP map.

Unbiased estimator for the unknown state, and for $\mathrm{tr}(\rho O_\alpha)$. What is the variance of the estimator, and how likely is a large deviation from the mean?

Example: Random Pauli measurement

Measure each qubit in a random Pauli basis: X, Y, or Z.

$$|s_i\rangle = \otimes_{a=1}^{n} |s_i^a\rangle, \quad |s_i^a\rangle \in \{|0\rangle, |1\rangle, |+\rangle, |-\rangle, |+i\rangle, |-i\rangle\}$$

$$\rho \approx \frac{1}{N}\sum_{i=1}^{N} \sigma_i^1 \otimes \cdots \otimes \sigma_i^n, \quad \sigma_i^a = 3|s_i^a\rangle\langle s_i^a| - \mathbb{I}$$

$$\text{variance} \leq \| O \|_{\text{shadow}}^2 \leq 4^{\text{weight}} \| O \|_{\infty}^2$$

where O is a *bounded-degree local observable*.

Example: Random Pauli measurement

Learn M observables by measuring N copies where …

$$N = O(B \log M / \epsilon^2)$$

$$B = \max_\alpha \| O_\alpha \|_{\text{shadow}}^2 \leq 4^{\text{weight}} \| O \|_{\infty,\max}^2$$

(Median of means estimation ensures that large deviations are exponentially rare.)

Learn *all* observables of constant weight and constant spectral norm from number of copies:

$$N = O\left(\log n / \epsilon^2\right)$$

Huang, Kueng, Preskill 2020

“Measure first, ask questions later.”

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

Local vs. global scrambling

Low depth circuits for predicting local observables.
Example: local Clifford transformations.

Deep circuits for predicting (some) global observables.
Example: global Clifford transformations.

As depth increases, access higher-weight observables.
Example: Chaotic dynamics for a constant time interval.

Hu, Choi, You 2021

What about noise?

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

Learning *observables*

Learning states: ρ is an arbitrary state, predict expectation for a restricted set of input observables.

Learning observables: O is an arbitrary observable, predict its expectation for a restricted class of input states, e.g. those drawn from a fixed distribution.

Goal: Learn a function on states that estimates tr(ρO) with a small *average* error.

$$\mathbb{E}_{\rho\sim\mathcal{D}} \,|\, h^{O}(\rho) - \mathrm{tr}(\rho O) \,|^{2} \leq \epsilon$$

How much training data do we need?

Learning observables

Idea:

(1) For a smooth distribution on input states, it suffices to learn a truncated low-weight approximation to O.

(2) Because of (1), it suffices to know only the low-weight reduced density operators of the input ρ (which we can learn efficiently from classical shadows).

To achieve average error ε, truncate O to weight $k = O(\log(1/\varepsilon)$ if the input distribution is *flat*.

$N=O(\log n)$ samples suffice, and the computational cost of learning and predicting is $O(Nn^k)$.

Huang, Chen, Preskill 2022

Learning *processes*

Learn *any* state, predict bounded-degree local observables.

Learn *any* observable, predict with small average error over a flat distribution of states.

Learn *any* process, predict expectations of local observables $\{O_\alpha\}$ in output with a small average error over a flat distribution of input states.

$$\mathbb{E}_{\rho\sim\mathcal{D}} \, | h^{\mathcal{E}}(\rho, O_\alpha) - \mathrm{tr}(\mathcal{E}(\rho) O_\alpha) |^2 \leq \epsilon$$

Learn a low-weight approximation to the unknown observable $\mathcal{E}^*(O_\alpha)$

Classical shadows of $\mathcal{E}(\rho)$ provide the training data, where ρ is a product state.

Huang, Chen, Preskill 2022

Learning processes

We can efficiently learn the process even if it is exponentially complex.

To do so, it suffices to input only product states, even if the input distribution has support on highly entangled states.

We are guaranteed accurate predictions of output *local* observables, not for worst case inputs, but on *average* and only if the input distribution is *flat*.

Learning states, observables, and processes

Using classical and quantum machine learning to learn new physics?

For a smoothly parametrized family of gapped local Hamiltonians, train on classical shadows of ground states of sampled Hamiltonians, and use classical ML for *efficient prediction of local properties of other ground states* in the same phase of matter.

Classify quantum phases of matter using classical ML by converting states to their classical shadows, and learning to classify the shadows. (Efficient if reduced density operators for subsystems of constant size suffice for phase identification.)

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

Quantum gravity: how experiments might help

Probe bulk geometry by measuring *boundary entanglement* structure.

Probe *bulk locality*, e.g. by studying boundary linear response.

Probe *fast scrambling*, Lyapunov spectrum.

Measure higher-order *quantum gravity corrections*.

Simulate very-high-energy bulk scattering.

Holographic dictionaries *beyond anti-de Sitter*.

Use gravitational intuition to understand emergent phenomena.

QG = QM