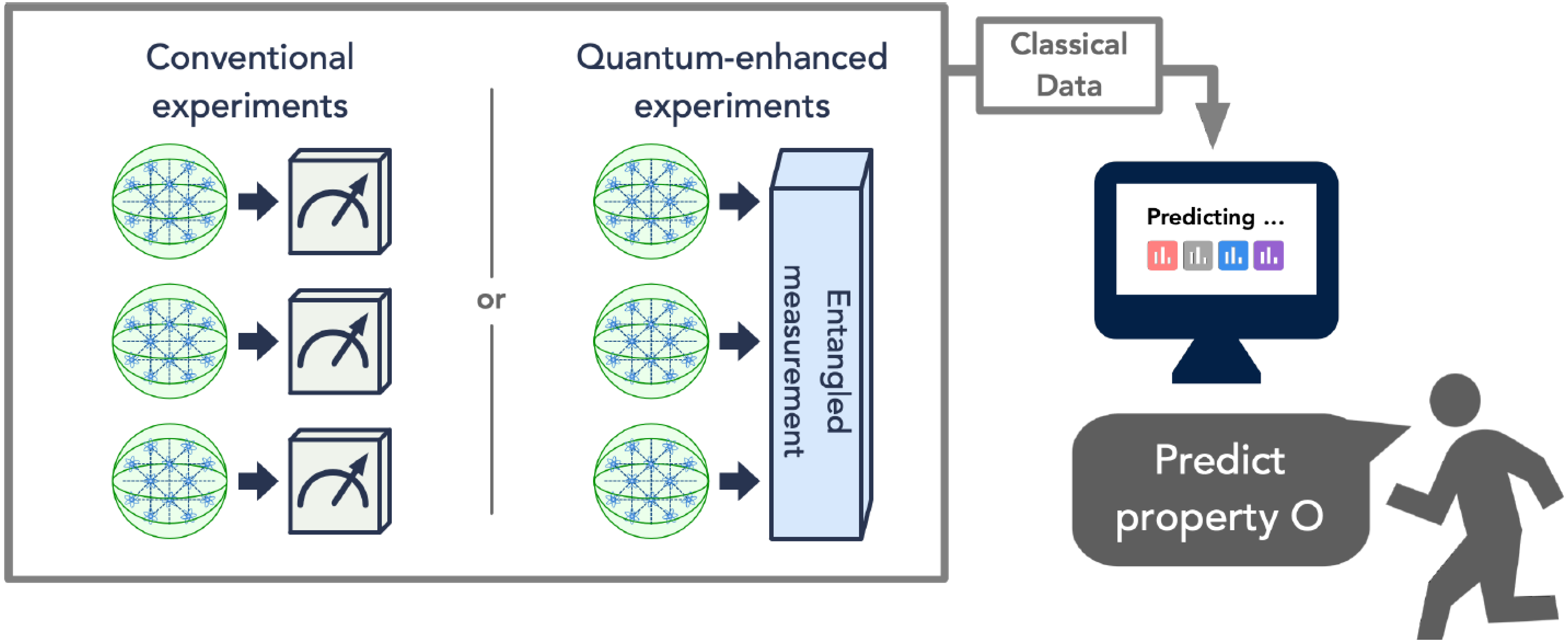
Learning in a Quantum World



John Preskill
Dirac Memorial Lecture
30 May2023

INSTITUTE FOR QUANTUM INFORMATION AND MATTER

@preskill

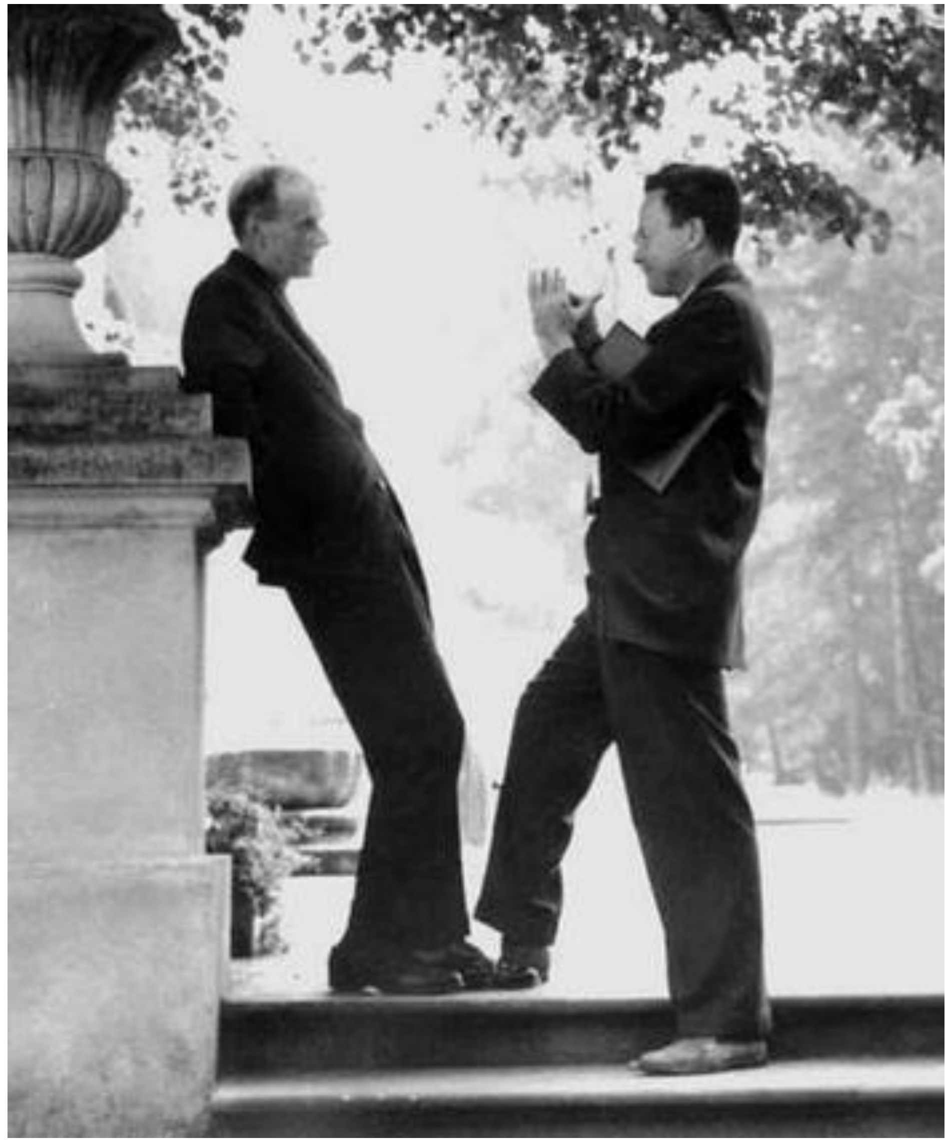
Paul A. M. Dirac
In 1933 (age 31)

$$\boxed{\Phi_0 = \frac{2\pi\hbar c}{e}}$$

"Under these circumstances one would be surprised if Nature had made no use of it."

Paul A. M. Dirac, Quantized Singularities in the Electromagnetic Field, Proceedings of the Royal Society, 1931

"When I was a young man, Dirac was my hero. He made a breakthrough, a new method of doing physics. He had the courage to simply guess at the form of an equation, the equation we now call the Dirac equation, and to try to interpret it afterwards."



Richard P. Feynman
The Reason for Antiparticles
Dirac Lecture, 1986

Overheard at the 1961 Solvay Conference on Physics

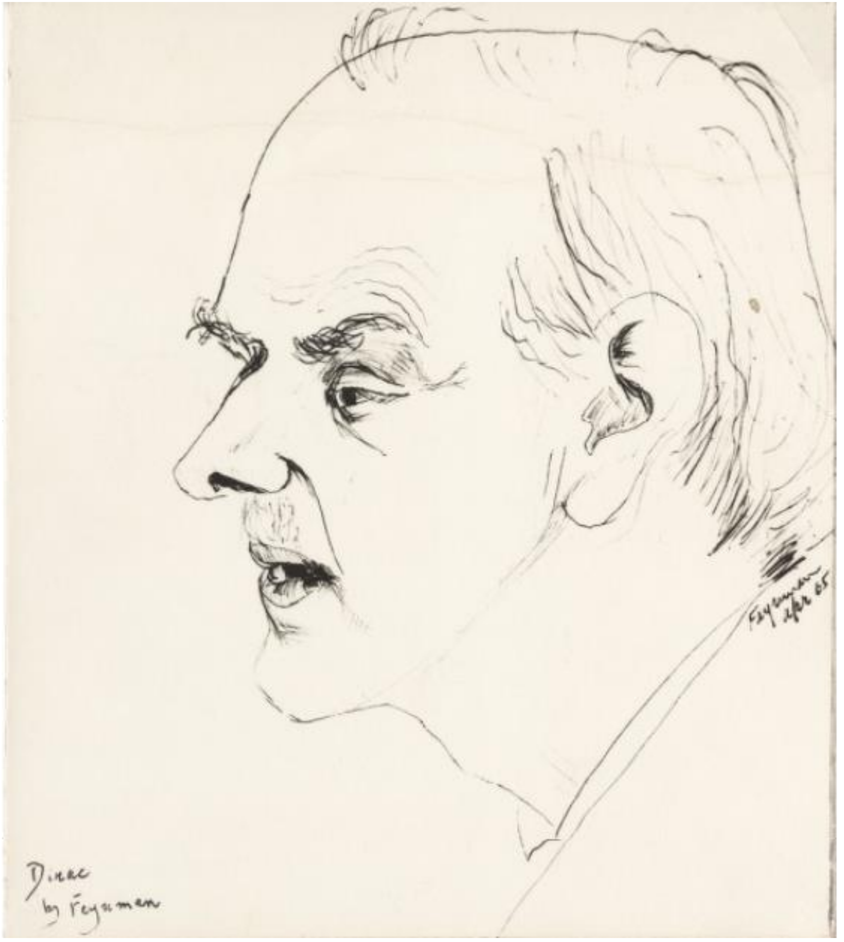
Feynman: I am Feynman.
Dirac: I am Dirac.

[*Silence*]

F: It must be wonderful to be the discoverer of that equation.
D: That was a long time ago.

[*Pause*]

D: What are you working on?
F: Mesons
D: Are you trying to discover an equation for them?
F: It is very hard.
D: One must try.



Feynman's sketch of Dirac – Caltech Archives

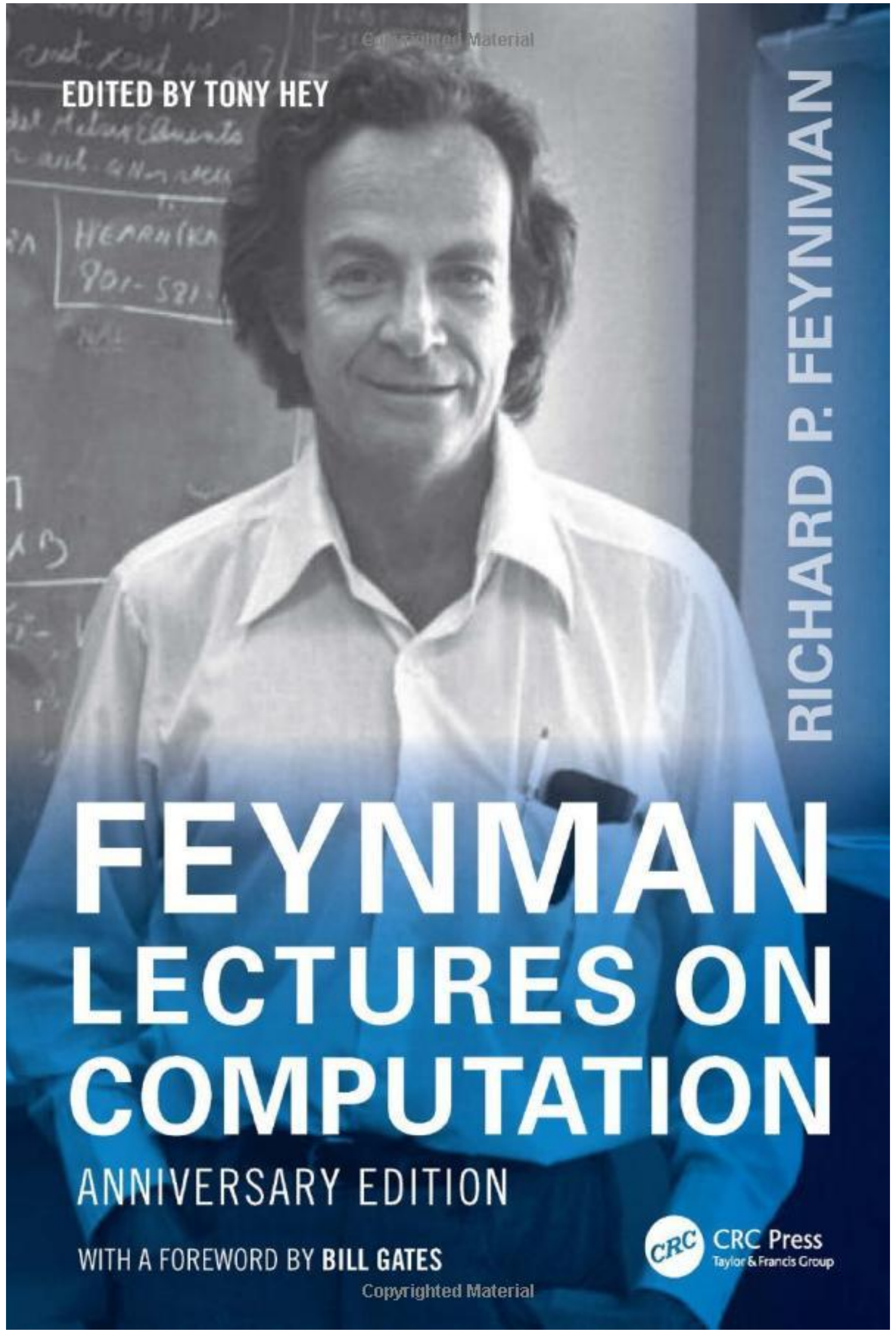
Abraham Pais, Inward Bound (1986)

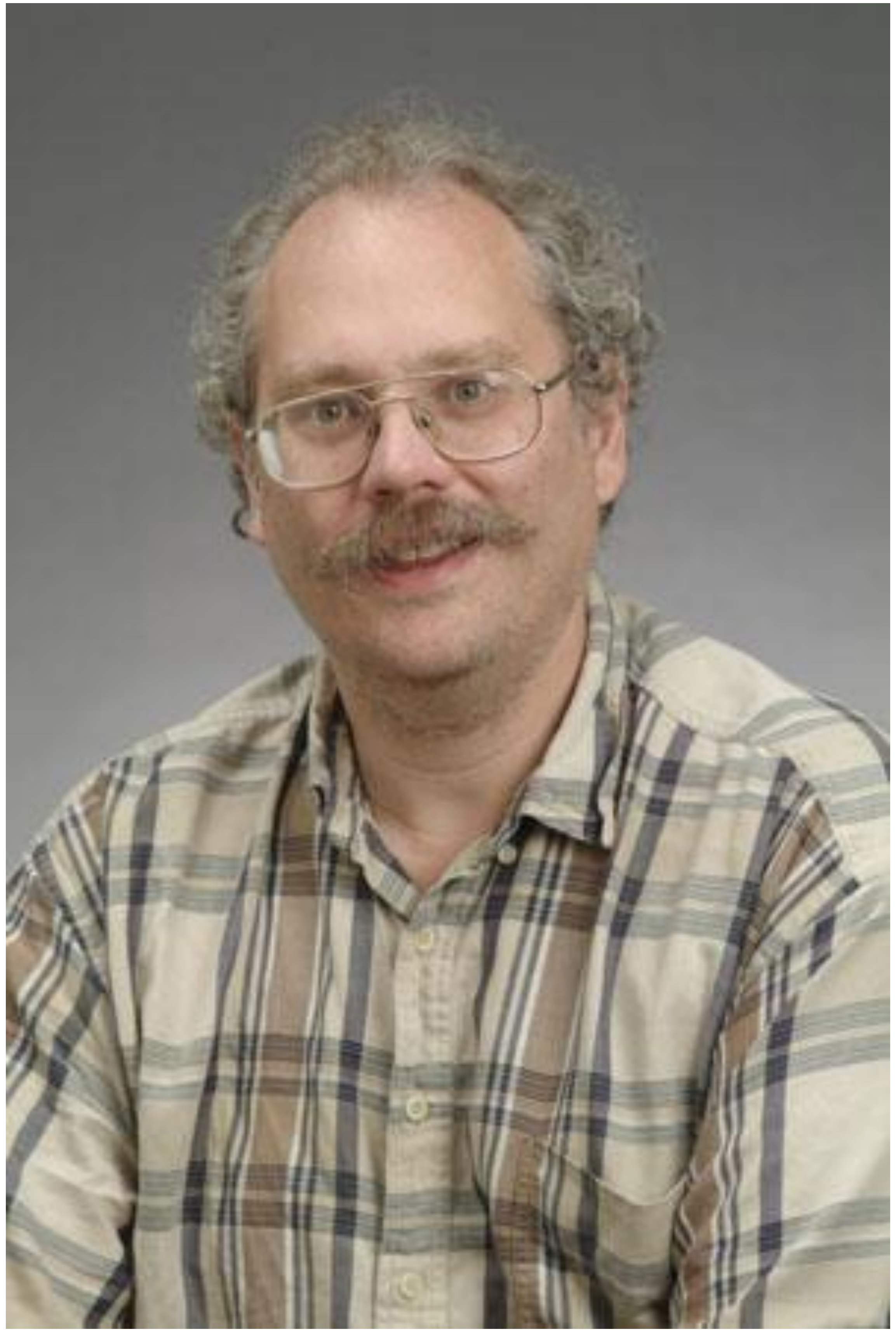
The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.

Paul A. M. Dirac, Quantum Mechanics of Many-Electron Systems, Proceedings of the Royal Society, 1929

Richard Feynman
(1981)

"You can simulate this with a quantum system, with quantum computer elements. It's not a Turing machine, but a machine of a different kind."



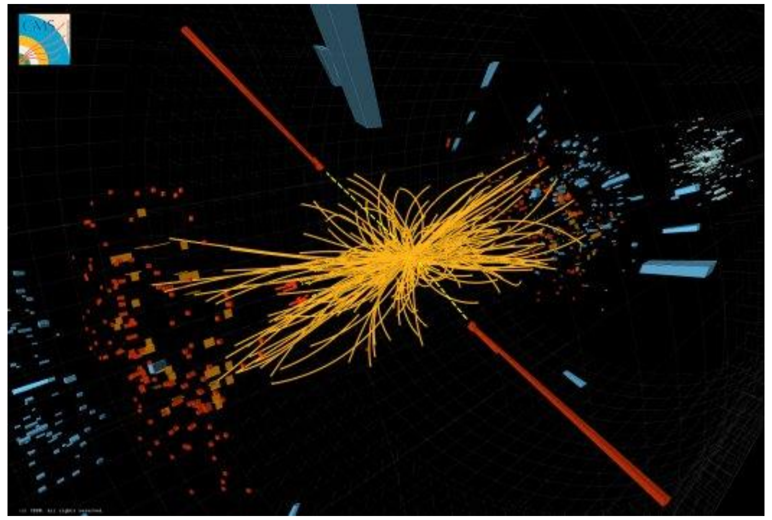
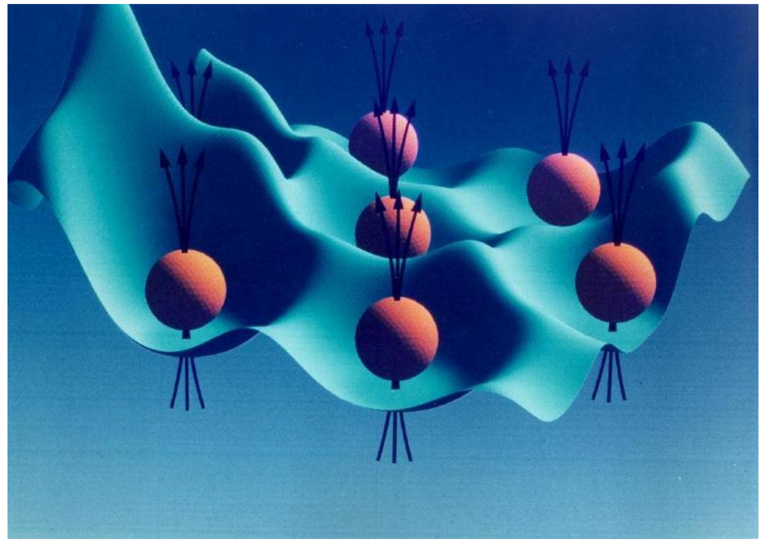


Peter Shor

(1994)

“These algorithms take a number of steps polynomial in the input size, for example, the number of digits of the integer to be factored.”

Frontiers of Physics

short distance	long distance	complexity
		
Higgs boson	Large scale structure	“More is different”
Neutrino masses	Cosmic microwave background	Many-body entanglement
Supersymmetry	Dark matter	Phases of quantum matter
Quantum gravity	Dark energy	Quantum computing
String theory	Gravitational waves	Quantum spacetime

Two fundamental ideas

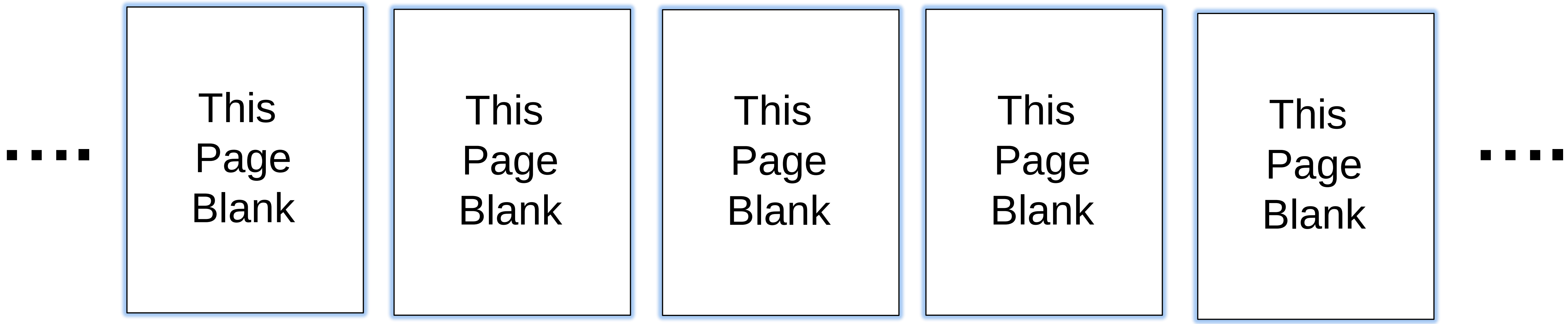
(1) *Quantum complexity*

Why we think quantum computing is powerful.

(2) *Quantum error correction*

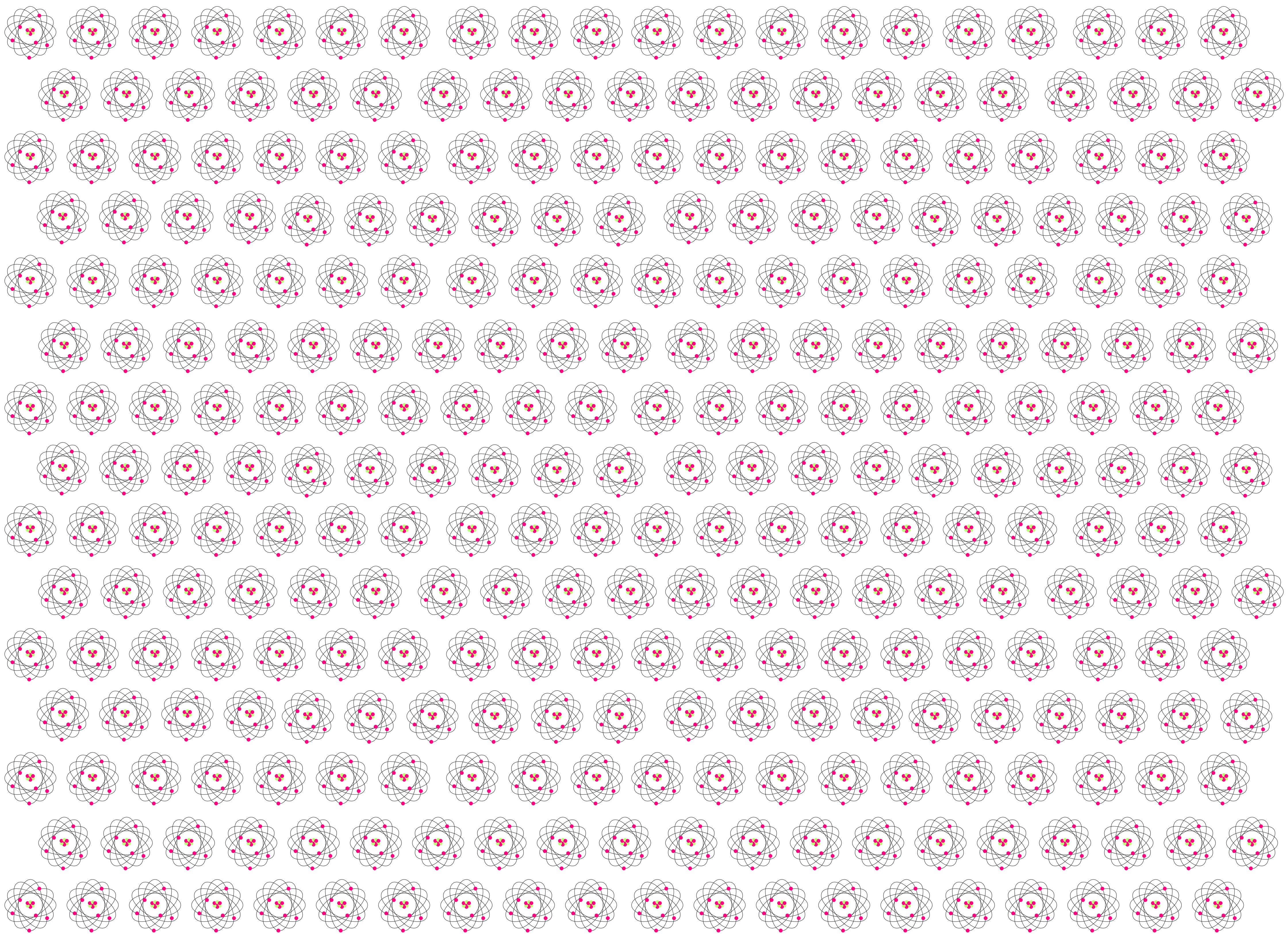
Why we think quantum computing is scalable.

Quantum entanglement



Nearly all the information in a typical entangled “quantum book” is encoded in the correlations among the “pages”.

You can't access the information if you read the book one page at a time.



A complete description of a typical quantum state of just 300 qubits requires more bits than the number of atoms in the visible universe.

Why we think quantum computing is powerful

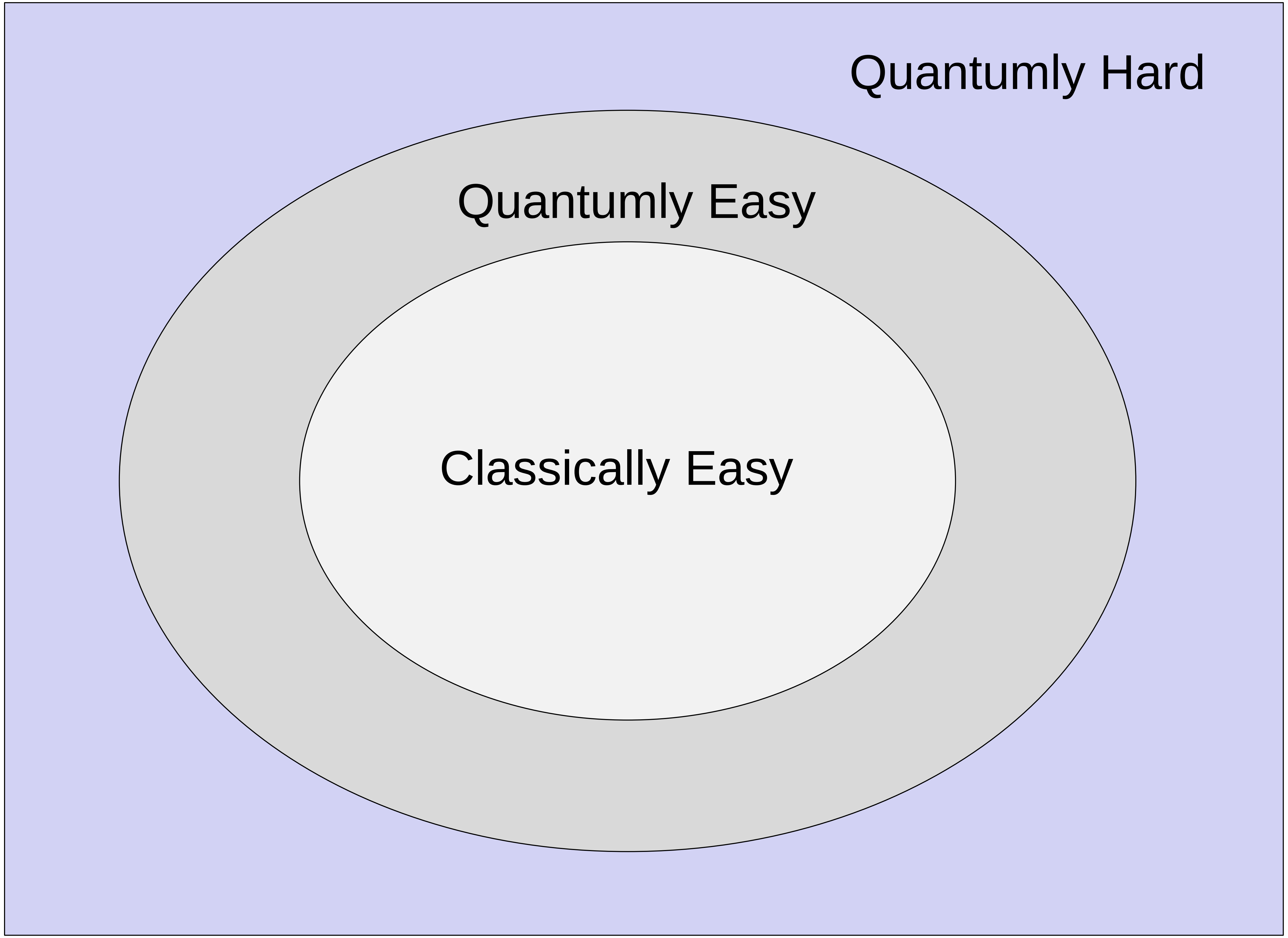
(1) Problems believed to be hard classically, which are easy for quantum computers. Factoring is the best known example.

(2) Complexity theory arguments indicating that quantum computers are hard to simulate classically.

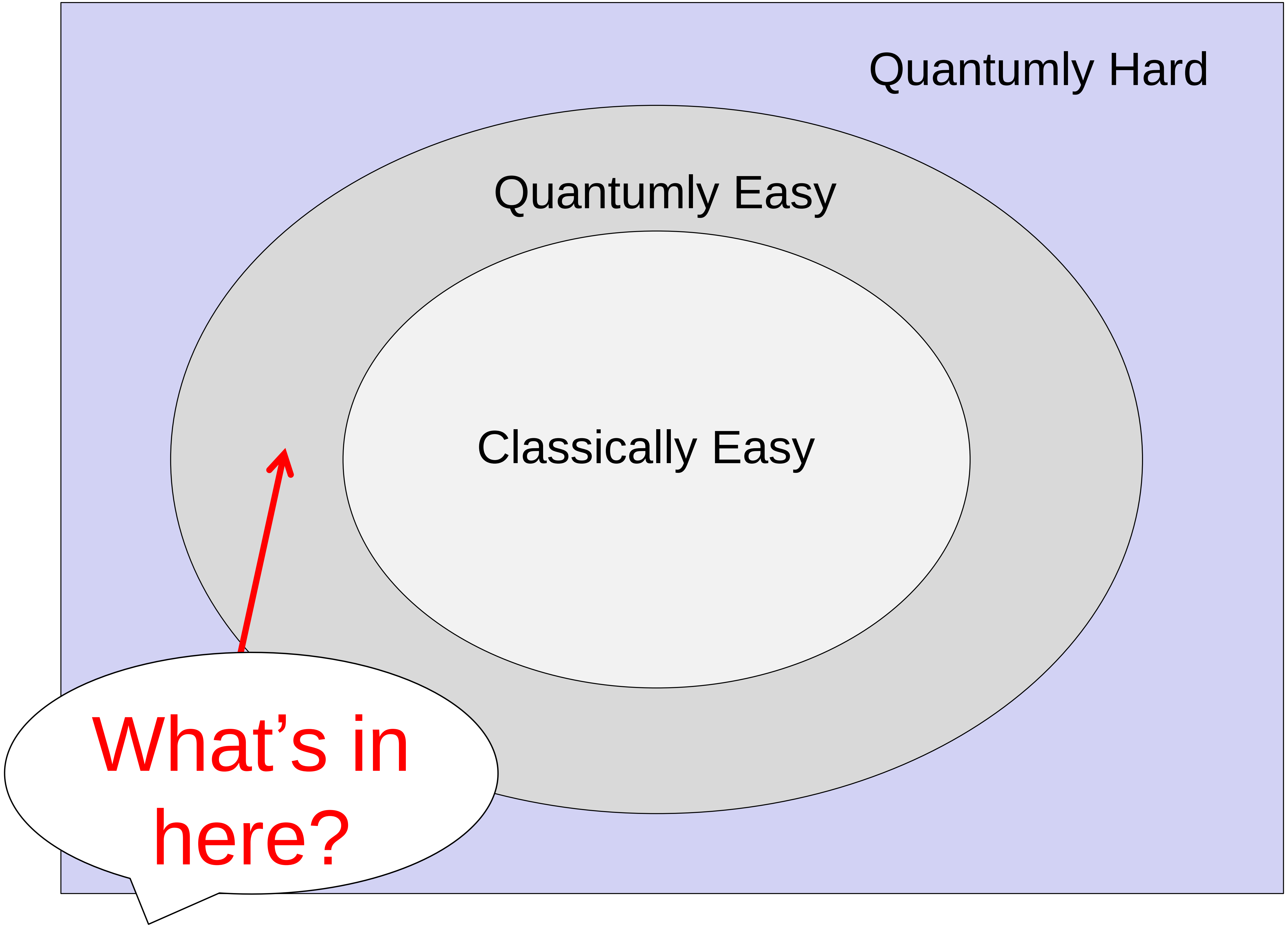
(3) We don’t know how to simulate a quantum computer efficiently using a digital (“classical”) computer.

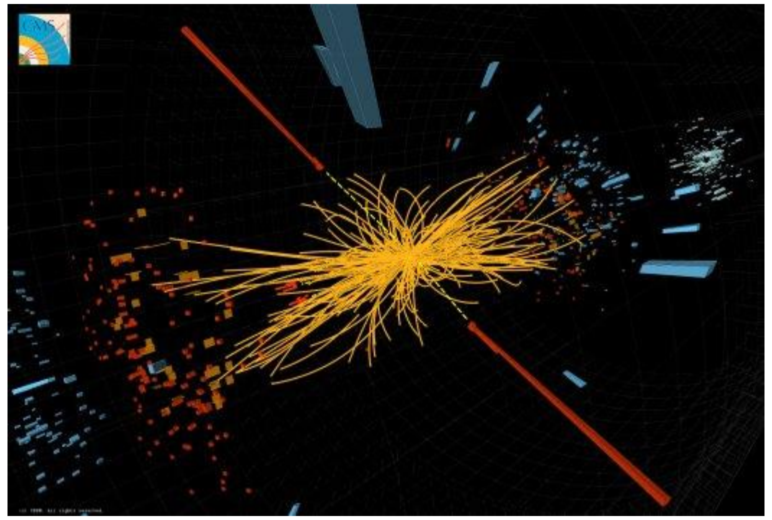
But … the power of quantum computing is limited. For example, we don’t believe that quantum computers can efficiently find exact solutions to worst-case instances of NP-hard optimization problems (e.g., the traveling salesman problem).

Problems

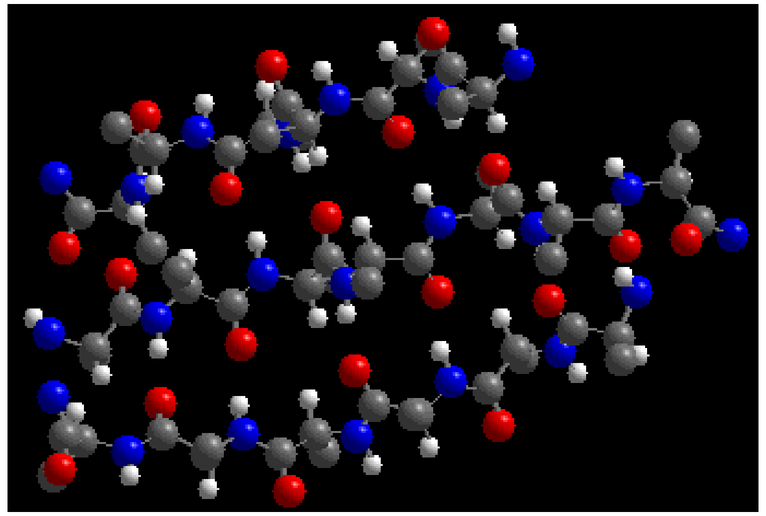


Problems

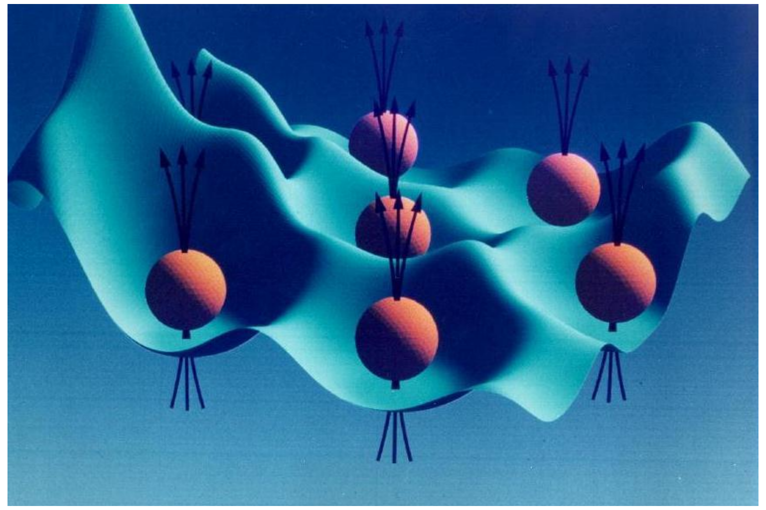

Quantumly Hard
Quantumly Easy
Classically Easy
What’s in here?



particle collision



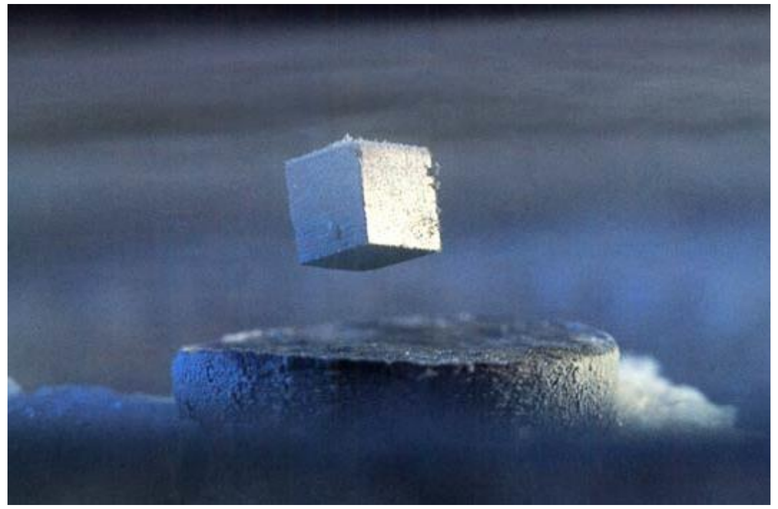
molecular chemistry



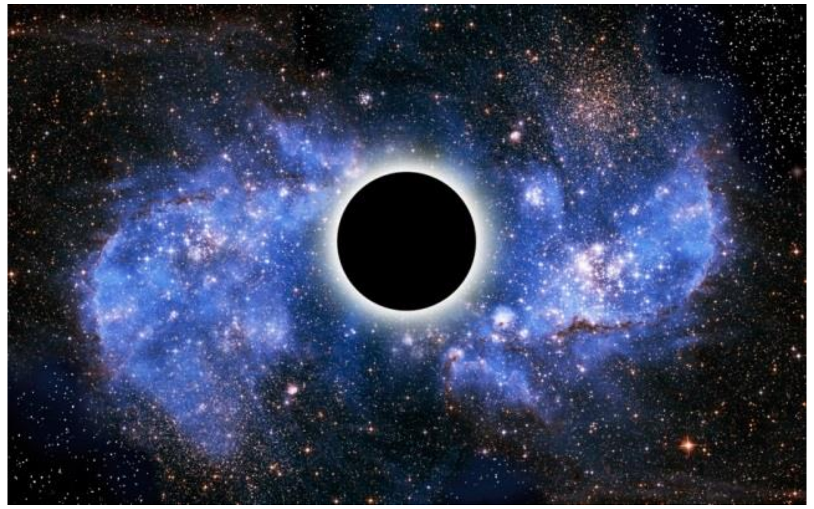
entangled electrons

A quantum computer can simulate efficiently any physical process that occurs in Nature.

(Maybe. We don’t actually know for sure.)



superconductor



black hole

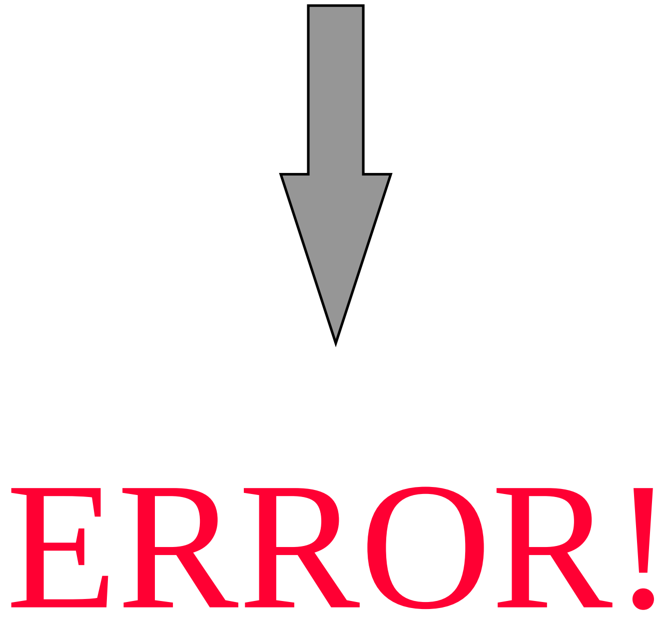
early universe

Why quantum computing is hard

We want qubits to interact strongly with one another.

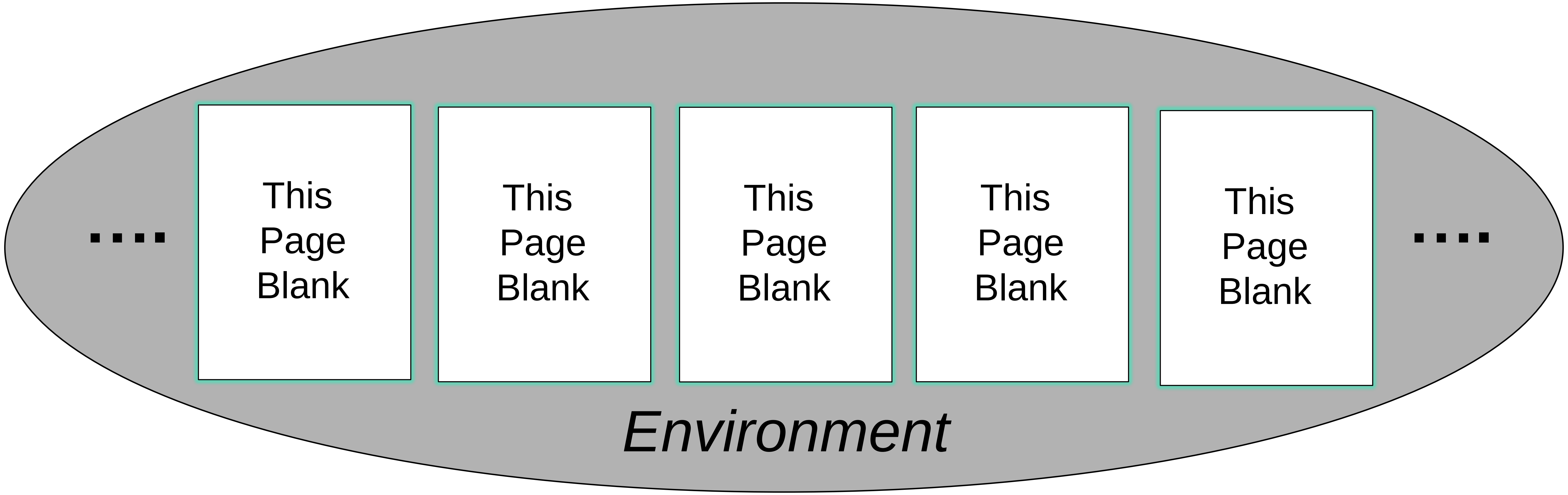
We don’t want qubits to interact with the environment.

Except when we control or measure them.



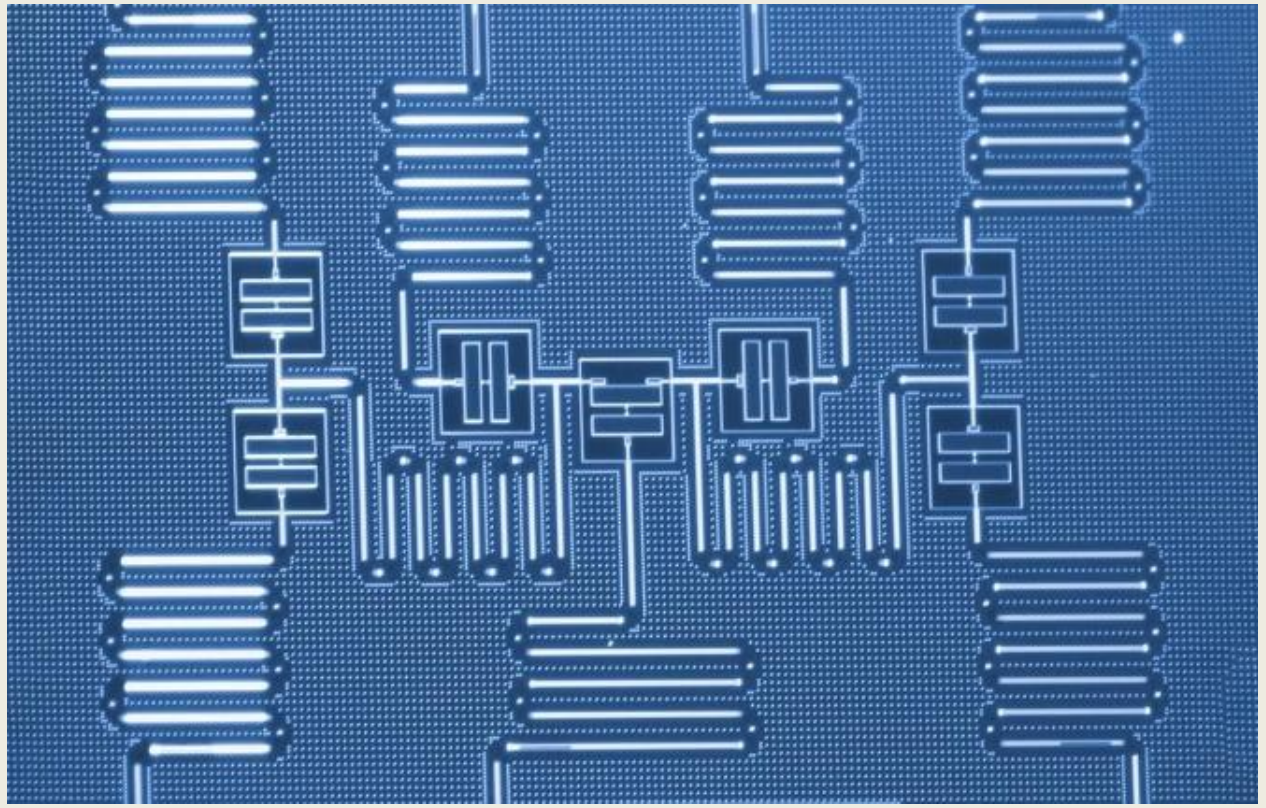
To resist decoherence, we must prevent the environment from "learning" about the state of the quantum computer during the computation.

Quantum error correction

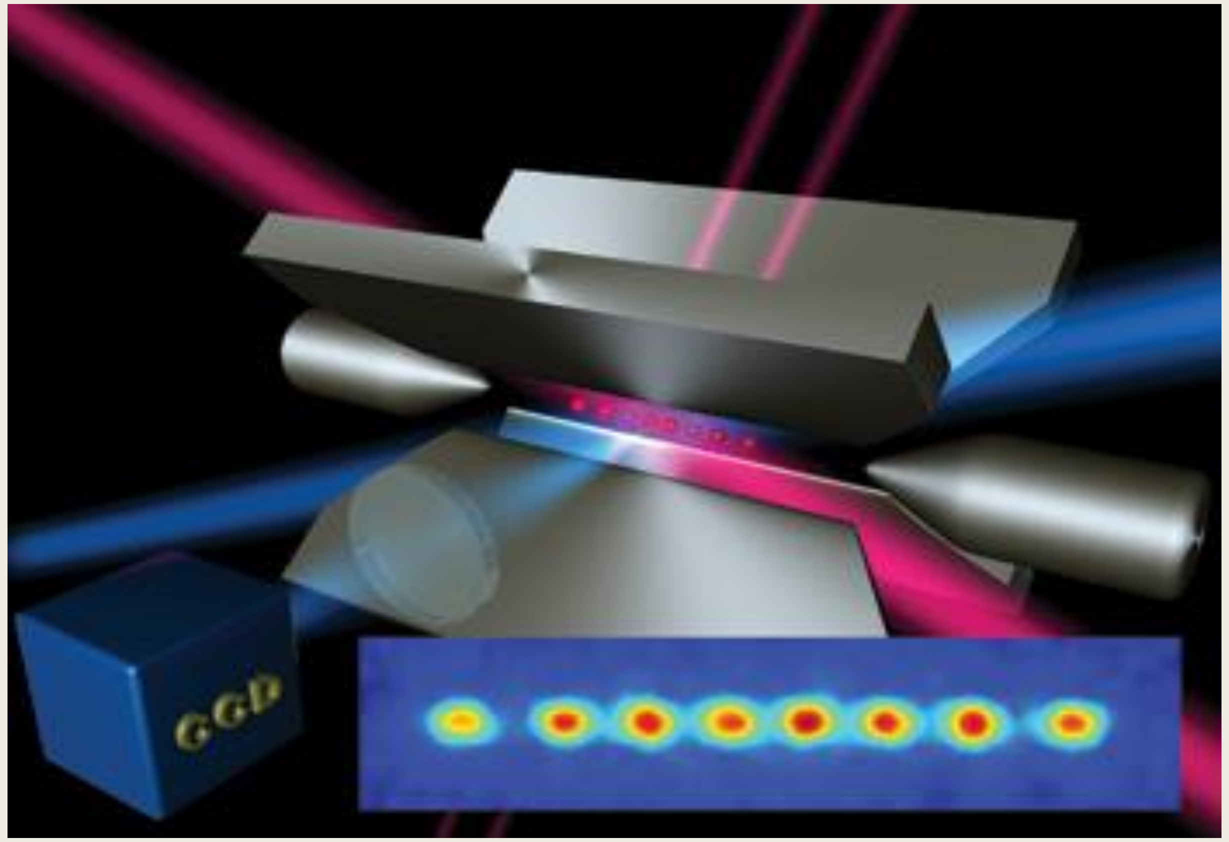


The protected "logical" quantum information is encoded in a highly entangled state of many physical qubits.

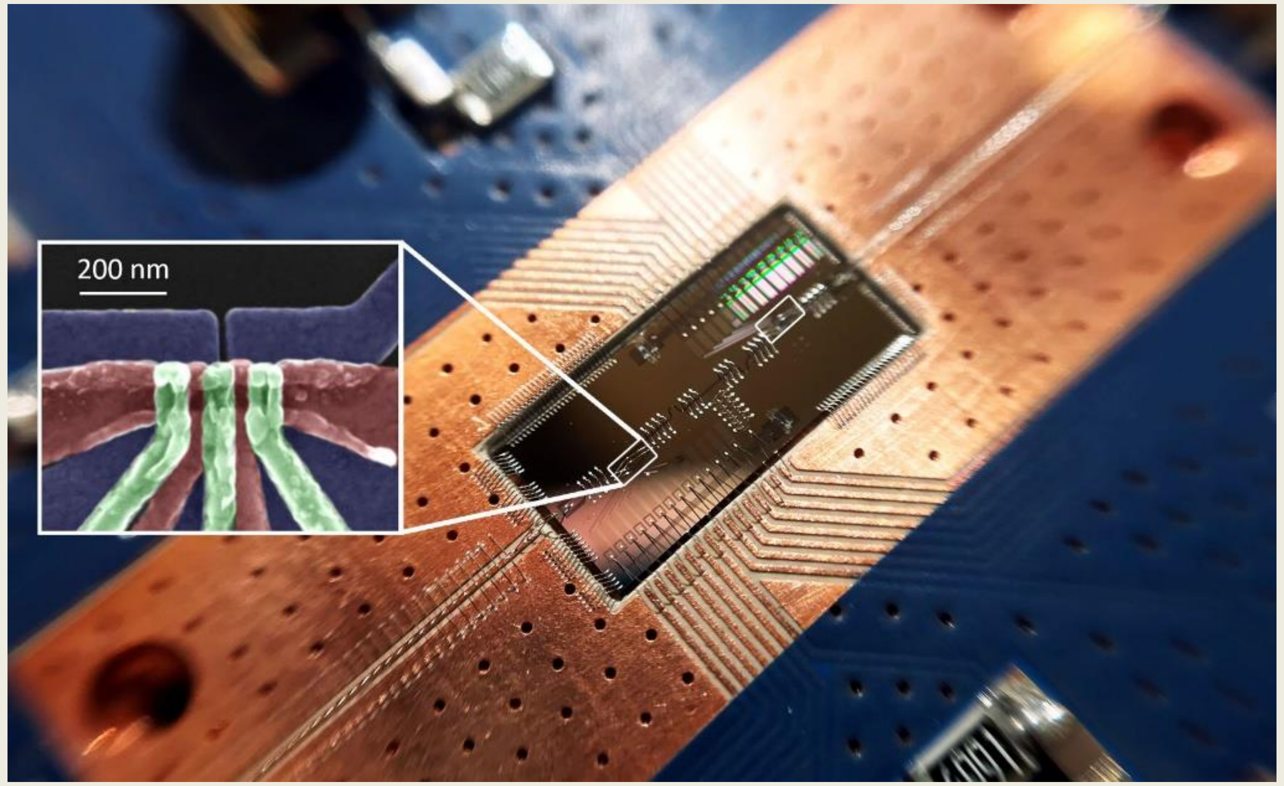
The environment can't access this information if it interacts locally with the protected system.



superconducting qubits



trapped ions



silicon spin qubits

photonics

Quantum computing in the NISQ Era

The (noisy) 50-100 qubit quantum computer has arrived.
(*NISQ* = noisy intermediate-scale quantum.)

NISQ devices cannot be simulated by brute force using the most powerful currently existing supercomputers.

Noise limits the computational power of NISQ-era technology.

NISQ will be an interesting tool for exploring physics. It *might* also have other useful applications. But we’re not sure about that.

NISQ will not change the world by itself. Rather it is a step toward more powerful quantum technologies of the future.

Potentially transformative scalable quantum computers may still be decades away. We’re not sure how long it will take.

Quantum 2, 79 (2018), arXiv:1801.00862

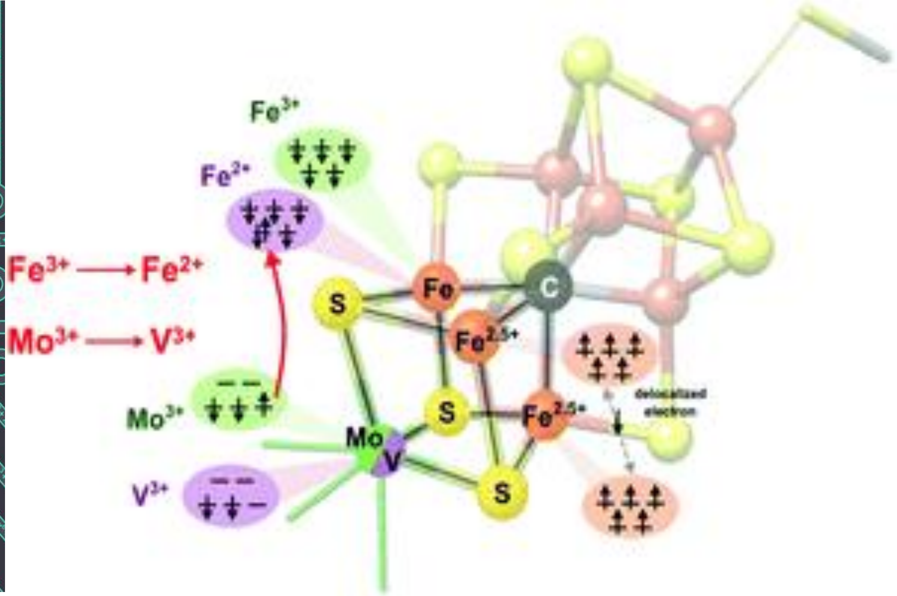
Applications of Quantum Computing

Cryptography

Break 2048 RSA

~ 6000 qubits

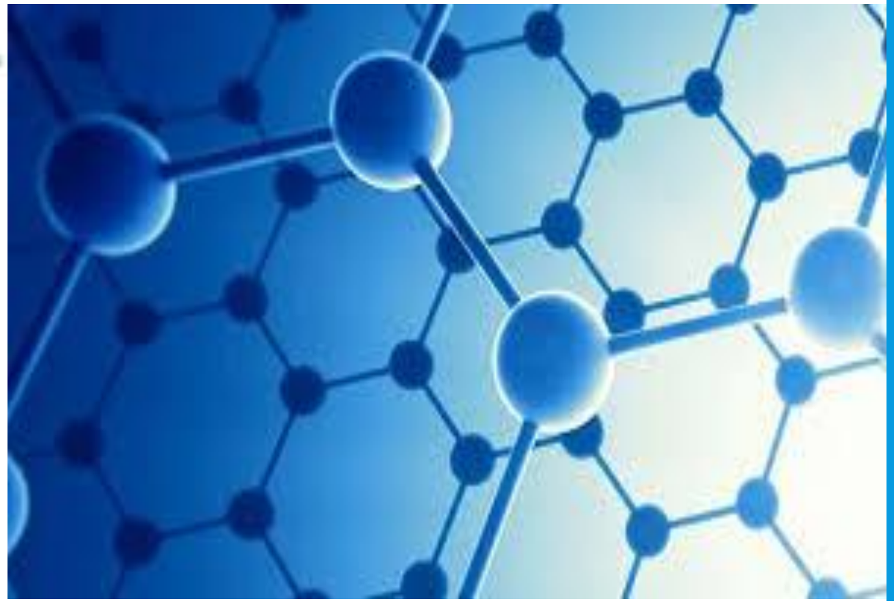
Physics/Chemistry



Simulate FeMoco

~ 200 qubits

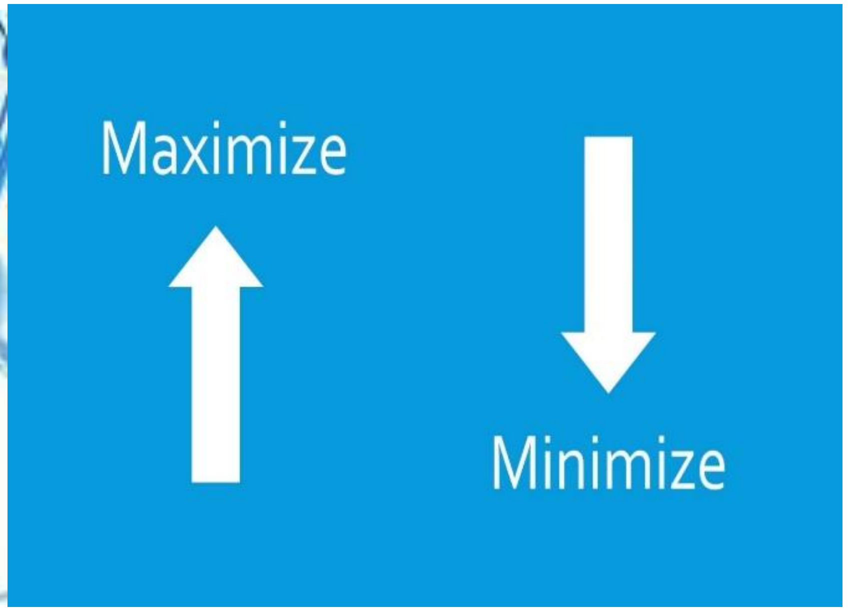
Materials Science



Simulate High-T_c superconductors

~ 70 qubits

Optimization



Scheduling, ranking, learning

~ 100 qubits (?)

Catch: perfect qubits with no noise

Applications of Quantum Computing

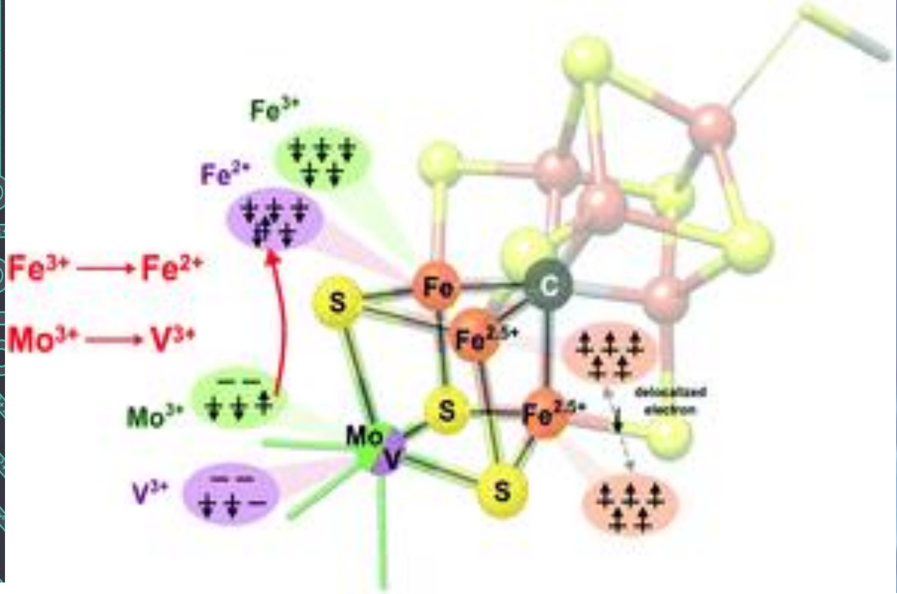
Cryptography

Break 2048 RSA

~

~10M qubits

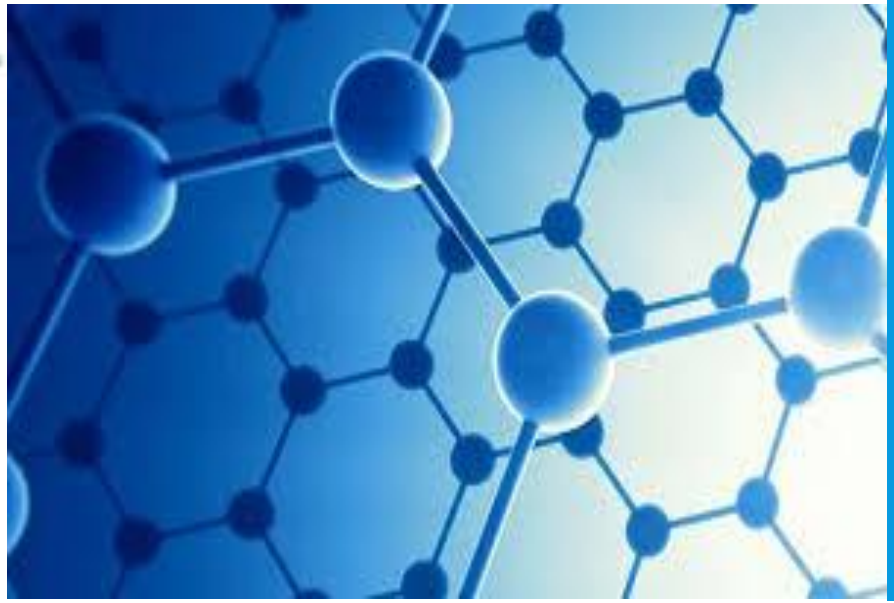
Physics/Chemistry



Simulate FeMoco

~ 1M qubits

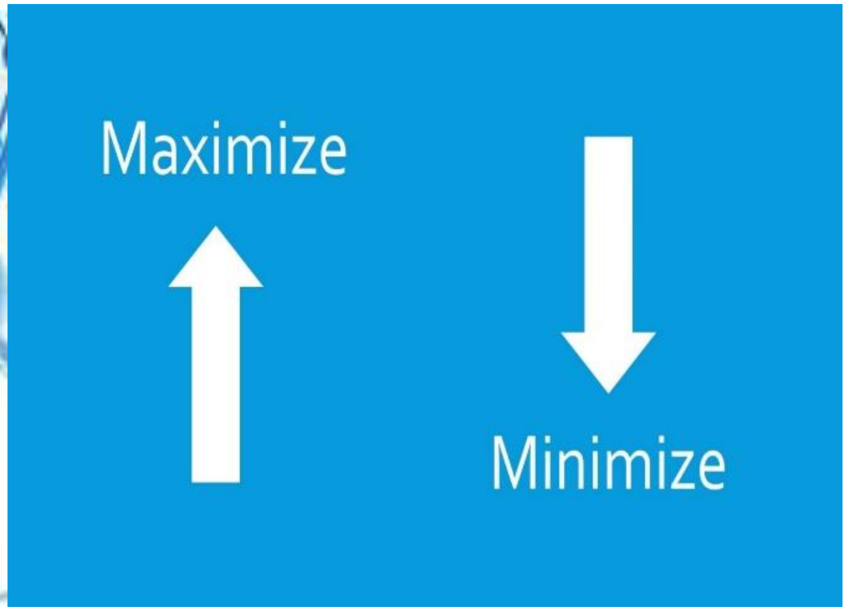
Materials Science



Simulate High-T_c superconductors

~ 100k qubits

Optimization



Scheduling, ranking, learning

~ 1M qubits (?)

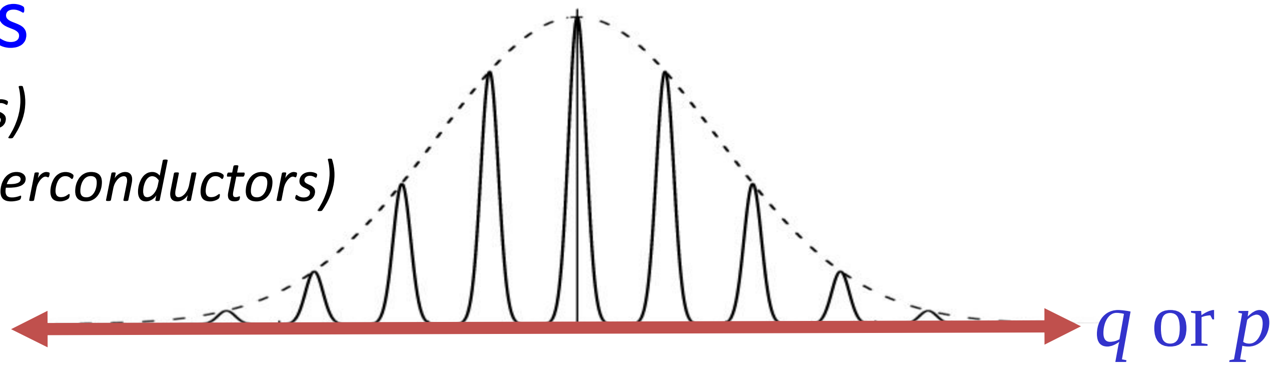
Qubits with 0.1% error rate

(Much) better gate error rates?

GKP codes

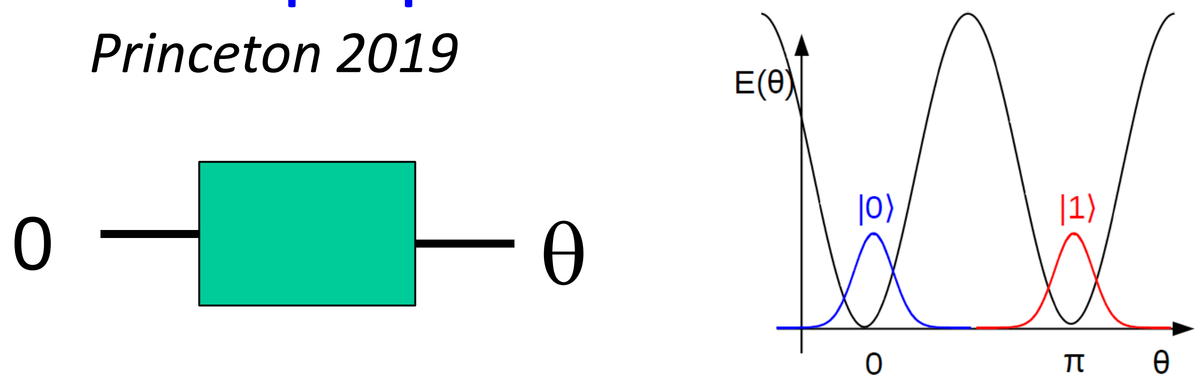
ETH 2018 (ions)

Yale 2019 (superconductors)



Zero-pi qubit

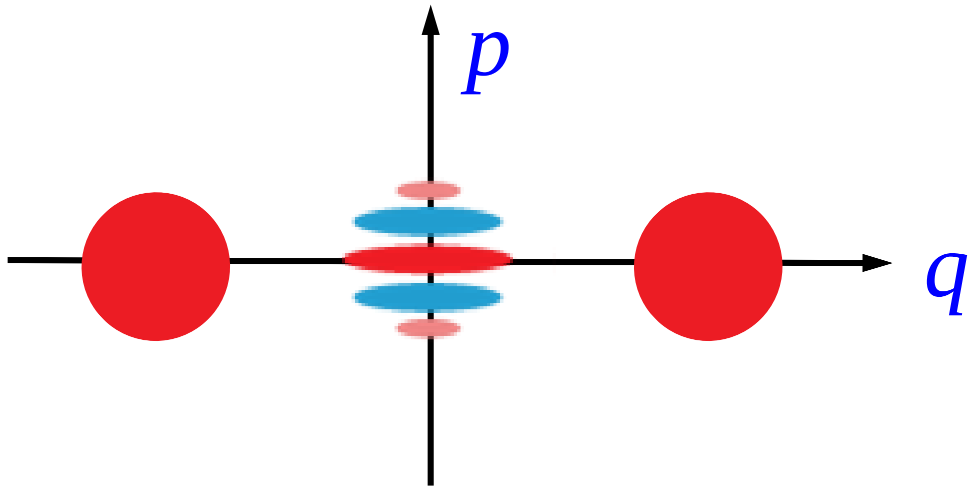
Princeton 2019



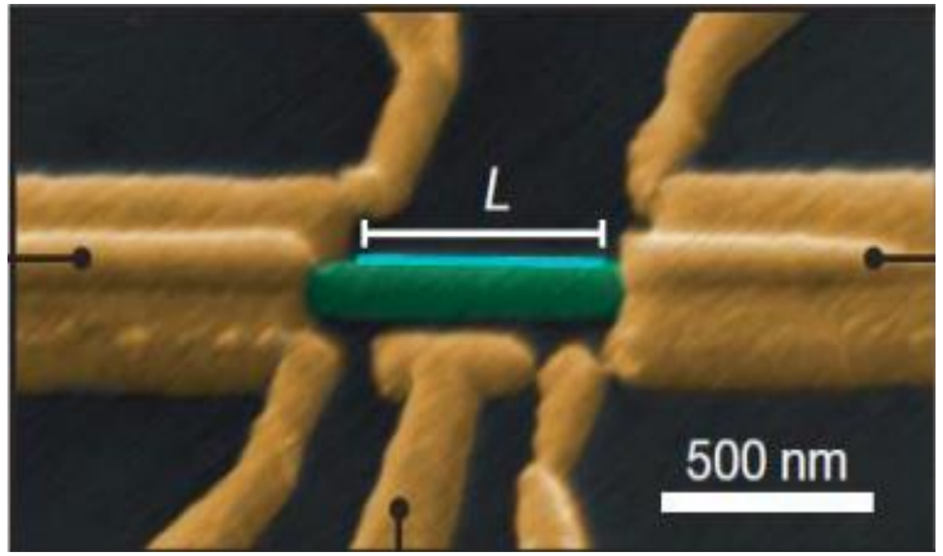
Concatenated cat codes

Yale 2019

Paris 2019



Majorana qubit



Open Questions

How will we scale up to quantum computing systems that can solve hard problems?

What are the important applications for science and for industry?

Prospects for the next 5 years

Encouraging progress toward scalable fault-tolerant quantum computing.

Scientific discoveries enabled by programmable quantum simulators and circuit-based quantum computers.

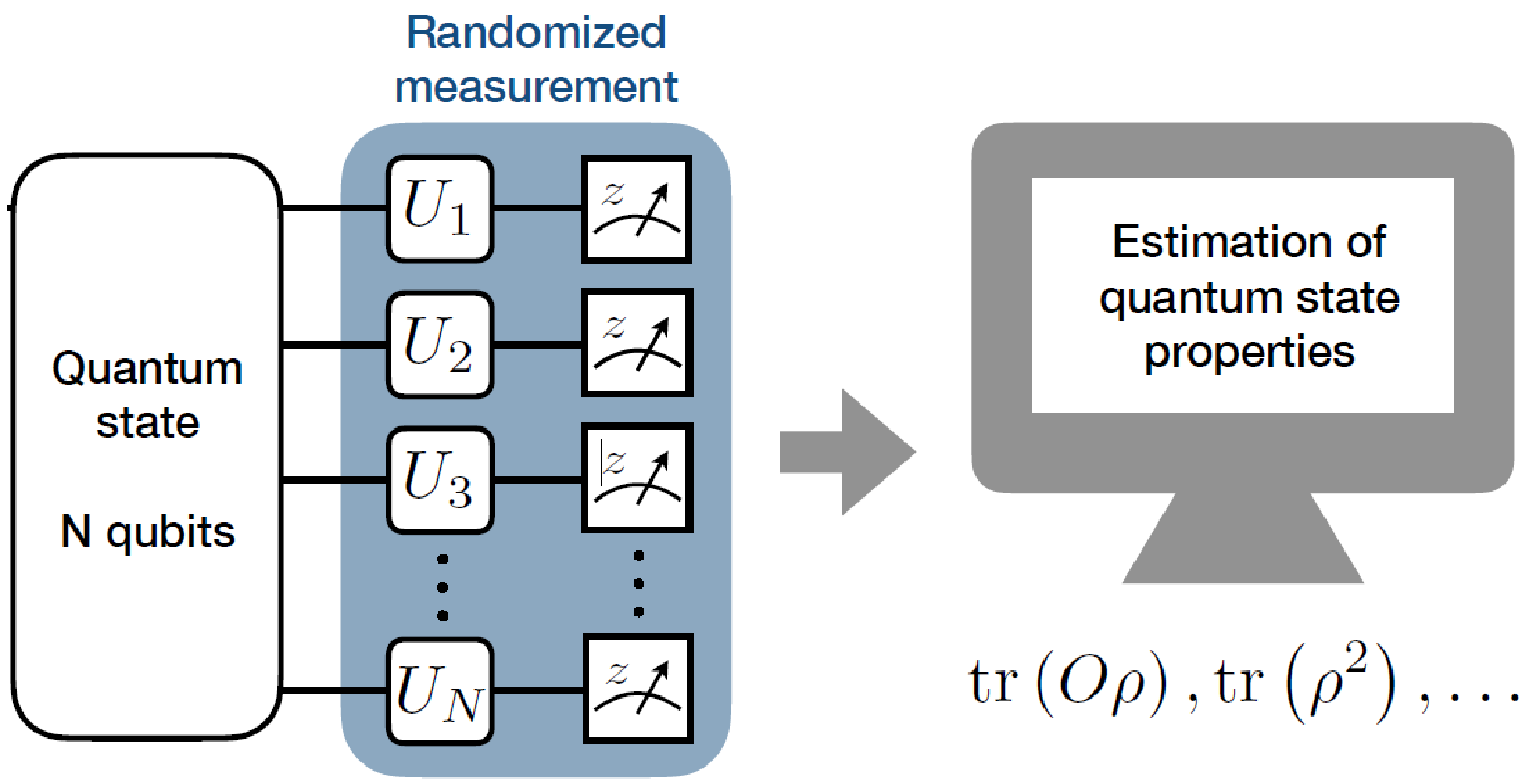
Making predictions in a quantum world

We live in a quantum world, yet we are classical beings.

Sometimes our classical nature impedes our ability to interact with, learn from, and understand the underlying quantum reality.

How can classical and quantum machines enhance our ability to learn about the quantum world?

Learning about the quantum world using classical machines



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

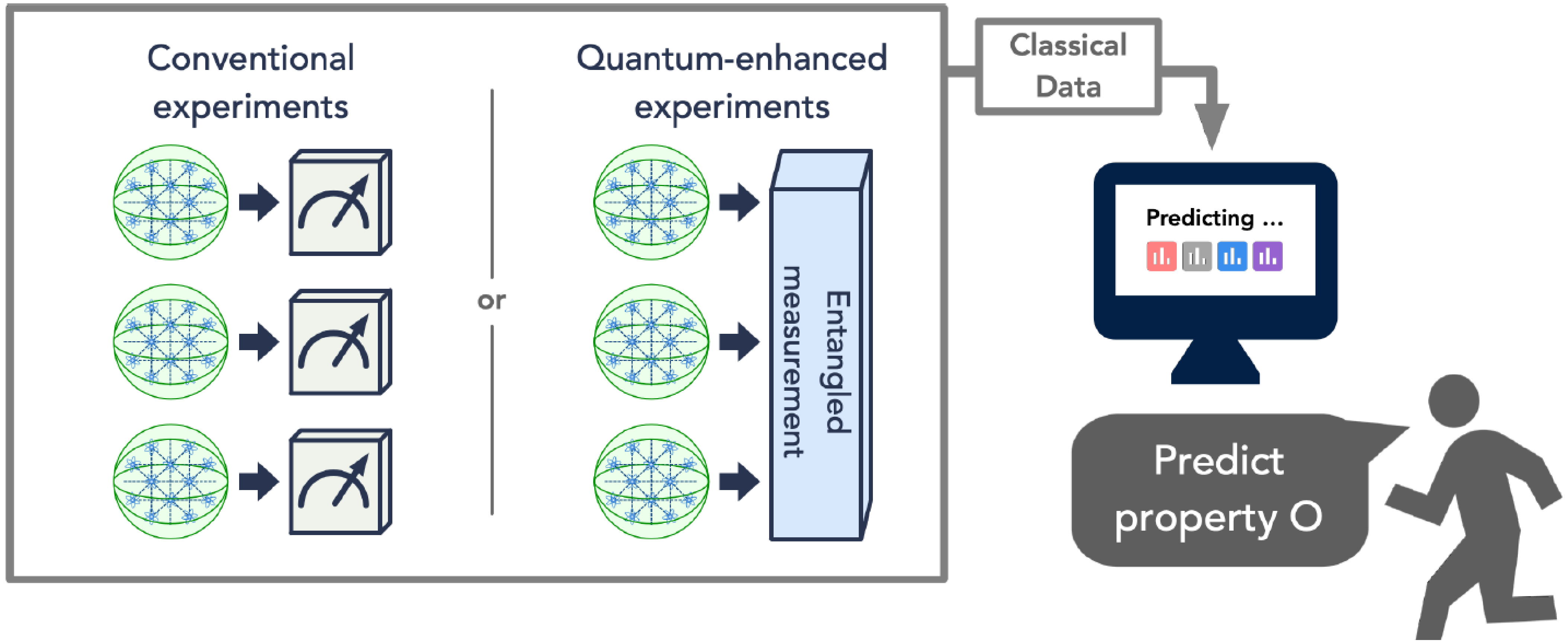
Apply classical processing (including machine learning) to the classical description.

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, Preskill 2022; Lewis, Huang,Tran, Lehner, Kueng, Preskill 2023

Learning about the quantum world using quantum machines



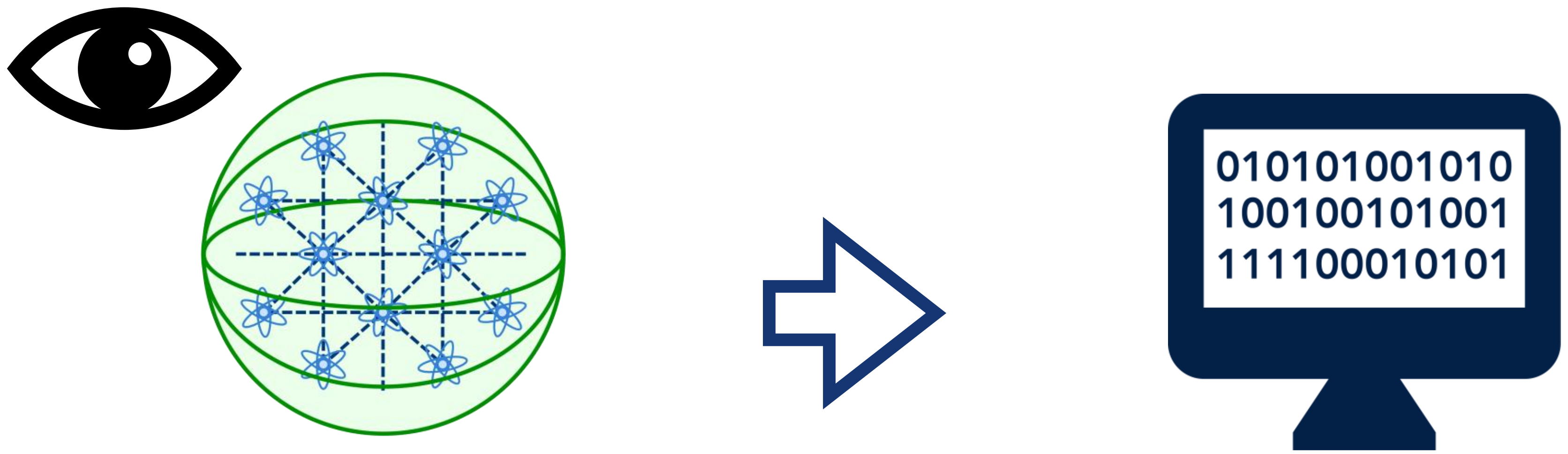
Quantum-enhanced measurement strategies: transduce detected quantum data to quantum memory and process it with a quantum computer.

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

Curse of dimensionality



Unknown Quantum System **Classical Description**

An n-qubit quantum system has Hilbert space dimension 2^n.

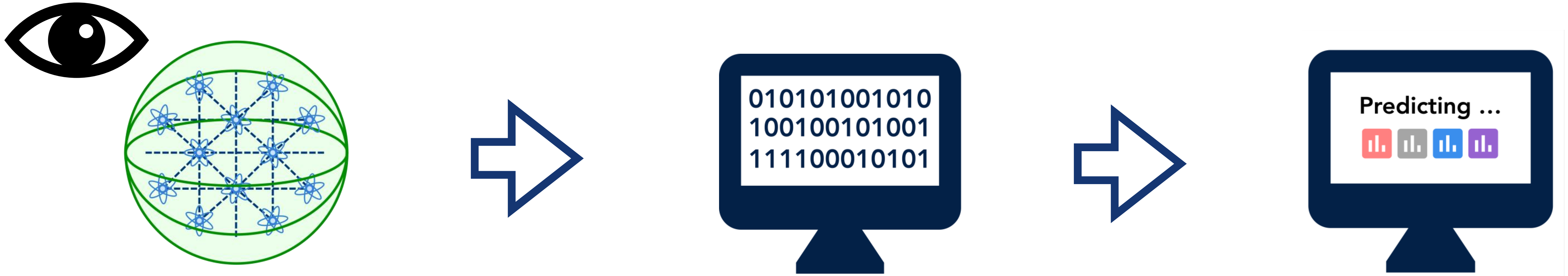
A complete classical description of a generic quantum state requires $O(4^n)$ bits.

In the worst case, acquiring such a description requires $O(4^n)$ experiments.

Using the classical description to predict a property (such as the expectation value of an observable) has a (worst-case) classical computation cost $O(4^n)$.

But a complete description might not be needed. We might be satisfied if we can predict many properties of the state.

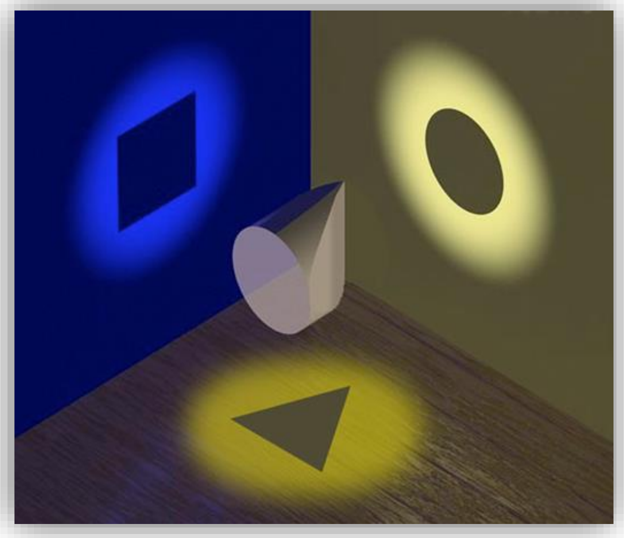
Classical shadows of quantum states



Unknown Quantum System | ***Efficient* Classical Description** | **Predicted Properties**

A tractable protocol backed by rigorous theory.

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

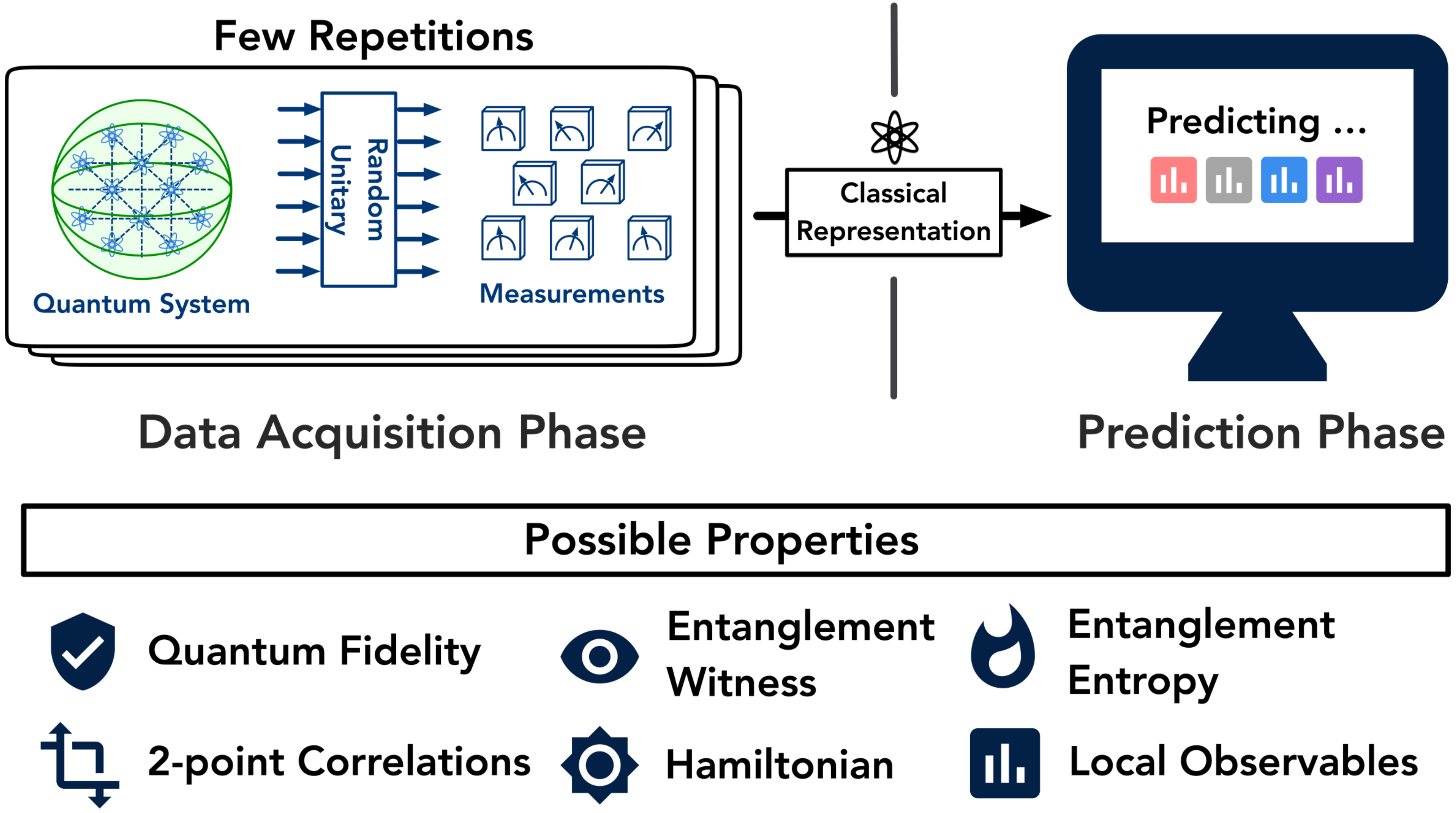


classical shadows

Huang, Kueng, Preskill 2020

Classical shadows of quantum states

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

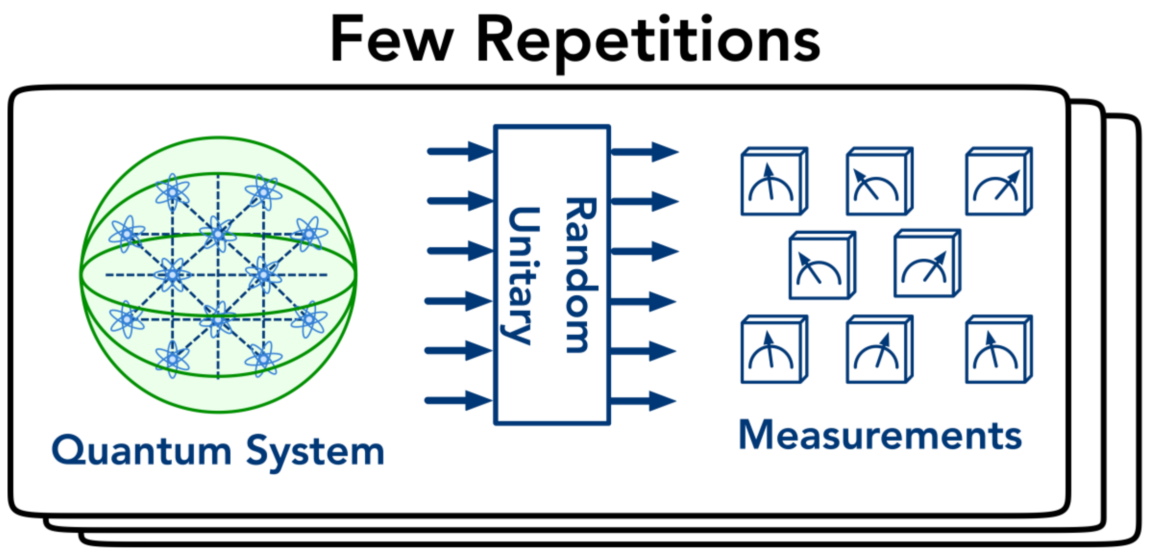


The Procedure: Data Acquisition Phase

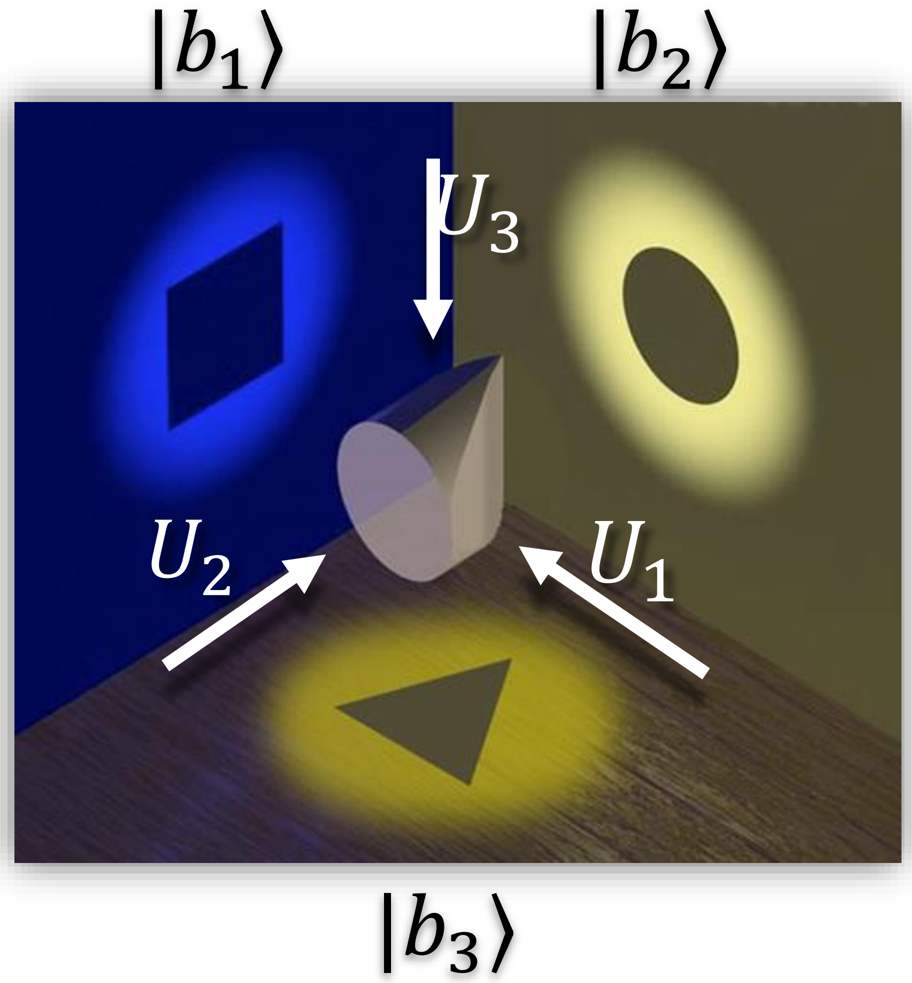
Given multiple copies of n-qubit quantum state ρ and an ensemble of unitary transformations $\{U_i\}$, repeat N times:

- 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



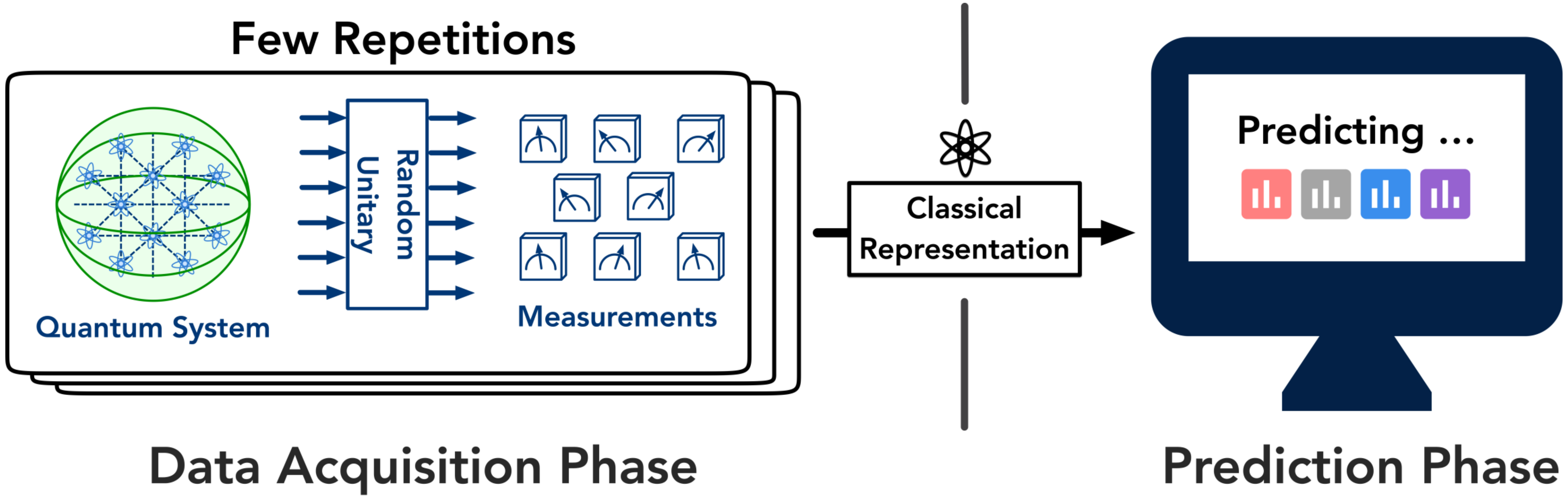
The Procedure: Prediction Phase

Given $S(\rho) = \{|s_1\rangle, \dots, |s_N\rangle\}$ (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
$$N = \mathcal{O}(B \log(M)/\epsilon^2) \text{ measurements.}$$
2. Subsequently, given any $O_1, \ldots, O_M$ with $B \geq \max \| O_i \|_{\text{shadow}}^2$, 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 $\| O \|_{\text{shadow}}^2$ 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: $\| O \|_{\text{shadow}}^2 \leq 3\, \mathrm{tr}(O^2)$ Application: Quantum fidelity $O = |\psi\rangle\langle\psi|$

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

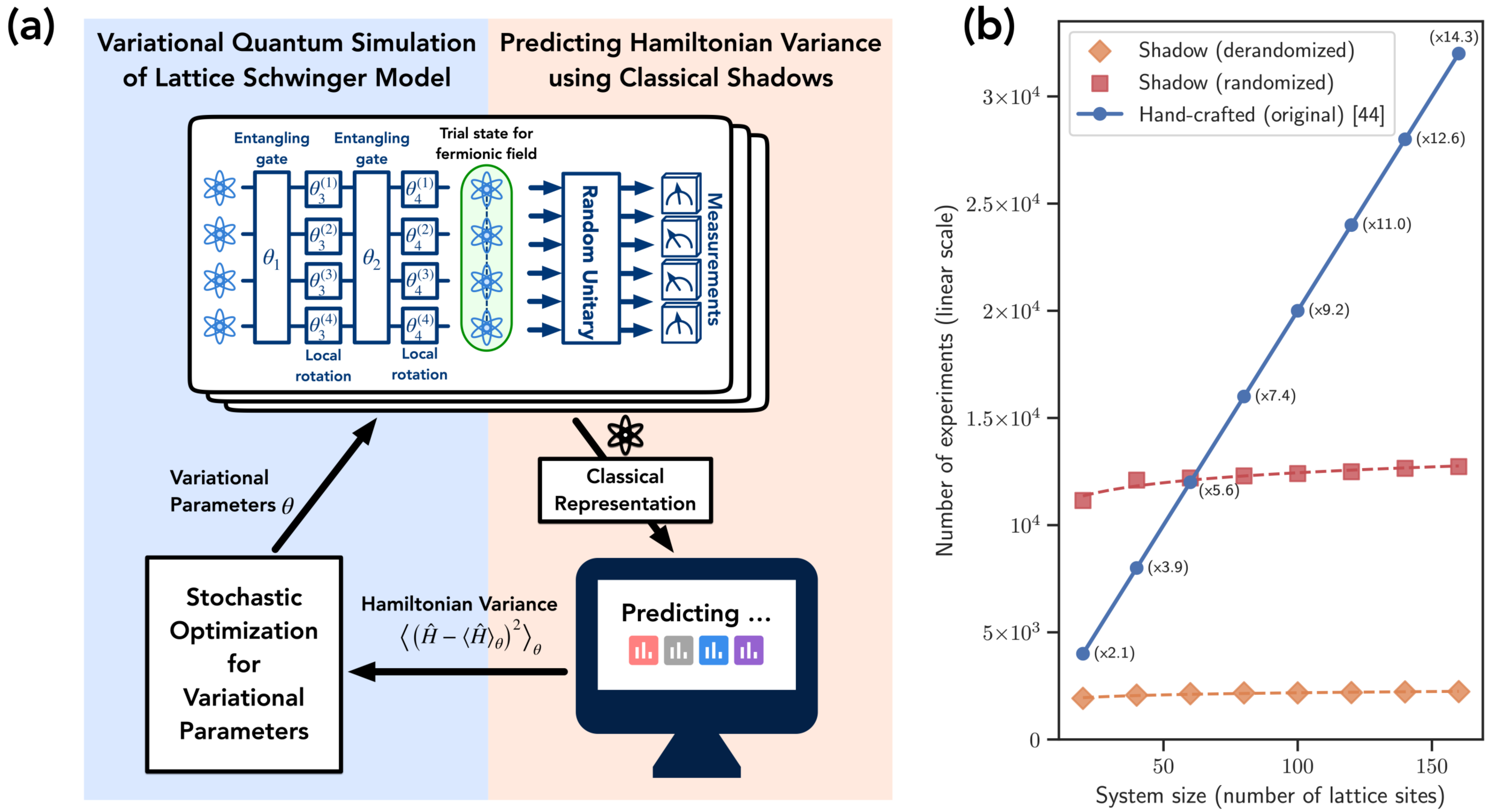
Observable O acts on w qubits

Huang, Kueng, Preskill 2020

“Measure first, ask questions later.”

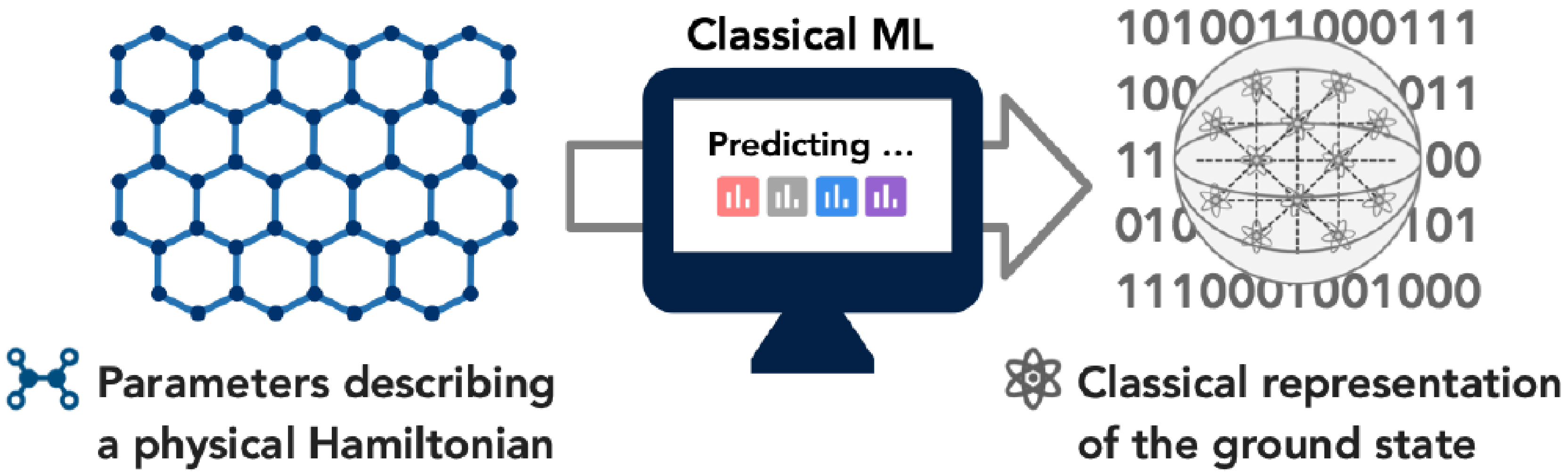
Elben, Flammia, Huang, Kueng, Preskill, Vermersch, Zoller,
The randomized measurement toolbox 2022

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



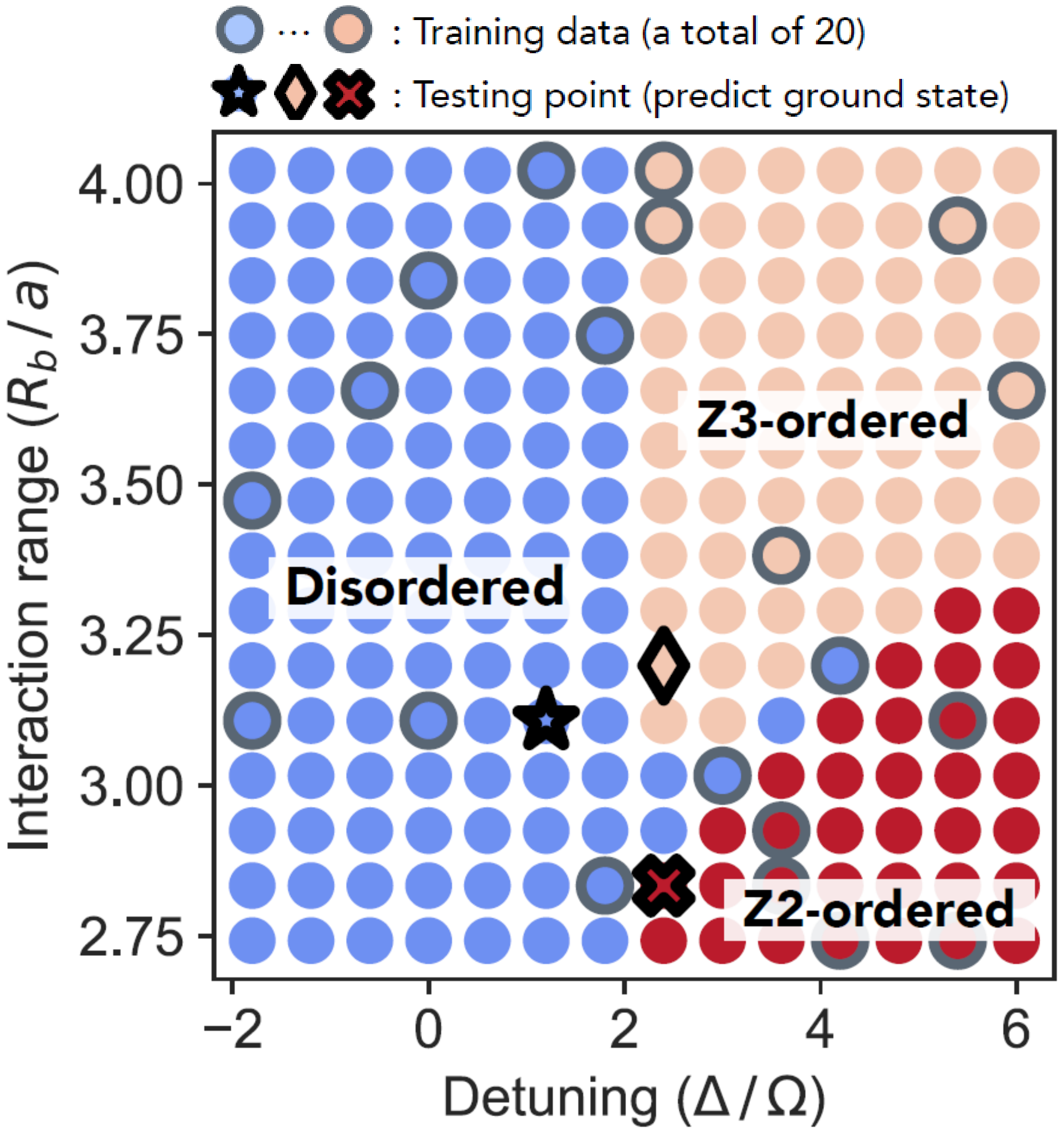
Theorem (*Learning properties of ground states*):
For any smooth family of local Hamiltonians $\{H(x), x \in [-1,1]^m\}$ in a finite spatial dimension with a constant spectral gap, a classical machine learning algorithm can learn to predict an efficient classical representation of the ground state $\rho(x)$ 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

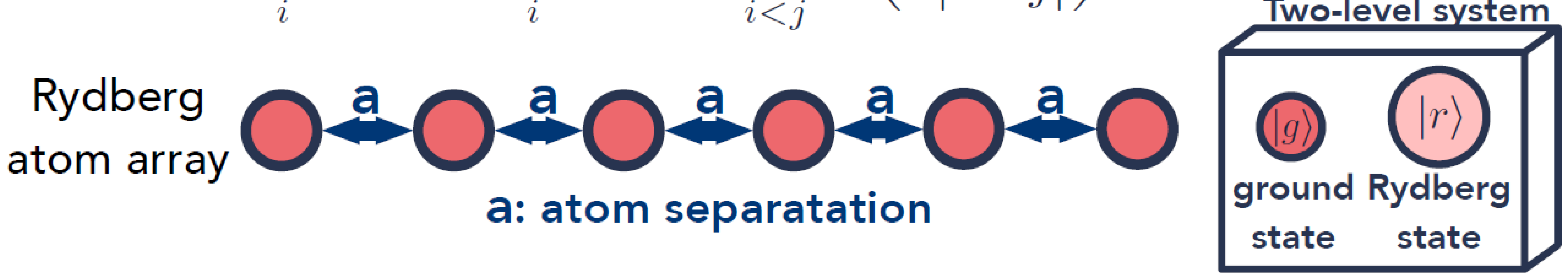


Chain of 51 atoms (as in Bernien *et al.* 2017). We can compute ground state properties using DMRG.

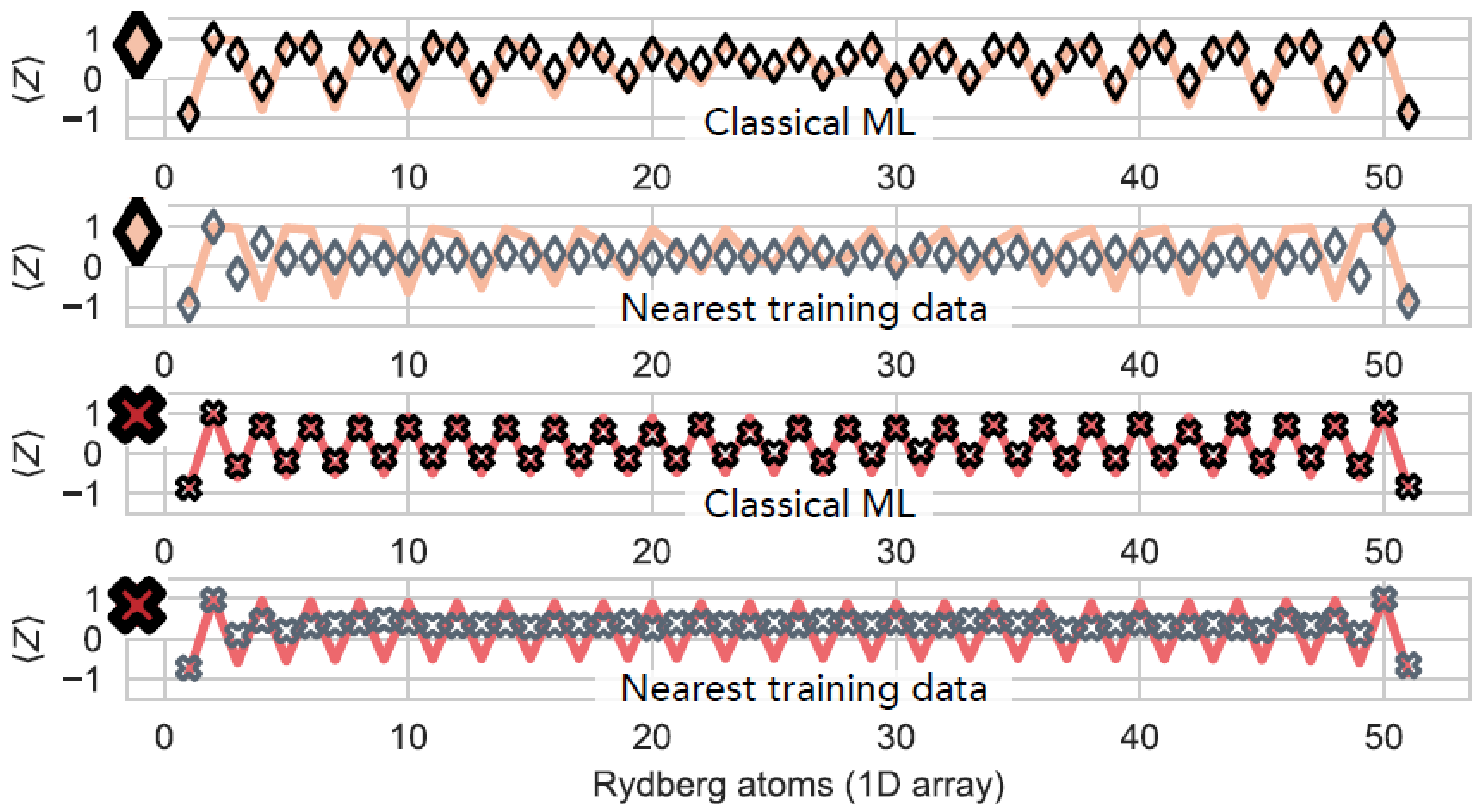
Our rigorous theory does not directly apply, because Hamiltonian is not gapped throughout the parameter regime considered. Yet predictions work well.

500 snapshots taken at each sampled value of x.

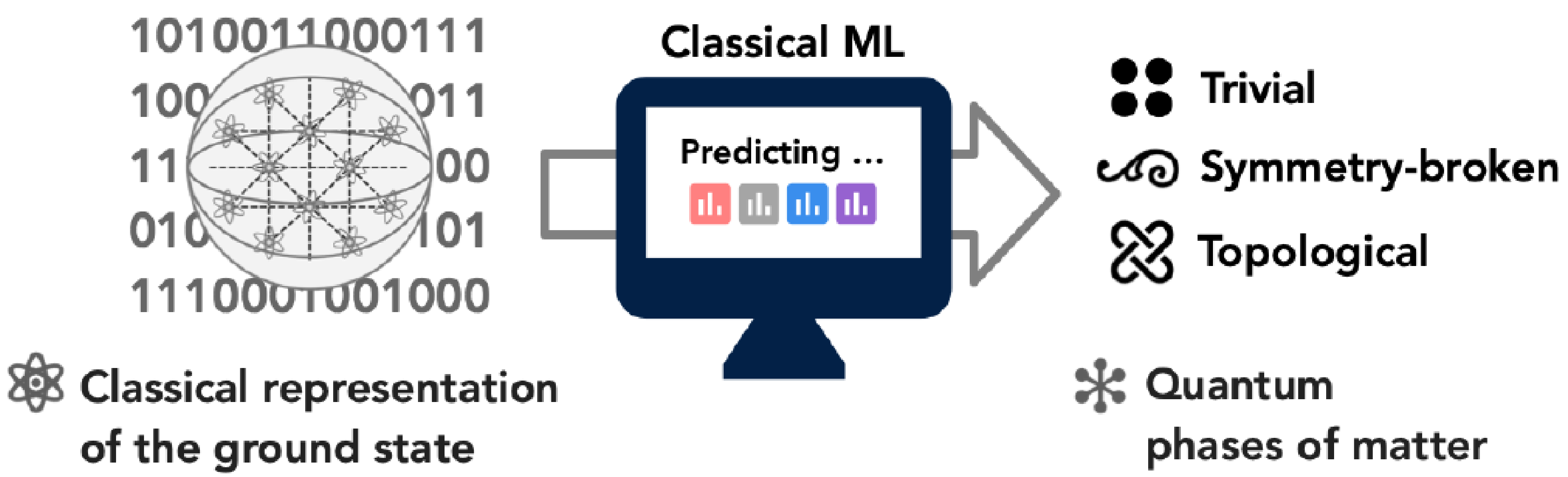
$$H = \sum_i \frac{\Omega}{2} X_i - \sum_i \Delta N_i + \sum_{i<j} \Omega \left(\frac{R_b}{a|i-j|} \right)^6 N_i N_j$$



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



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.

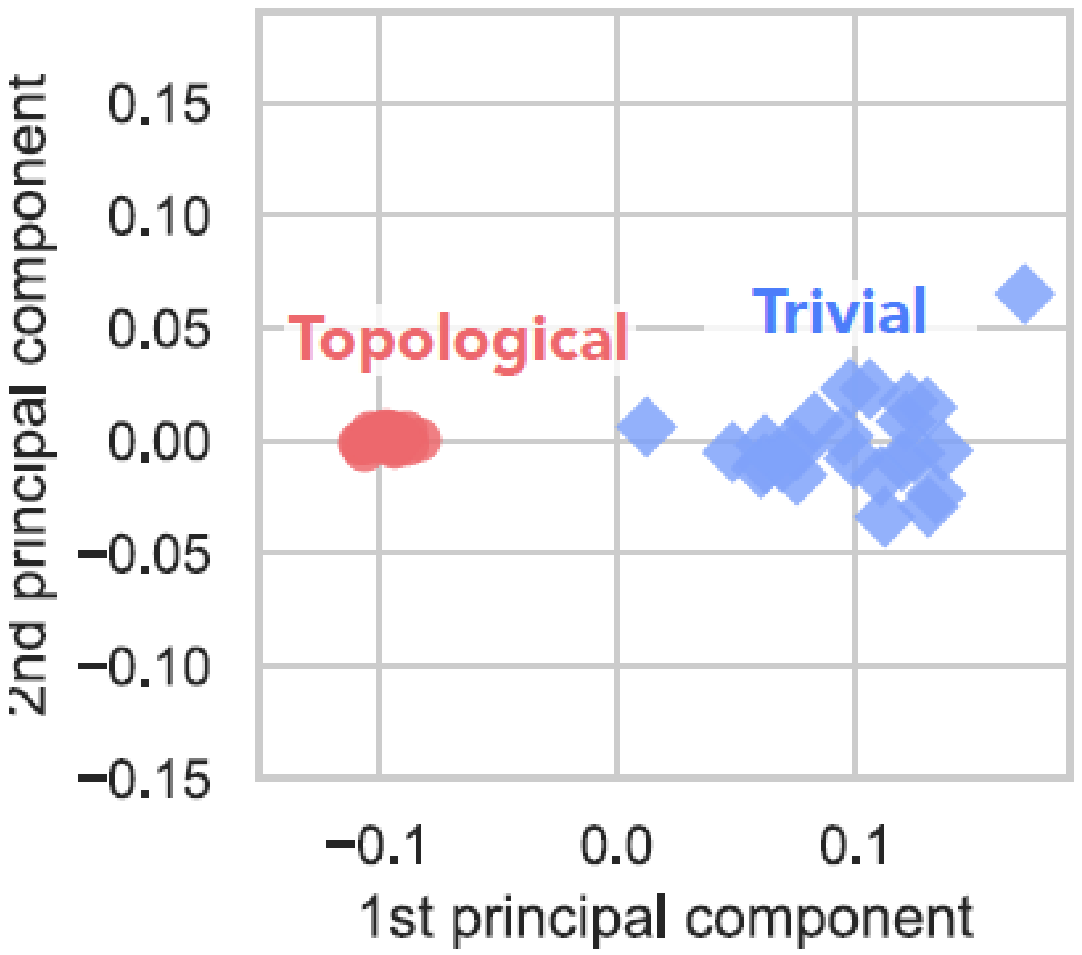
Idea: convert each quantum state to its classical shadow, and learn to classify these shadows.

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

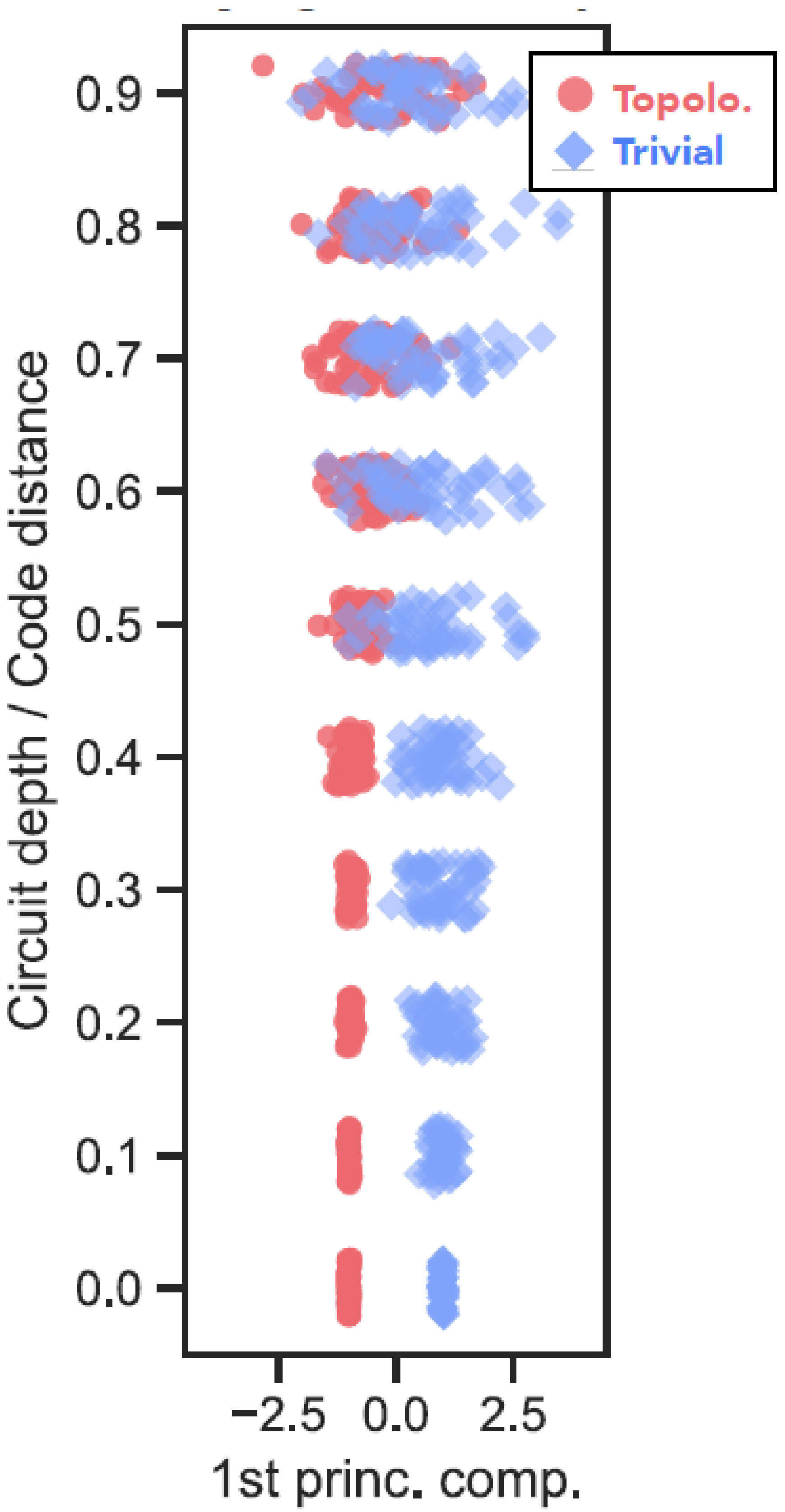
Example: Distinguishing the 2D toric code phase from the trivial phase



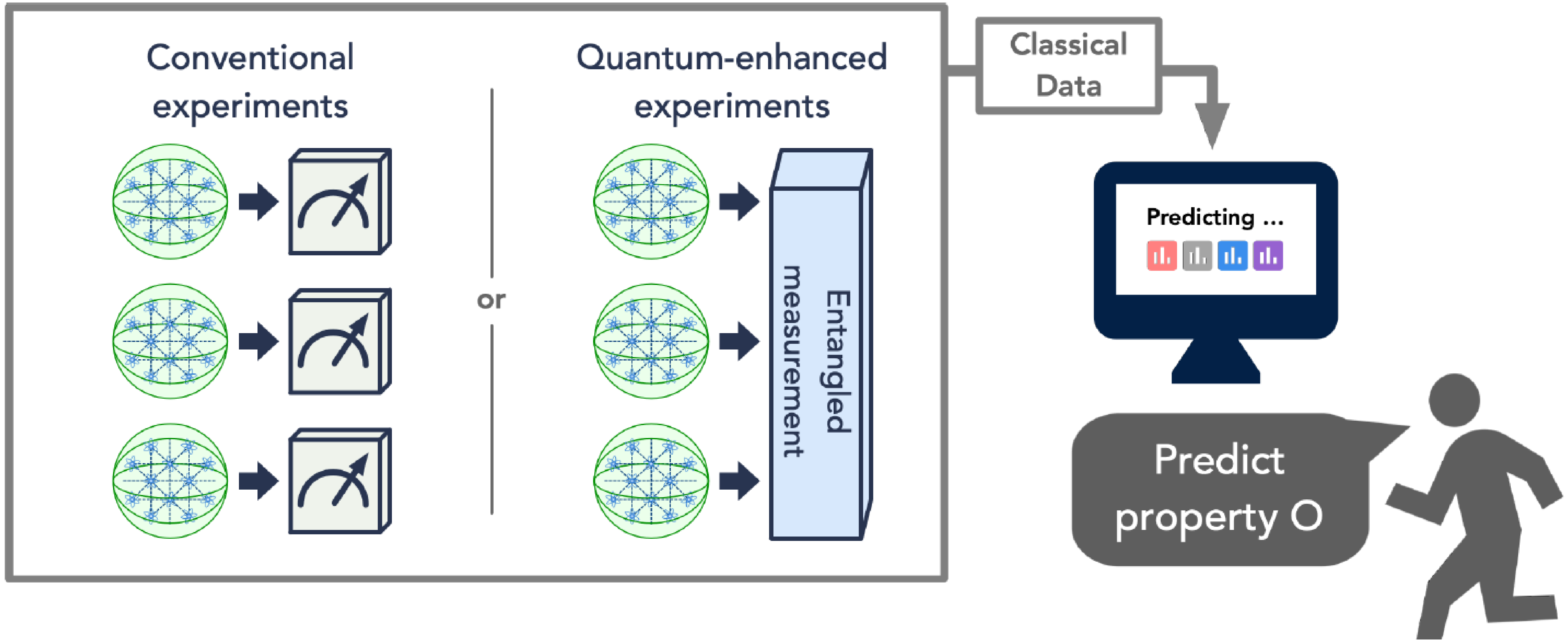
No local circuit of constant depth acting on a product state can reach a topologically ordered state.

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.

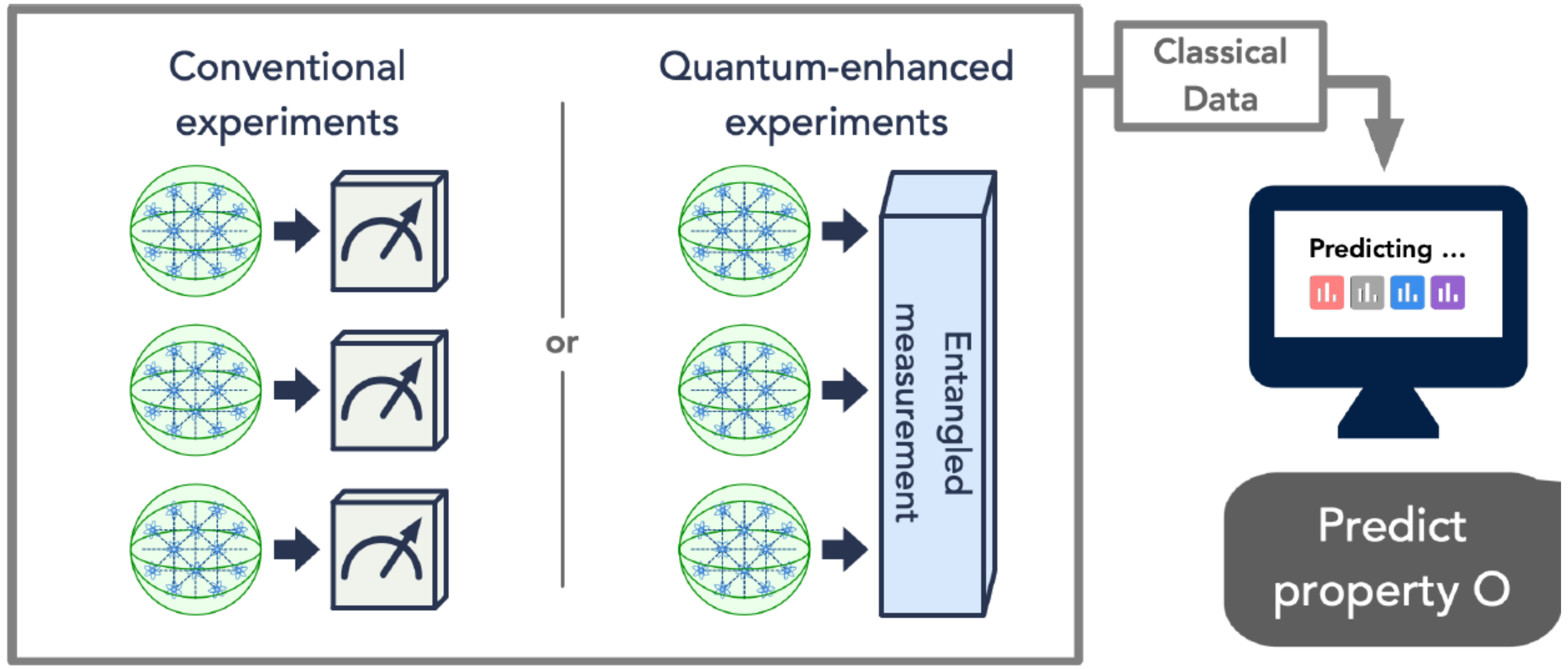
Huang, Broughton, Cotler, Chen, Li, Mohseni, Neven, Babbush, Kueng, Preskill, McClean 2022

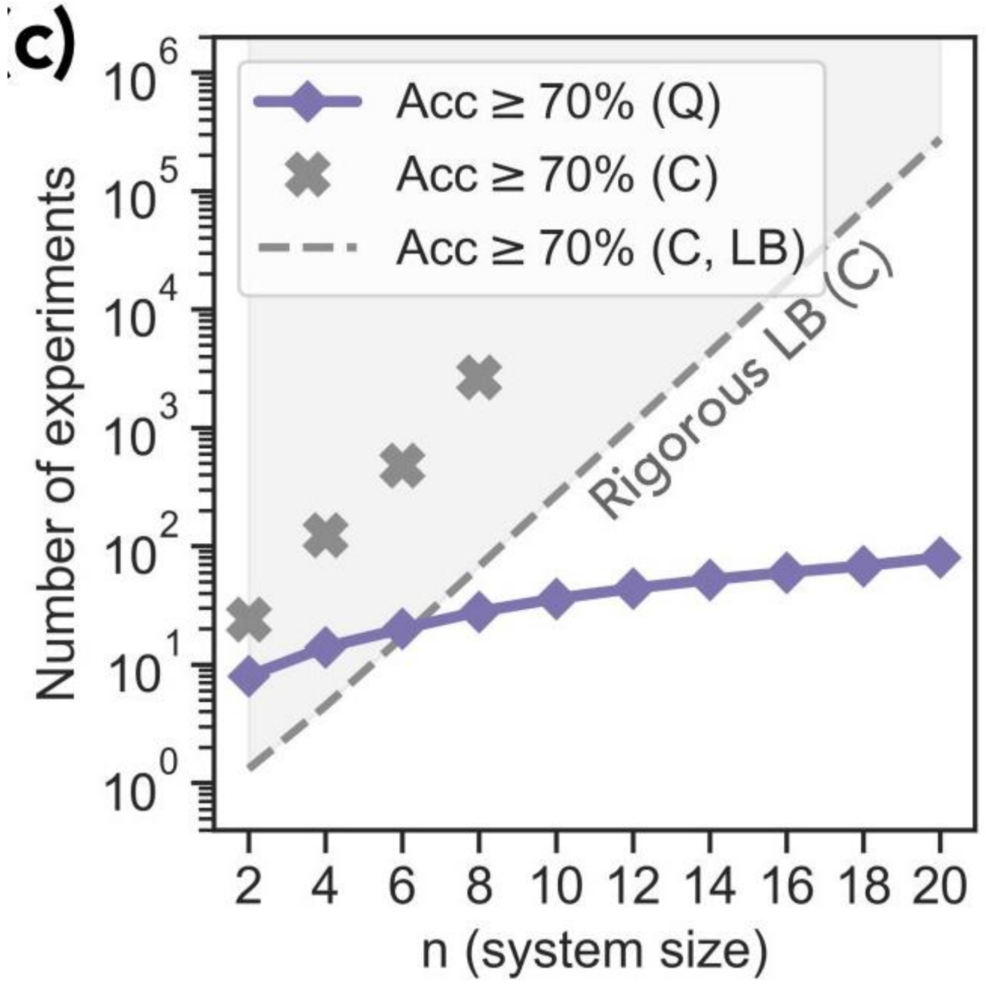
Conventional experiments vs. quantum-enhanced experiments

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

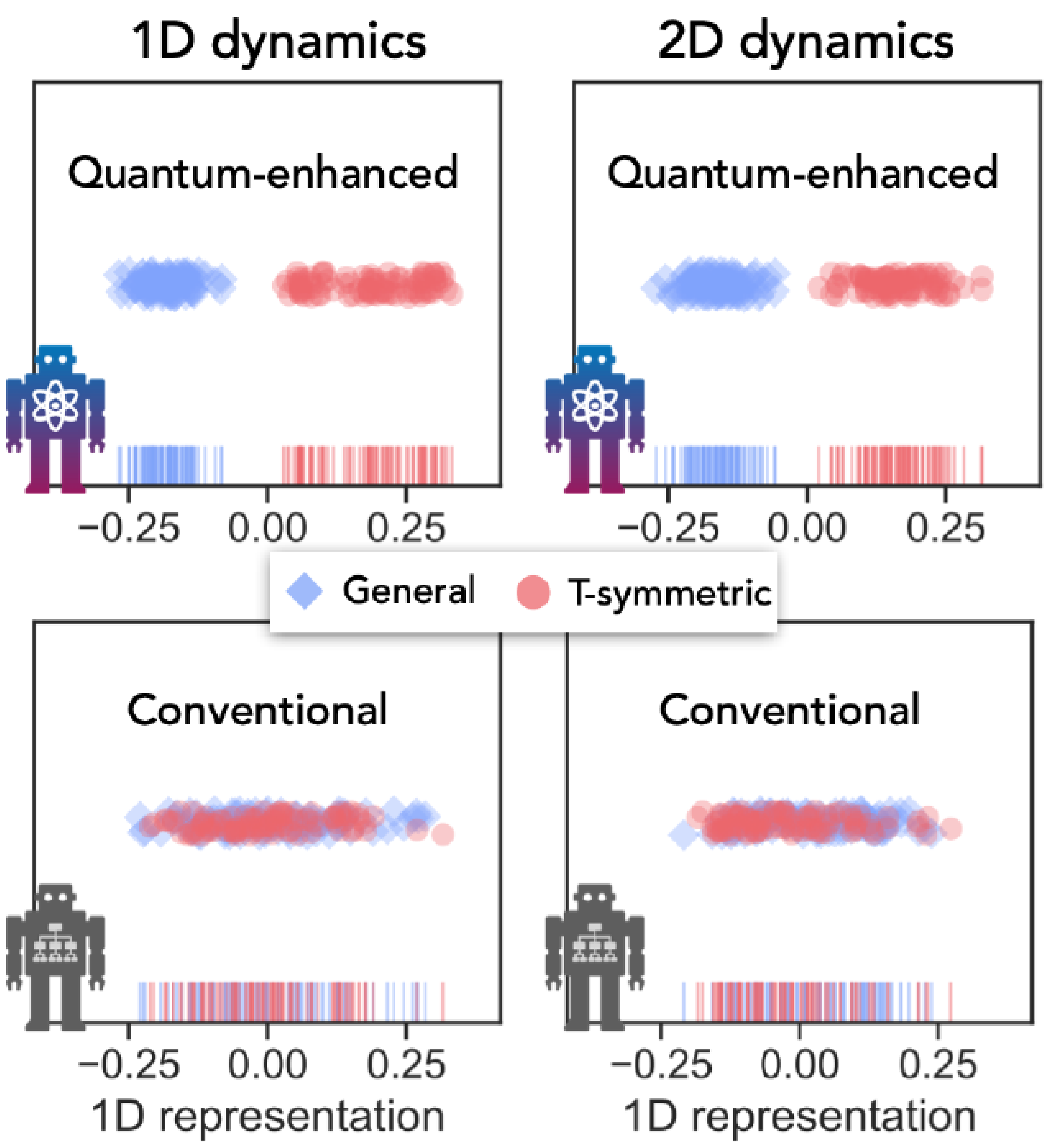
Conventional experiments vs. quantum-enhanced experiments

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

Learning in a quantum world

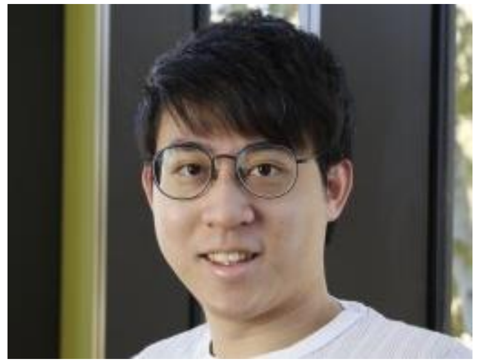
Broadly useful applications of quantum computing may still be a ways off, and quantum error correction is most likely the key to getting there. But existing quantum platforms already provide unprecedented opportunities for exploring exotic properties of quantum matter.

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.”

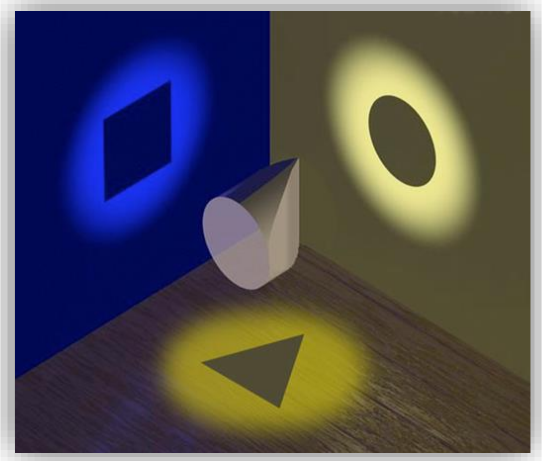
Access to data from quantum experiments may enable classical machine learning to solve quantum problems that would be too hard to solve without access to data.

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

The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.

Paul A. M. Dirac, Quantum Mechanics of Many-Electron Systems, Proceedings of the Royal Society, 1929