Fundamental aspects of solving quantum problems with machine learning

Hsin-Yuan (Robert) Huang, Richard Kueng, Michael Broughton, Masoud Mohseni, Ryan Babbush, Sergio Boixo, Hartmut Neven, Jarrod McClean and John Preskill

Google AI Quantum
Institute of Quantum Information and Matter (IQIM), Caltech
Johannes Kepler University Linz

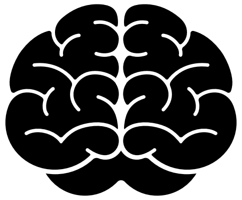
[1] **Information-theoretic bounds on quantum advantage in machine learning**, *arXiv:2101.02464*.
[2] **Power of data in quantum machine learning**, *arXiv:2011.01938*.
[3] **Provable machine learning algorithms for quantum many-body problems**, *In preparation*.

Motivation

- Machine learning (ML) has received great attention in the quantum community these days.

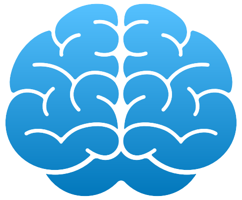
Classical ML for quantum physics/chemistry

The goal:
Solve challenging quantum many-body problems **better** than traditional classical algorithms



Enhancing ML with quantum computers

The goal:
Design quantum ML algorithms that yield significant **advantage** over any classical algorithm



"Solving the quantum many-body problem with artificial neural networks." *Science* 355.6325 (2017): 602-606.
"Learning phase transitions by confusion." *Nature Physics* 13.5 (2017): 435-439.
"Supervised learning with quantum-enhanced feature spaces." *Nature* 567.7747 (2019): 209-212.

Motivation

- Yet, many fundamental questions remain to be answered.

Classical ML
for quantum physics/chemistry

The question:
How can ML be more useful
than non-ML algorithms?

Enhancing ML
with quantum computers

The question:
What are the advantages of
quantum ML in general?

"Solving the quantum many-body problem with artificial neural networks." *Science* 355.6325 (2017): 602-606.
"Learning phase transitions by confusion." *Nature Physics* 13.5 (2017): 435-439.
"Supervised learning with quantum-enhanced feature spaces." *Nature* 567.7747 (2019): 209-212.

General Setting

- In this work, we focus on training an ML model to predict

$$x \mapsto f_{\mathscr{E}}(x) = \mathrm{Tr}(O\mathscr{E}(|x\rangle\langle x|)),$$

  where x is a classical input, $\mathscr{E}$ is an **unknown** CPTP map, and O is an observable.

- This is **very general**: includes any function computable by a quantum computer.

Example 1

Predicting outcomes of physical experiments

x : parameters describing the experiment

$\mathscr{E}$: the physical process in the experiment

O : what the scientist measure

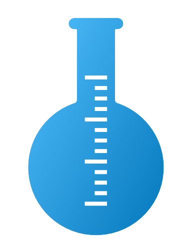
Example 2

Predicting ground state properties of a physical system

x : parameters describing a physical system

$\mathscr{E}$: a process for preparing ground state

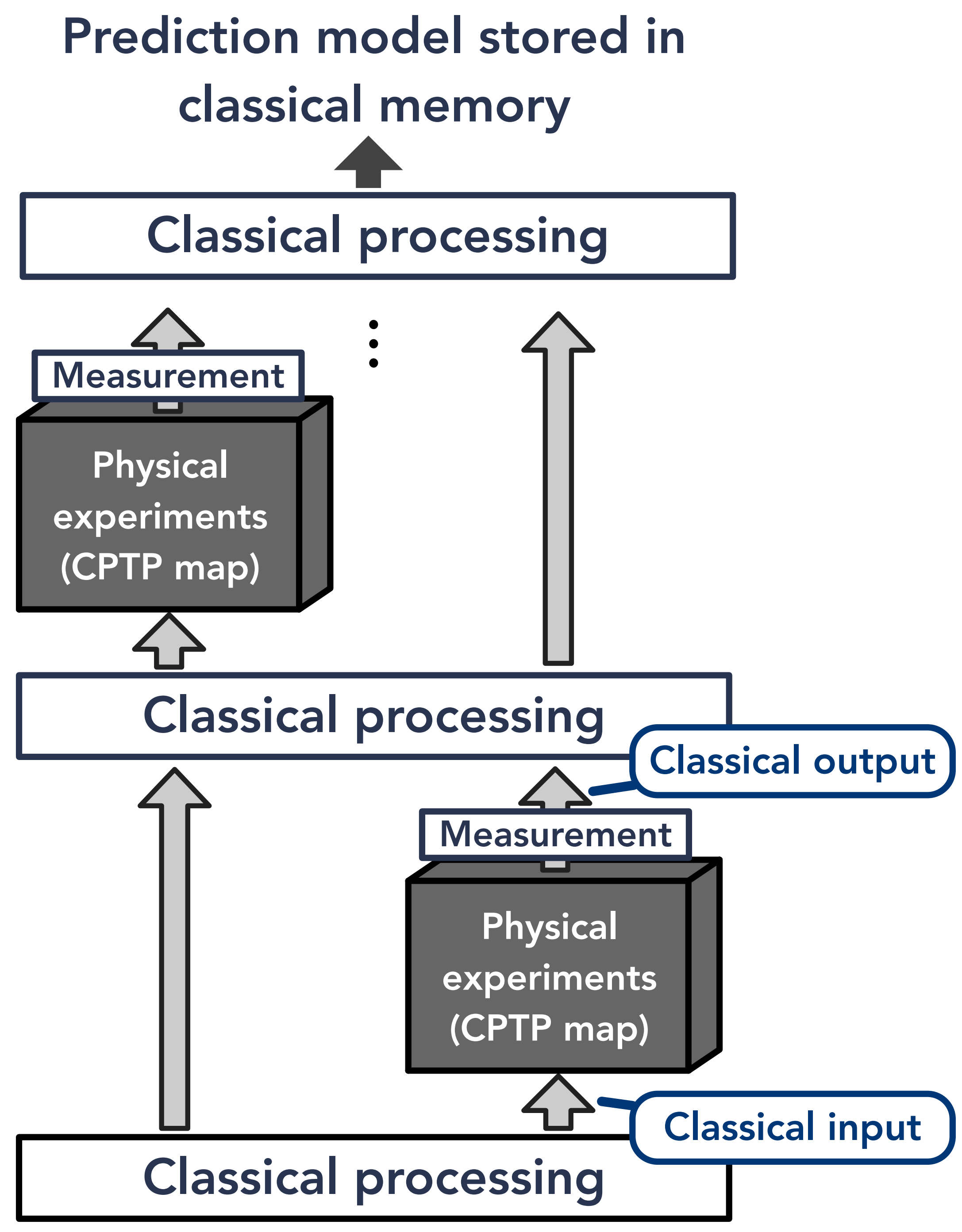
O : the property we want to predict



General Setting

Classical machine learning

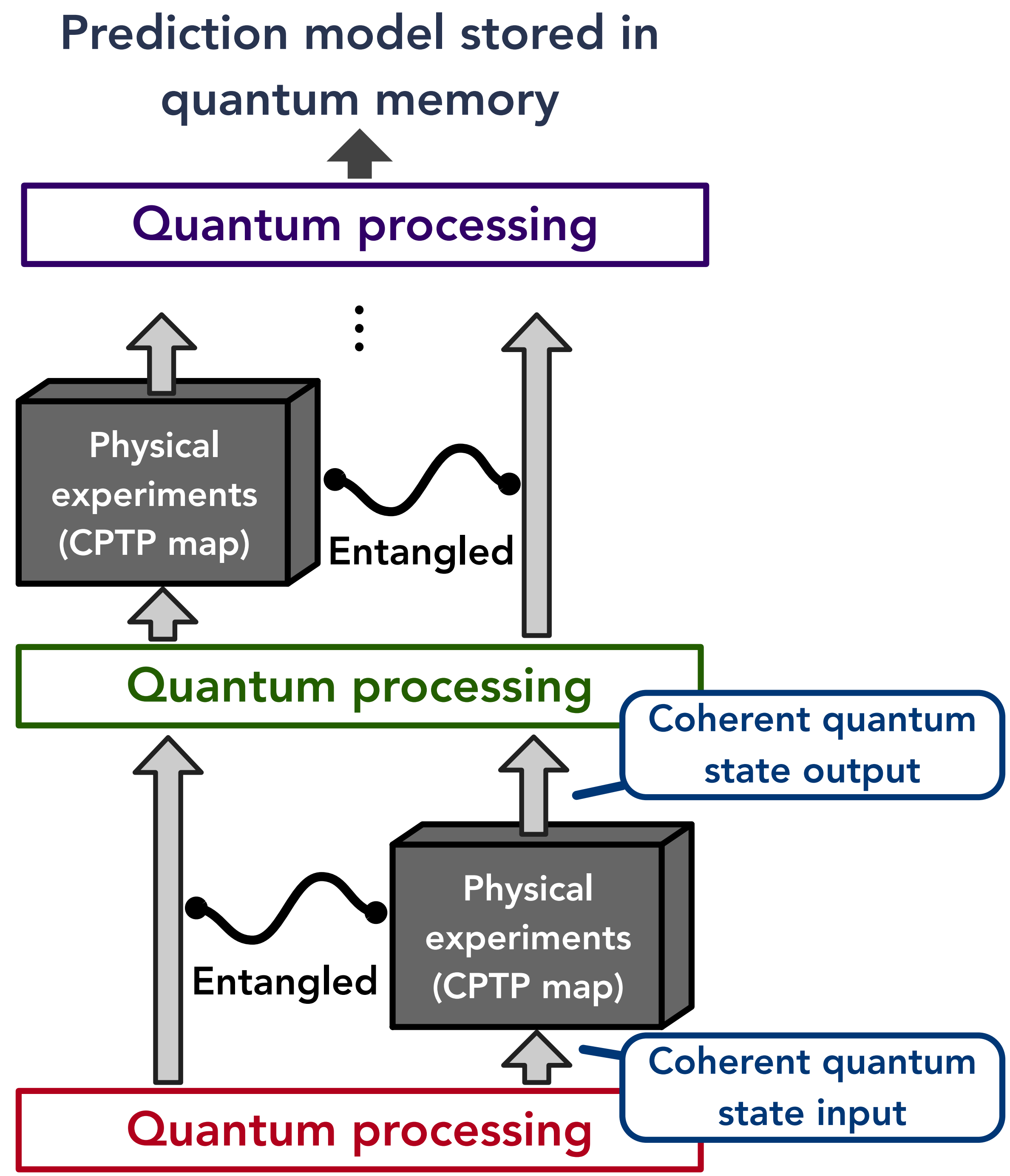
- Learning agents can actively perform experiments to learn a prediction model.
- Each query begins with a choice of classical input x and ends with an arbitrary POVM measurement.
- A prediction model $h(x) \approx f_{\mathscr{E}}(x)$ is created after learning.



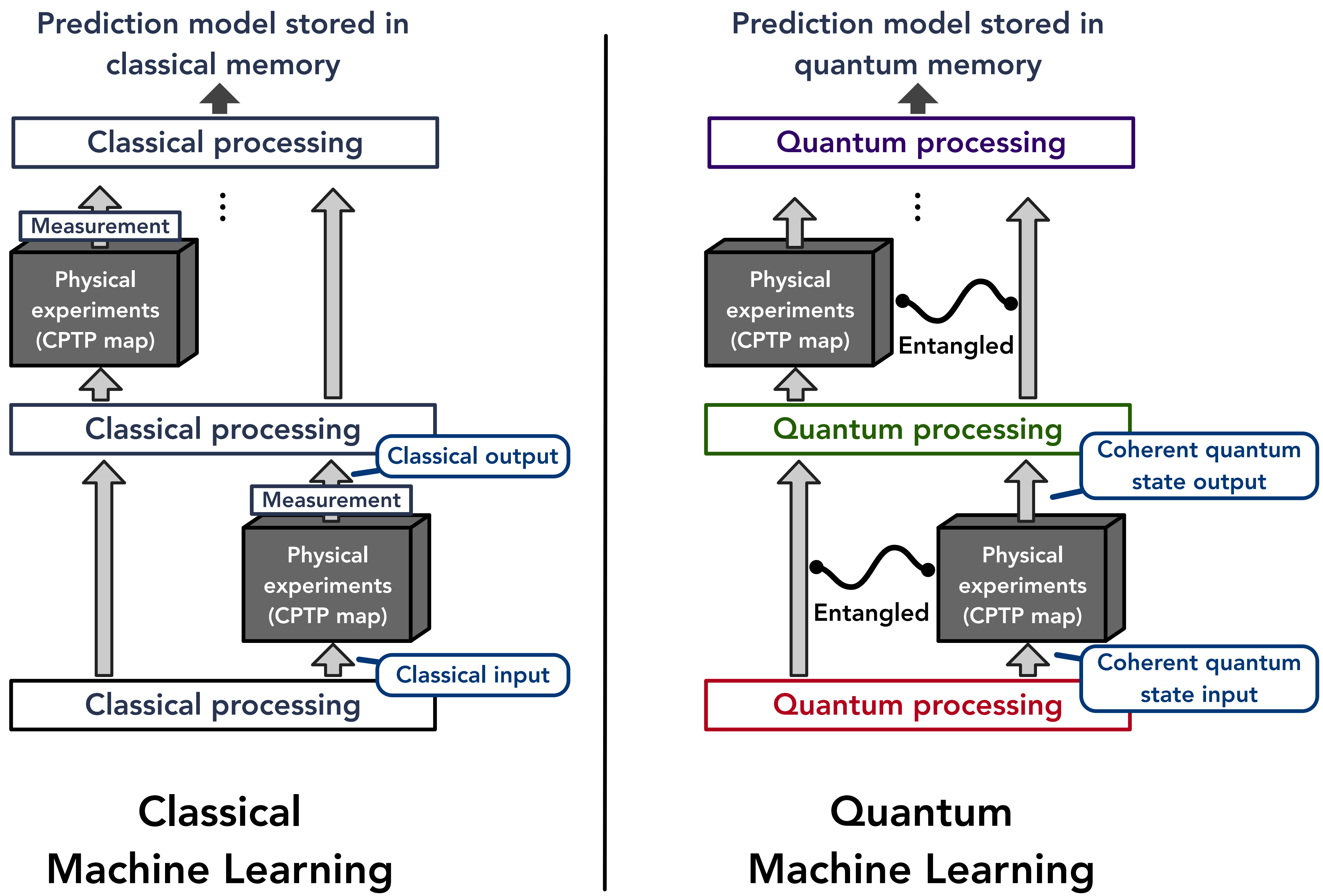
General Setting

Quantum machine learning

- Similar to classical ML setting.
- Each query consists of an arbitrary access to the CPTP map $\mathscr{E}$ (the input can be entangled, and no measurement at the end).
- A prediction model $h(x) \approx f_{\mathscr{E}}(x)$ is stored in a quantum memory instead of a classical memory.



General Setting



The setup is closely related to Quantum Algorithmic Measurements by Aharonov, Cotler, Qi

Main Questions

Information-theoretic aspect:

Do classical ML need significantly more experiments (query complexity) than quantum ML to predict $f_{\mathscr{E}}(x) = \mathrm{Tr}(O\mathscr{E}(|x\rangle\langle x|))$?

[1] **Information-theoretic bounds on quantum advantage in machine learning**, *arXiv:2101.02464*.

Computational aspect:

Could classical ML use data to efficiently compute $f_{\mathscr{E}}(x) = \mathrm{Tr}(O\mathscr{E}(|x\rangle\langle x|))$ even if $f_{\mathscr{E}}(x)$ is hard to compute with a classical computer?

[2] **Power of data in quantum machine learning**, *arXiv:2011.01938*.

Application:

Could we train classical ML to predict ground state properties?

[3] **Provable machine learning algorithms for quantum many-body problems**, *In preparation*.

Information-theoretic aspect

Theorem (Huang, Kueng, Preskill; 2021 [1])

Consider any observable O, any family of CPTP maps $\mathscr{F} = \{\mathscr{E}\}$ with n-qubit input and m-qubit output, and any input distribution $\mathscr{D}$.

Suppose a quantum ML uses N_{Q} queries to the unknown CPTP map $\mathscr{E}$ to learn a prediction model $h_{\mathrm{Q}}(x)$ that achieves a prediction error of

$$\mathbb{E}_{x\sim\mathscr{D}} \left| h_{\mathrm{Q}}(x) - f_{\mathscr{E}}(x) \right|^2 \leq \epsilon,$$

then there is a classical ML using $N_{\mathrm{C}} \leq \mathcal{O}(mN_{\mathrm{Q}}/\epsilon)$ to learn a prediction model $h_{\mathrm{C}}(x)$ that achieves a prediction error of

$$\mathbb{E}_{x\sim\mathscr{D}} \left| h_{\mathrm{C}}(x) - f_{\mathscr{E}}(x) \right|^2 \leq \mathcal{O}(\epsilon)\,.$$

[1] **Information-theoretic bounds on quantum advantage in machine learning**, *arXiv:2101.02464.*

Information-theoretic aspect

Theorem (Huang, Kueng, Preskill; 2021 [1])

Concept/hypothesis class in statistical learning theory

Consider any observable O, any family of CPTP maps $\mathscr{F} = \{\mathscr{E}\}$ with n-qubit input and m-qubit output, and any input distribution $\mathscr{D}$.

Suppose a quantum ML uses N_{Q} queries to the unknown CPTP map $\mathscr{E}$ to learn a prediction model $h_{\mathrm{Q}}(x)$ that achieves a prediction error of

$$\mathbb{E}_{x\sim\mathscr{D}} \left| h_{\mathrm{Q}}(x) - f_{\mathscr{E}}(x) \right|^2 \leq \epsilon,$$

then there is a classical ML using $N_{\mathrm{C}} \leq \mathcal{O}(mN_{\mathrm{Q}}/\epsilon)$ to learn a prediction model $h_{\mathrm{C}}(x)$ that achieves a prediction error of

$$\mathbb{E}_{x\sim\mathscr{D}} \left| h_{\mathrm{C}}(x) - f_{\mathscr{E}}(x) \right|^2 \leq \mathcal{O}(\epsilon)\,.$$

[1] **Information-theoretic bounds on quantum advantage in machine learning**, *arXiv:2101.02464.*

Information-theoretic aspect

Theorem (Huang, Kueng, Preskill; 2021 [1])

Consider any observable O, any family of CPTP maps $\mathscr{F} = \{\mathscr{E}\}$ with n-qubit input and m-qubit output, and any input distribution $\mathscr{D}$.

Suppose a quantum ML uses N_{Q} queries to the unknown CPTP map $\mathscr{E}$ to learn a prediction model $h_{\mathrm{Q}}(x)$ that achieves a predict

Average prediction error

$$\mathbb{E}_{x\sim\mathscr{D}} \left| h_{\mathrm{Q}}(x) - f_{\mathscr{E}}(x) \right|^2 \leq \epsilon,$$

then there is a classical ML using $N_{\mathrm{C}} \leq \mathcal{O}(mN_{\mathrm{Q}}/\epsilon)$ to learn a prediction model $h_{\mathrm{C}}(x)$ that achieves a prediction error of

$$\mathbb{E}_{x\sim\mathscr{D}} \left| h_{\mathrm{C}}(x) - f_{\mathscr{E}}(x) \right|^2 \leq \mathcal{O}(\epsilon)\,.$$

[1] **Information-theoretic bounds on quantum advantage in machine learning**, *arXiv:2101.02464.*

Information-theoretic aspect

Good news for classical ML,
bad news for quantum ML.

- No large separation in query complexity for small average prediction error even on quantum problems.
- Other measures of prediction error (e.g., worst-case) admits provable exponential advantage; see [1] for an example based on shadow tomography.

[1] **Information-theoretic bounds on quantum advantage in machine learning**, *arXiv:2101.02464*.

Main Questions

Information-theoretic aspect:

Do classical ML need significantly more experiments (query complexity) than quantum ML to predict $f_{\mathscr{E}}(x) = \mathrm{Tr}(O\mathscr{E}(|x\rangle\langle x|))$?

[1] Information-theoretic bounds on quantum advantage in machine learning, *arXiv:2101.02464*.

Computational aspect:

Could classical ML use data to efficiently compute $f_{\mathscr{E}}(x) = \mathrm{Tr}(O\mathscr{E}(|x\rangle\langle x|))$ even if $f_{\mathscr{E}}(x)$ is hard to compute with a classical computer?

[2] Power of data in quantum machine learning, *arXiv:2011.01938*.

Application:

Could we train classical ML to predict ground state properties?

[3] Provable machine learning algorithms for quantum many-body problems, *In preparation*.

Main Questions

Information-theoretic aspect:

Do classical ML need significantly more experiments (query complexity) than quantum ML to predict $f_{\mathscr{E}}(x) = \mathrm{Tr}(O\mathscr{E}(|x\rangle\langle x|))$? **No!**

[1] Information-theoretic bounds on quantum advantage in machine learning, *arXiv:2101.02464.*

Computational aspect:

Could classical ML use data to efficiently compute $f_{\mathscr{E}}(x) = \mathrm{Tr}(O\mathscr{E}(|x\rangle\langle x|))$ even if $f_{\mathscr{E}}(x)$ is hard to compute with a classical computer?

[2] Power of data in quantum machine learning, *arXiv:2011.01938.*

Application:

Could we train classical ML to predict ground state properties?

[3] Provable machine learning algorithms for quantum many-body problems, *In preparation.*

Main Questions

Information-theoretic aspect:

Do classical ML need significantly more experiments (query complexity) than quantum ML to predict $f_{\mathcal{E}}(x) = \mathrm{Tr}(O\mathcal{E}(|x\rangle\langle x|))$? **No!**

[1] Information-theoretic bounds on quantum advantage in machine learning, *arXiv:2101.02464*.

Computational aspect:

Could classical ML use data to efficiently compute $f_{\mathcal{E}}(x) = \mathrm{Tr}(O\mathcal{E}(|x\rangle\langle x|))$ even if $f_{\mathcal{E}}(x)$ is hard to compute with a classical computer?

[2] Power of data in quantum machine learning, *arXiv:2011.01938*.

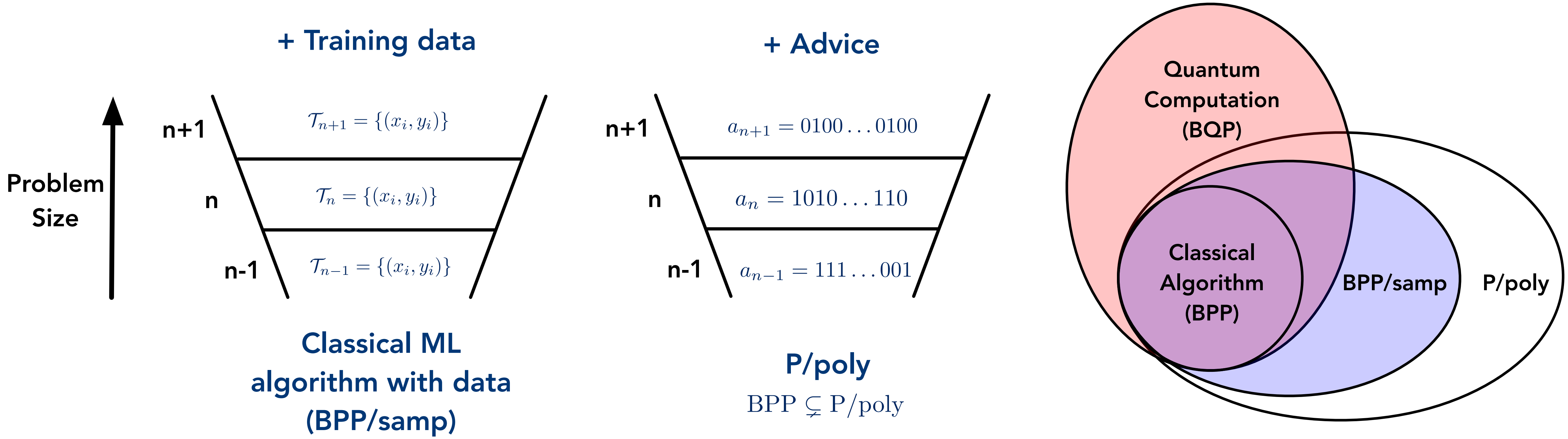
Application:

Could we train classical ML to predict ground state properties?

[3] Provable machine learning algorithms for quantum many-body problems, *In preparation*.

Computational aspect

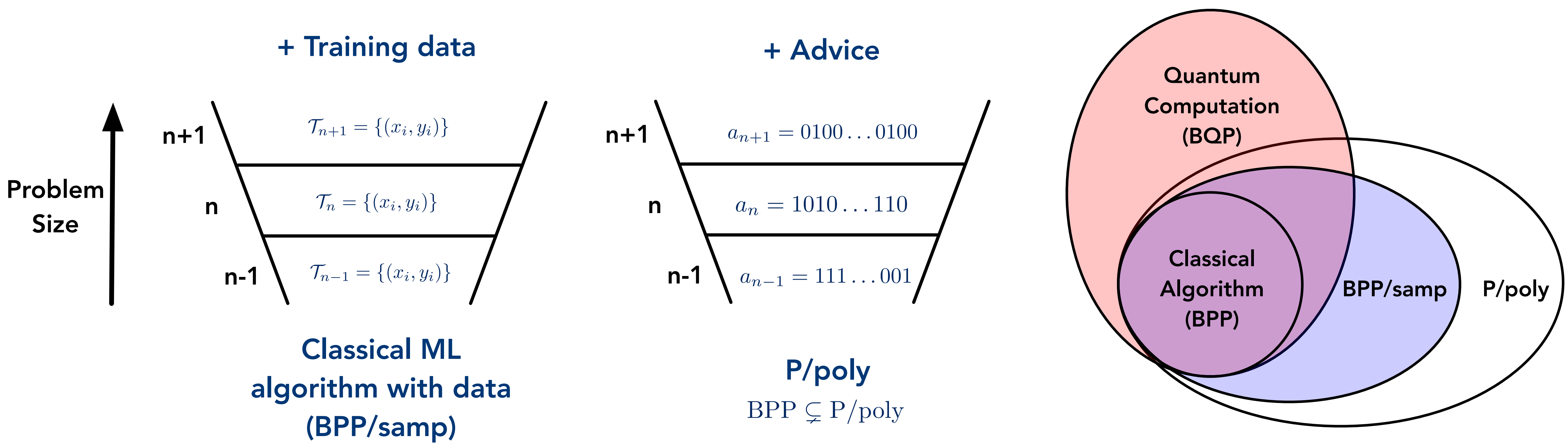
- The formal difference between classical ML and non-ML algorithm is that ML algorithm can learn from data.
- We define a complexity class for classical algorithm that could learn from sampled data (BPP/samp).
- BPP/samp is a restricted class of P/poly.



[2] **Power of data in quantum machine learning**, *arXiv:2011.01938.*

Computational aspect

- Classical algorithms learning from data could solve problems that can not be solved by non-ML algorithms.
- This is only true when data can not be computed in BPP. (such as data from quantum experiments)



[2] **Power of data in quantum machine learning**, *arXiv:2011.01938.*

Computational aspect

- A sufficient condition for solving quantum problems with computationally efficient classical ML algorithms.
- Consider a training data of $\{(x_i, y_i = f_{\mathscr{E}}(x_i))\}_{i=1}^{N_{\mathrm{C}}}$, then there is an efficient classical ML algorithm producing $h_{\mathrm{C}}(x)$ with

$$\mathbb{E}_{x \sim \mathscr{D}} \left| h_{\mathrm{C}}(x) - f_{\mathscr{E}}(x) \right|^2 \leq \mathcal{O}\left(\sqrt{\frac{s}{N_{\mathrm{C}}}} \right).$$

- The classical ML algorithm is based on a kernel matrix K and

$$s = \sum_{i,j} (K^{-1})_{ij} f_{\mathscr{E}}(x_i) f_{\mathscr{E}}(x_j).$$

[2] Power of data in quantum machine learning, *arXiv:2011.01938*.

Main Questions

Information-theoretic aspect:

Do classical ML need significantly more experiments (query complexity) than quantum ML to predict $f_{\mathscr{E}}(x) = \mathrm{Tr}(O\mathscr{E}(|x\rangle\langle x|))$? **No!**

[1] **Information-theoretic bounds on quantum advantage in machine learning**, *arXiv:2101.02464*.

Computational aspect:

Could classical ML use data to efficiently compute $f_{\mathscr{E}}(x) = \mathrm{Tr}(O\mathscr{E}(|x\rangle\langle x|))$ even if $f_{\mathscr{E}}(x)$ is hard to compute with a classical computer?

[2] **Power of data in quantum machine learning**, *arXiv:2011.01938*.

Application:

Could we train classical ML to predict ground state properties?

[3] **Provable machine learning algorithms for quantum many-body problems**, *In preparation*.

Main Questions

Information-theoretic aspect:

Do classical ML need significantly more experiments (query complexity) than quantum ML to predict $f_{\mathscr{E}}(x) = \mathrm{Tr}(O\mathscr{E}(|x\rangle\langle x|))$? **No!**

[1] **Information-theoretic bounds on quantum advantage in machine learning**, *arXiv:2101.02464*.

Computational aspect:

Could classical ML use data to efficiently compute $f_{\mathscr{E}}(x) = \mathrm{Tr}(O\mathscr{E}(|x\rangle\langle x|))$ even if $f_{\mathscr{E}}(x)$ is hard to compute with a classical computer? **Yes!**

[2] **Power of data in quantum machine learning**, *arXiv:2011.01938*.

Application:

Could we train classical ML to predict ground state properties?

[3] **Provable machine learning algorithms for quantum many-body problems**, *In preparation*.

Main Questions

Information-theoretic aspect:

Do classical ML need significantly more experiments (query complexity) than quantum ML to predict $f_{\mathscr{E}}(x) = \mathrm{Tr}(O\mathscr{E}(|x\rangle\langle x|))$? **No!**

[1] Information-theoretic bounds on quantum advantage in machine learning, *arXiv:2101.02464*.

Computational aspect:

Could classical ML use data to efficiently compute $f_{\mathscr{E}}(x) = \mathrm{Tr}(O\mathscr{E}(|x\rangle\langle x|))$ even if $f_{\mathscr{E}}(x)$ is hard to compute with a classical computer? **Yes!**

[2] Power of data in quantum machine learning, *arXiv:2011.01938*.

Application:

Could we train classical ML to predict ground state representation?

[3] Provable machine learning algorithms for quantum many-body problems, *In preparation*.

Application

- Consider in the experimental lab, we can synthesize physical systems in the ground state $\rho(x)$ of some local Hamiltonian $H(x)$.

- Here, $x \in [-1,1]^n$ is the known controllable parameters, while $x \mapsto H(x)$ is not perfectly known.

- Could we design a classical ML algorithm that can learn from physical experiments to predict efficient representation of $\rho(x)$ for new input x?

[3] **Provable machine learning algorithms for quantum many-body problems**, *In preparation.*

Application

We give a classical ML algorithm with rigorous guarantees.

Learning phase: for each of N experiments,

1. Samples a random controllable parameter $x_i \sim \mathrm{Unif}[-1,1]^n$.
2. Performs a randomized measurement on the ground state $\rho(x_i)$ to construct a single-shot classical shadow $S(\rho(x_i))$ [4].

Prediction phase: predicts the ground state representation for new x to be

$$\sigma_\Lambda(x) = \sum_{i=1}^{N} \kappa_\Lambda(x, x_i) S(\rho(x_i)),$$

where $\kappa_\Lambda(x,y) = \sum_{k\in\mathbb{Z}^n} \mathbf{1}\left\{\|k\|_2 \leq \Lambda\right\} \exp\left(\mathrm{i}\pi\langle k, x-y\rangle\right)$ (l_2-Dirichlet kernel).

[3] **Provable machine learning algorithms for quantum many-body problems**, *In preparation*.

[4] **Predicting many properties of a quantum system from very few measurements,** *Nat. Phys.*

Application

Theorem (Huang, Kueng, Preskill; 2021 [3])

For any smooth class of local Hamiltonians in a finite spatial dimension with a constant spectral gap, given the number of experiments $N = \mathrm{poly}(n)$,

$$\mathbb{E}_{x\sim[-1,1]^n} \left| \mathrm{Tr}(O\sigma_\Lambda(x)) - \mathrm{Tr}(O\rho(x)) \right|^2 \leq \epsilon,$$

for any sum of local observables $O = \sum_{j=1}^{L} O_j$ with $\sum_{j=1}^{L} \|O_j\| = \mathcal{O}(1)$ and ϵ: const.

- NP-complete for computing 1-body local observables to constant error in 2D local Hamiltonians with a constant spectral gap [5].
- Proof relies on classical shadow formalism, quasi-adiabatic evolution, and generalization bounds of kernel methods.

[3] **Provable machine learning algorithms for quantum many-body problems**, *In preparation*.

[5] **Sub-exponential algorithm for 2D frustration-free spin systems with gapped subsystems**.

Application

Theorem (Huang, Kueng, Preskill; 2021 [3])

For any smooth class of local Hamiltonians in a finite spatial dimension with a constant spectral gap, given the number of experiments $N = \mathrm{poly}(n)$,

$$\mathbb{E}_{x\sim[-1,1]^n} |\operatorname{Tr}(O\sigma_\Lambda(x)) - \operatorname{Tr}(O\rho(x))|^2 \leq \epsilon,$$

for any sum of local observables $O = \sum_{j=1}^{L} O_j$ with $\sum_{j=1}^{L} \|O_j\| = \mathcal{O}(1)$ and ϵ: const.

- A matching lower bound could be proved for any classical ML algorithm.
- Using information-theoretic bounds [1], a near-matching lower bound holds for any quantum ML algorithm (no large quantum advantage in this generality).

[1] **Information-theoretic bounds on quantum advantage in machine learning**, *arXiv:2101.02464.*

[3] **Provable machine learning algorithms for quantum many-body problems**, *In preparation.*

Main Questions

Information-theoretic aspect:

Do classical ML need significantly more experiments (query complexity) than quantum ML to predict $f_{\mathscr{E}}(x) = \mathrm{Tr}(O\mathscr{E}(|x\rangle\langle x|))$? **No!**

[1] **Information-theoretic bounds on quantum advantage in machine learning**, *arXiv:2101.02464*.

Computational aspect:

Could classical ML use data to efficiently compute $f_{\mathscr{E}}(x) = \mathrm{Tr}(O\mathscr{E}(|x\rangle\langle x|))$ even if $f_{\mathscr{E}}(x)$ is hard to compute with a classical computer? **Yes!**

[2] **Power of data in quantum machine learning**, *arXiv:2011.01938*.

Application:

Could we train classical ML to predict ground state representation?

[3] **Provable machine learning algorithms for quantum many-body problems**, *In preparation*.

Main Questions

Information-theoretic aspect:

Do classical ML need significantly more experiments (query complexity) than quantum ML to predict $f_{\mathscr{E}}(x) = \mathrm{Tr}(O\mathscr{E}(|x\rangle\langle x|))$? **No!**

[1] **Information-theoretic bounds on quantum advantage in machine learning**, *arXiv:2101.02464*.

Computational aspect:

Could classical ML use data to efficiently compute $f_{\mathscr{E}}(x) = \mathrm{Tr}(O\mathscr{E}(|x\rangle\langle x|))$ even if $f_{\mathscr{E}}(x)$ is hard to compute with a classical computer? **Yes!**

[2] **Power of data in quantum machine learning**, *arXiv:2011.01938*.

Application:

Could we train classical ML to predict ground state representation? **Yes!**

[3] **Provable machine learning algorithms for quantum many-body problems**, *In preparation*.

Numerics verify the theoretical predictions

Conclusion

Exponential separation for worst-case error.

Information-theoretic aspect:

Do classical ML need significantly more experiments (query complexity) than quantum ML to predict $f_{\mathscr{E}}(x) = \mathrm{Tr}(O\mathscr{E}(|x\rangle\langle x|))$? **No!**

[1] Information-theoretic bounds on quantum advantage in machine learning, *arXiv:2101.02464*.

Quantum ML is still computationally more powerful.

Computational aspect:

Could classical ML use data to efficiently compute $f_{\mathscr{E}}(x) = \mathrm{Tr}(O\mathscr{E}(|x\rangle\langle x|))$ even if $f_{\mathscr{E}}(x)$ is hard to compute with a classical computer? **Yes!**

[2] Power of data in quantum machine learning, *arXiv:2011.01938*.

We should expect more applications with quantum ML!

Application:

Could we train classical ML to predict ground state representation? **Yes!**

[3] Provable machine learning algorithms for quantum many-body problems, *In preparation*.