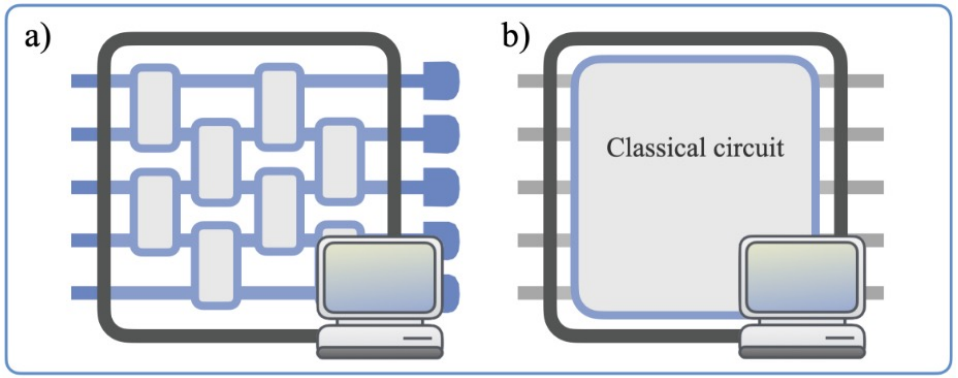
Making quantum machine learning separations more physical/practical

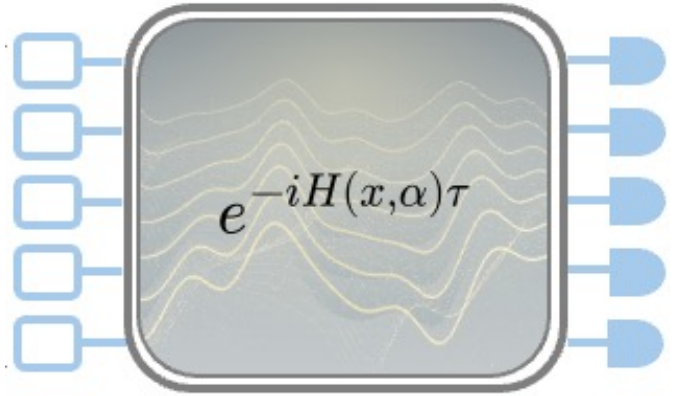
Niklas Pirnay, Sofiene Jerbi, Jean-Pierre Seifert, Jens Eisert

npj Quantum Information (2026)



Rahul Bandyopadhyay, Riccardo Molteni, Jens Eisert, Vedran Dunjko, Sofiene Jerbi

arXiv: 2607.06472



Universiteit Leiden

HZB Helmholtz Zentrum Berlin

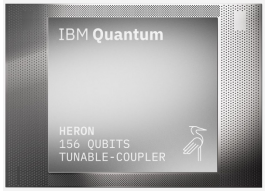
Freie Universität Berlin

TECHNISCHE UNIVERSITÄT BERLIN

AQTIVATE 2026
30/09/2026

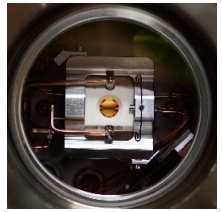
Motivation

Quantum computers are increasing in scale, contorollability, error tolerance:



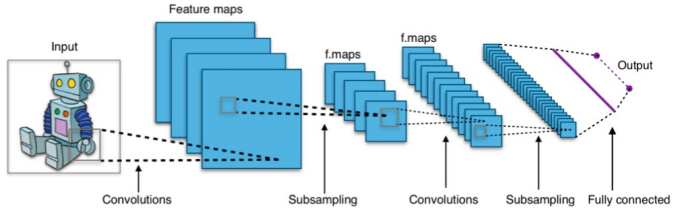
Superconducting qubits
~150 qubits, 5k gates

Neutral atom arrays
~280 physical qubits, ~50 logical qubits

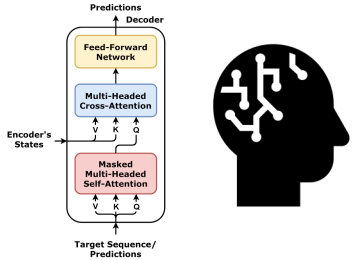


Trapped ions
~50 ions, → 2D error-corrected architecture

Major advances in the development of machine learning models, with increased applicability:



Convolutional neural networks



Transformers and AI assistants (LLMs)



AlphaGo

Deep reinforcement learning

How can quantum computing enhance machine learning / AI?

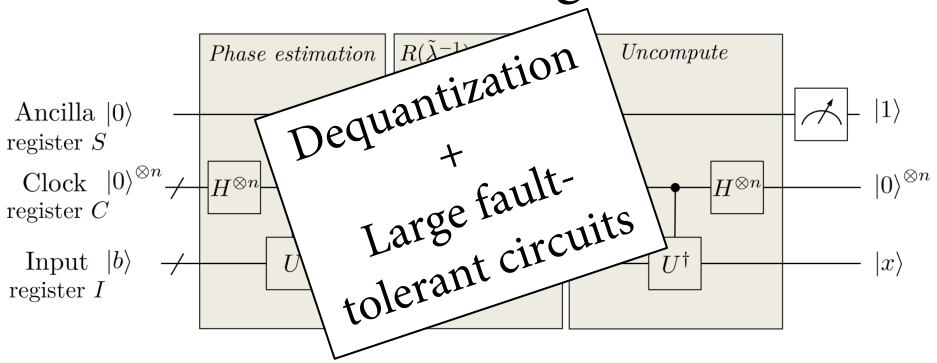
How can (quantum) machine learning help us understand the quantum world?

Paradigm shift in quantum machine learning

Pre-2018:

Grover-based algorithms
&
HHL-based algorithms:

Large fault-tolerant circuits



Published: 14 September 2017

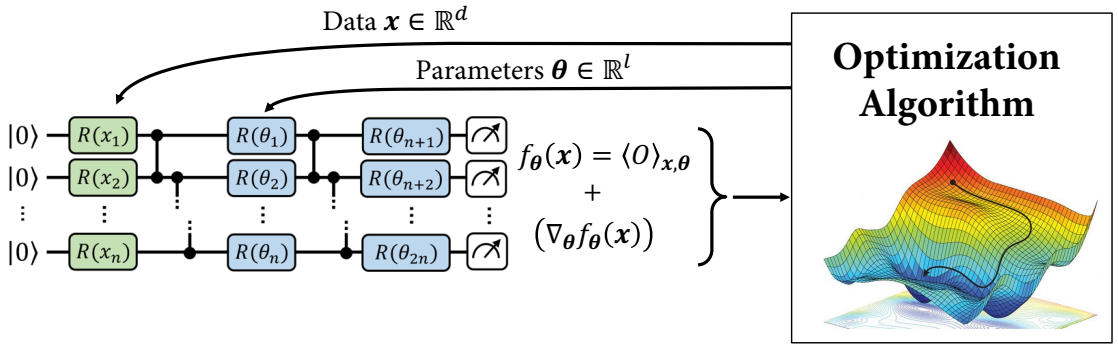
Quantum machine learning

Jacob Biamonte ✉, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe & Seth Lloyd

Nature **549**, 195–202 (2017) | Cite this article

Post-2018:

Parametrized quantum circuits (PQC)

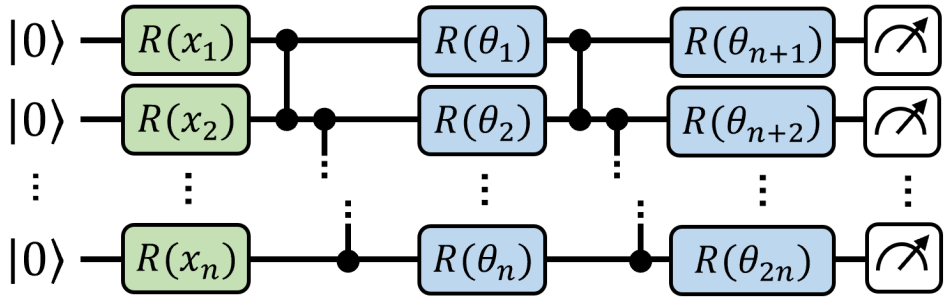


Farhi & Neven, **Classification** with quantum neural networks on near term processors, arXiv:**1802**.06002

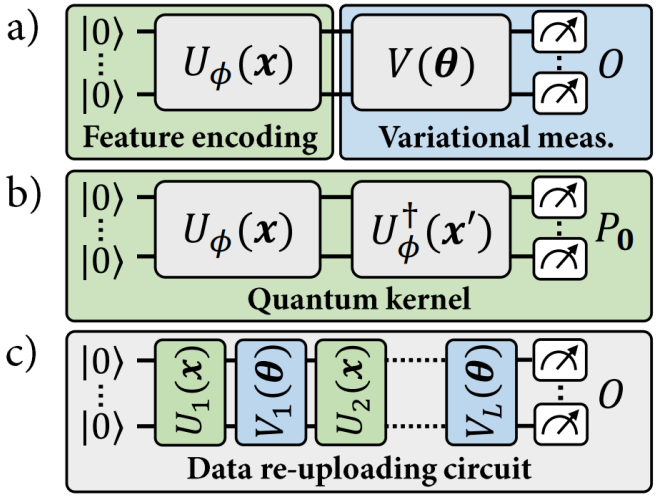
Benedetti *et al.*, A **generative modeling** approach for benchmarking and training shallow quantum circuits, arXiv:**1801**.07686 (NPJ QI)

Mitarai, Negoro, Kitagawa & Fujii, Quantum circuit learning, arXiv:**1803**.00745 (PRA) (**Regression tasks**)

Quantum machine learning models



Various models:



Various names:

Parametrized quantum circuits

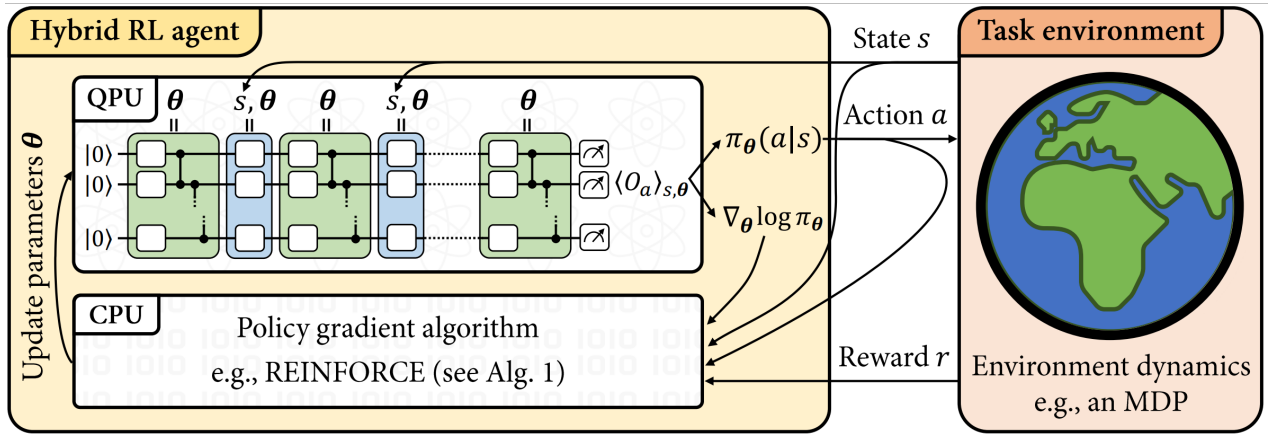
Variational quantum circuits

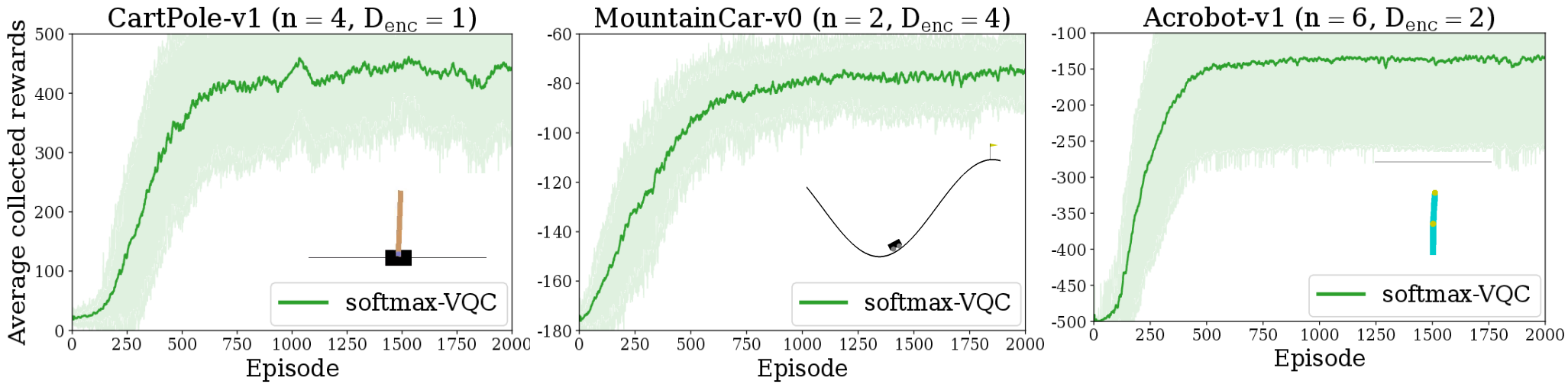
Quantum neural networks

Various applications:

- Classification:
  - ✓ Havlíček *et al.*, *Nature* 2019
  - ✓ Schuld *et al.*, *PRL* 2019
- Regression:
  - ✓ Mitarai *et al.*, *PRA* 2018
- Generative modeling:
  - ✓ Liu *et al.*, *PRA* 2018
- Reinforcement learning:
  - ✓ Chen *et al.*, *IEEE* 2020
  - ✓ Jerbi *et al.*, *NeurIPS* 2021

Promising empirical performance

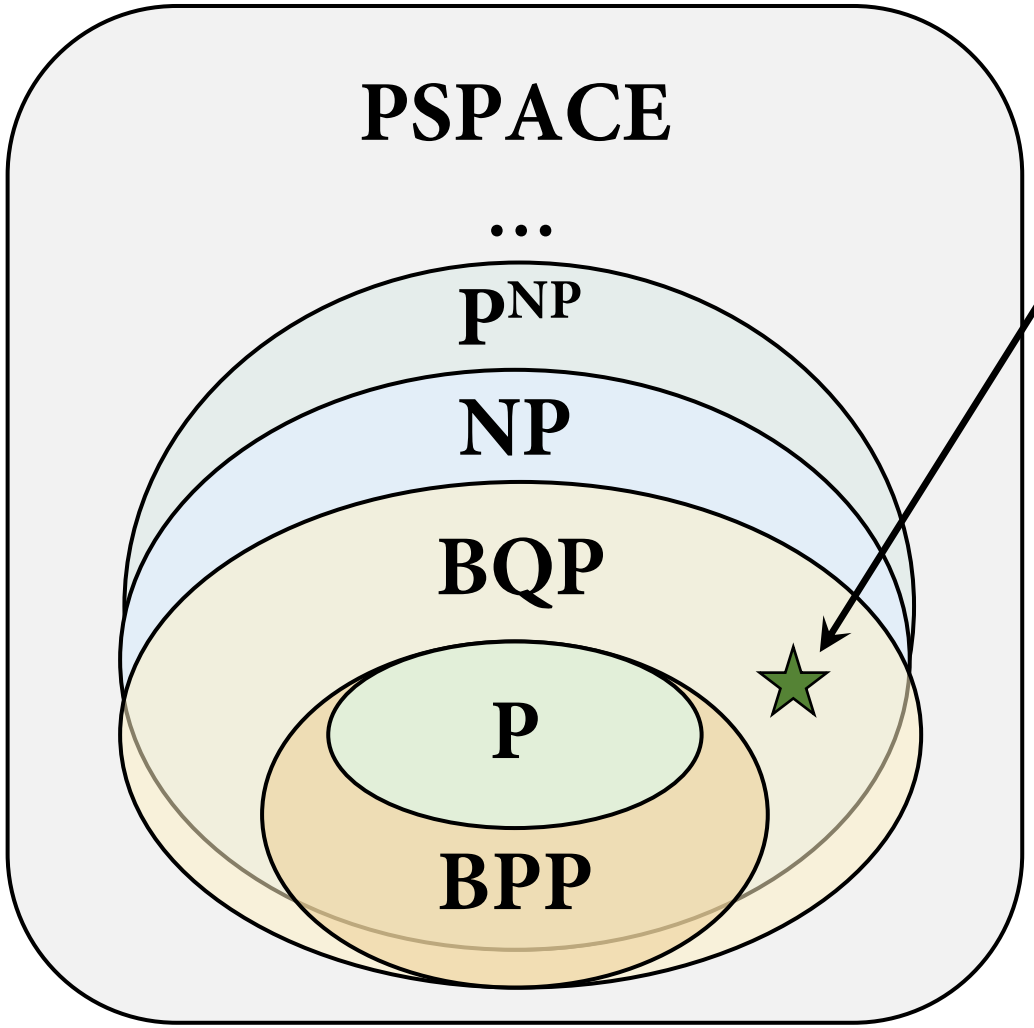




Jerbi, Gyurik, Marshall, Briegel & Dunjko, **Parametrized quantum policies for reinforcement learning**, NeurIPS 2021

What can QC theory tell us about PQCs prospects of quantum advantages?

BQP v.s. BPP

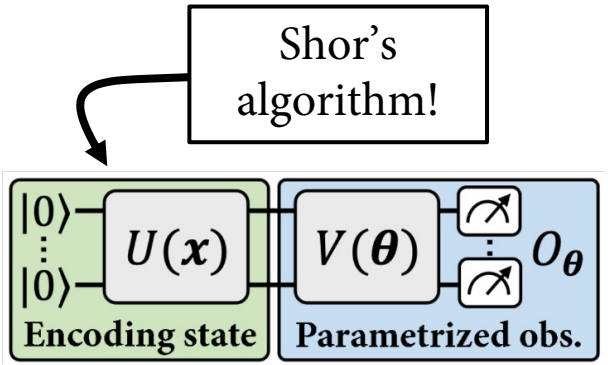


No unconditional separations known

Based on some assumptions, e.g.:

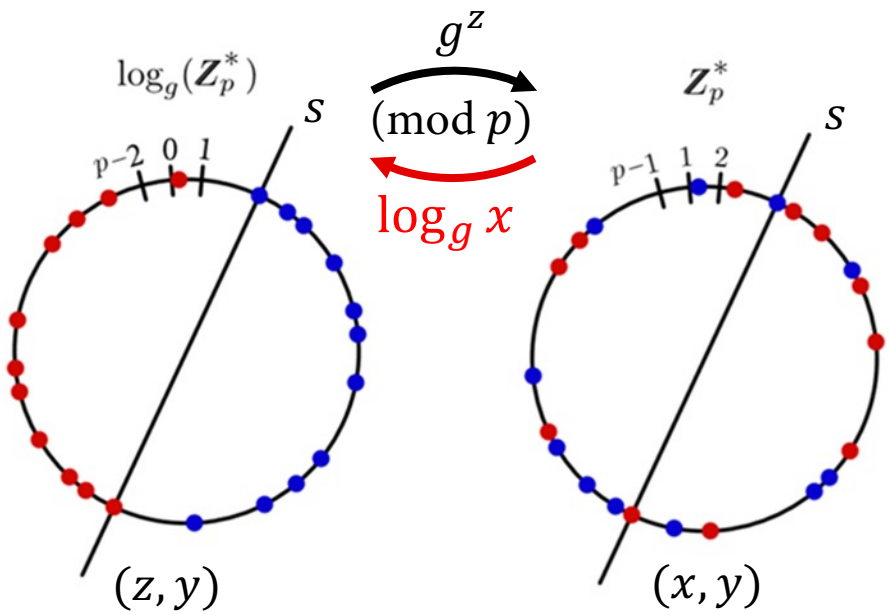
- Factoring is classically hard
- Polynomial hierarchy does not collapse

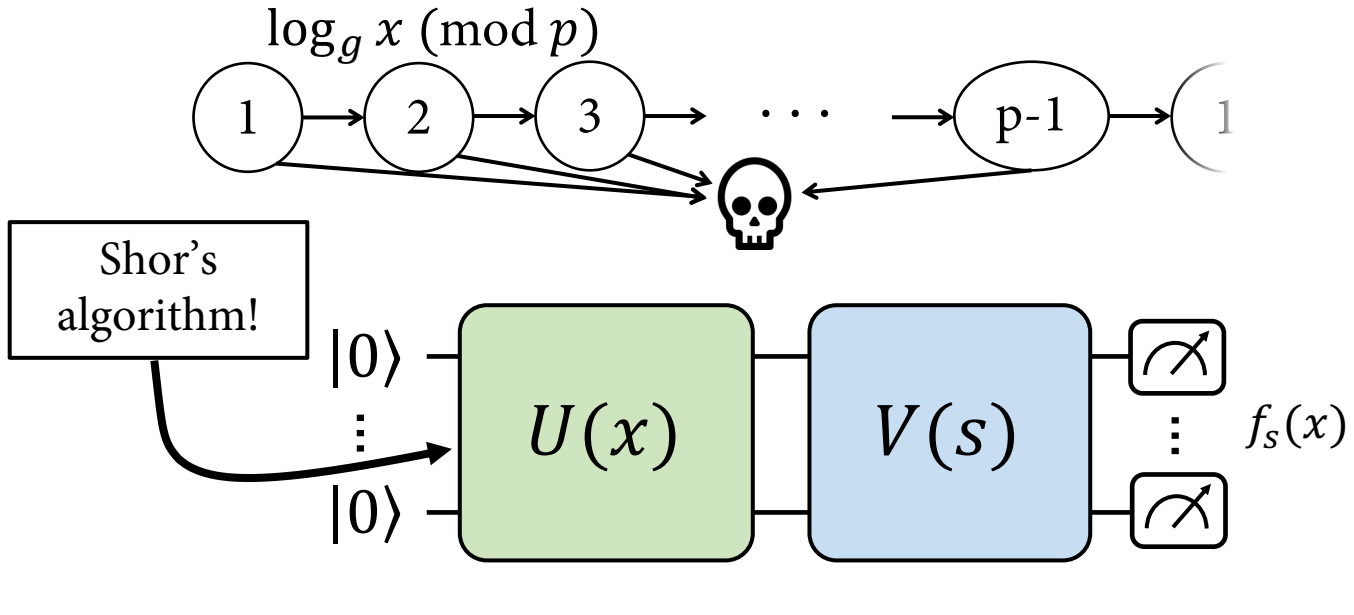
Idea: construct learning problems out of the tasks assumed in BQP but not in BPP



Provable quantum advantage

- Learning tasks with a **provable** quantum advantage over any classical learner:





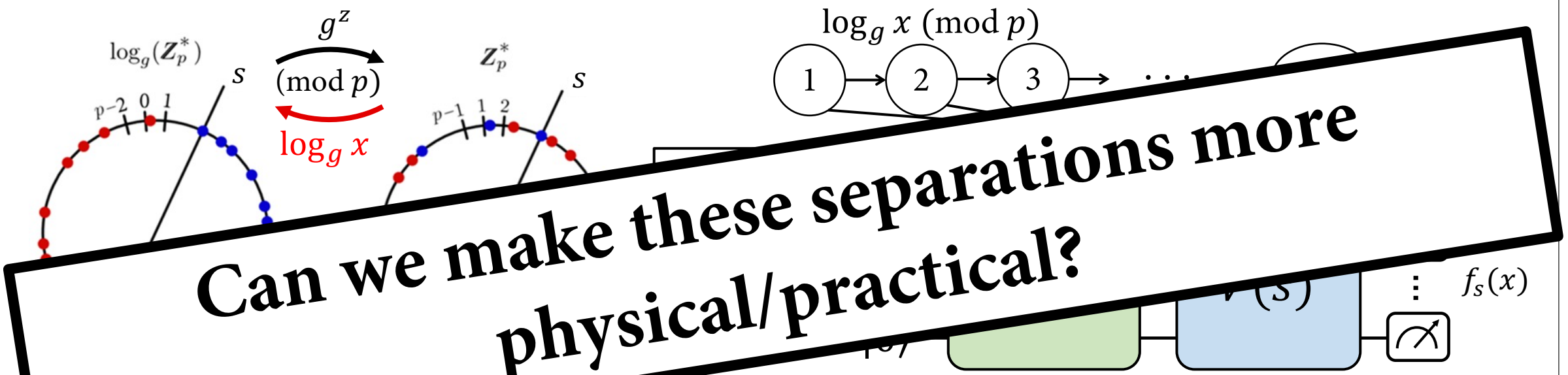
Theorem (informal). There exist learning tasks where:

1. Quantum agents can achieve close-to-optimal performance with high prob.
2. Classical agents cannot achieve a performance (much) better than random, under hardness of $\log_g$

Liu, Arunachalam & Temme, **A rigorous and robust quantum speed-up in supervised machine learning**, Nat. Phys. (2021)

Provable quantum advantage

- Learning tasks with a **provable** quantum advantage over any classical learner:



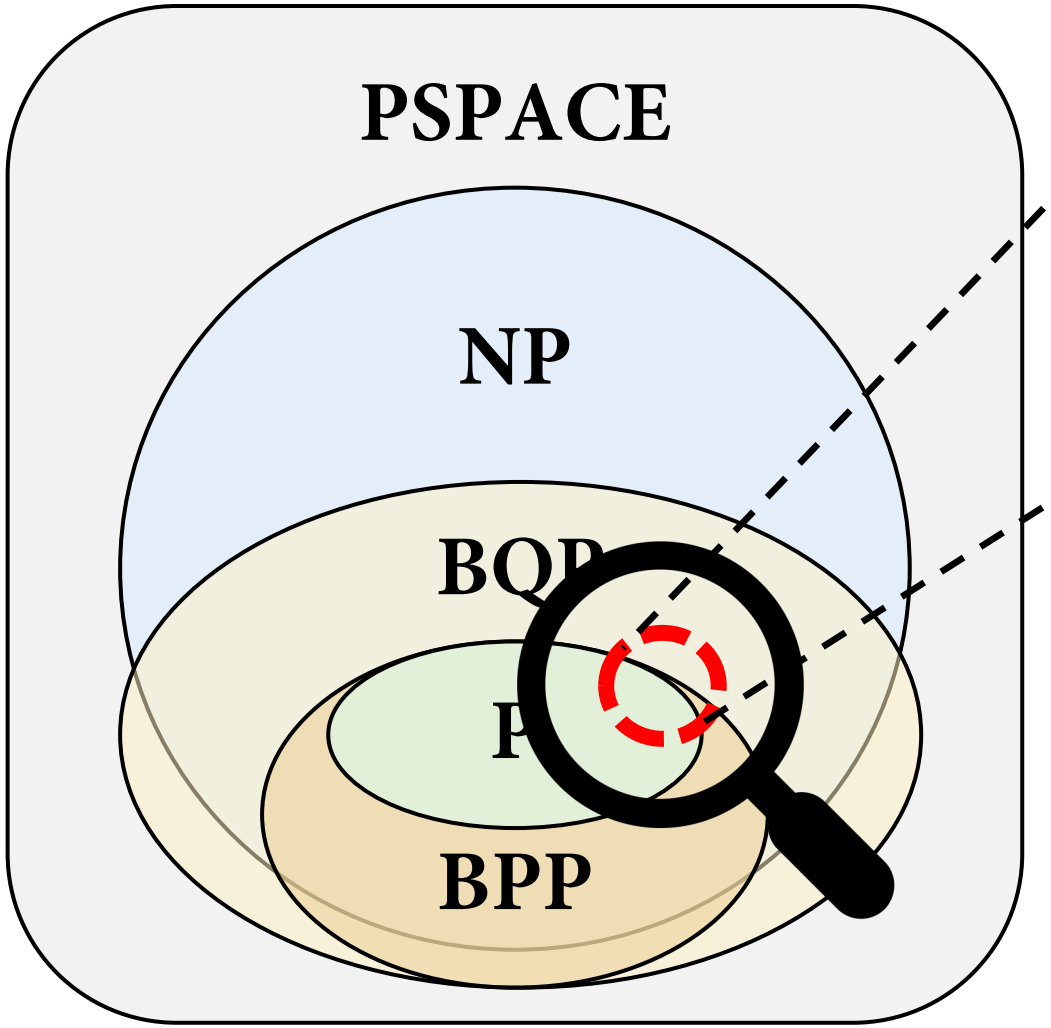
Can we make these separations more physical/practical?

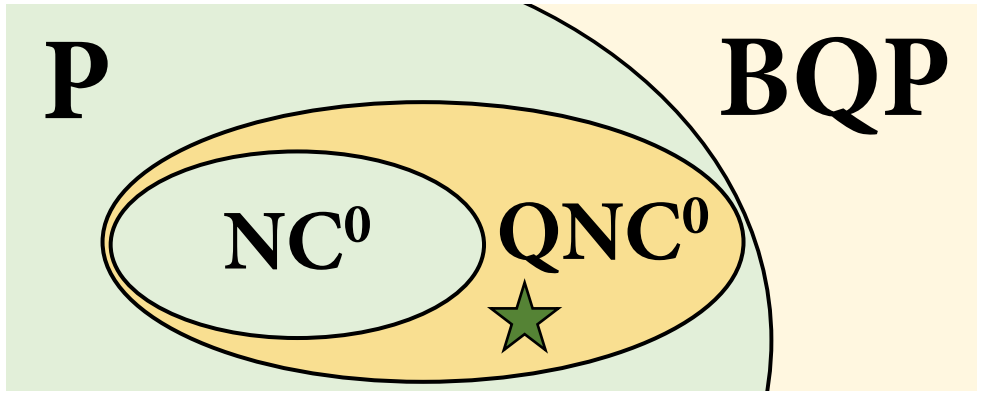
2). There exist learning tasks where:

1. Quantum agents can achieve close-to-optimal performance with high prob.
2. Classical agents cannot achieve a performance (much) better than random, under hardness of $\log_g$

Liu, Arunachalam & Temme, **A rigorous and robust quantum speed-up in supervised machine learning**, Nat. Phys. (2021)

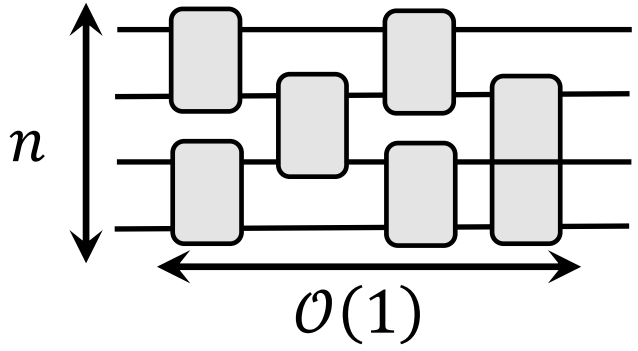
More near-term: shallow-depth separations





NC^0: constant-depth classical circuits with bounded fan-in gates

QNC^0: constant-depth quantum circuits with 1- and 2-qubit gates (no meas. & FF)



Generative learning

- Concept class:

$$\mathcal{C} = \{\mathrm{p}_i\}_i$$

- Training data:

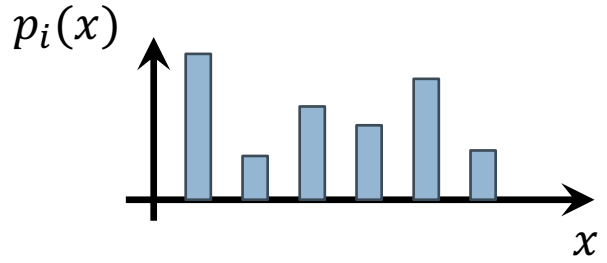
$$\mathcal{D} = \{x_1, \dots, x_M \sim p_i\}$$

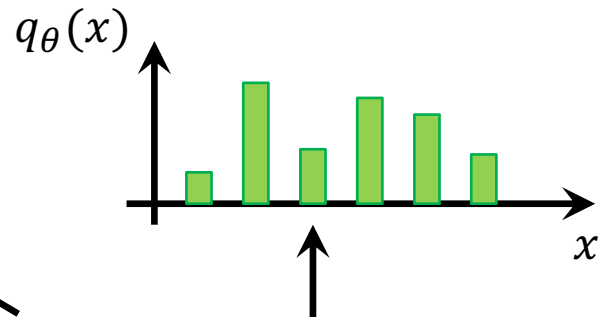
- Hypothesis class:

$$\mathcal{H} = \{q_\theta\}_\theta$$

- Learning performance:

$$\mathrm{TV}(q_\theta, p_i) = \sum_x |q_\theta(x) - p_i(x)|$$



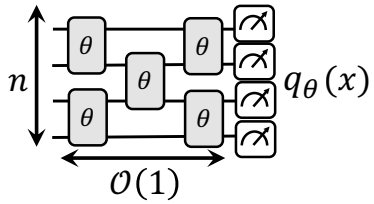


Generated by a restricted computational model
E.g., circuits in NC^0 or QNC^0

Learning separations

Quantum learnability

A class of quantum circuits
$\{U(\theta)\}_\theta \subset \mathrm{QNC}^0$



A learning algorithm that,

for a target TV-distance ϵ,
for a prob. of error δ,

returns a θ such that $\mathrm{TV}(q_\theta, p_i) \leq \epsilon$
with prob. $1 - \delta$

using $\mathrm{poly}(n, \delta^{-1}, \epsilon^{-1})$ time and data

Classical hardness

Hardness to learn

<

Expressibility

Watts & Parham (2023):

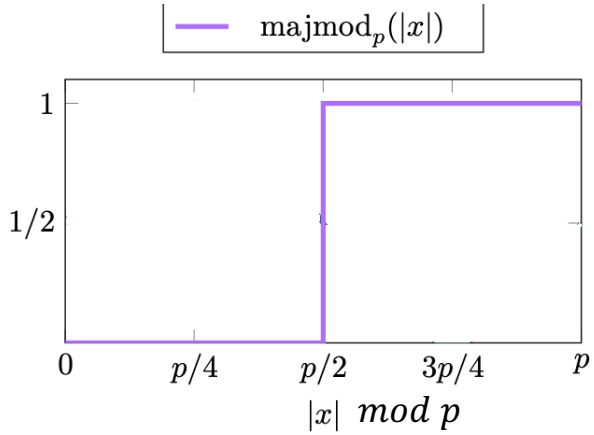
Unconditional Quantum Advantage for Sampling with Shallow Circuits

- A family of distributions:

for all $n \geq 1$

Slightly more complex than: $(x, \quad majmod_p(|x|))$

for $x \sim \mathcal{U}(\{0,1\}^n)$

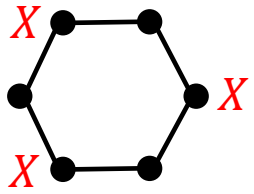


- **Quantum expressibility:**

Mostly relies on non-locality of $|GHZ_n\rangle$ as a ressource

But: $|GHZ_n\rangle$ cannot be prepared in QNC^0

Solution:



Poor-man GHZ

$\frac{1}{\sqrt{2}}(|z\rangle + |\bar{z}\rangle)$

Problem: Only show existence of circuits (no explicit construction)

Classical hardness:

Extension of a previous result:

Viola, *The complexity of distributions*, SIAM Journal on Computing (2012)

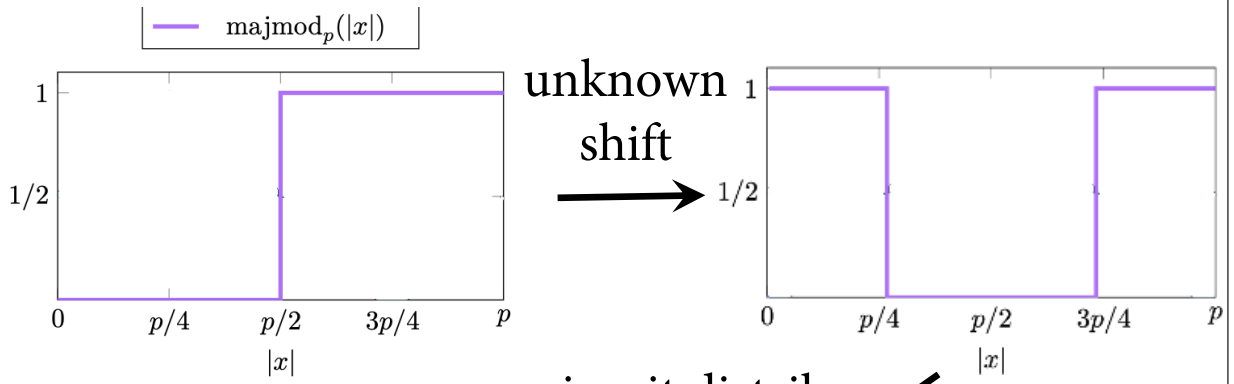
Classical circuits need $\mathcal{O}(\text{loglog}(n))$ depth to express these distributions

Based on "locality" (few bodyness) of correlations as function of depth

Showing a learning advantage

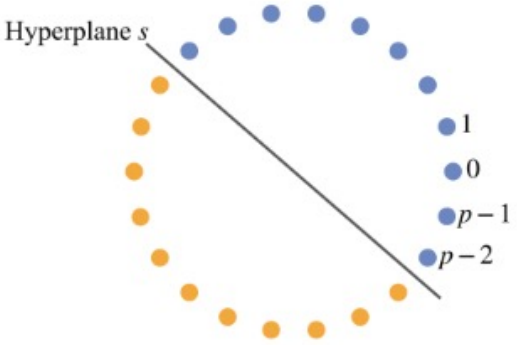
- **Non-trivial extension to a learning problem:**

for all $n \geq 1$
$(x, majmod_p(|x|))$ ⟶ for all $n \geq 1$
$(x, majmod_p(|x| + s))$



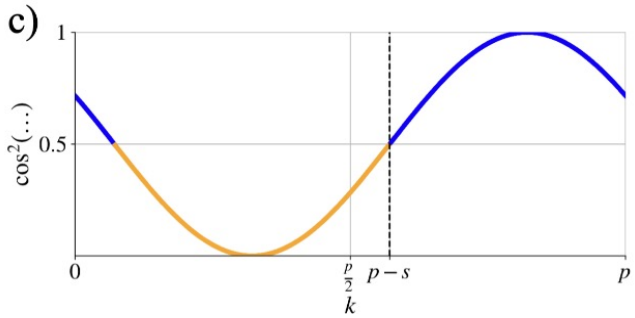
- **Quantum learnability:**

Learning algorithm for finding secret shift s



Points almost uniformly sampled on the ring with probabilistic labels

circuit distrib.

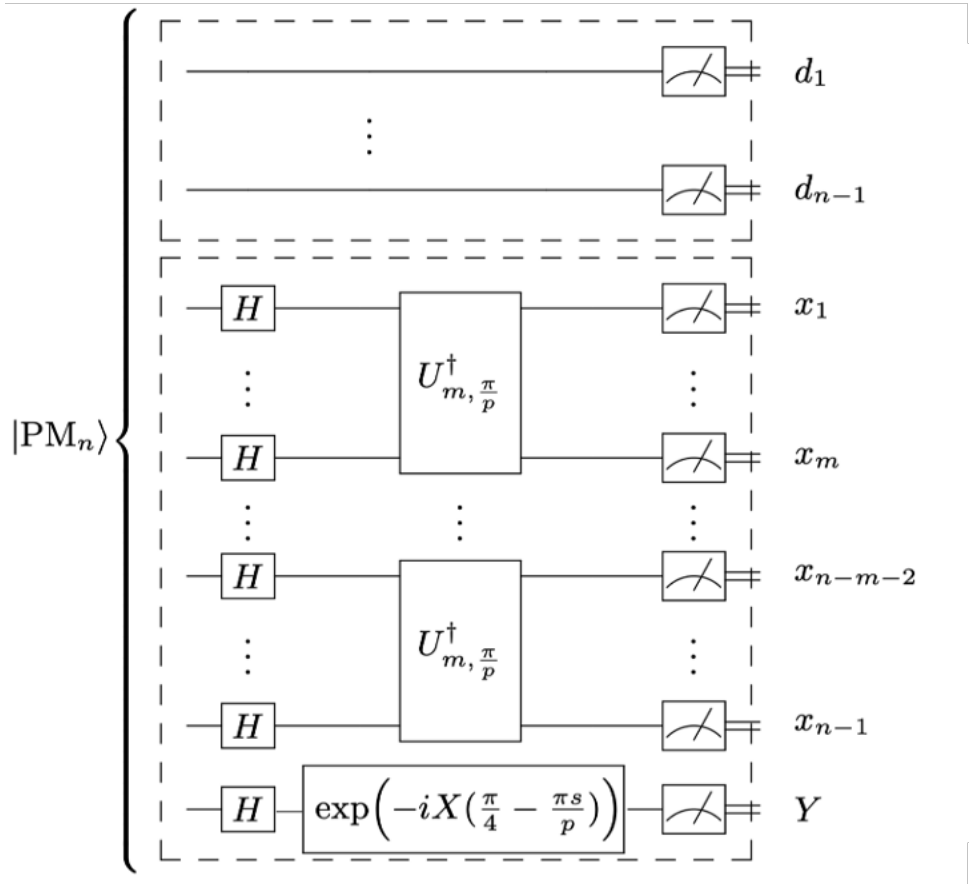


We approx. this curve and identify s exactly with $\tilde{\mathcal{O}}(p^4)$ samples

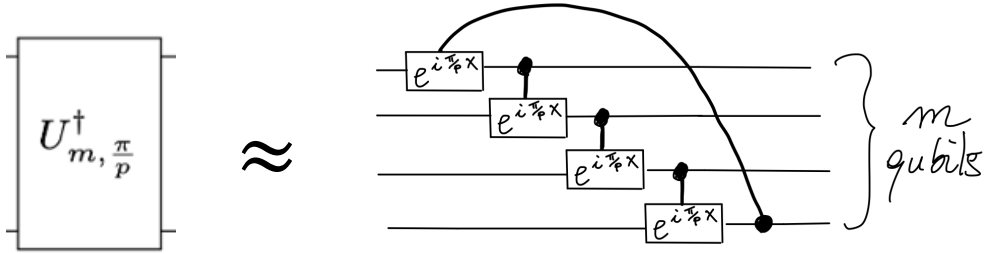
"Noisy SVM"

Open problem

No explicit (simple) construction for the shallow quantum circuits.



$|PM_n\rangle$: Poor-man's GHZ state

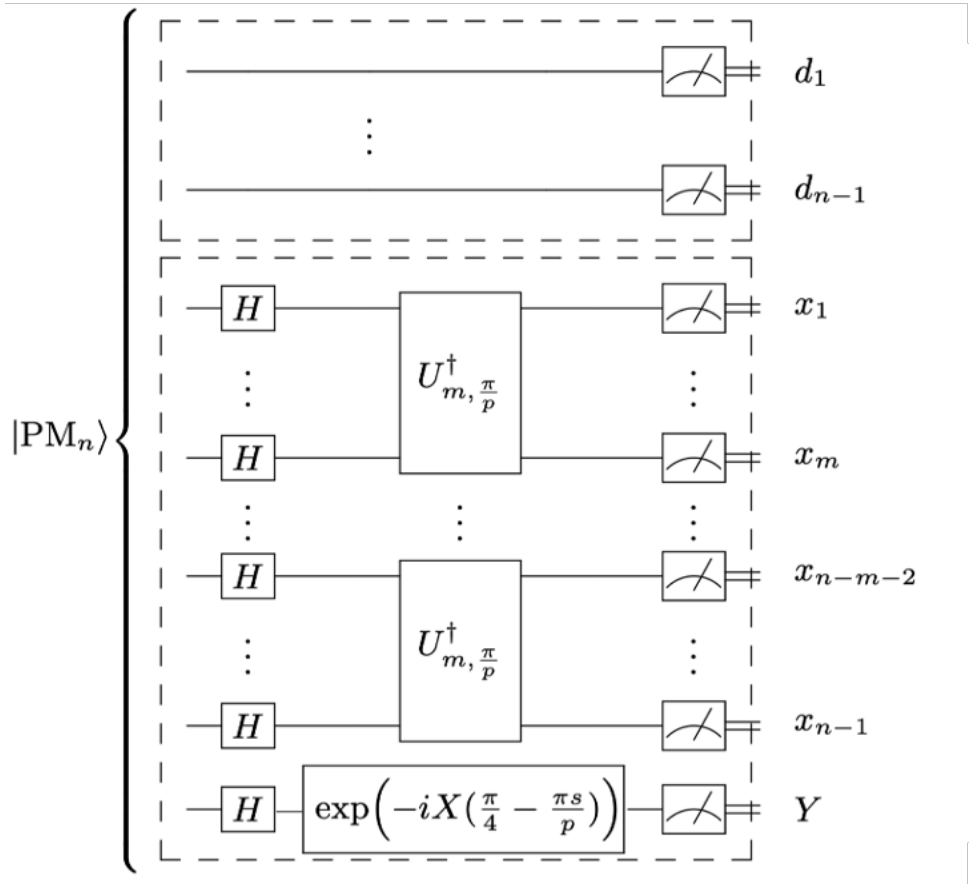


Can be compiled for a given m (independent of n), using at most $\mathcal{O}(4^m m^3)$ gates.

We would like a simpler (shallower) circuit.

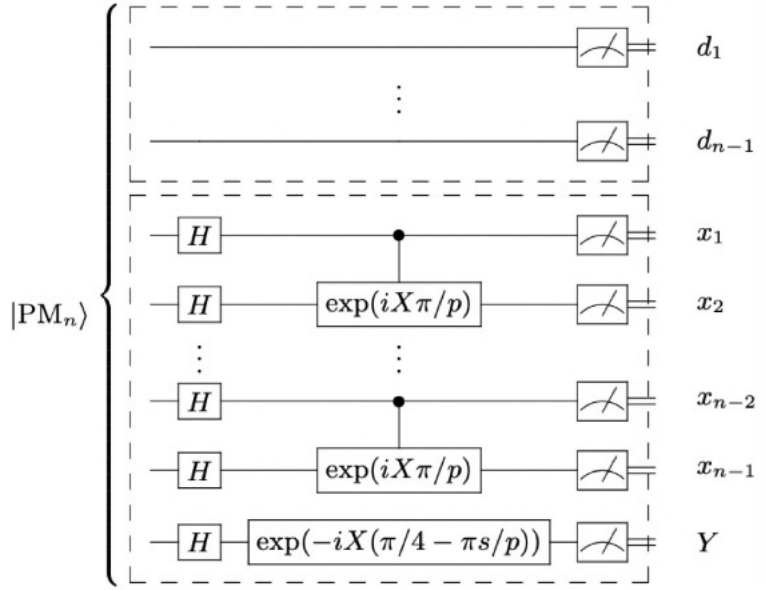
Open problem

No explicit (simple) construction for the shallow quantum circuits.

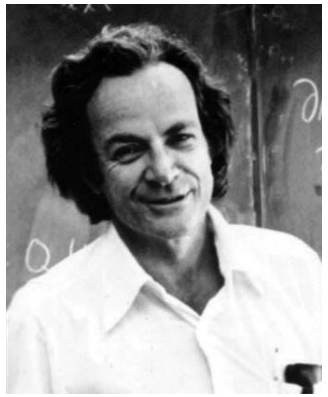


Conjecture:

This circuit leads to a distribution that is not in NC^0 (even approximately).



More physical: properties of quantum systems



Feynman, 1982

Widespread belief:

Quantum computers should be capable of simulating
quantum systems more efficiently than classical computers



Lloyd, 1996

Our intuition:

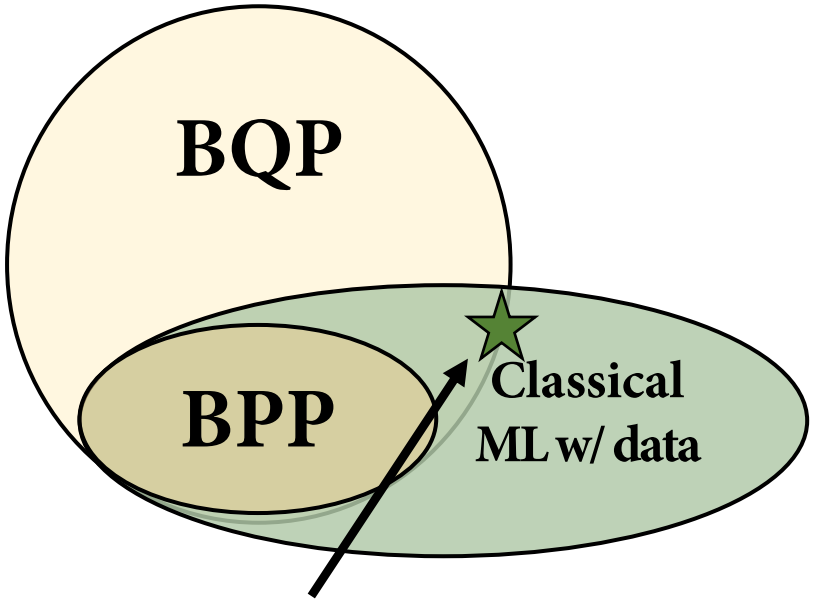
Advantage should generalize to learning properties of quantum systems

Crucial remark:

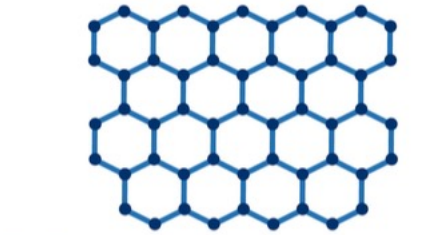
We consider learning tasks where data is classical (no quantum states as input)
but generated from quantum processes

Power of data

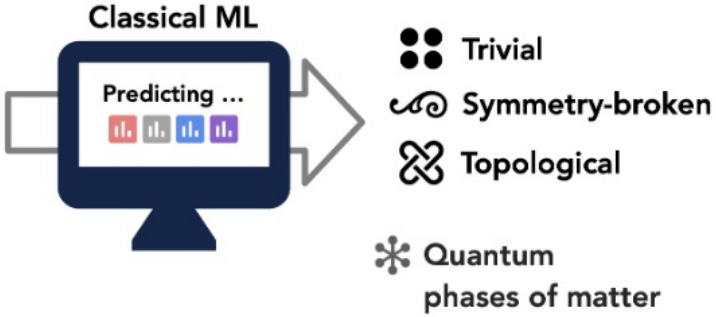
What are we actually comparing?



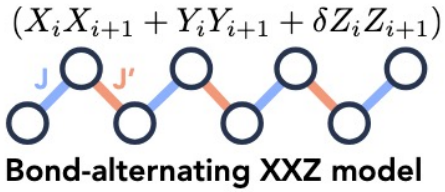
Predicting ground-state properties of gapped Hamiltonians from examples

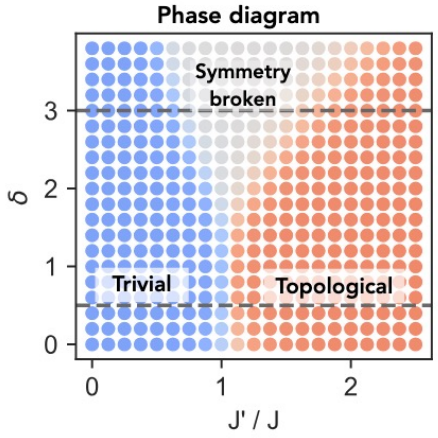


Parameters describing a physical Hamiltonian



Hamiltonian $H(x)$, $x \in [0,1]^m$ with ground states $\psi(x)$

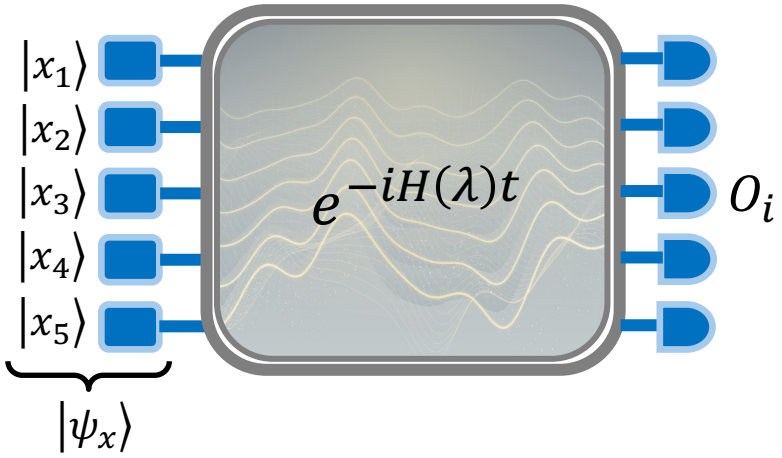


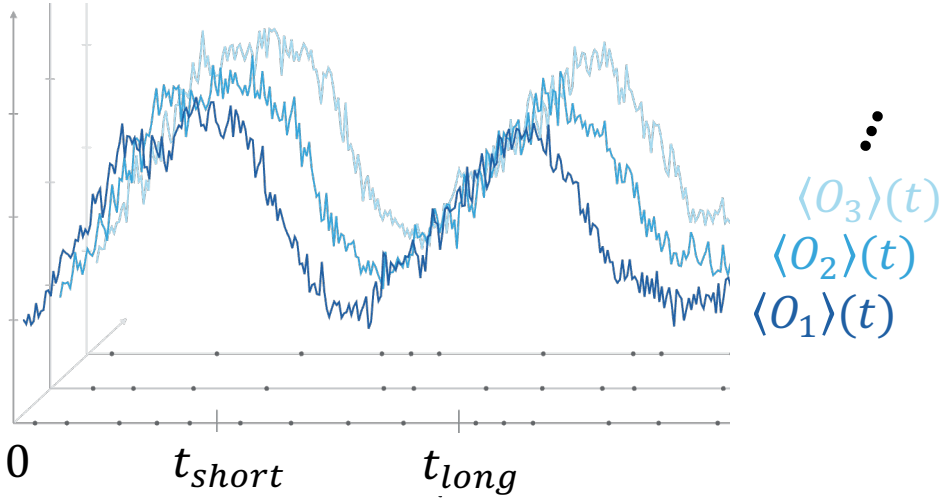


Huang, Küng, Torlai, Albert & Preskill. **Provably efficient machine learning for quantum many-body problems**. *Science* (2022)

Gyurik & Dunjko. **Exponential separations between classical and quantum learners**. *arXiv:2306.16028*

Learning task based on quantum dynamics





Target functions (concepts):

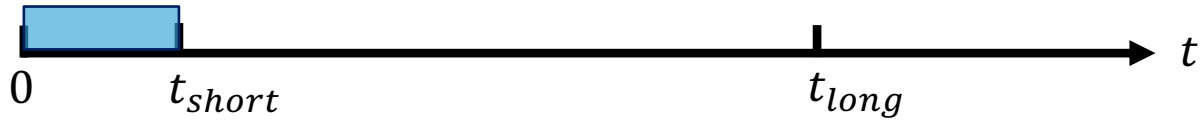
Hamiltonian family: $\{H(\lambda) = \sum_a \lambda_a P_a\}_\lambda$ → Function family: $\{f_\lambda\}_\lambda$ where $f_\lambda(x,t) = \begin{pmatrix} \mathrm{Tr}\left[e^{-i\mathrm{H}(\lambda)t}\psi_x e^{i\mathrm{H}(\lambda)t}O_1\right], \\ \mathrm{Tr}[\quad \ldots \quad O_2], \\ \ldots \quad \end{pmatrix}$

PAC learning:

From training data $\mathcal{D} = \left\{\left(x^{(i)}, t^{(i)}, f_\lambda(x^{(i)}, t^{(i)})\right)\right\}_i$

output function h, s.t. $\mathbb{E}_{x,t}[\|h(x,t) - f_\lambda(x,t)\|_\infty] \leq \varepsilon$ with prob. $1-\delta$

A Hamiltonian learning algorithm

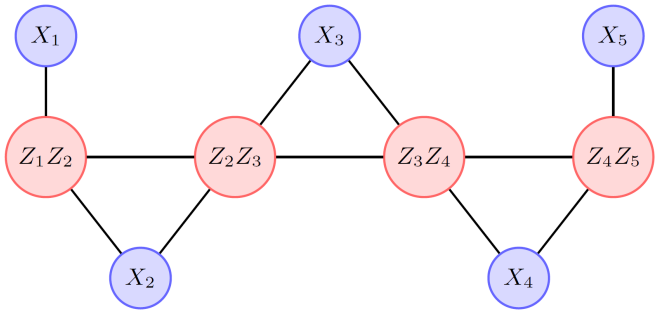


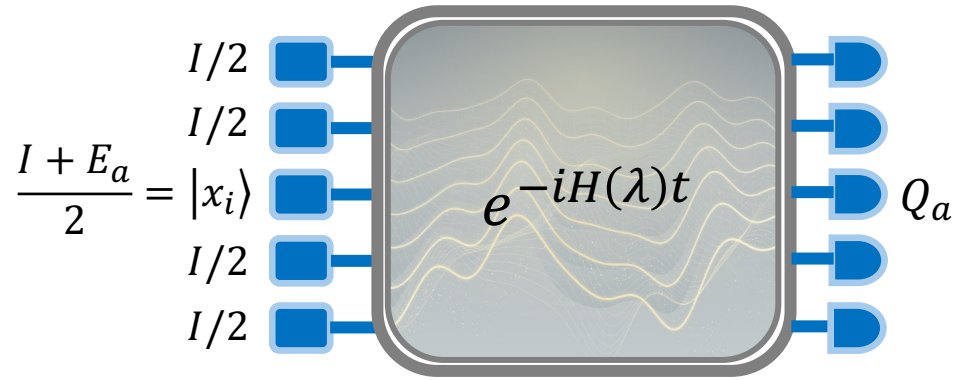
Unknown Hamiltonian:

$$H(\lambda) = \sum_{a\in[M]} \lambda_a P_a$$

Goal: learn λ' s.t. $\|\lambda' - \lambda\|_\infty \leq \varepsilon$

P_a: local (few-body) Paulis w/ <u>low-intersection</u>





$\{P_a, E_a\} = 0$

$Q_a = 2i[P_a, E_a]$

Taylor expansion of time evolution:

$$\mathrm{Tr}\left[Q_a e^{-iH(\lambda)t} E_a e^{iH(\lambda)t}\right] = 4\lambda_a t + O(t^2)$$

↳ As in $\lambda_a P_a$

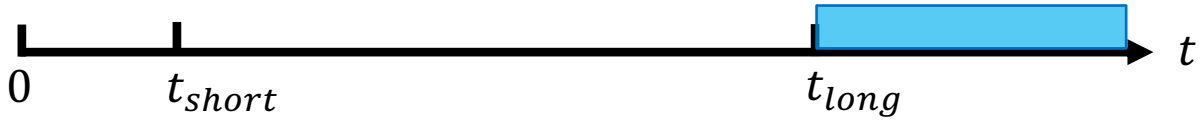
Generalizing:

Once λ' learned, $e^{iH(\lambda')t}$ approximates $e^{iH(\lambda)t}$ with guarantee:

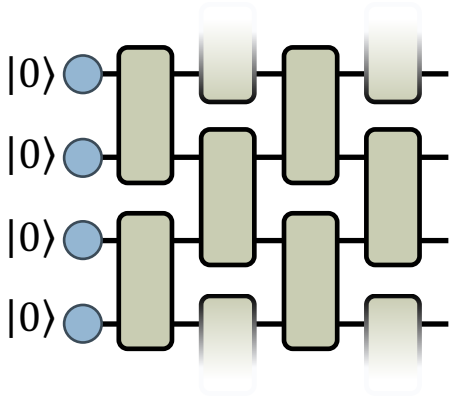
$$|\langle O\rangle_{\lambda'} - \langle O\rangle_\lambda| \leq 2\|O\|_\infty M|t|\varepsilon$$

Haah, Kothari, Tang. **Optimal learning of quantum Hamiltonians from high-temperature Gibbs states**, Nat. Phys. (2024)

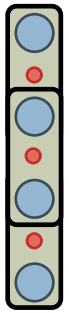
BQP-Hard instances



Universal BQP circuit



Hamiltonian



Feynman-Kitaev clock construction

: work qubit

: clock qudit

Feynman-Kitaev clock Hamiltonian

$$H_{BQP} = \sum_t \underbrace{|t+1\rangle\langle t|}_{\text{Clock transition}} \otimes \underbrace{U_t}_{\text{Gate}} + |t\rangle\langle t+1| \otimes U_t^\dagger$$

For unary clock encoding:

$$|t\rangle = \left|1^t 0^{T-t}\right\rangle$$

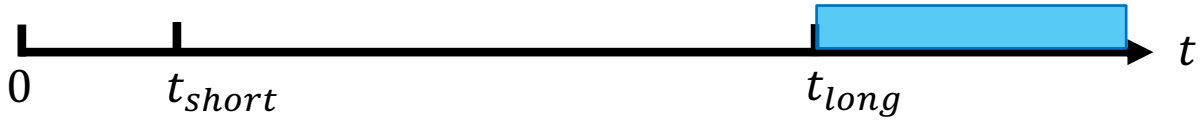
And universal gate set:

$$U_t \in \{\text{CNOT}, \text{H}, \text{T}\}$$

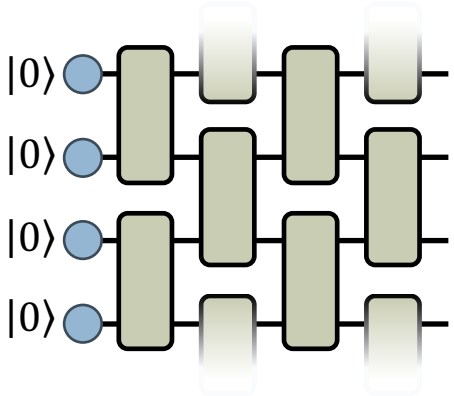
H_{BQP} is 5-local (in the Pauli basis)

$e^{iH_{BQP}T}$ simulates orginal circuit
for $T \in O(\text{poly}(n))$

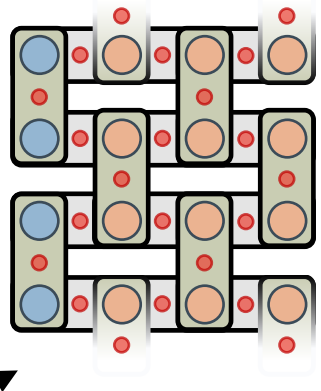
BQP-Hard instances



Universal BQP circuit



2D low-intersection Hamiltonian



: work qubit
: aux. qubit
: clock qubit
: SWAP terms

Feynman-Kitaev clock construction

Feynman-Kitaev clock Hamiltonian

$$H_{BQP} = \sum_t \underbrace{|t+1\rangle\langle t|}_{\text{Clock transition}} \otimes \underbrace{U_t}_{\text{Gate}} + |t\rangle\langle t+1| \otimes U_t^\dagger$$

For unary clock encoding:

$$|t\rangle = \left|1^t 0^{T-t}\right\rangle$$

And universal gate set:

$$U_t \in \{\text{CNOT}, \text{H}, \text{T}\} \cup \{\textbf{SWAP}\}$$

H_{BQP} is 5-local (in the Pauli basis) **& low intersection**

$e^{iH_{BQP}T}$ simulates orginal circuit
for $T \in O(\text{poly}(n))$

Putting it all together



Short-time samples allow to learn:

$$H(\lambda) = \sum_{a\in[M]} \lambda_a P_a$$

such that $e^{iH(\lambda')t}$ generalizes to t_{long}

Long-time samples can encode BQP-hard computations:

$$H_{BQP} = \sum_t |t+1\rangle\langle t| \otimes U_t + |t\rangle\langle t+1| \otimes U_t^\dagger$$

& can be quantumly simulated

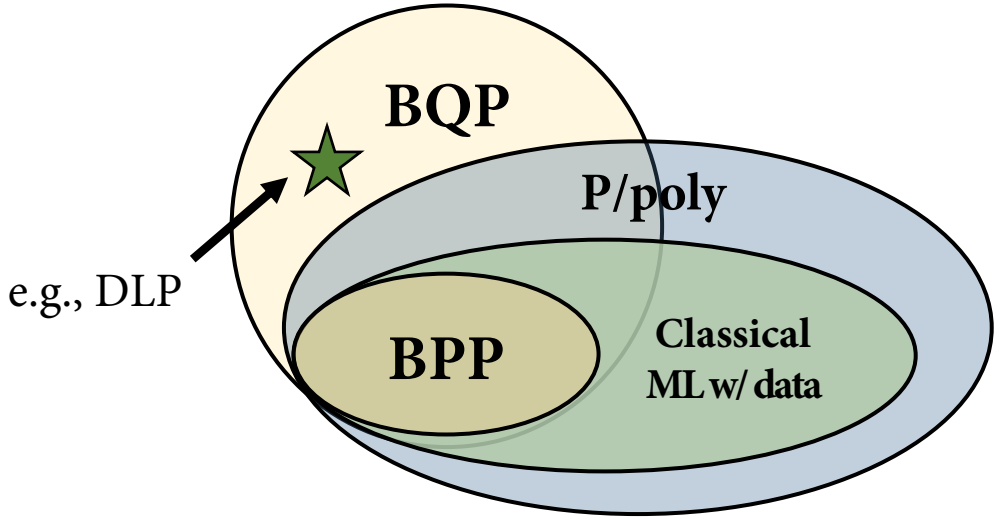
PAC learning:

From training data $\mathcal{D} = \{(x^{(i)}, t^{(i)}, f_\lambda(x^{(i)}, t^{(i)}))\}_i$

output function h, s.t. $\mathbb{E}_{x,t}[\|h(x,t) - f_\lambda(x,t)\|_\infty] \leq \varepsilon$ with prob. $1-\delta$

↳ We constructed an efficient quantum learner $\forall\ \varepsilon, \delta$

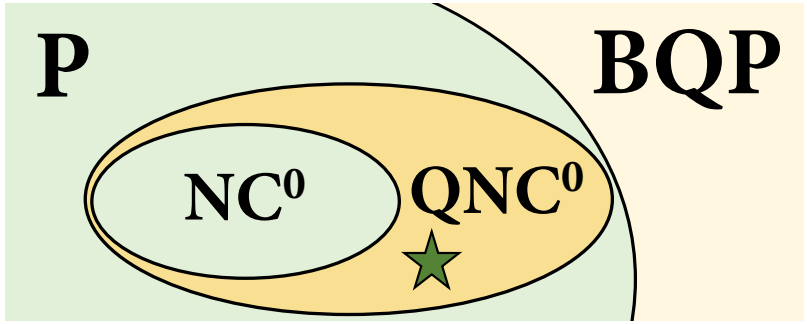
↳ An efficient classical learner with non-trivial performance would need to perform well at long times, and therefore could simulate BQP computations



Summary (1)

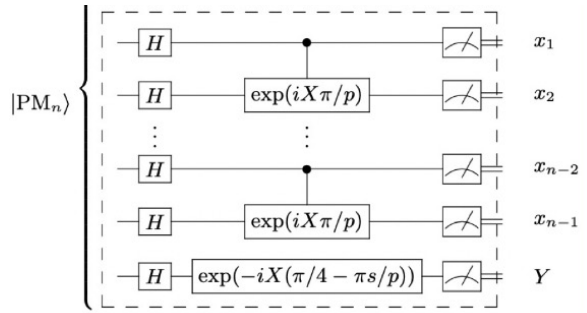
Main question:

Can we show unconditional learning advantages using near-term, shallow-depth circuits?



Outlook (1)

- Can we show that simpler explicit constructions also lead to an advantage?
- Can we show robustness to noise?
- More physically-motivated tasks?



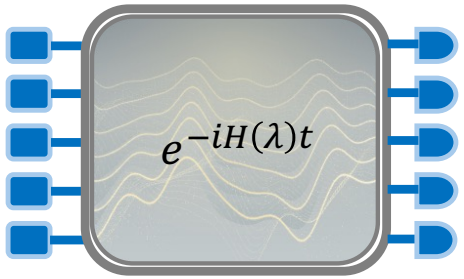
Grier et al. (arXiv:2510.07808) -> Explicit construction for a related task but classical hardness only applies for $n \ggg$ # atoms in obs. universe

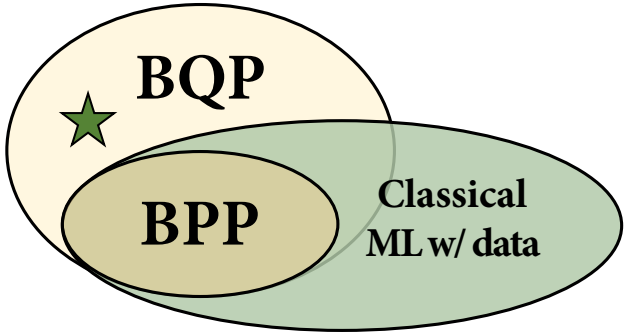
$2^{2^{\cdot^{\cdot^{2}}}}_{30}$ v.s. $2^{2^{2^{2}}}_{4}$

Summary (2)

Main question:

Can we show (conditional) learning advantages for physically motivated learning tasks?





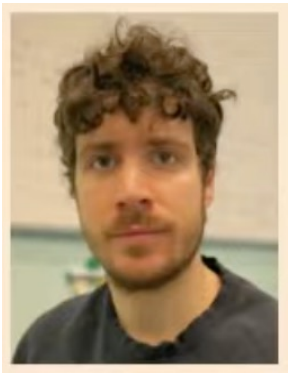
Outlook (2)

- Getting rid of short-time evolutions? → Hamiltonian learning from long times
- Extension to other physical settings? Gibbs states? Lindbladian evolution?

Special thanks



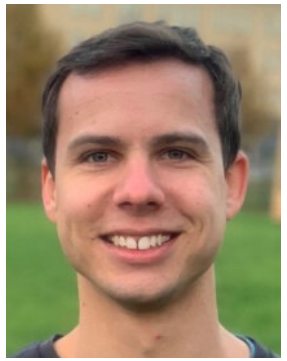
Rahul Bandyopadhyay



Riccardo Molteni



Vedran Dunjko



Niklas Pirnay



Jens Eisert



J.-P. Seifert

Universiteit Leiden

Freie Universität Berlin

TECHNISCHE UNIVERSITÄT BERLIN