Quantum AI: Toward the next revolution driven by quantum computing and AI

Keisuke Fujii

**Professor, Graduate School of Engineering Science/
Deputy director, Center for Quantum Information and Quantum Biology
The University of Osaka**

大阪大学
OSAKA UNIVERSITY

大阪大学
公式マスコットキャラクタ
「ワニ博士」

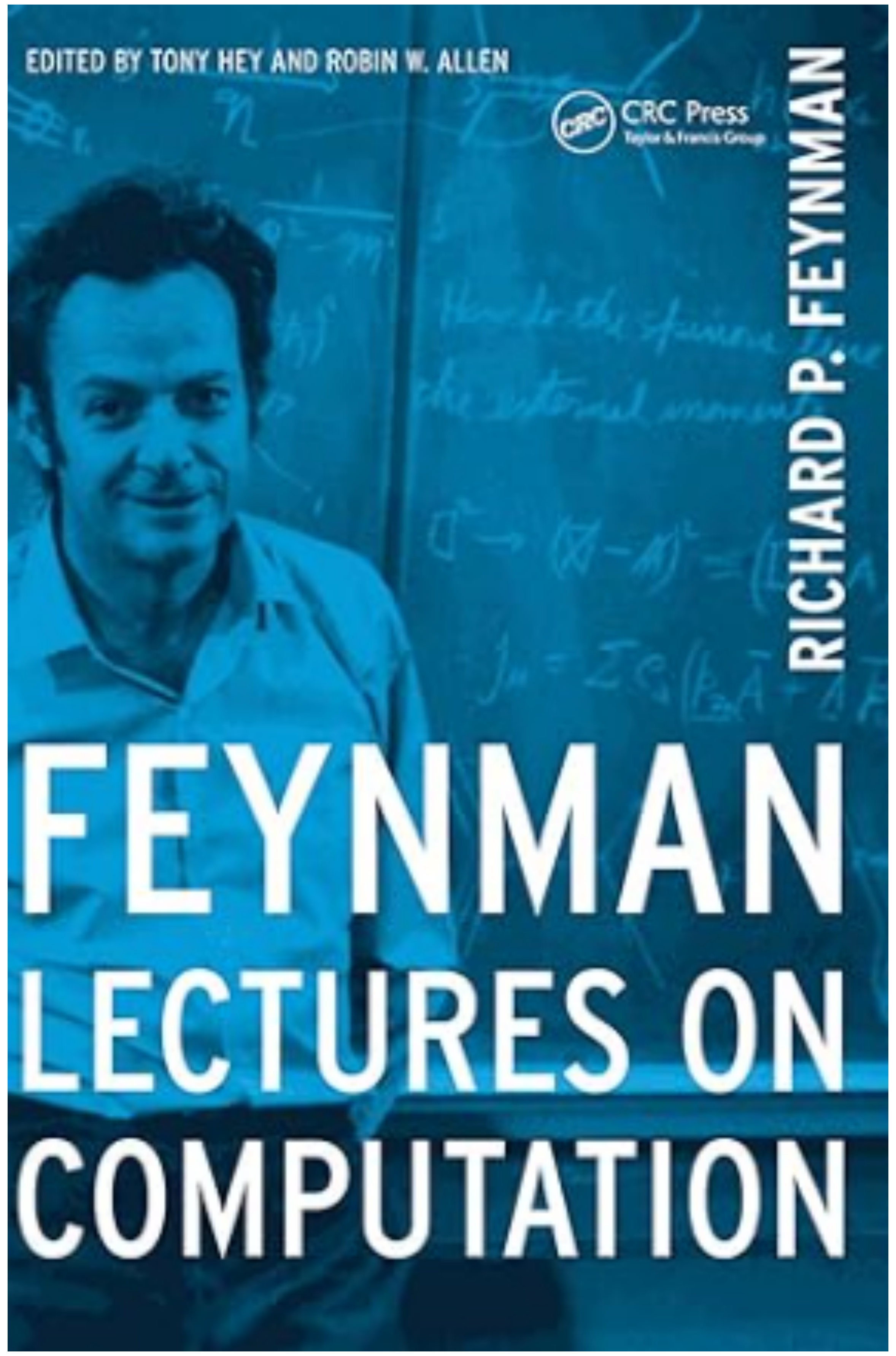
QIQB

RIKEN
QUANTUM
COMPUTING

Introduction

Historical link between Quantum Computing and AI

- R. Feynman, J. Hopfield, and C. Mead pioneering interdisciplinary lecture course "*The Physics of Computation*" at Caltech (1981–1984).
- In 1981, Feynman proposed quantum computers, bridging quantum physics and information theory seeking fundamental limits on computing.
- In 1982, Hopfield introduced the Hopfield network, pioneer of neural networks connecting physics and computation.

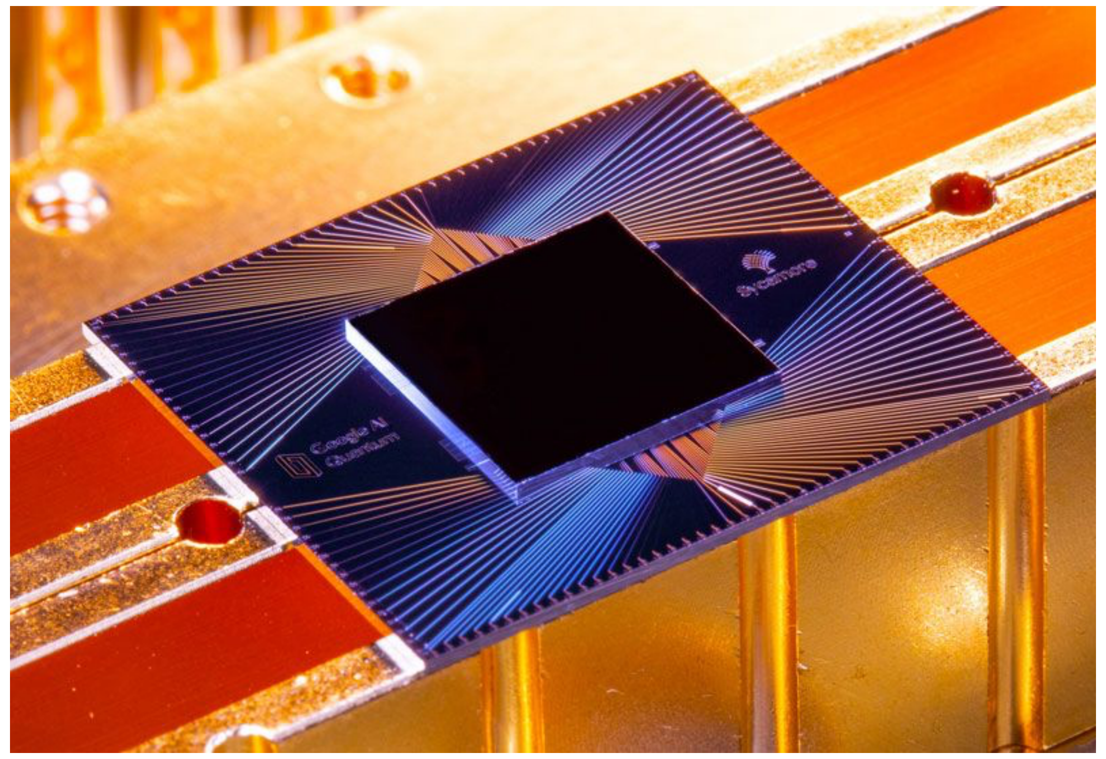


Since then, Quantum Computing and AI took separate paths — until they met again around 2018.

We are now in the NISQ era

NISQ: Noisy Intermediate-scale Quantum technologies/computers
J. Preskill *“Quantum Computing in the NISQ era and beyond”* Quantum 2018

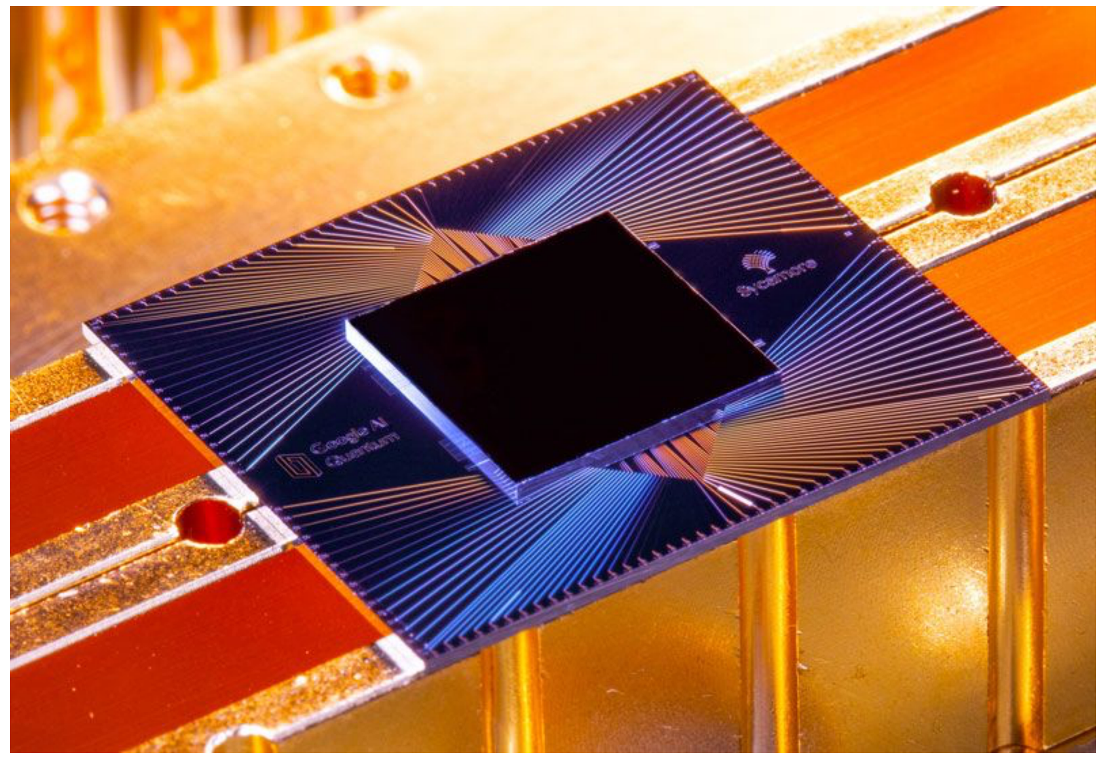
- Quantum computers of over 100 qubits are now available.
- # of qubits is still too small for FTQC, i.e. error-corrected quantum computers, which typically needs 1M qubits.
- Quantum-classical hybrid algorithms are thought to be effective with shallow quantum circuits.



Google team, Nature **574** 505 (2019)

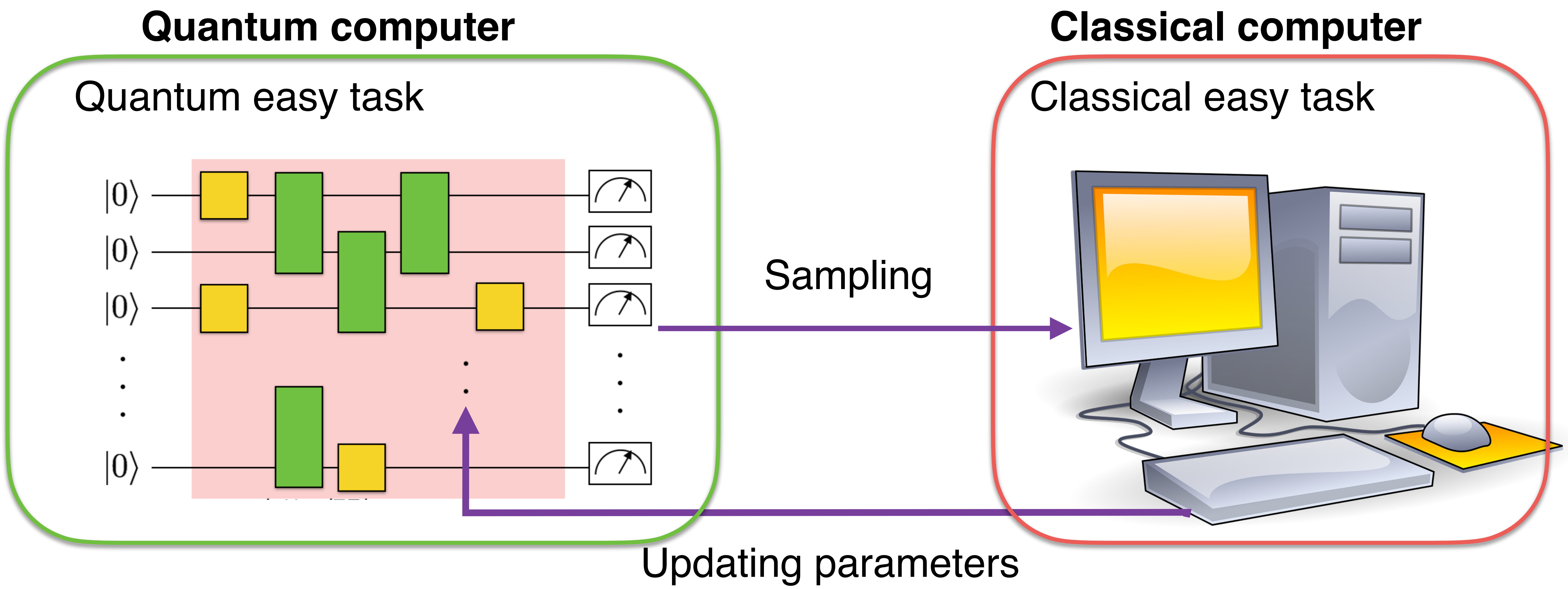
We are now in the NISQ era

NISQ: Noisy Intermediate-scale Quantum technologies/computers
J. Preskill *“Quantum Computing in the NISQ era and beyond”* Quantum 2018



Google team, Nature **574** 505 (2019)

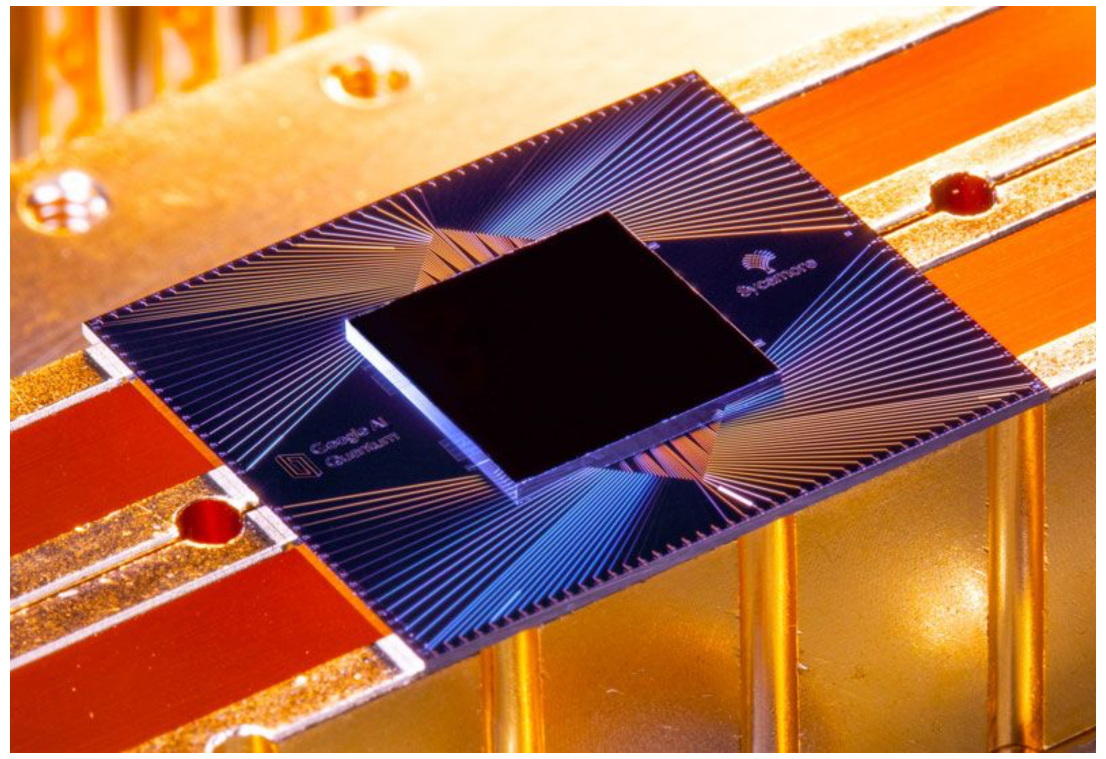
- Quantum computers of over 100 qubits are now available.
- # of qubits is still too small for FTQC, i.e. error-corrected quantum computers, which typically needs 1M qubits.
- Quantum-classical hybrid algorithms are thought to be effective with shallow quantum circuits.



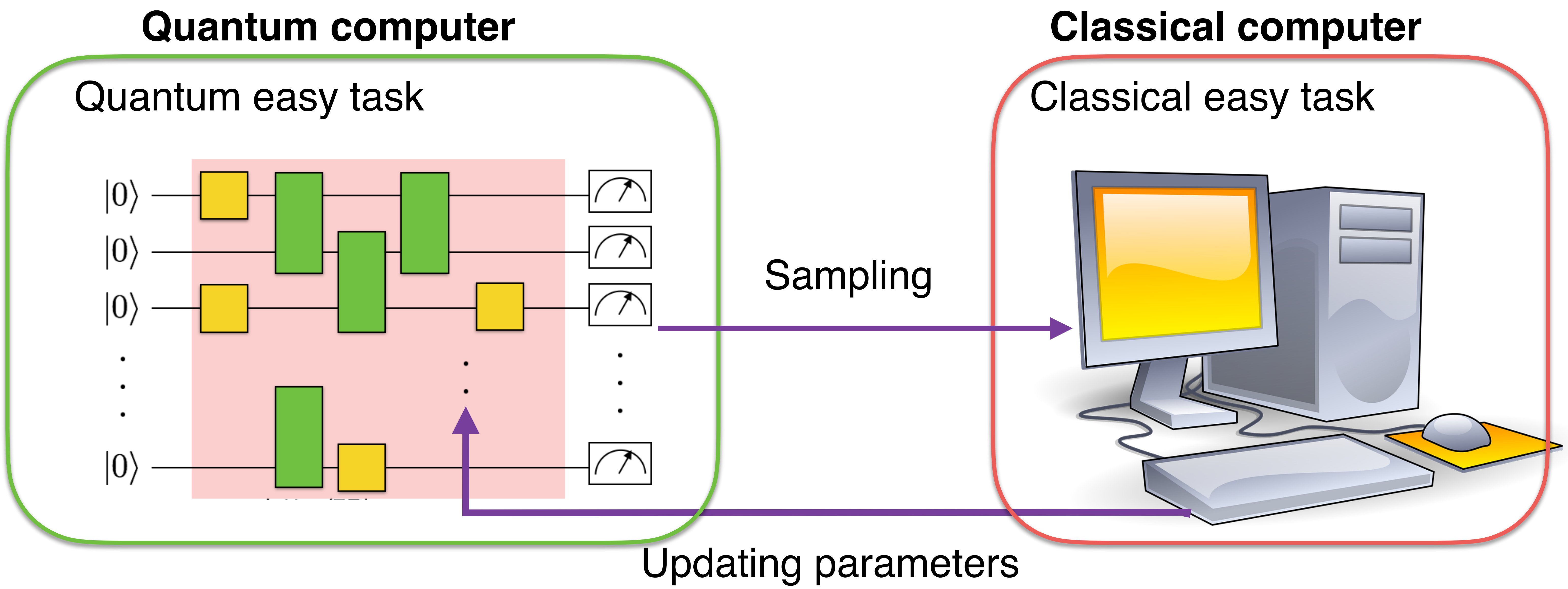
We are now in the NISQ era

NISQ: Noisy Intermediate-scale Quantum technologies/computers
J. Preskill *"Quantum Computing in the NISQ era and beyond"* Quantum 2018

- Quantum computers of over 100 qubits are now available.
- # of qubits is still too small for FTQC, i.e. error-corrected quantum computers, which typically needs 1M qubits.
- Quantum-classical hybrid algorithms are thought to be effective with shallow quantum circuits.



Google team, Nature **574** 505 (2019)



Quantum Machine Learning

Quantum Computer for ML

&

ML for Quantum Computer

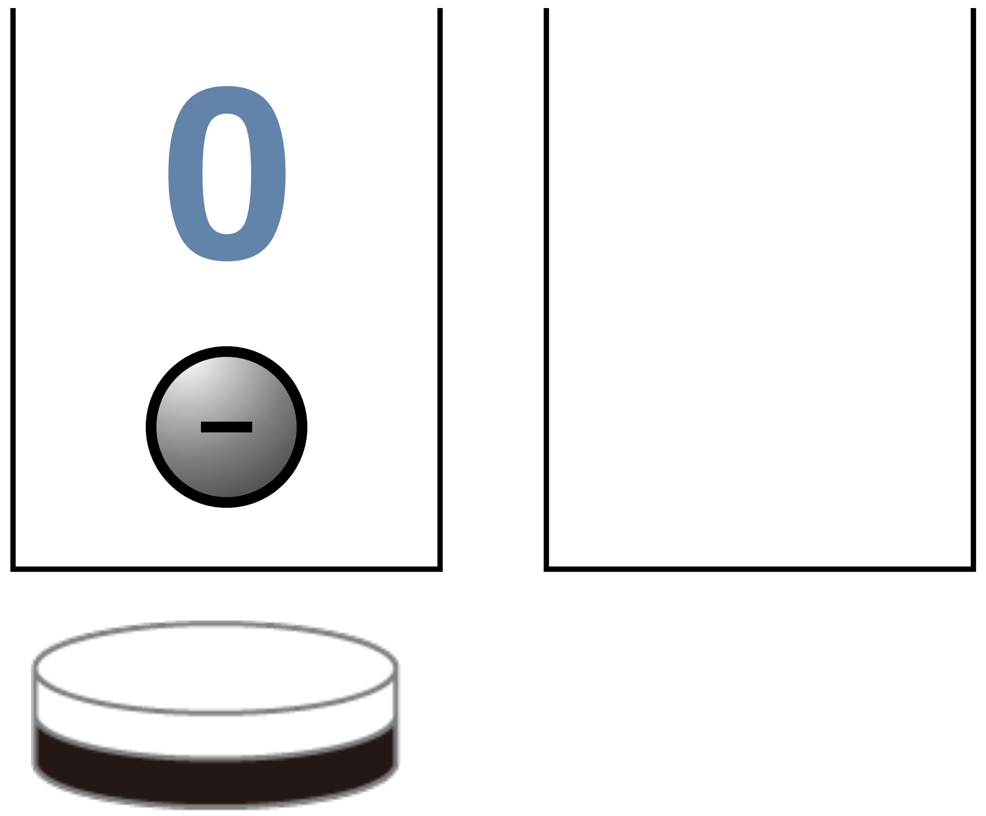
How Quantum Computer and Quantum Machine Learning work

Classical bit and Quantum bit

The minimum unit of information in the “classical" or “quantum” world

Classical bit

$$x \in \{0, 1\}$$

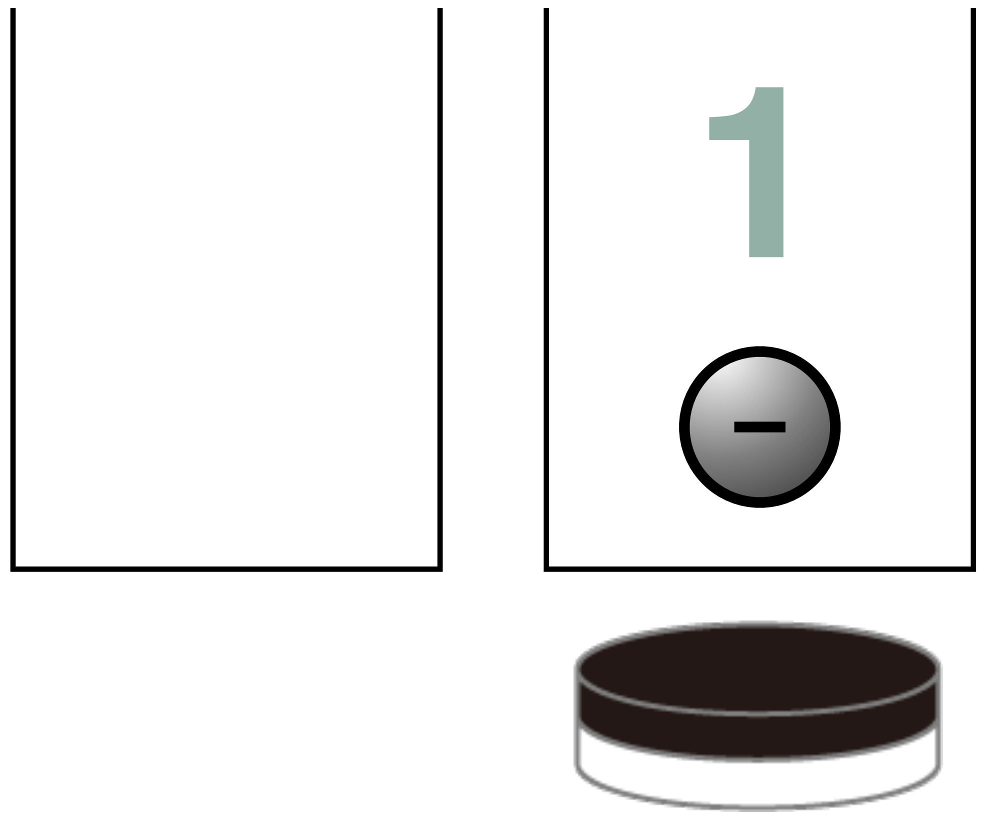


Classical bit and Quantum bit

The minimum unit of information in the “classical" or “quantum” world

Classical bit

↑ **0** or **1** ↓ $x \in \{0, 1\}$

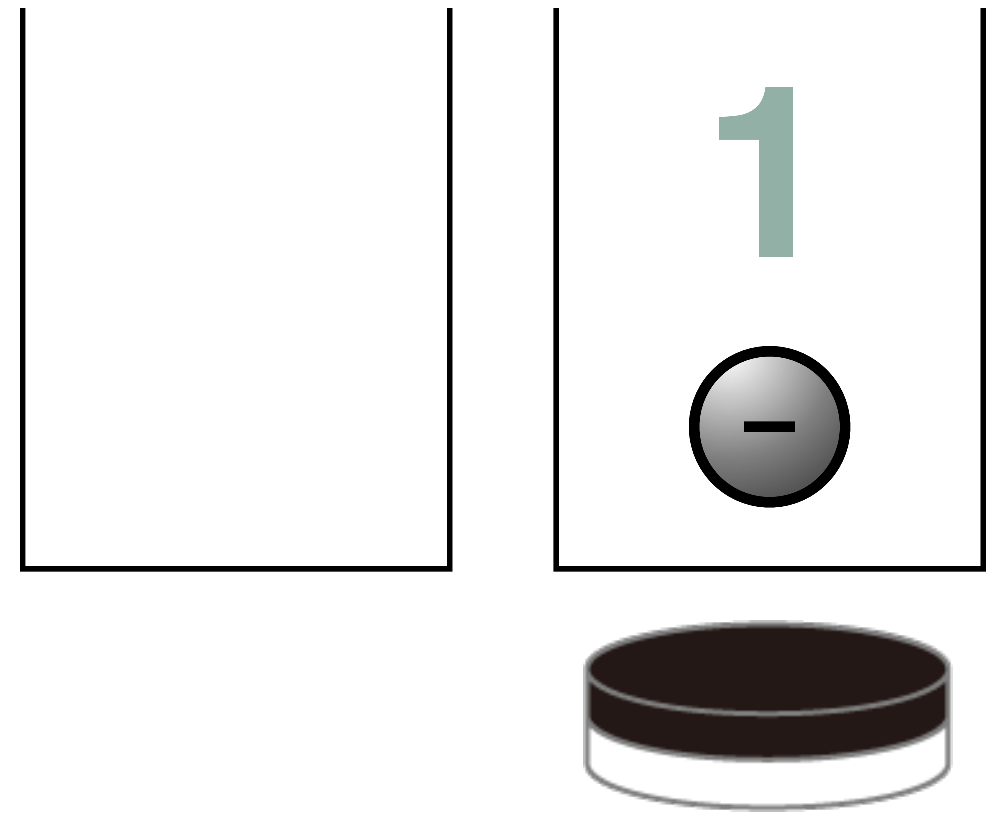


Classical bit and Quantum bit

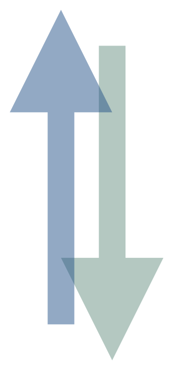
The minimum unit of information in the “classical" or “quantum” world

Classical bit

0 or 1 $x \in \{0, 1\}$



Quantum bit

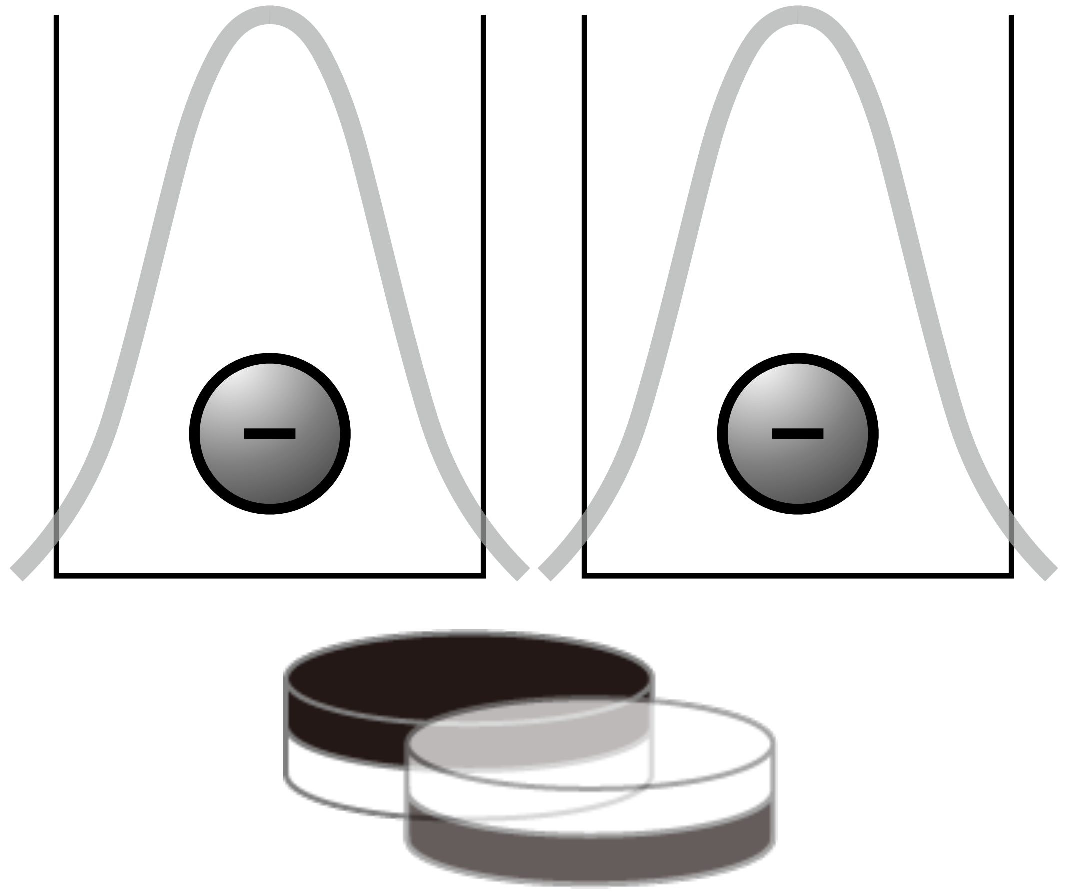


$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

Superimposed state of 0 and 1

complex vector space

$$|\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$



measurement 0 or 1 ?

$$p_0 = |\alpha|^2$$

$$p_1 = |\beta|^2$$

Probability is given by squared abs. of complex Amplitudes.

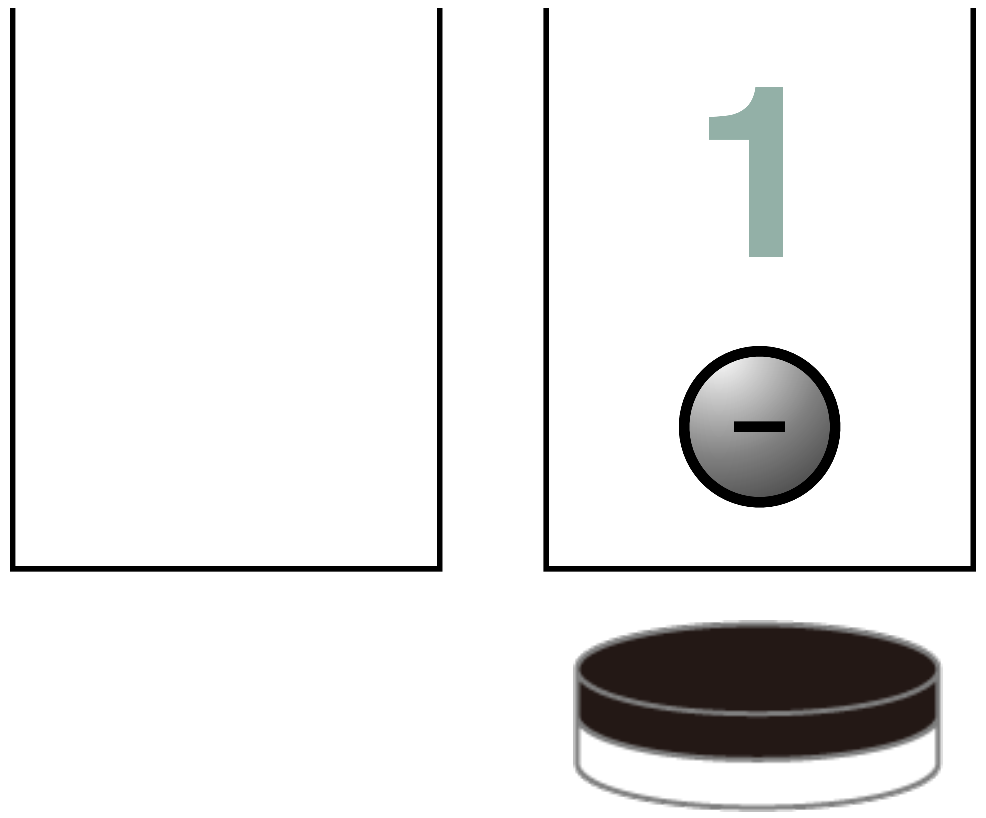
Classical bit and Quantum bit

The minimum unit of information in the “classical" or “quantum” world

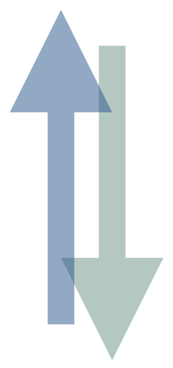
Classical bit

0 or **1**

$$x \in \{0, 1\}$$



Quantum bit

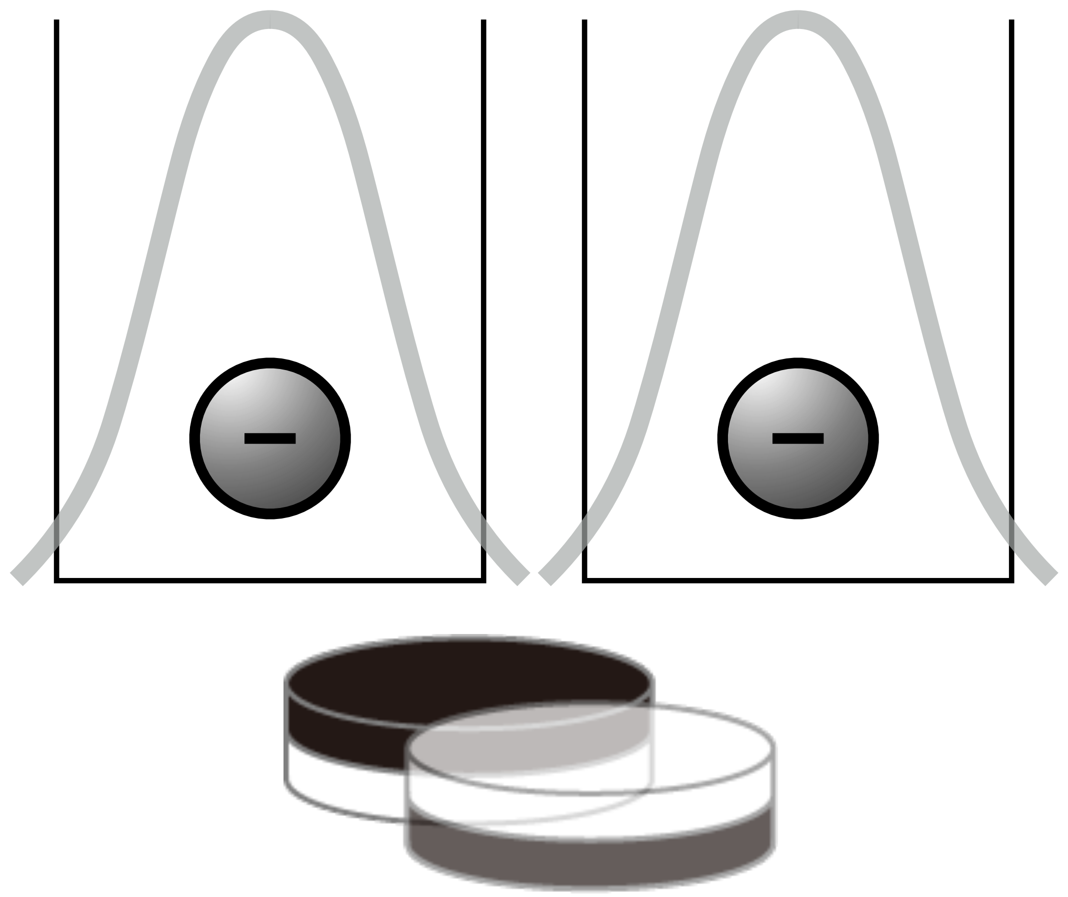


$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

Superimposed state of 0 and 1

complex vector space

$$|\psi\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$



measurement 0 or 1 ?

$$p_0 = |\alpha|^2$$

$$p_1 = |\beta|^2$$

Whether it is 0 or 1 has yet to be determined.

Probability is given by squared abs. of complex Amplitudes.

Basics of Quantum Computing

Basics of Quantum Computing

n-qubit system: a state vector $|\psi\rangle \in \mathbf{C}^{2^n}$ $\quad |\psi\rangle = \begin{pmatrix} \psi_{0\ldots0} \\ \psi_{0\ldots1} \\ \vdots \\ \psi_{1\ldots1} \end{pmatrix}$ or density matrix $\rho = |\psi\rangle\langle\psi|$

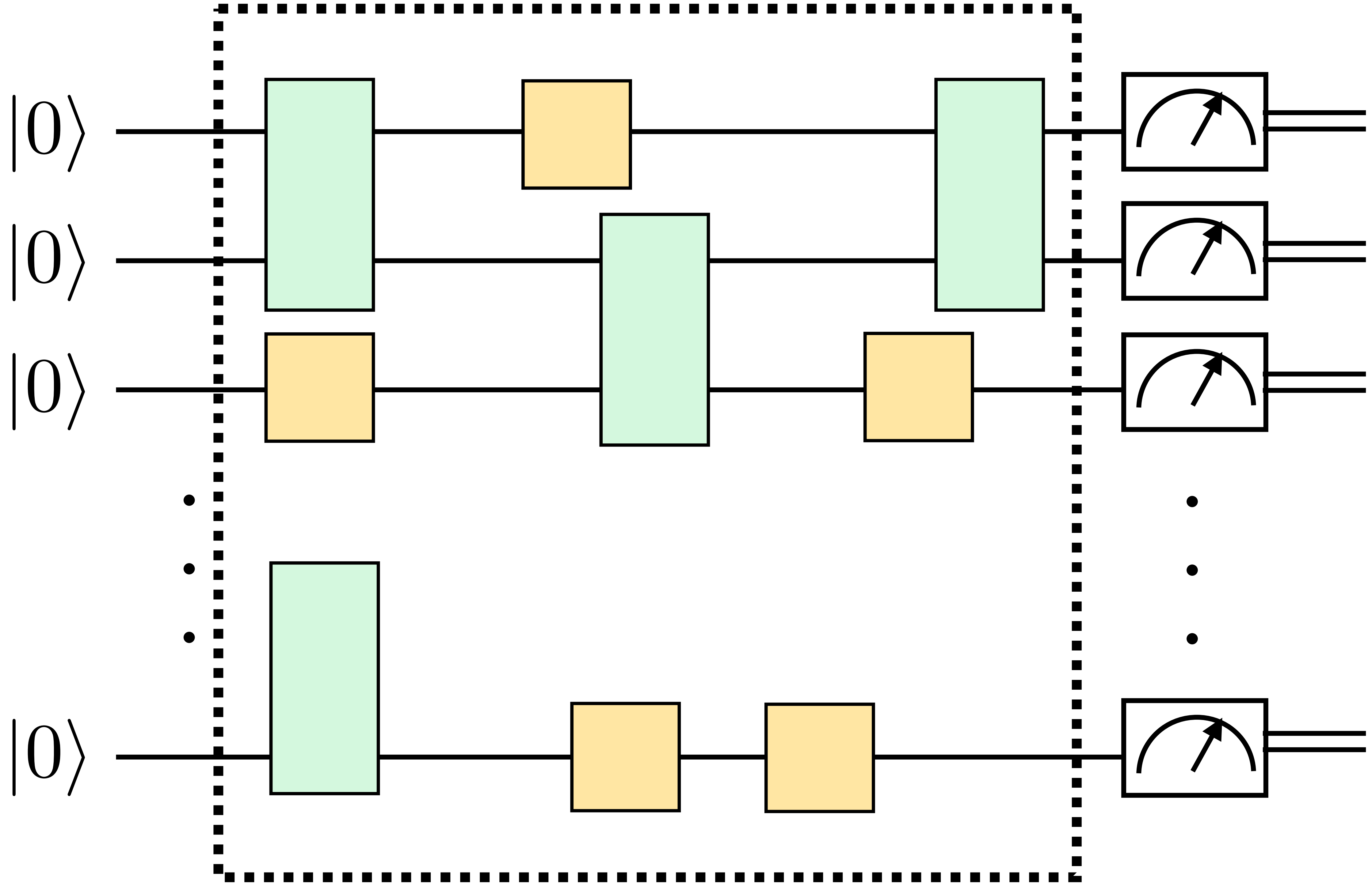
(quantum analog of prob. dist.)

Basics of Quantum Computing

n-qubit system: a state vector $|\psi\rangle \in \mathbf{C}^{2^n}$ $\quad |\psi\rangle = \begin{pmatrix} \psi_{0\ldots0} \\ \psi_{0\ldots1} \\ \vdots \\ \psi_{1\ldots1} \end{pmatrix}$ or density matrix $\rho = |\psi\rangle\langle\psi|$

(quantum analog of prob. dist.)

$\rho_{\text{in}} = |0\ldots0\rangle\langle 0\ldots0|$



$\rho_{\text{out}} = U\rho_{\text{in}}U^\dagger$(Schrodinger pic.)

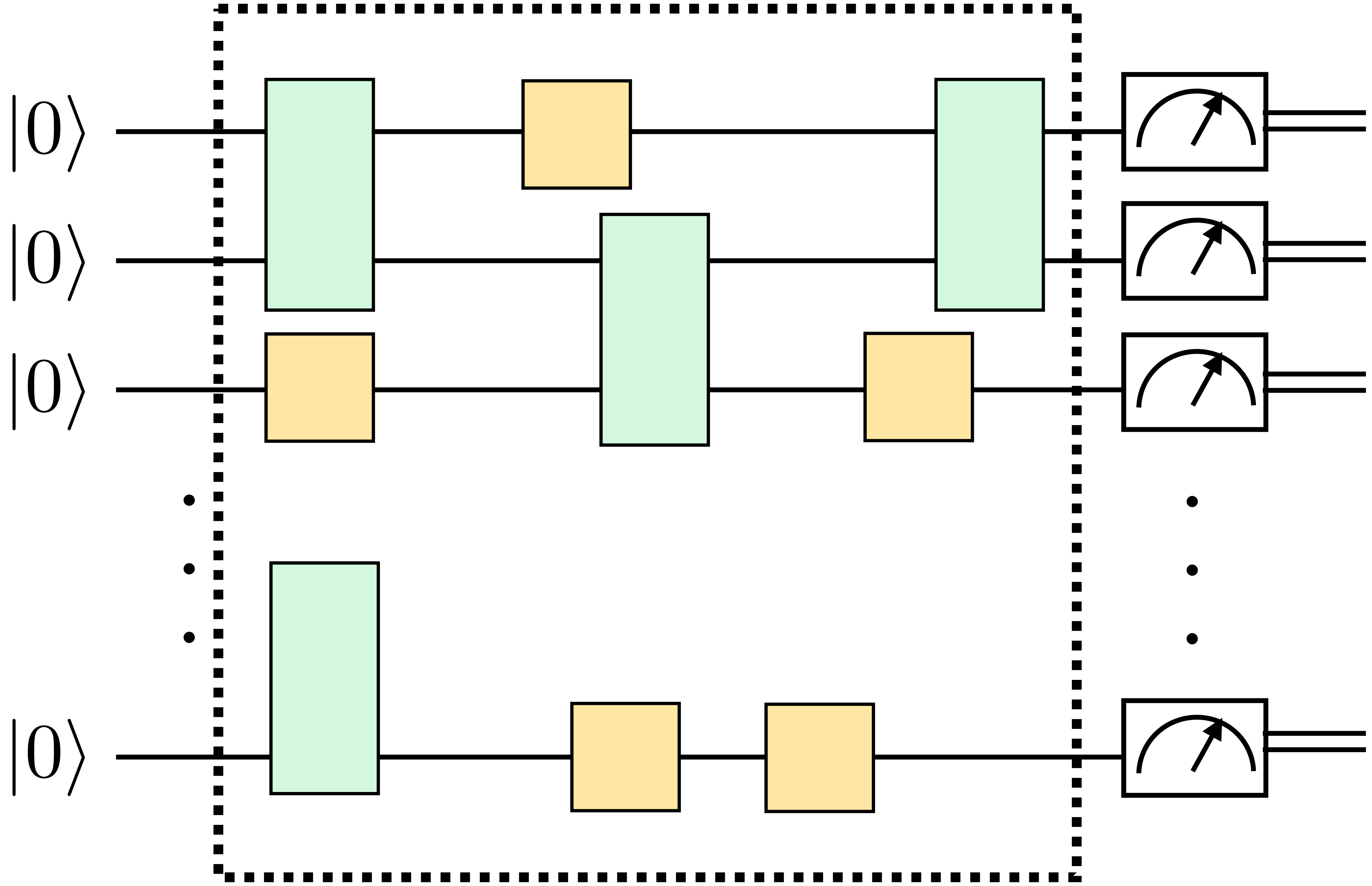
quantum circuit U written as a $2^n \times 2^n$ unitary matrix

Basics of Quantum Computing

n-qubit system: a state vector $|\psi\rangle \in \mathbf{C}^{2^n}$ $\quad |\psi\rangle = \begin{pmatrix} \psi_{0...0} \\ \psi_{0...1} \\ \vdots \\ \psi_{1...1} \end{pmatrix}$ or density matrix $\rho = |\psi\rangle\langle\psi|$

(quantum analog of prob. dist.)

$\rho_{\mathrm{in}} = |0...0\rangle\langle 0...0|$



Output is associated with an Hermitian matrix, called observable M

The value obtained from the outputs of quantum computer:

$$\langle 0...0|U^\dagger M U|0...0\rangle = \mathrm{Tr}[U\rho_{\mathrm{in}}U^\dagger M]$$
$$= \mathrm{Tr}[\rho_{\mathrm{in}}\tilde{M}]$$

$$\tilde{M} = U^\dagger M U$$

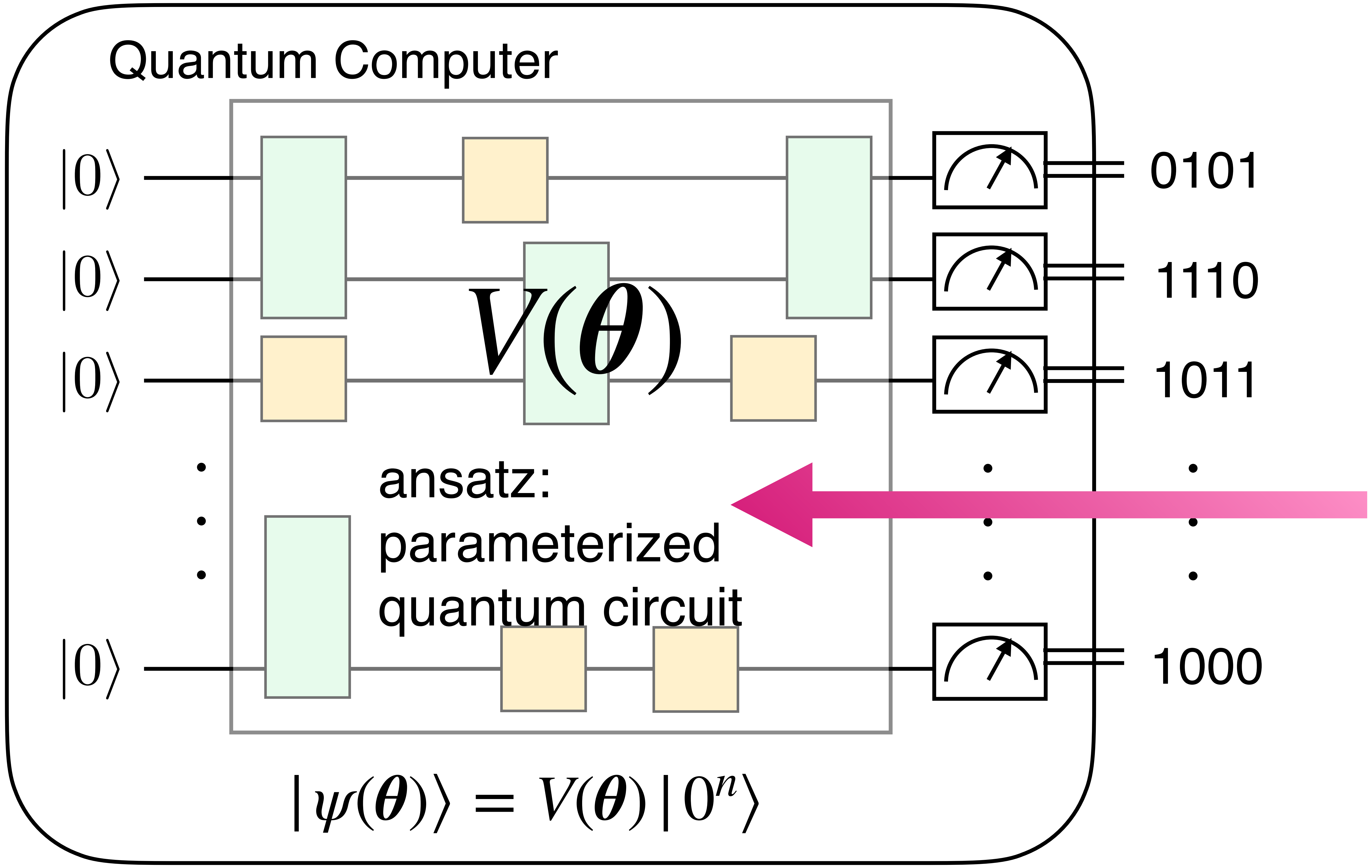
(Heisenberg pic.)

$\rho_{\mathrm{out}} = U\rho_{\mathrm{in}}U^\dagger$(Schrodinger pic.)

quantum circuit U written as a $2^n \times 2^n$ unitary matrix

Variational Quantum Algorithms

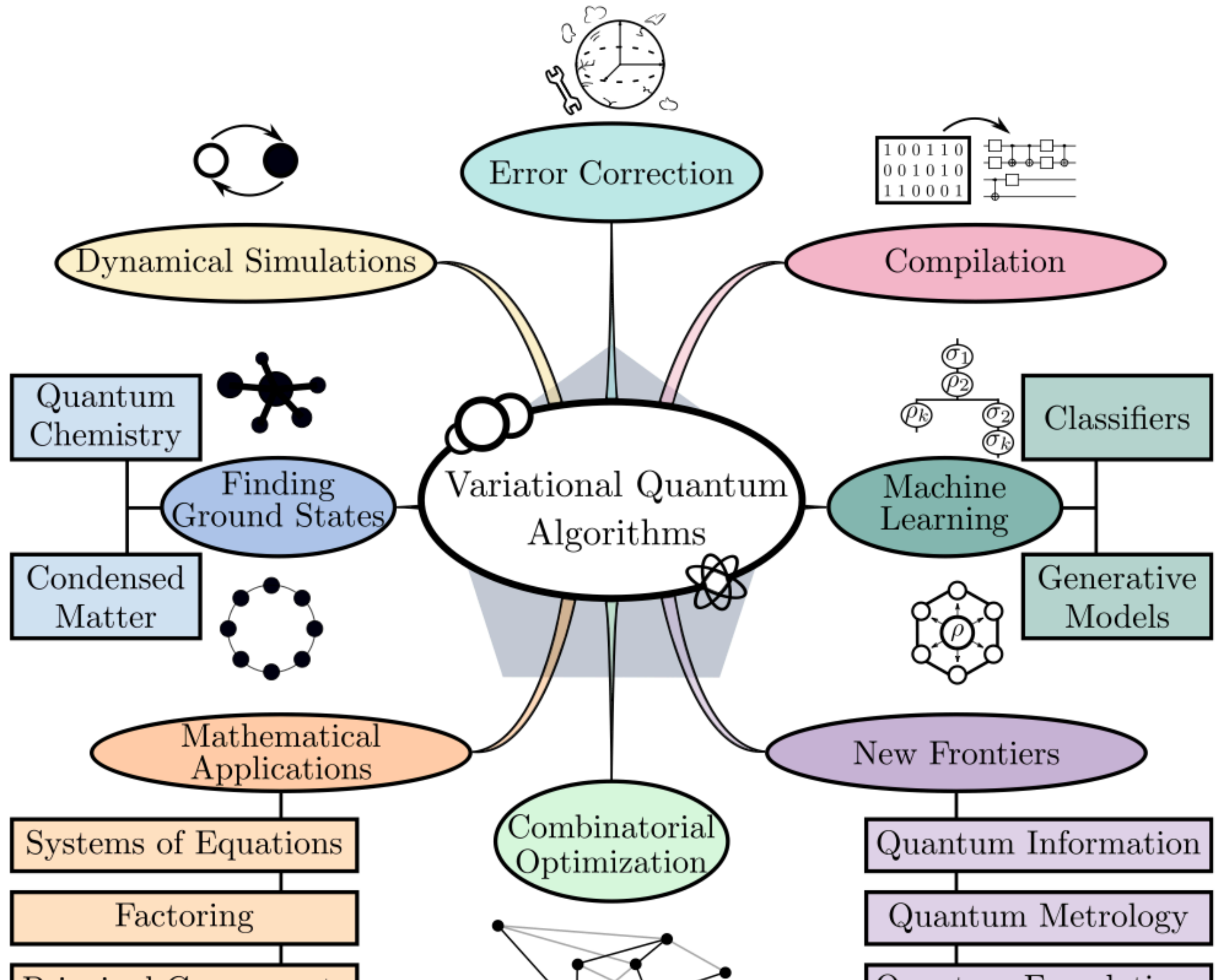
M. Cerezo et al. *"Variational Quantum Algorithms"*, Nature Review Physics **3**, 625 (2021).



Expectation value w.r.t. an observable

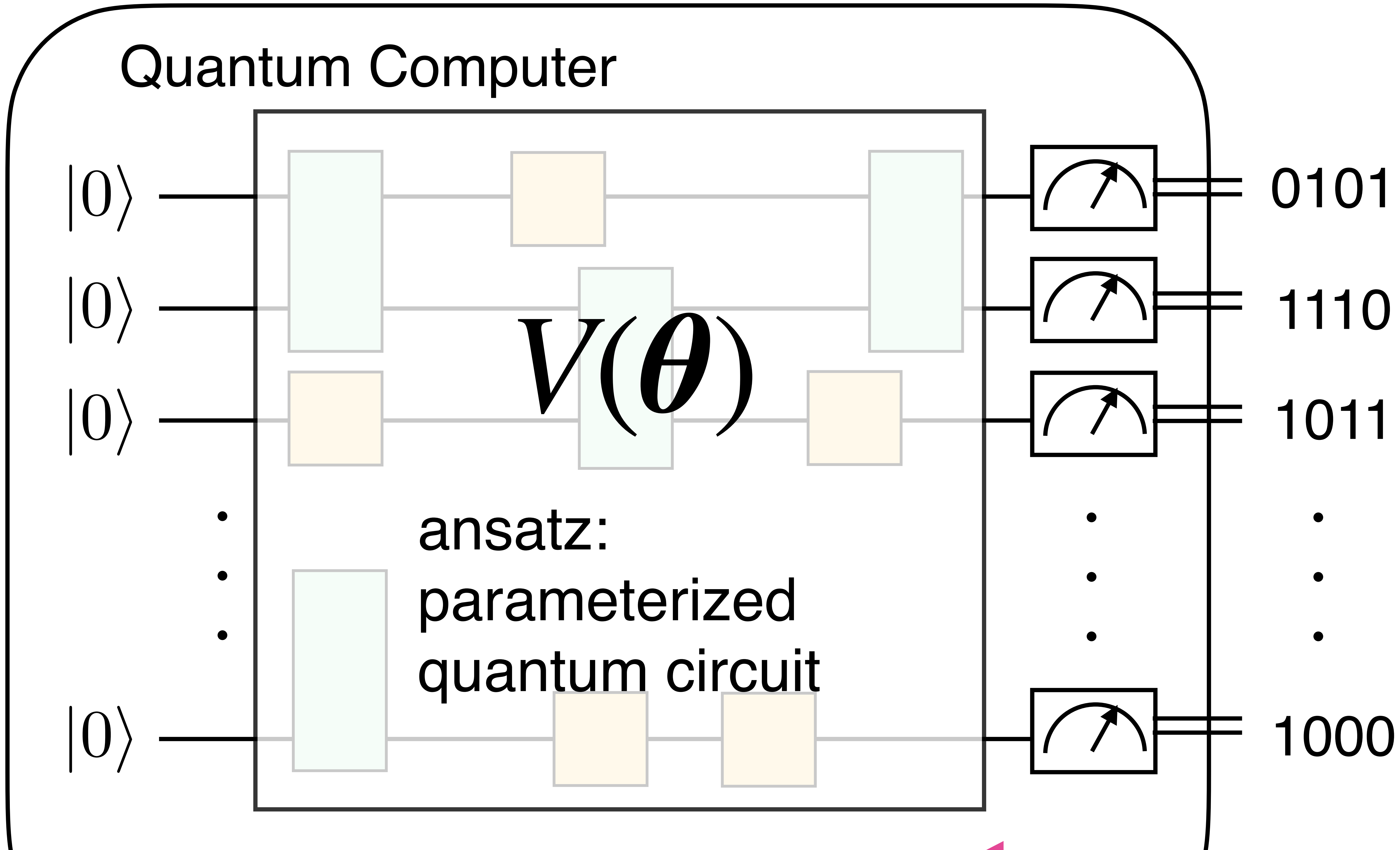
$$\mathscr{L}(\boldsymbol{\theta}) = \langle \psi(\boldsymbol{\theta})\,|\,O\,|\,\psi(\boldsymbol{\theta})\rangle$$

Update parameters so as to minimize the cost function



Variational Quantum Algorithms

M. Cerezo et al. *“Variational Quantum Algorithms”*, Nature Review Physics **3**, 625 (2021).



$$|\psi(\boldsymbol{\theta})\rangle = V(\boldsymbol{\theta})\,|0^n\rangle$$

$$\mathscr{L}(\boldsymbol{\theta}) = \langle\psi(\boldsymbol{\theta})\,|\,H\,|\,\psi(\boldsymbol{\theta})\rangle$$

Update parameters so as to minimize the cost function

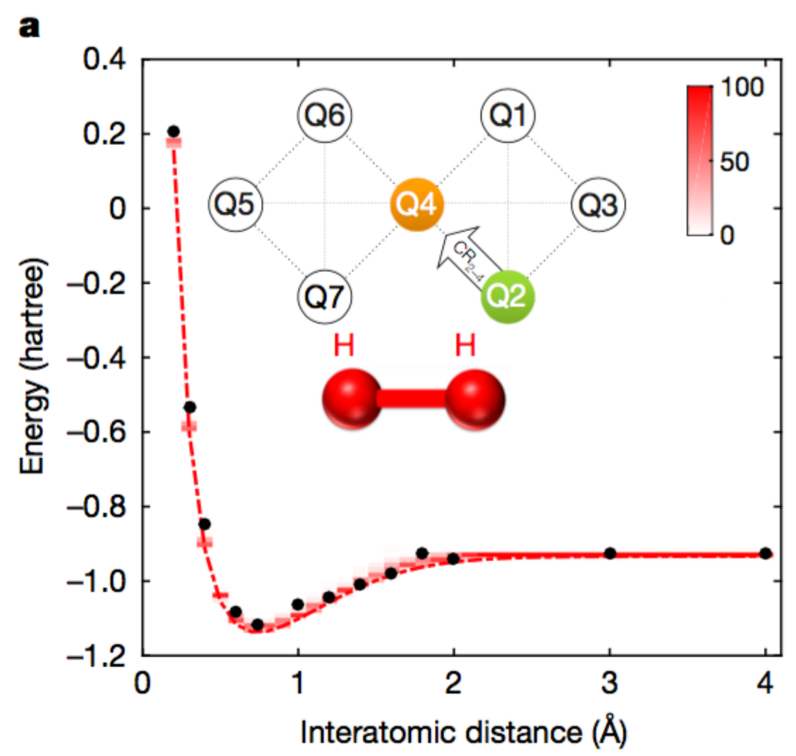
VQE: variational quantum eigensolver

A. Peruzzo, J. McClean, et al., Nature Communications 5, 4213 (2014).

$$H = \sum_{\nu\mu} J_{\nu\mu} c_\nu^\dagger c_\mu + \sum_{ijkl} J_{ijkl} c_i^\dagger c_j^\dagger c_k c_l$$

$$\downarrow$$

$$H = \sum_i h_i \sigma_i + \sum_{ij} h_{ij} \sigma_i \otimes \sigma_j + \ldots$$



Kandala, Mezzacapo *et al*, Nature **549** 242 (2017)

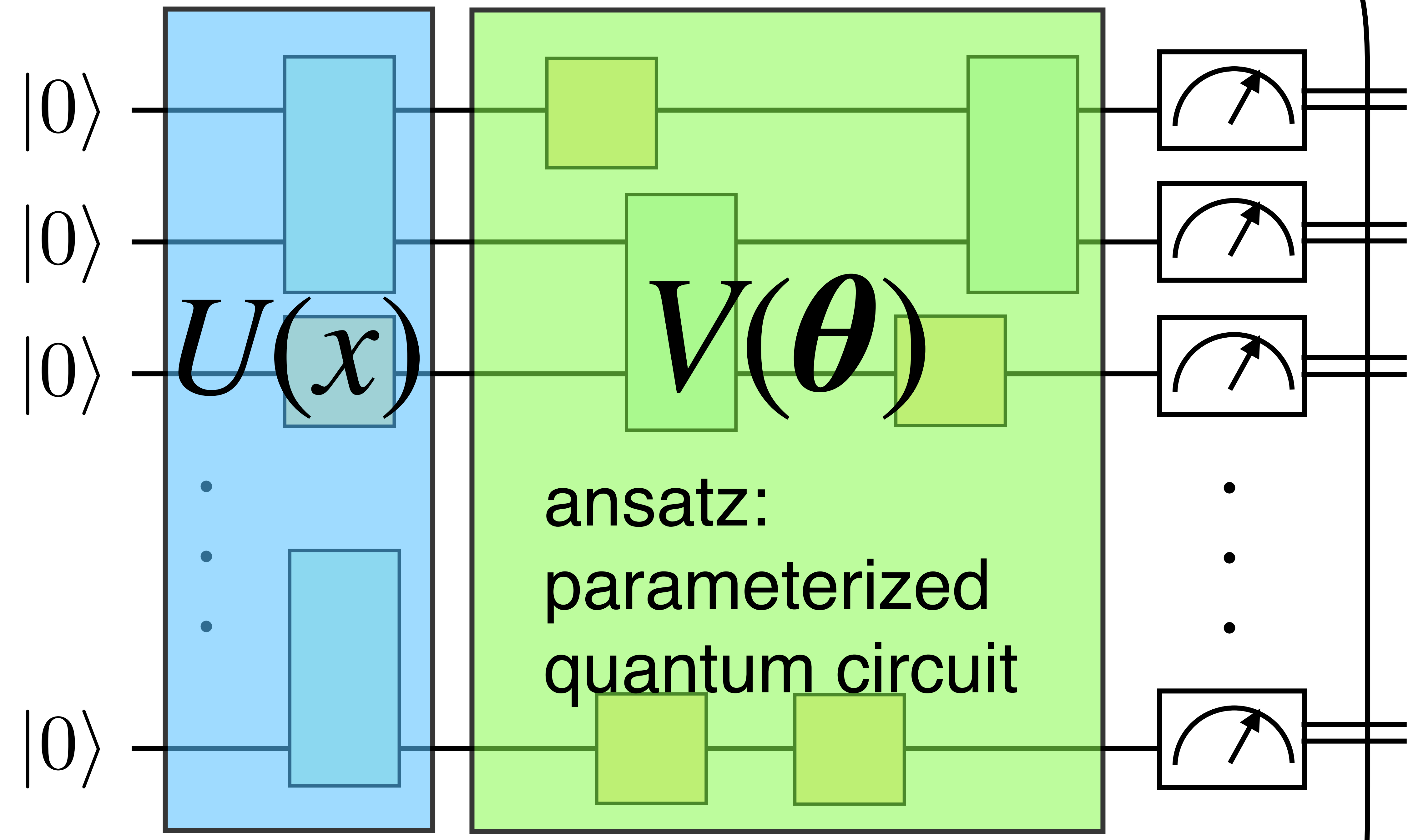
Quantum Circuit Learning

Quantum circuit learning: supervised machine learning with trainable parameterized quantum circuits

K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii “Quantum Circuit Learning”, Phys. Rev. A **98**, 032309 (2018).
E. Farhi and H. Neven, arXiv:1802.06002. “Classification with Quantum Neural Networks on Near Term Processors”

Quantum Computer

quantum feature map

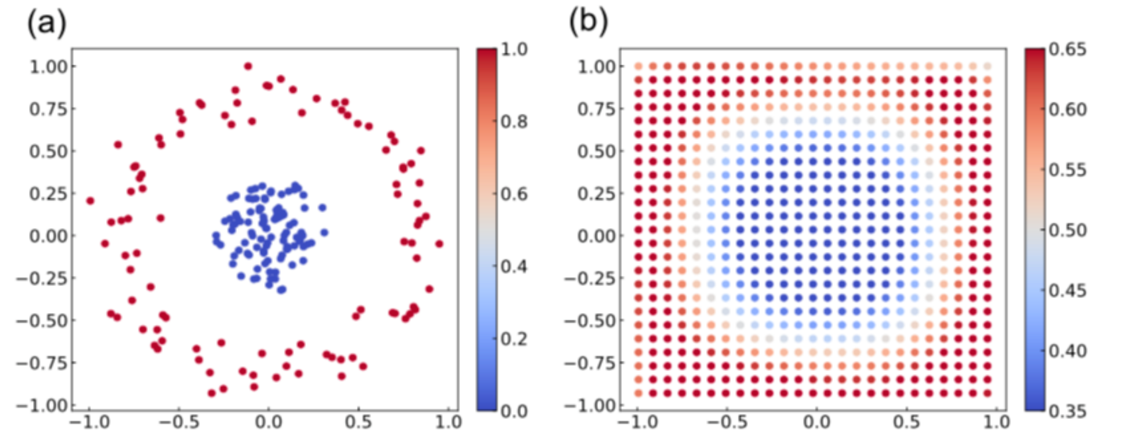


$$|\psi(x,\boldsymbol{\theta})\rangle = V(\boldsymbol{\theta})U(x)|0^n\rangle$$

$$f(x,\boldsymbol{\theta}) = \langle \psi(x,\boldsymbol{\theta})|M|\psi(x,\boldsymbol{\theta})\rangle = \mathrm{Tr}[M(\theta)\rho(x)]$$

$$M(\theta) = V^\dagger(\theta)MV(\theta)$$

$$L(\boldsymbol{\theta}) = \sum_i \left(y_i - f(x_i,\boldsymbol{\theta})\right)^2$$



parameter-shift rule:

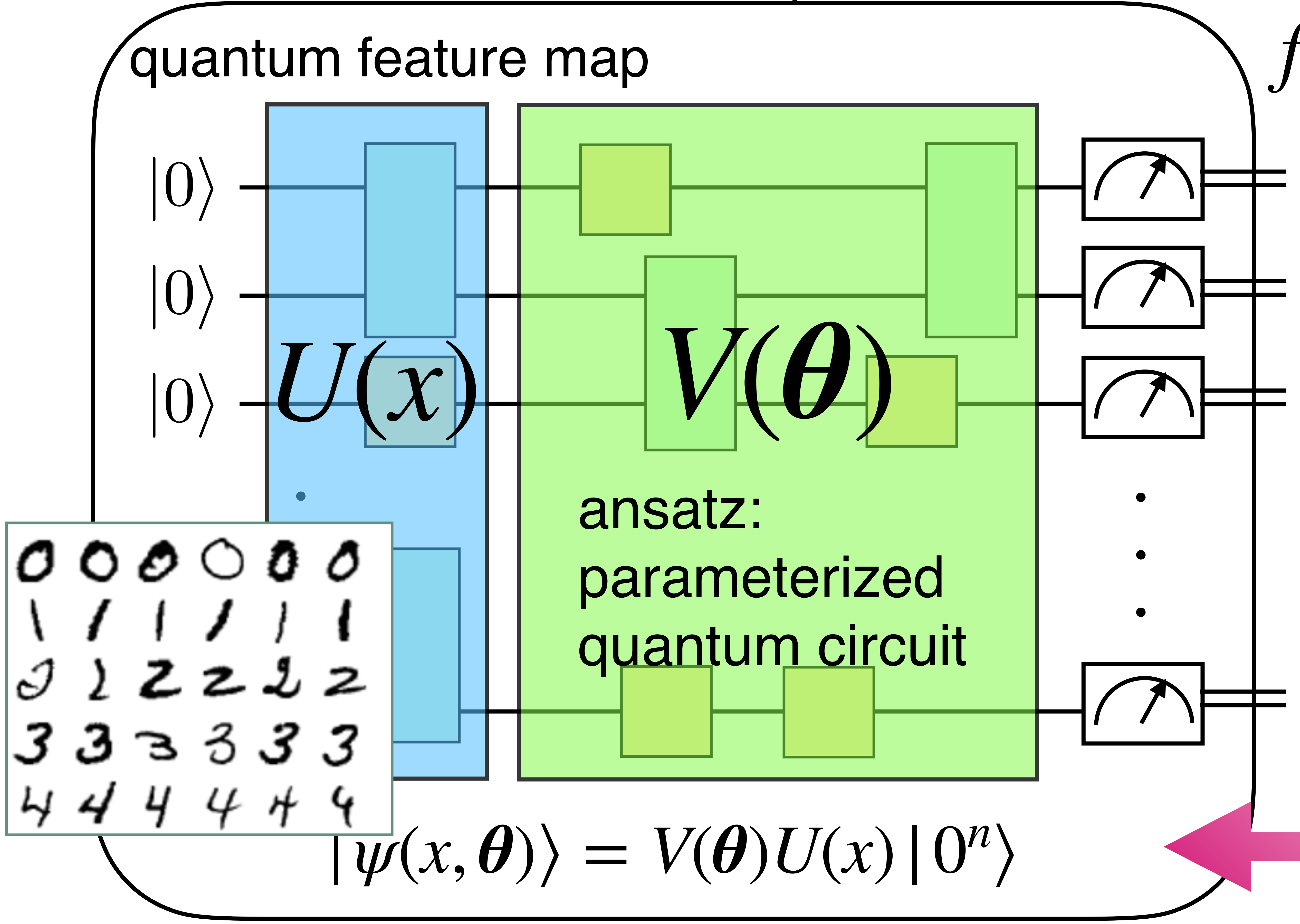
$$\frac{\partial}{\partial\theta_i}y(x,\boldsymbol{\theta}) = \frac{1}{2}\left[\left\langle O\left(\boldsymbol{\theta}+\frac{\pi}{2}\boldsymbol{e}_i\right)\right\rangle - \left\langle O\left(\boldsymbol{\theta}-\frac{\pi}{2}\boldsymbol{e}_i\right)\right\rangle\right]$$

Quantum Circuit Learning

Quantum circuit learning: supervised machine learning with trainable parameterized quantum circuits

K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii “Quantum Circuit Learning”, Phys. Rev. A **98**, 032309 (2018).
E. Farhi and H. Neven, arXiv:1802.06002. “Classification with Quantum Neural Networks on Near Term Processors”

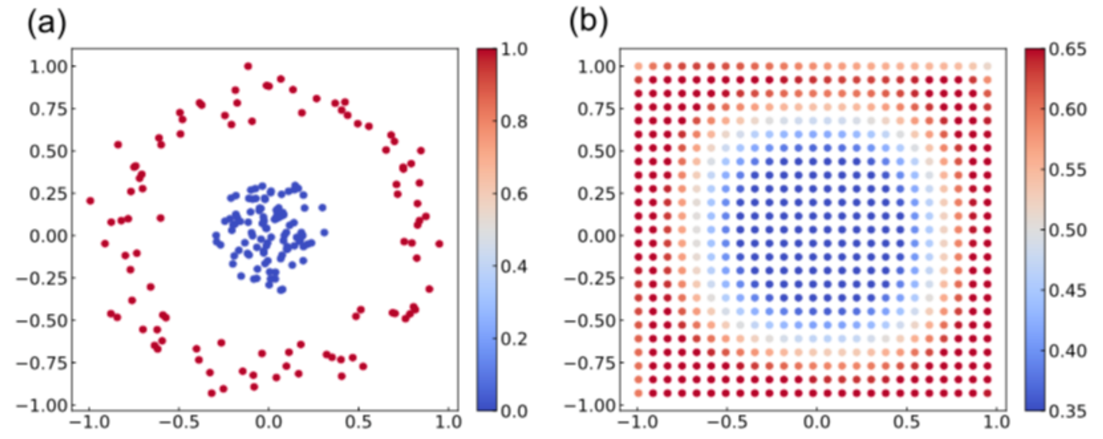
Quantum Computer



$$f(x,\boldsymbol{\theta}) = \langle \psi(x,\boldsymbol{\theta})|M|\psi(x,\boldsymbol{\theta})\rangle$$
$$= \mathrm{Tr}[M(\theta)\rho(x)]$$

$$M(\theta) = V^\dagger(\theta) M V(\theta)$$

$$L(\boldsymbol{\theta}) = \sum_i \left(y_i - f(x_i,\boldsymbol{\theta})\right)^2$$



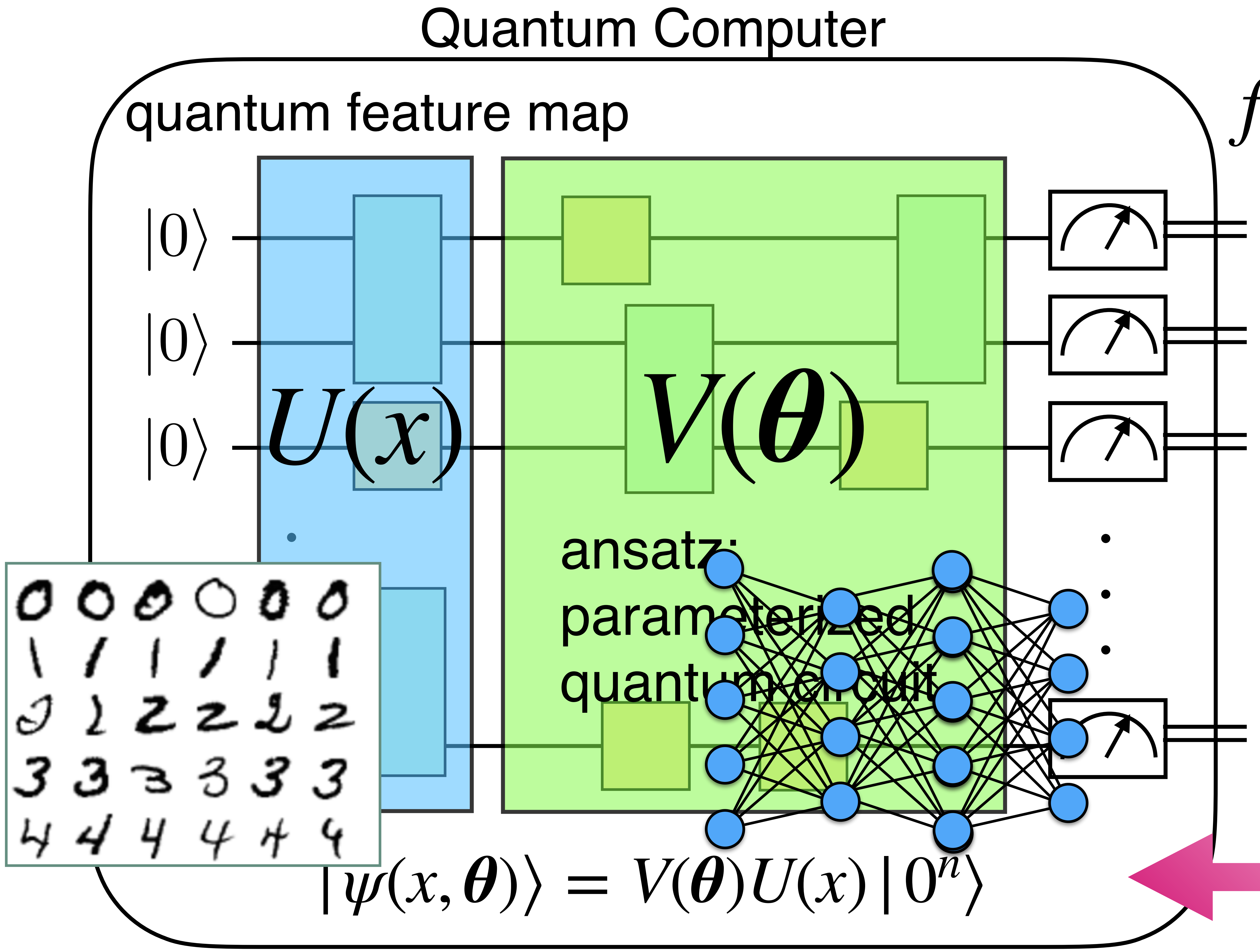
parameter-shift rule:

$$\frac{\partial}{\partial\theta_i} y(x,\boldsymbol{\theta}) = \frac{1}{2}\left[\left\langle O\left(\boldsymbol{\theta} + \frac{\pi}{2}\boldsymbol{e}_i\right)\right\rangle - \left\langle O\left(\boldsymbol{\theta} - \frac{\pi}{2}\boldsymbol{e}_i\right)\right\rangle\right]$$

Quantum Circuit Learning

Quantum circuit learning: supervised machine learning with trainable parameterized quantum circuits

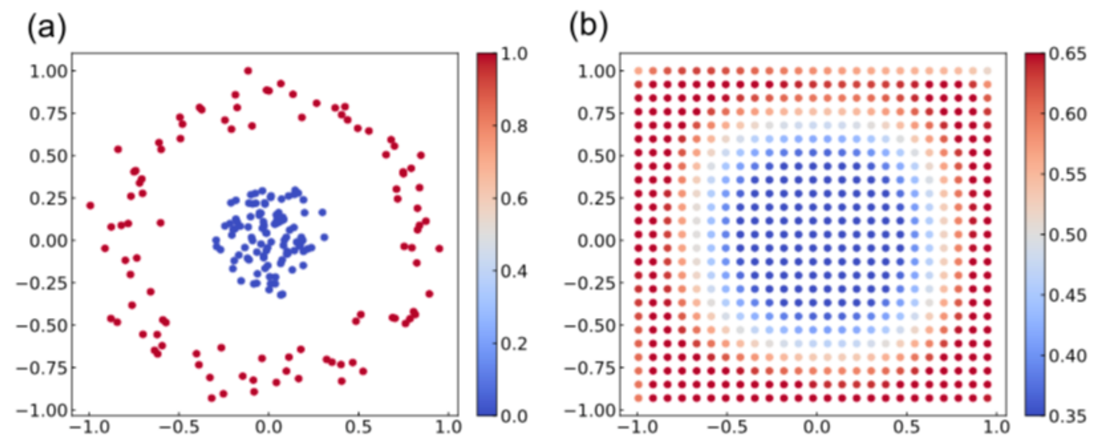
K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii “Quantum Circuit Learning”, Phys. Rev. A **98**, 032309 (2018).
E. Farhi and H. Neven, arXiv:1802.06002. “Classification with Quantum Neural Networks on Near Term Processors”



$$f(x,\boldsymbol{\theta}) = \langle \psi(x,\boldsymbol{\theta}) | M | \psi(x,\boldsymbol{\theta}) \rangle$$
$$= \mathrm{Tr}[M(\theta)\rho(x)]$$

$$M(\theta) = V^{\dagger}(\theta) M V(\theta)$$

$$L(\boldsymbol{\theta}) = \sum_i \left(y_i - f(x_i,\boldsymbol{\theta})\right)^2$$



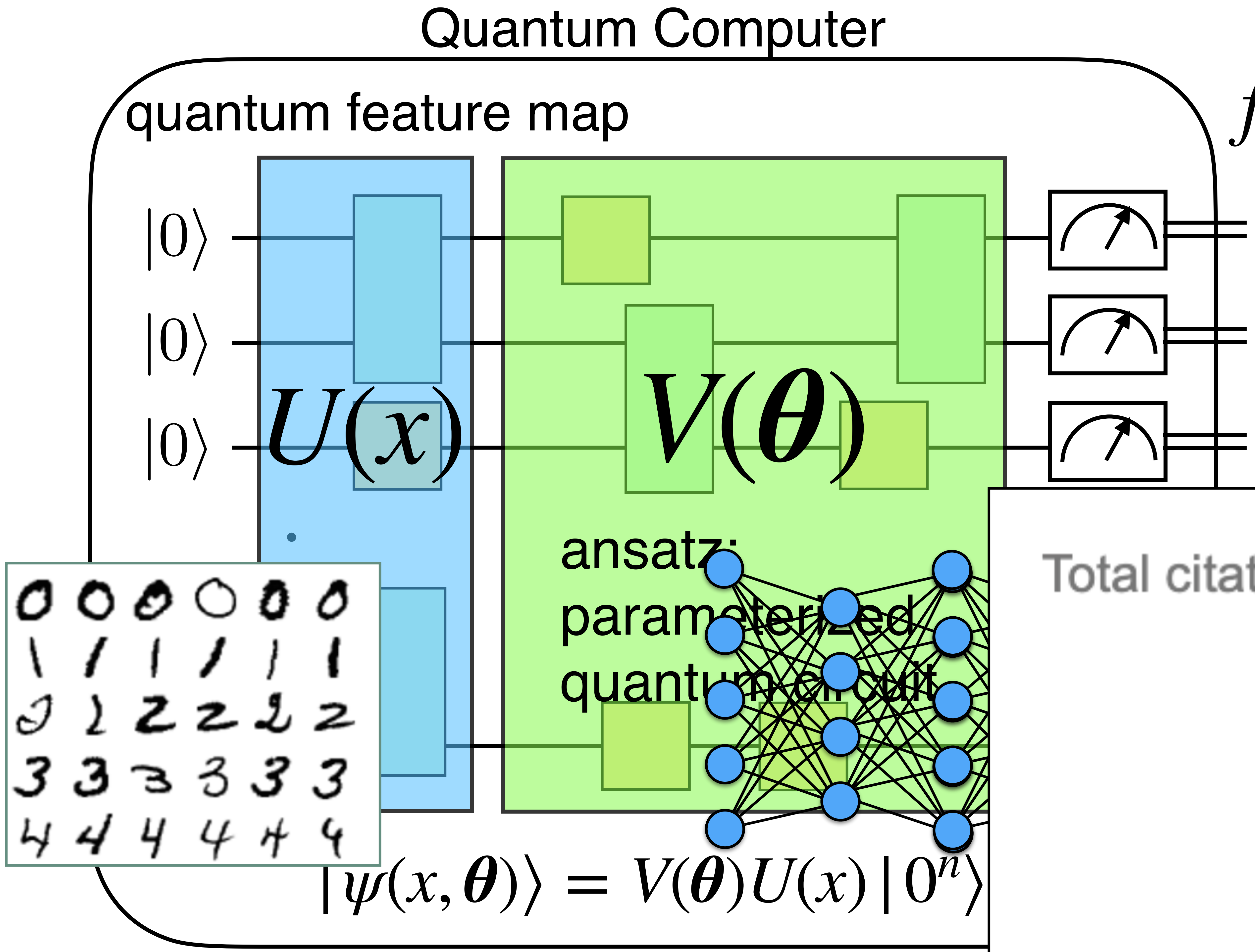
parameter-shift rule:

$$\frac{\partial}{\partial \theta_i} y(x,\boldsymbol{\theta}) = \frac{1}{2}\left[\left\langle O\left(\boldsymbol{\theta} + \frac{\pi}{2}\boldsymbol{e}_i\right)\right\rangle - \left\langle O\left(\boldsymbol{\theta} - \frac{\pi}{2}\boldsymbol{e}_i\right)\right\rangle\right]$$

Quantum Circuit Learning

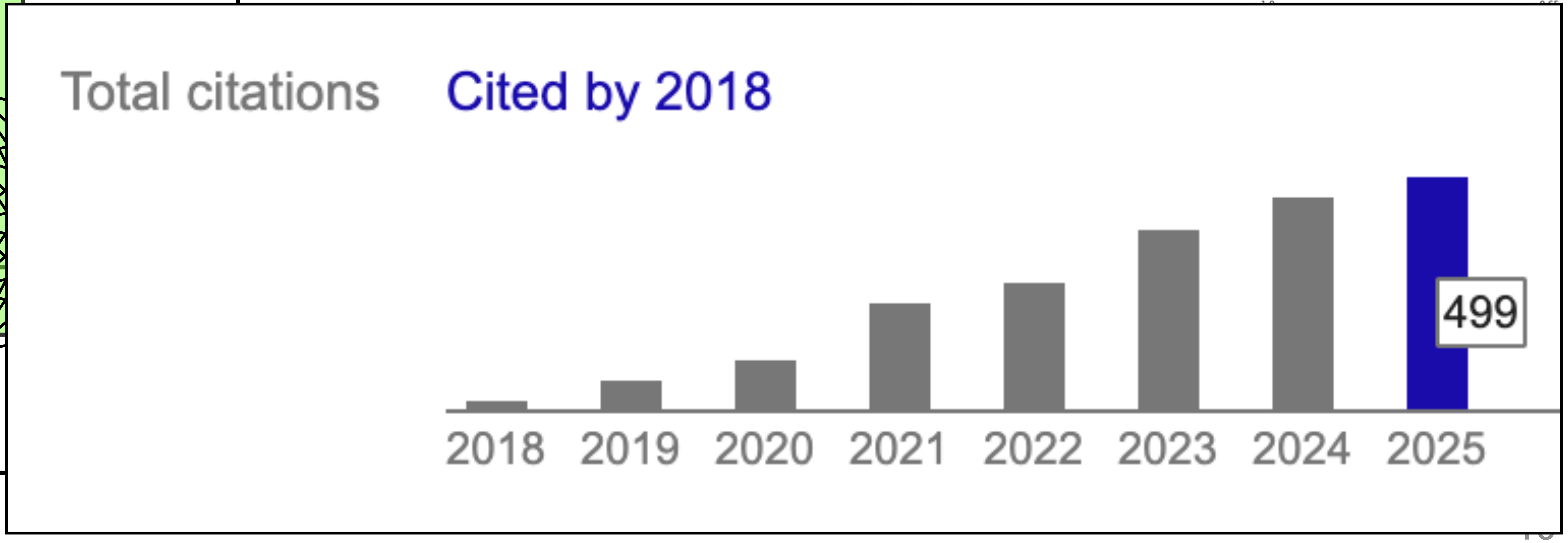
Quantum circuit learning: supervised machine learning with trainable parameterized quantum circuits

K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii “Quantum Circuit Learning”, Phys. Rev. A **98**, 032309 (2018).
E. Farhi and H. Neven, arXiv:1802.06002. “Classification with Quantum Neural Networks on Near Term Processors”



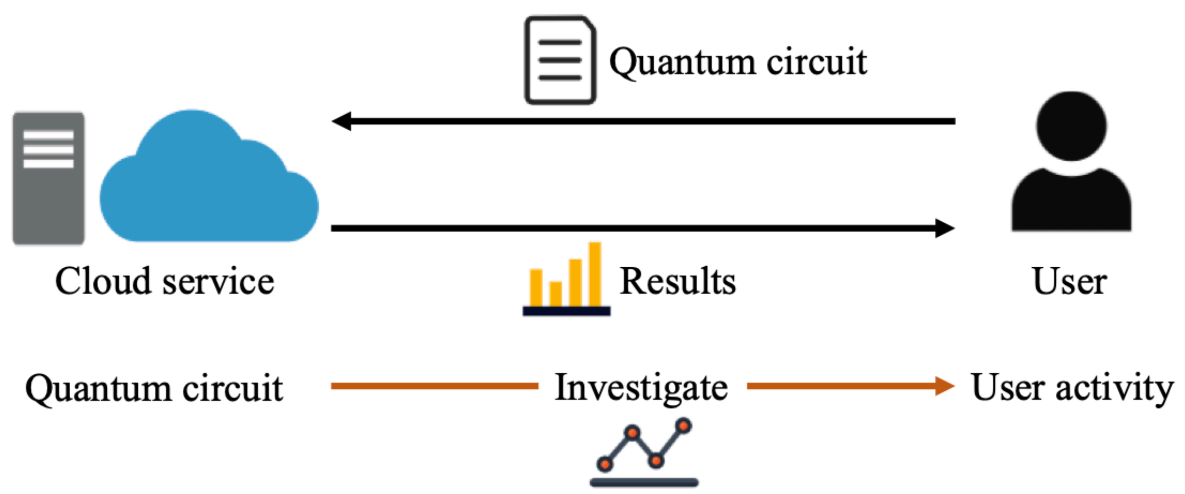
$$f(x,\boldsymbol{\theta}) = \langle \psi(x,\boldsymbol{\theta}) | M | \psi(x,\boldsymbol{\theta}) \rangle$$
$$= \mathrm{Tr}[M(\theta)\rho(x)]$$

$$M(\theta) = V^{\dagger}(\theta) M V(\theta)$$



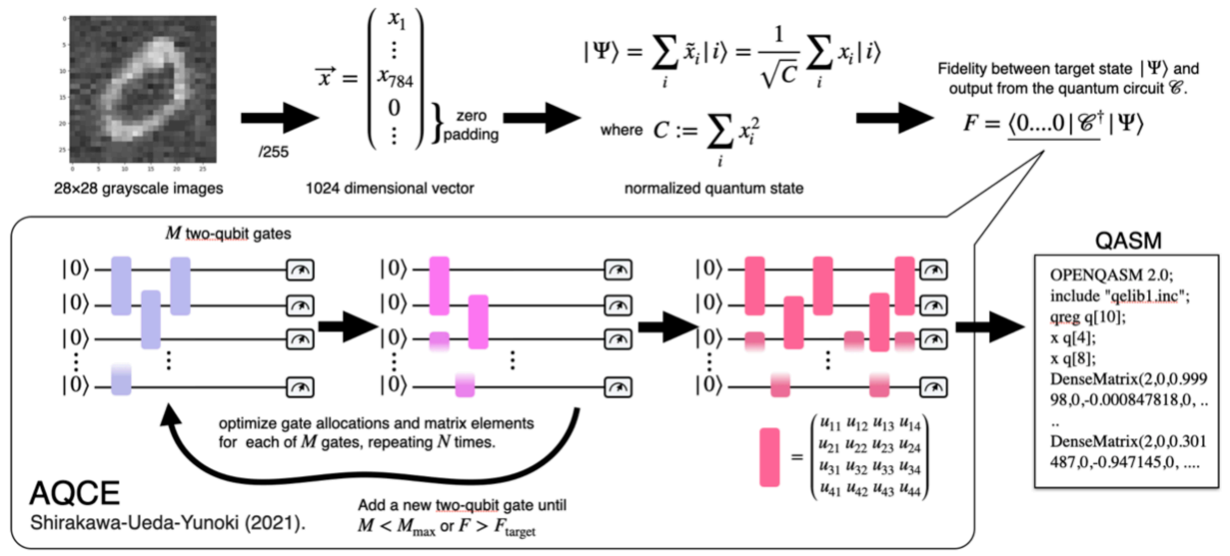
Our activities on Quantum Machine Learning

Quantum Datasets



VQE-generated dataset

Nakayama-Mitarai-Placidi-**KF**
Physical Review Research (2025).

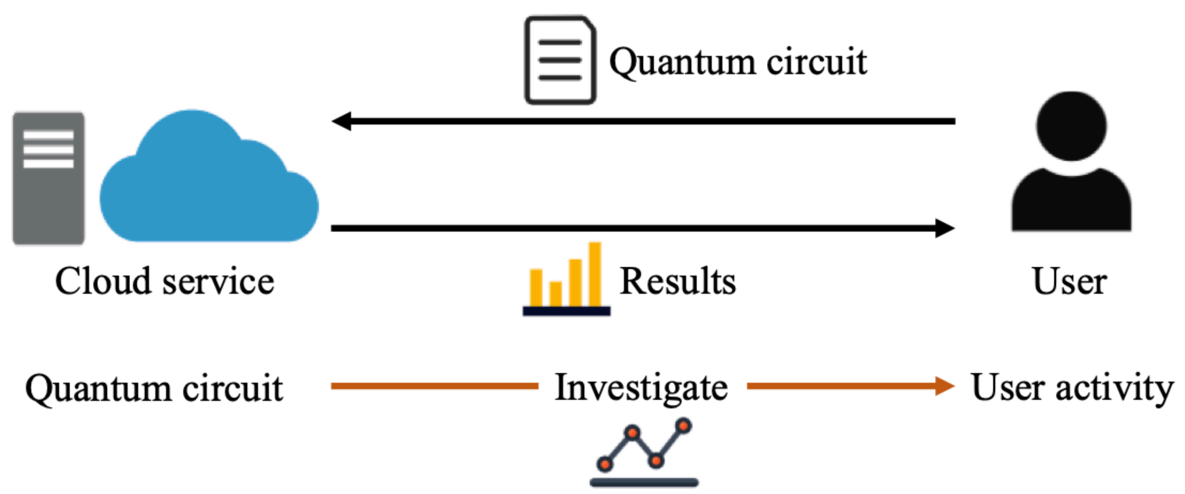


MNISQ dataset

Placidi *et al.*, arXiv:2306.16627.

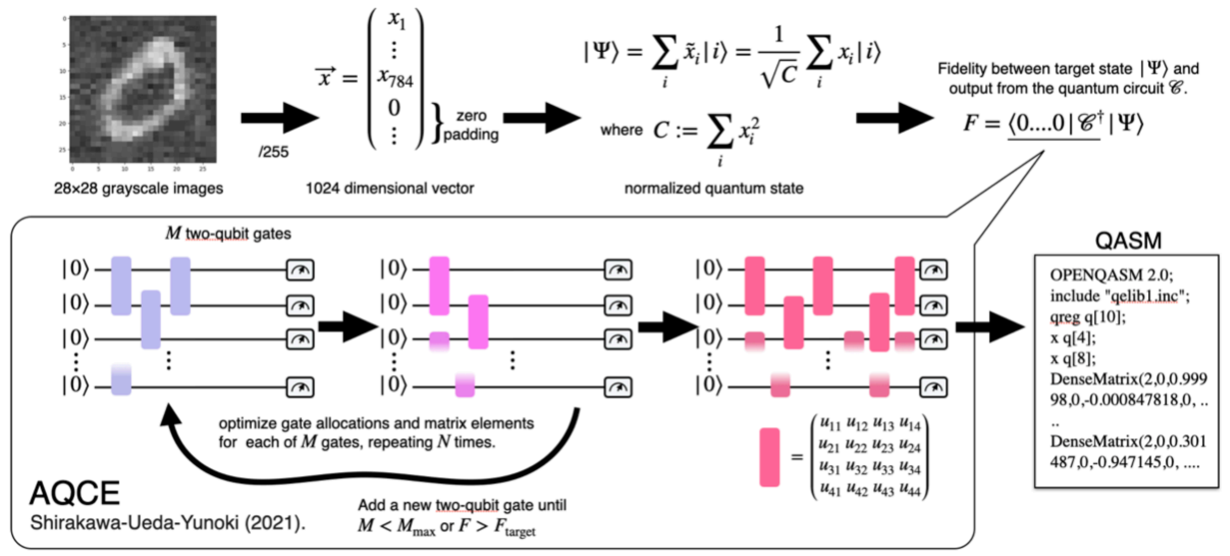
Our activities on Quantum Machine Learning

Quantum Datasets



VQE-generated dataset

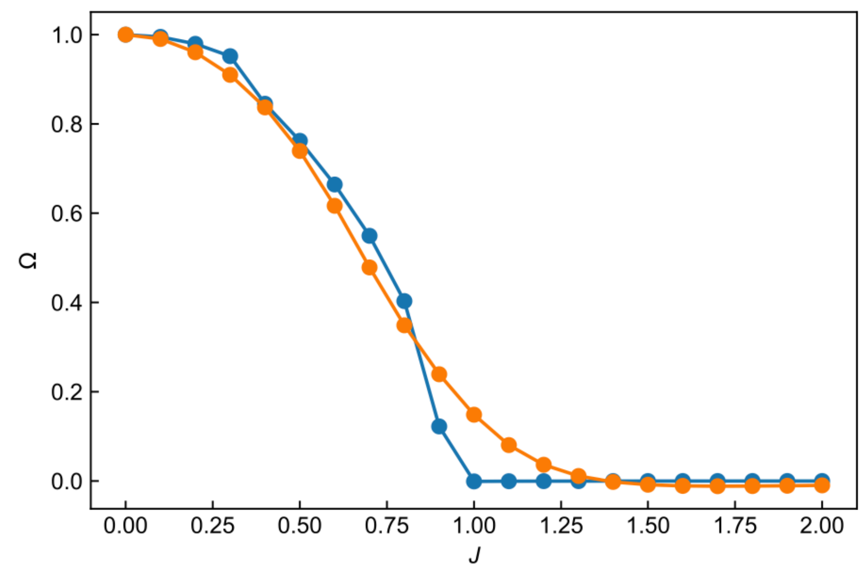
Nakayama-Mitarai-Placidi-**KF**
Physical Review Research (2025).

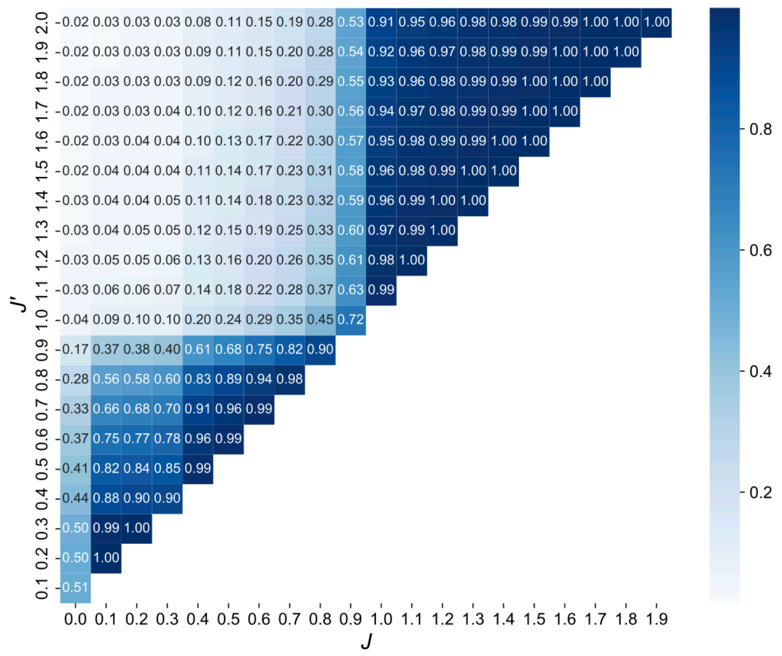


MNISQ dataset

Placidi *et al.*, arXiv:2306.16627.

QML for detecting quantum phase

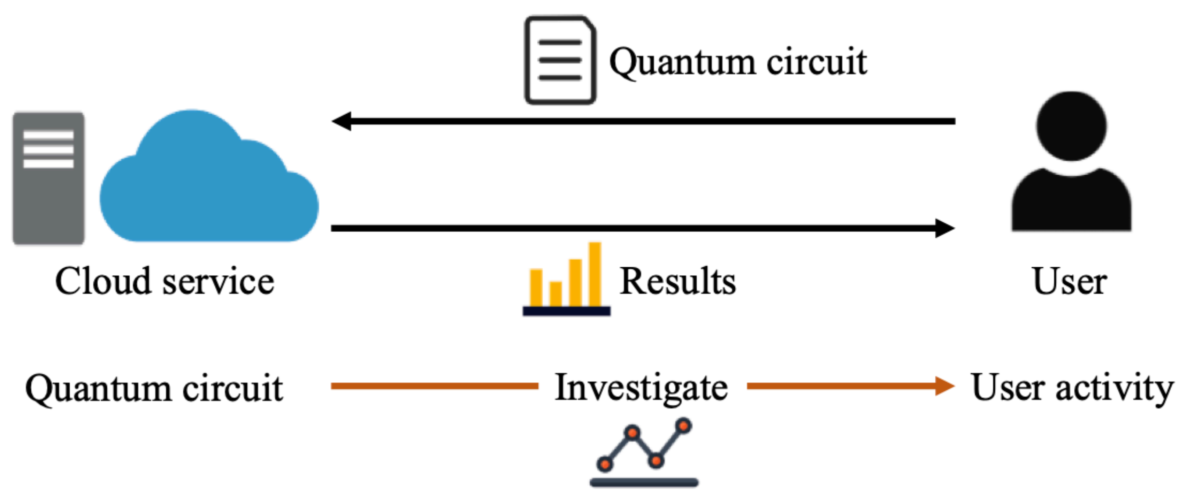




Okada-Osaki-Mitarai-**KF** Physical Review Research (2023)

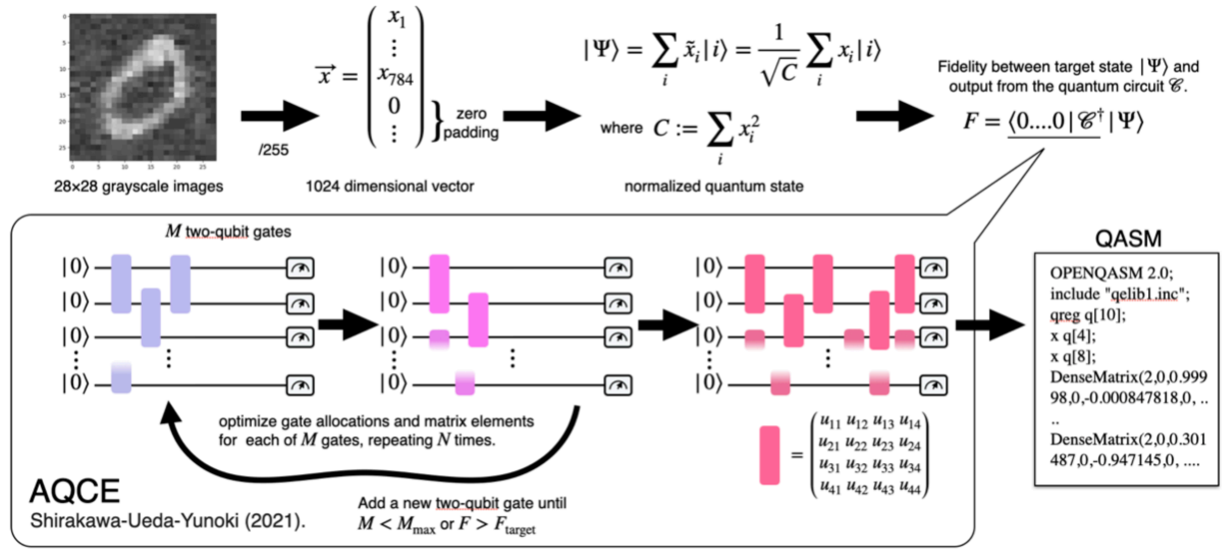
Our activities on Quantum Machine Learning

Quantum Datasets



VQE-generated dataset

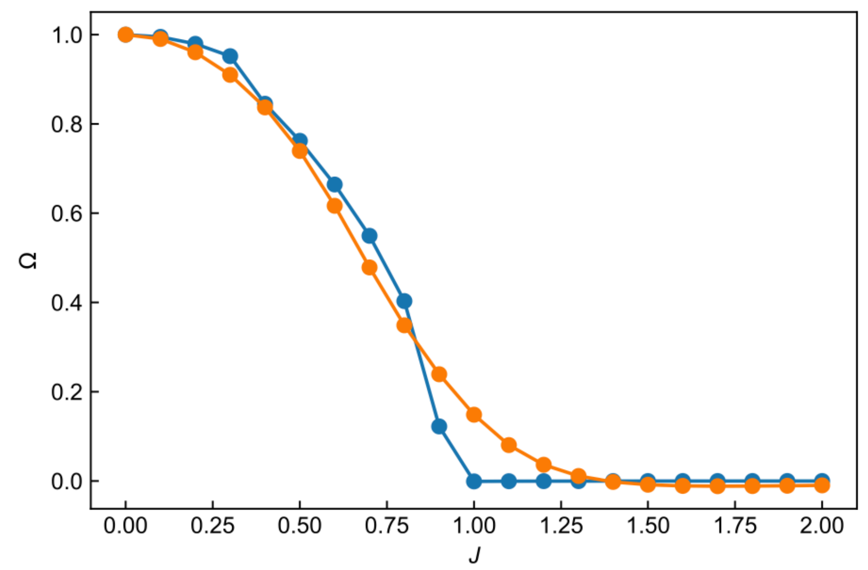
Nakayama-Mitarai-Placidi-**KF**
Physical Review Research (2025).

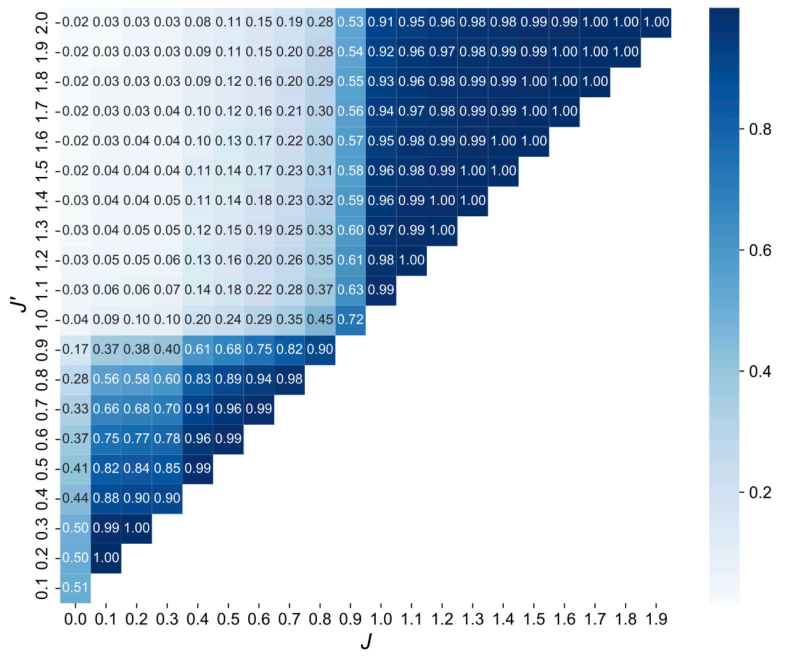


MNISQ dataset

Placidi *et al.*, arXiv:2306.16627.

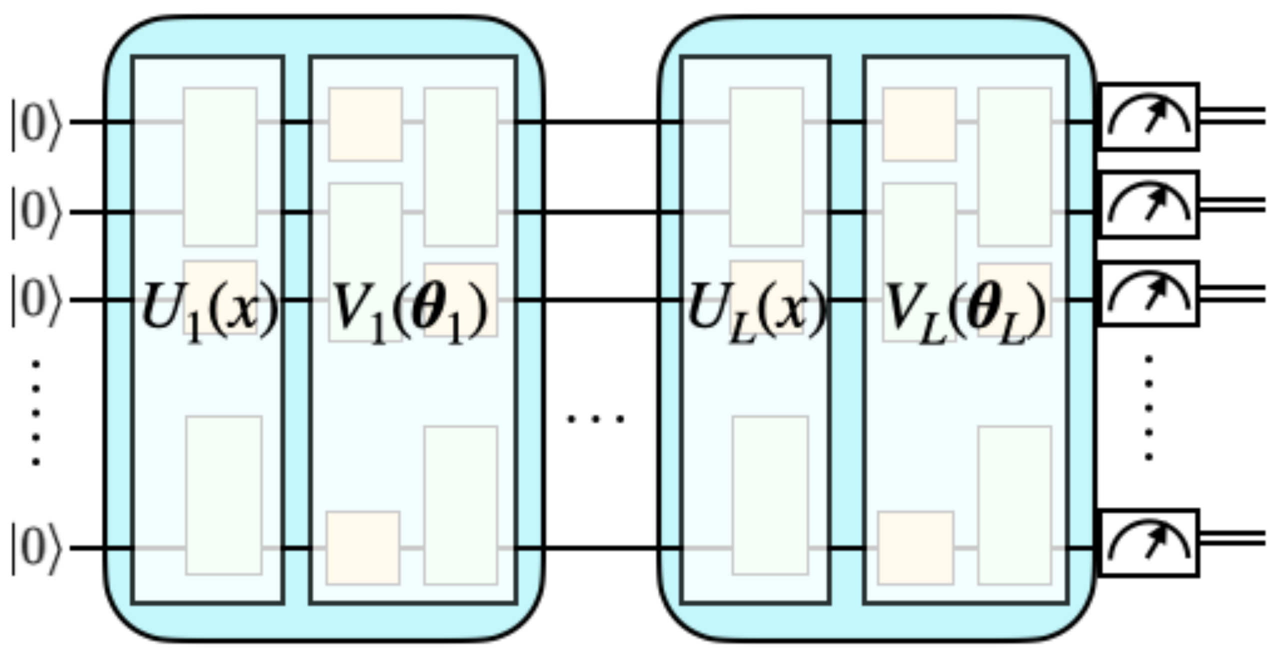
QML for detecting quantum phase

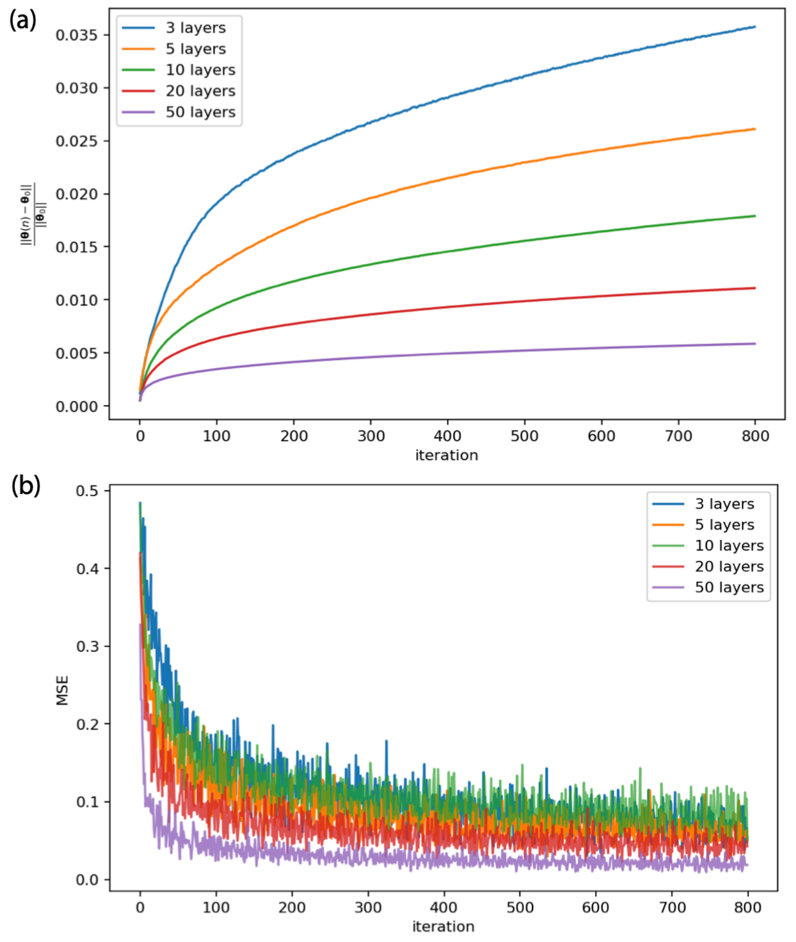




Okada-Osaki-Mitarai-**KF** Physical Review Research (2023)

Quantum tangent kernel



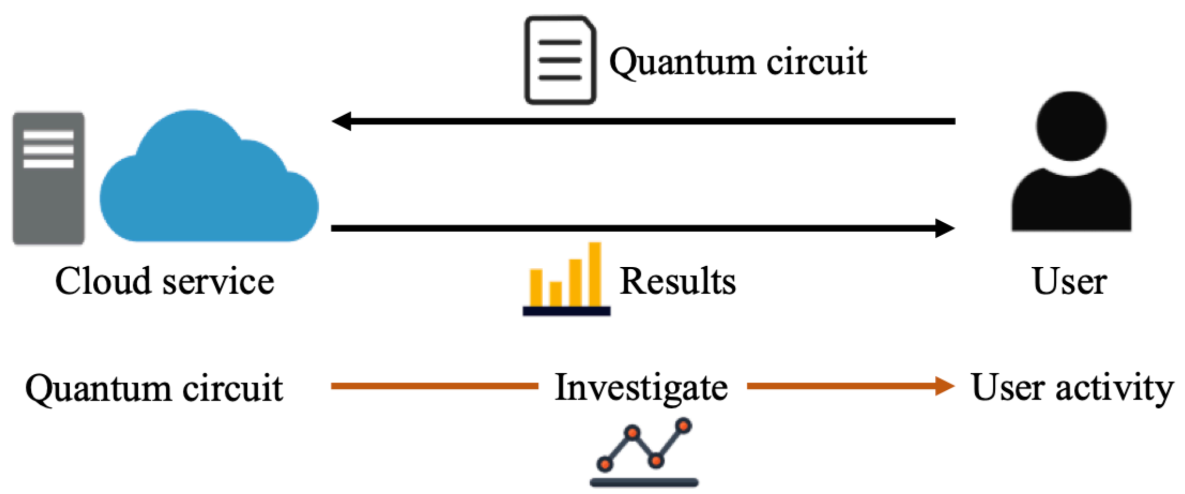


Overparameterization by data re-upload ansatz.

Shirai-Kubo-Mitarai-**KF** Physical Review Research (2024)

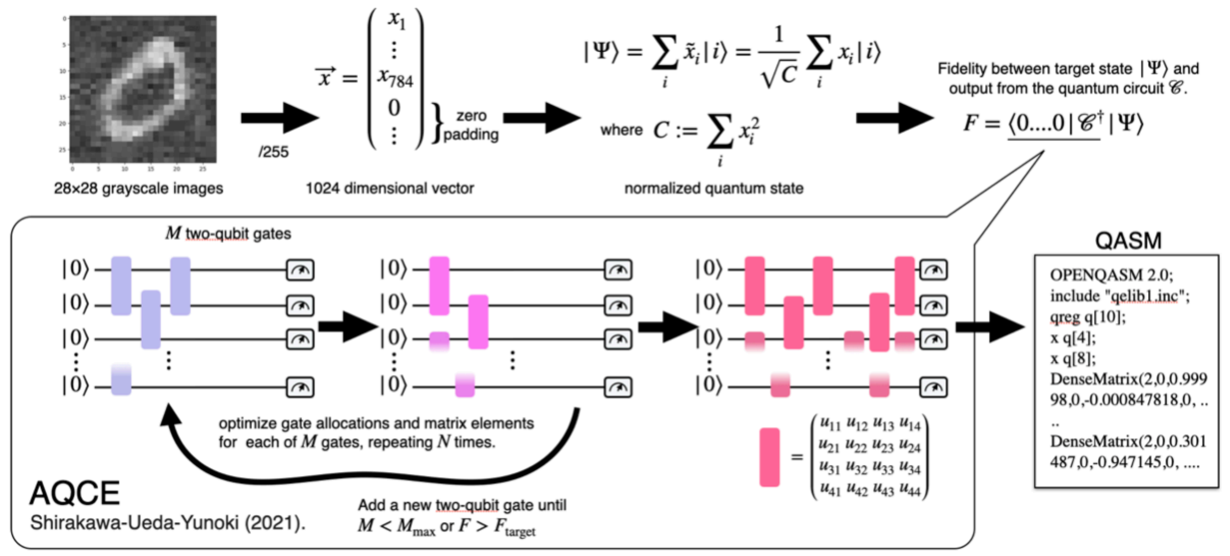
Our activities on Quantum Machine Learning

Quantum Datasets



VQE-generated dataset

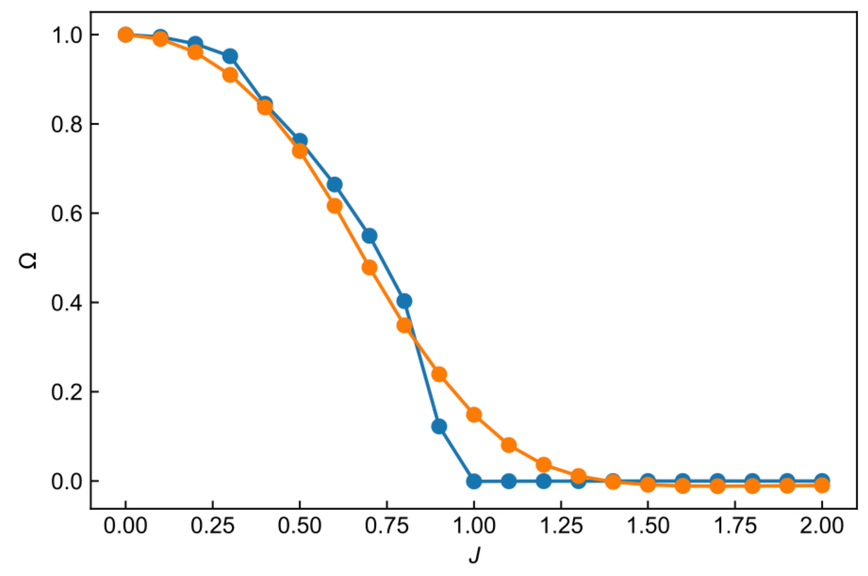
Nakayama-Mitarai-Placidi-**KF**
Physical Review Research (2025).

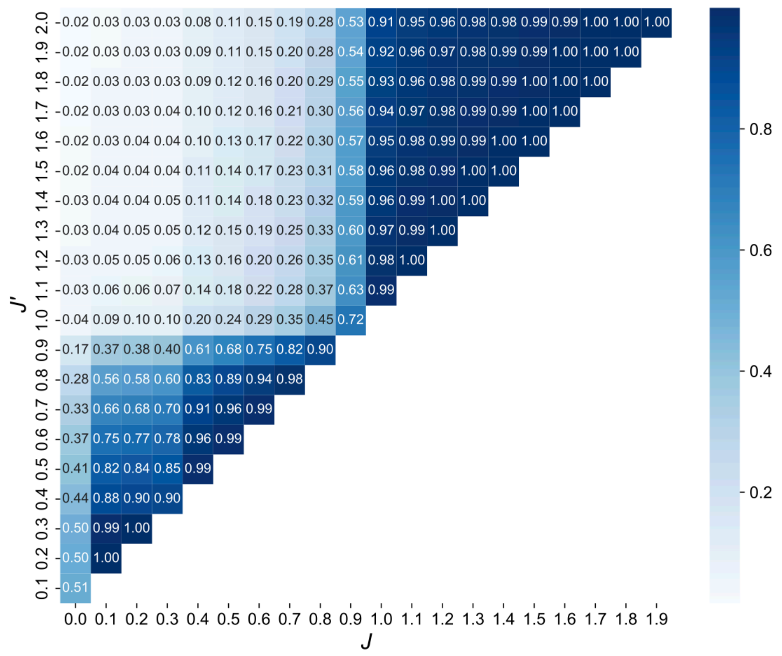


MNISQ dataset

Placidi *et al.*, arXiv:2306.16627.

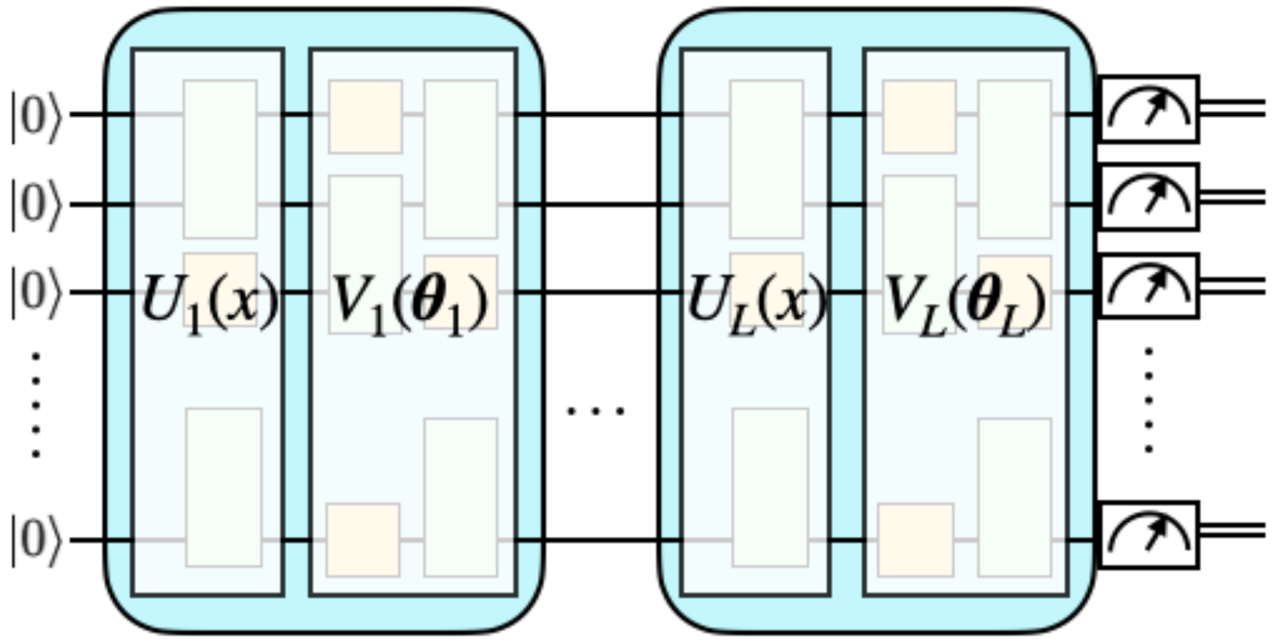
QML for detecting quantum phase

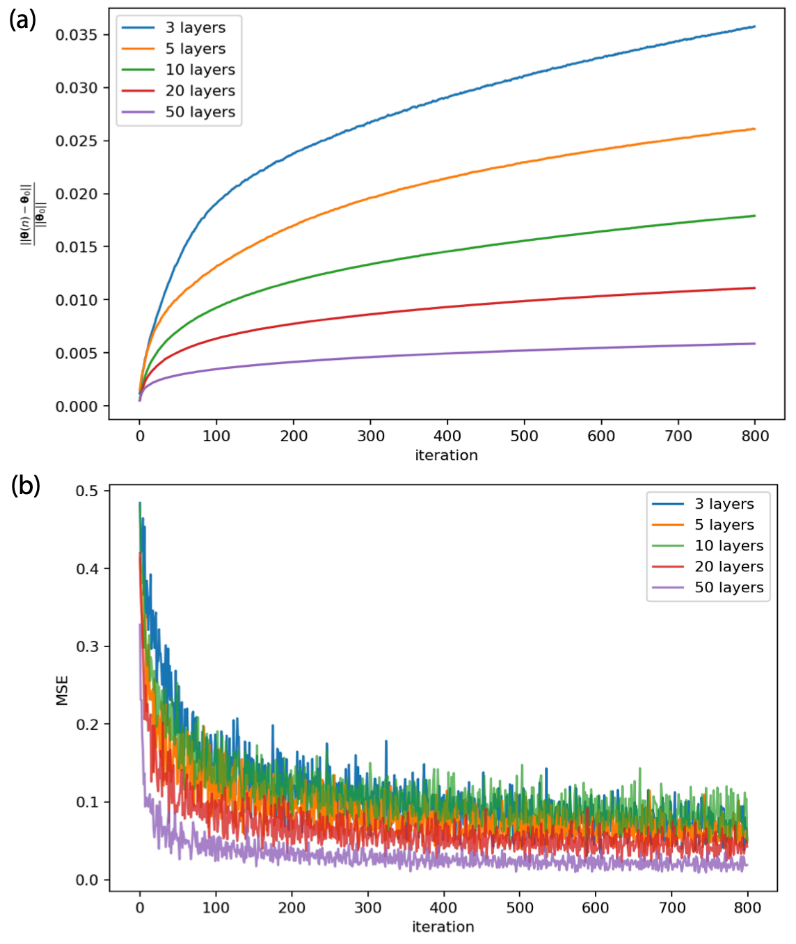




Okada-Osaki-Mitarai-**KF** Physical Review Research (2023)

Quantum tangent kernel

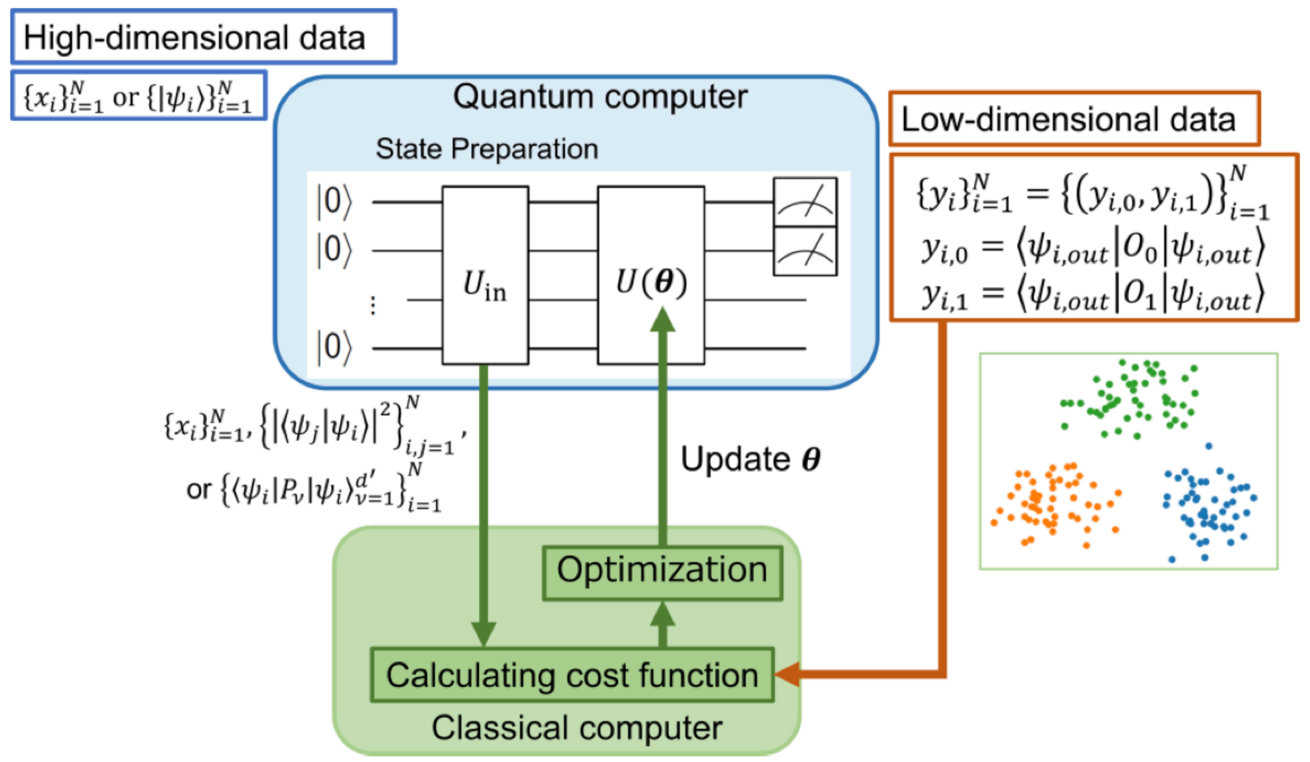


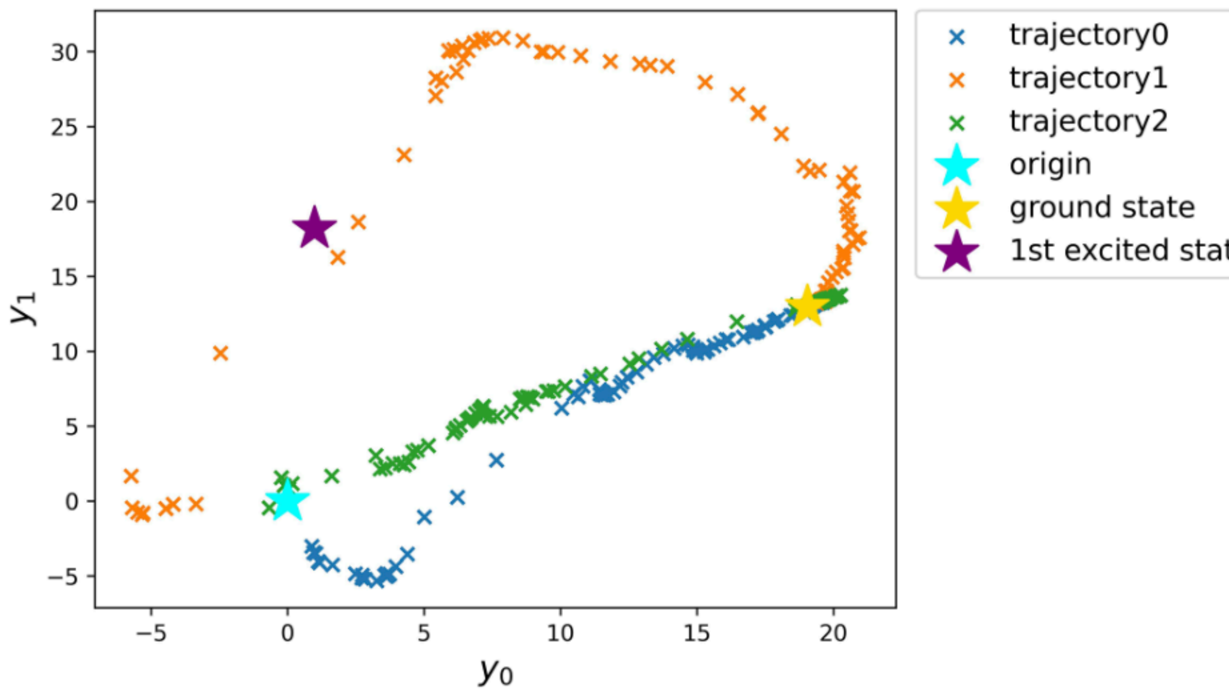


Overparameterization by data re-upload ansatz.

Shirai-Kubo-Mitarai-**KF** Physical Review Research (2024)

Low-dimensional visualization of Quantum States





Kawase-Mitarai-**KF,** Physical Review Research (2022)
Kawase-Mitarai-**KF,** Physical Review Research (2024)

What is potential advantages of Quantum Machine Learning?

Witness of Quantum Advantage in QML?

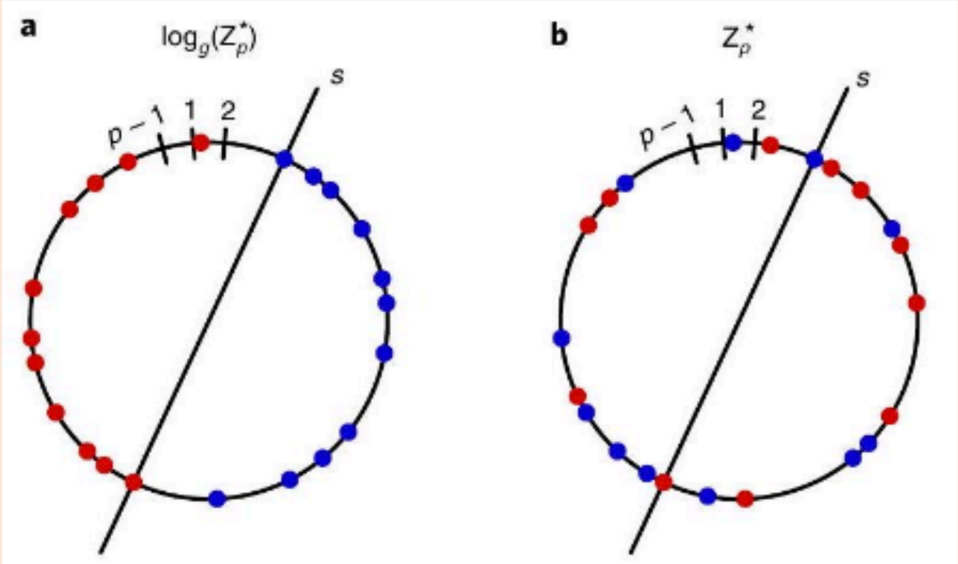
Artificially designed problem set up to be advantageous to quantum computers:

- Artificially designed problem set up to be advantageous to quantum computers:

Label 0 Label 0 Label 1 Label 1 Label 0 Label 0

[6] Huang, HY., Broughton, M., Mohseni, M. et al. Power of data in quantum machine learning. Nat Commun 12, 2631 (2021)

- Dataset generated from discrete logarithmic problem:

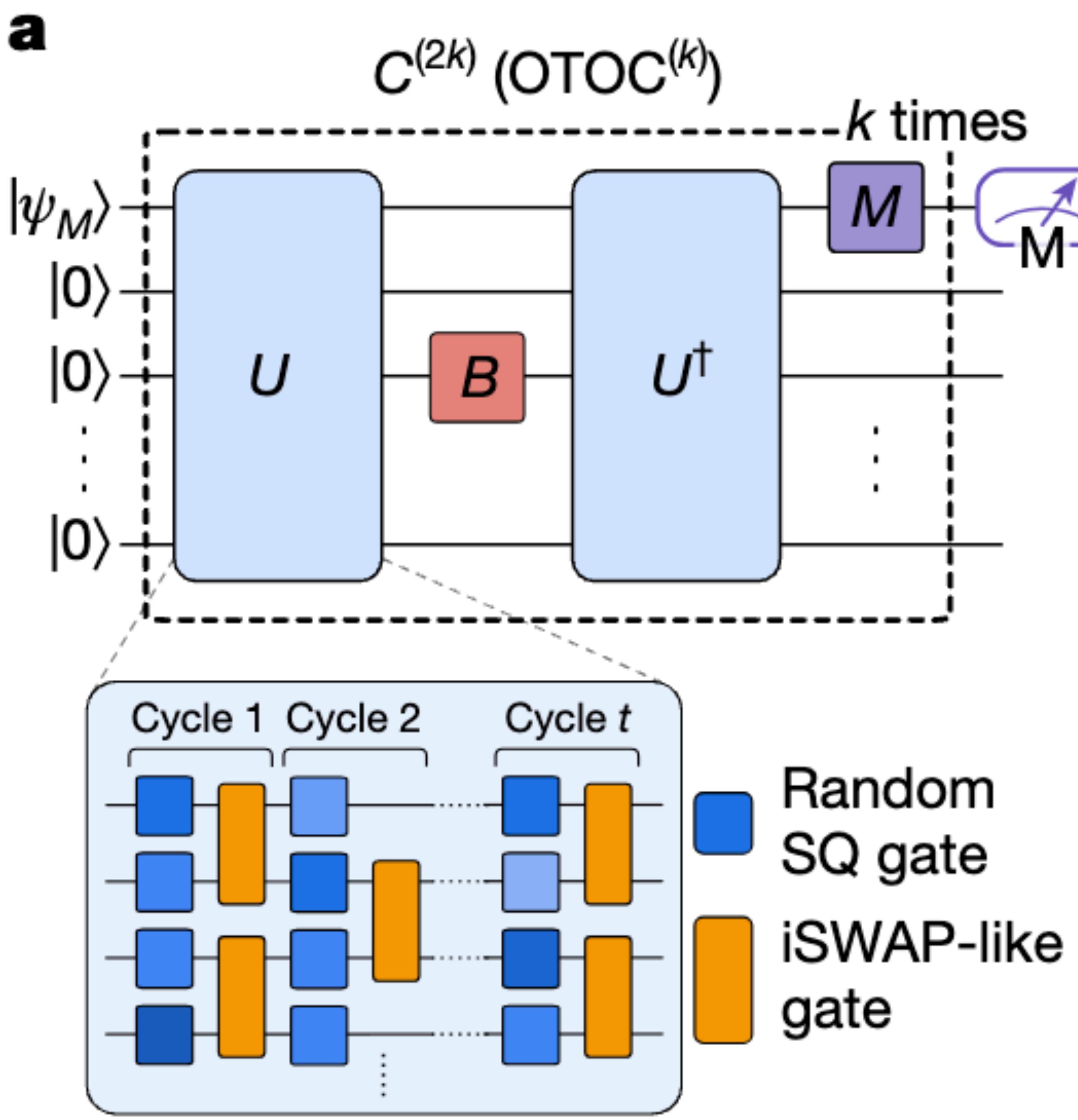


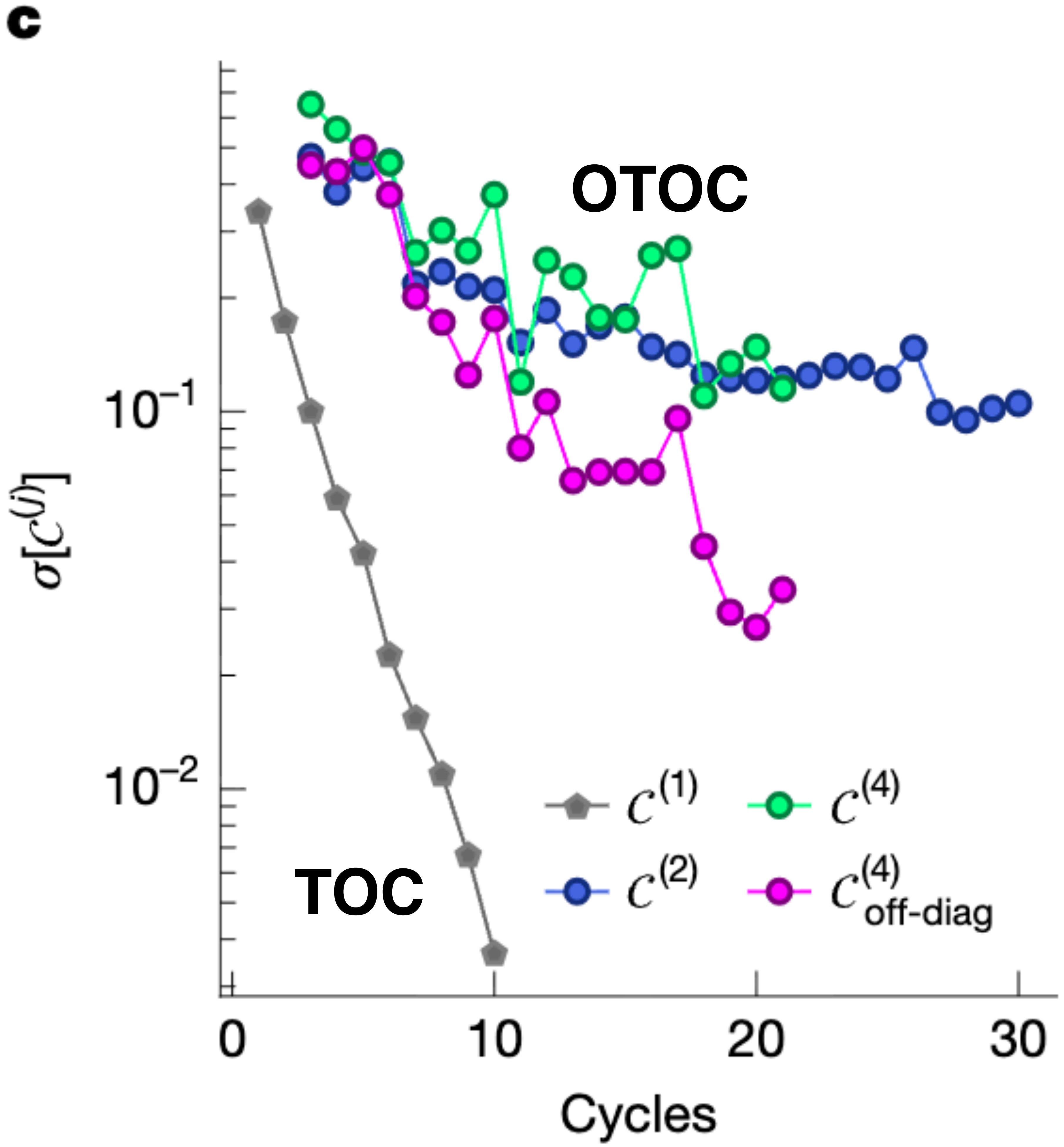
[7] Y. Liu, S. Arunachalam, and K. Temme, A rigorous and robust quantum speed-up in supervised machine learning, Nature Physics 17, 1013 (2021)

⇨ The data sets or problem settings where QML has a “natural” advantage are not well understood.

Verifiable Quantum Advantage at the Edge of Quantum Chaos

Google Quantum AI & collaborators, "*Observation of constructive interference at the edge of quantum ergodicity*" Nature (2025).





OTOC is a useful metric to probe information scrambling in quantum many-body systems, where the signature of Hamiltonian is reflected. →***Quantum Echoes***

Verifiable Quantum Advantage at the edge of Quantum Chaos

Google Quantum AI & collaborators, *"Quantum computation of molecular geometry via many-body nuclear spin echoes", arXiv:2510.19550.*

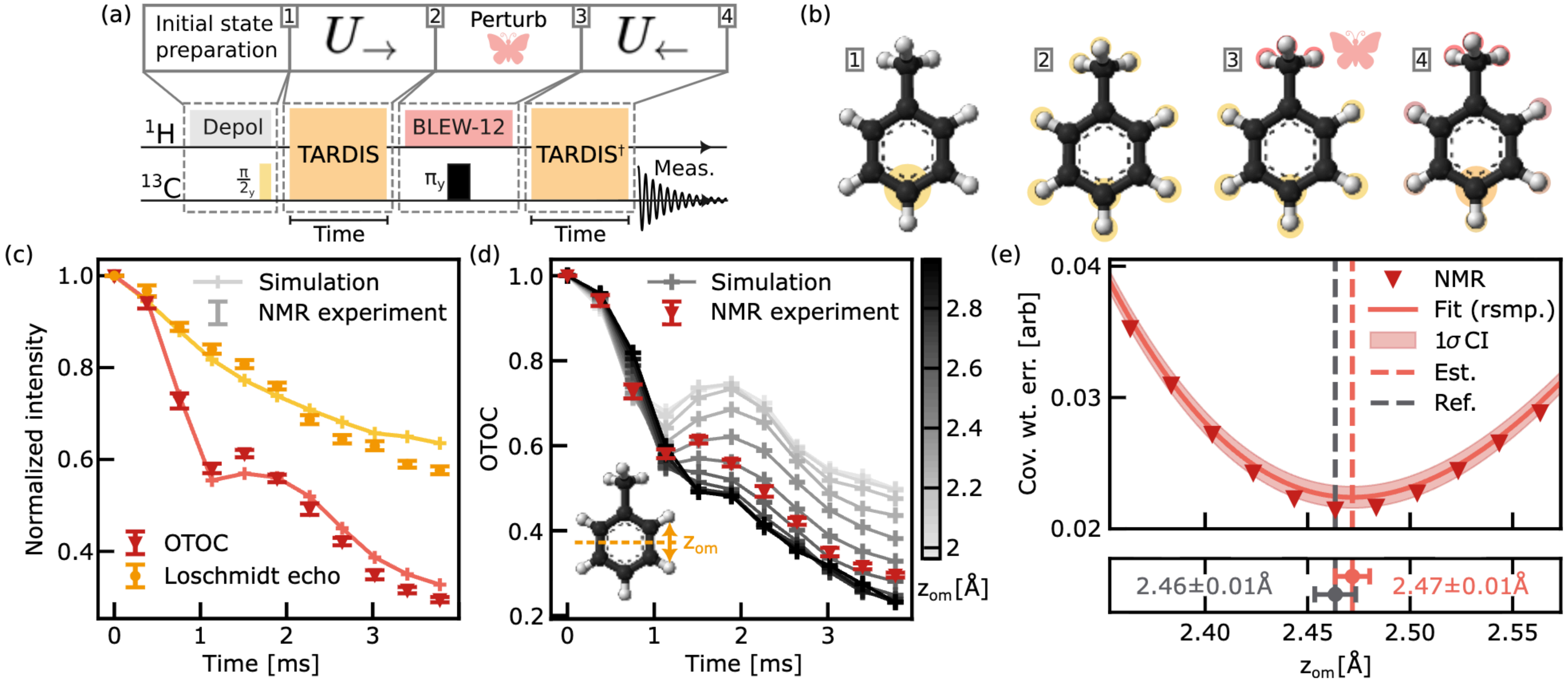


FIG. 2. **Benchmarking the structural sensitivity of out-of-time-ordered correlators (OTOCs) in [$4-{}^{13}$C]-toluene.**

Experimentally (OTOC NMR experiments) supervised learning with quantum computers with trainable parameters.

Verifiable Quantum Advantage at the edge of Quantum Chaos

Google Quantum AI & collaborators, *“Quantum computation of molecular geometry via many-body nuclear spin echoes”, arXiv:2510.19550.*

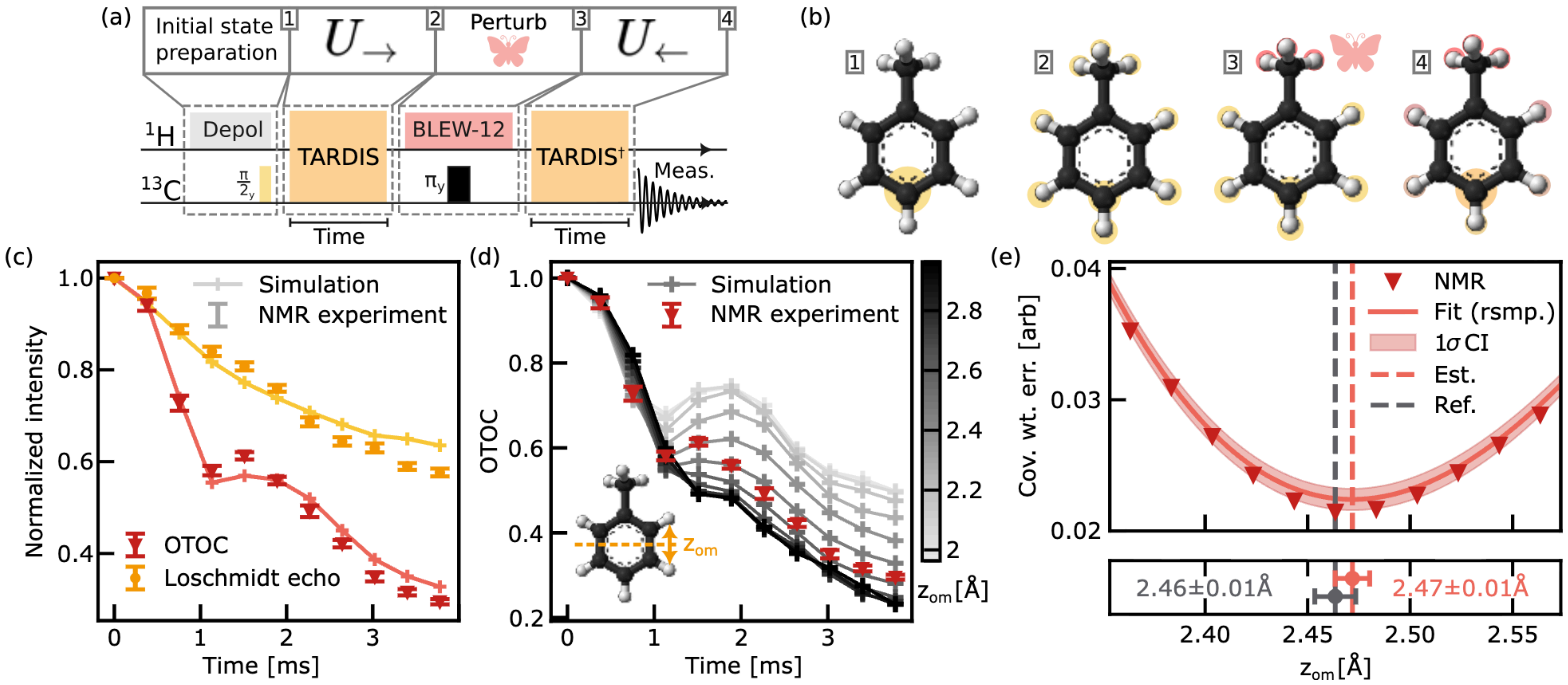


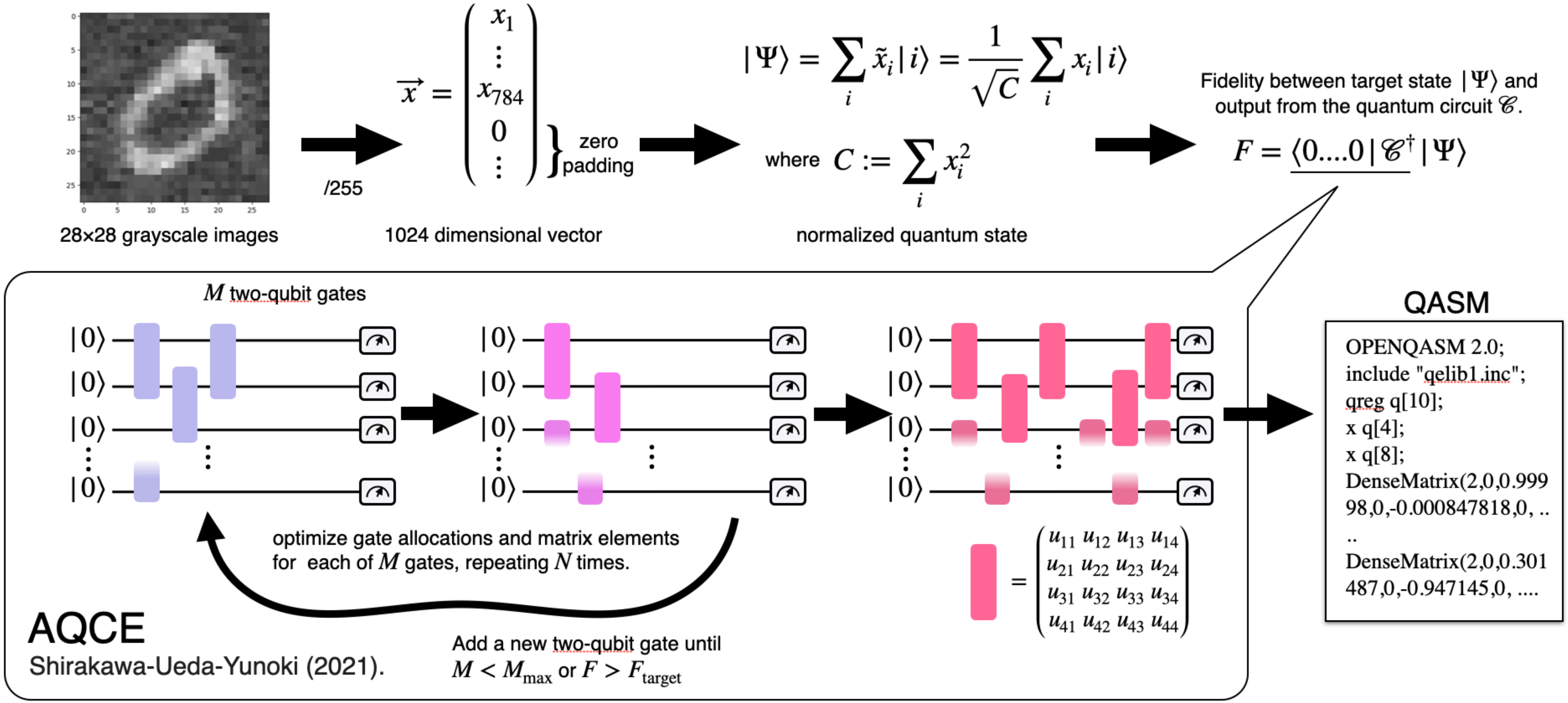
FIG. 2. **Benchmarking the structural sensitivity of out-of-time-ordered correlators (OTOCs) in [4−^{13}C]-toluene.**

Experimentally (OTOC NMR experiments) supervised learning with quantum computers with trainable parameters.
→”Quantum data” is important!

MNISQ dataset: World Largest Quantum “Circuit” Dataset

Placidi *et al.*, *“MNISQ: A Large-Scale Quantum Circuit Dataset for Machine Learning on/for Quantum Computers in the NISQ era.”* arXiv:2306.16627.

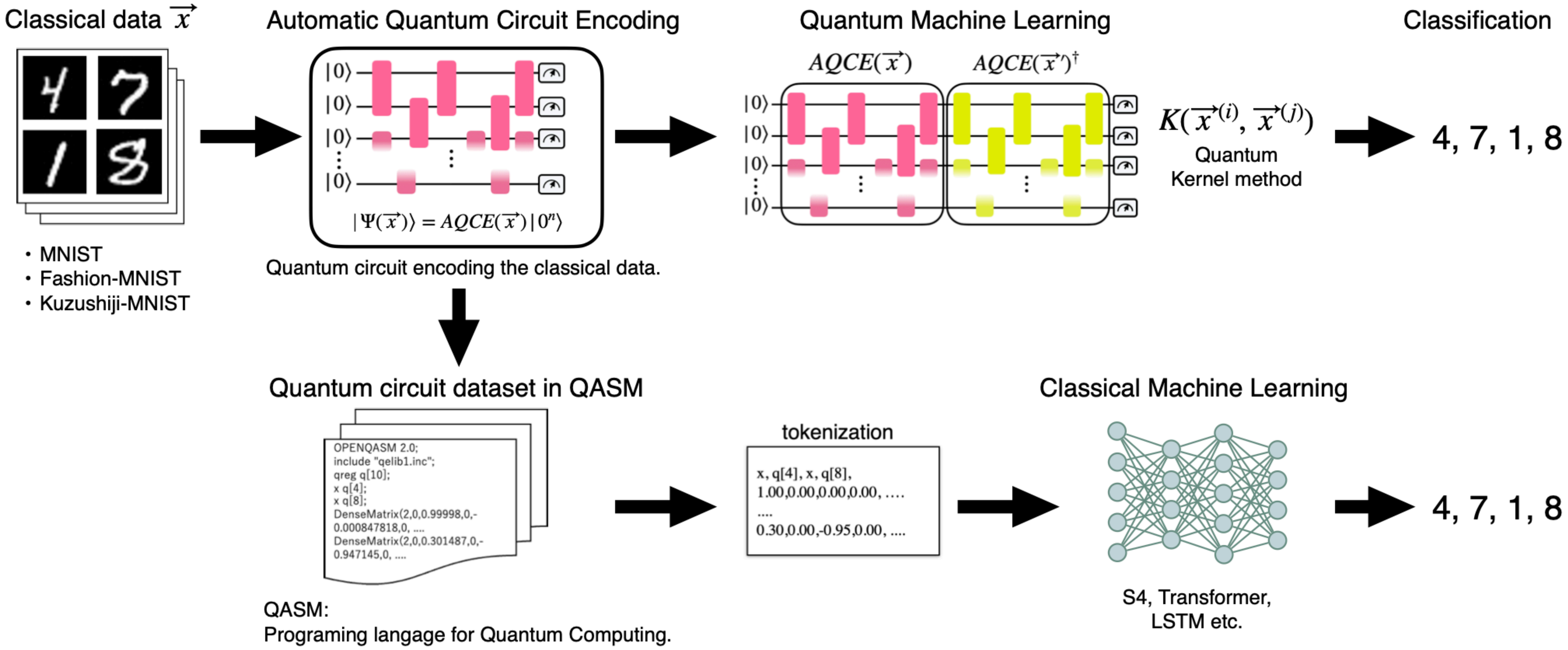
Embedding classical dataset (MNIST, Fashion-MNIST, Kuzushiji-MNIST) into quantum circuits generates a large (4,950,000) quantum circuit data set (MNISQ).



MNISQ dataset: World Largest Quantum “Circuit” Dataset

Placidi *et al., “MNISQ: A Large-Scale Quantum Circuit Dataset for Machine Learning on/for Quantum Computers in the NISQ era.”* arXiv:2306.16627.

Embedding classical dataset (MNIST, Fashion-MNIST, Kuzushiji-MNIST) into quantum circuits generates a large (4,950,000) quantum circuit data set (MNISQ).

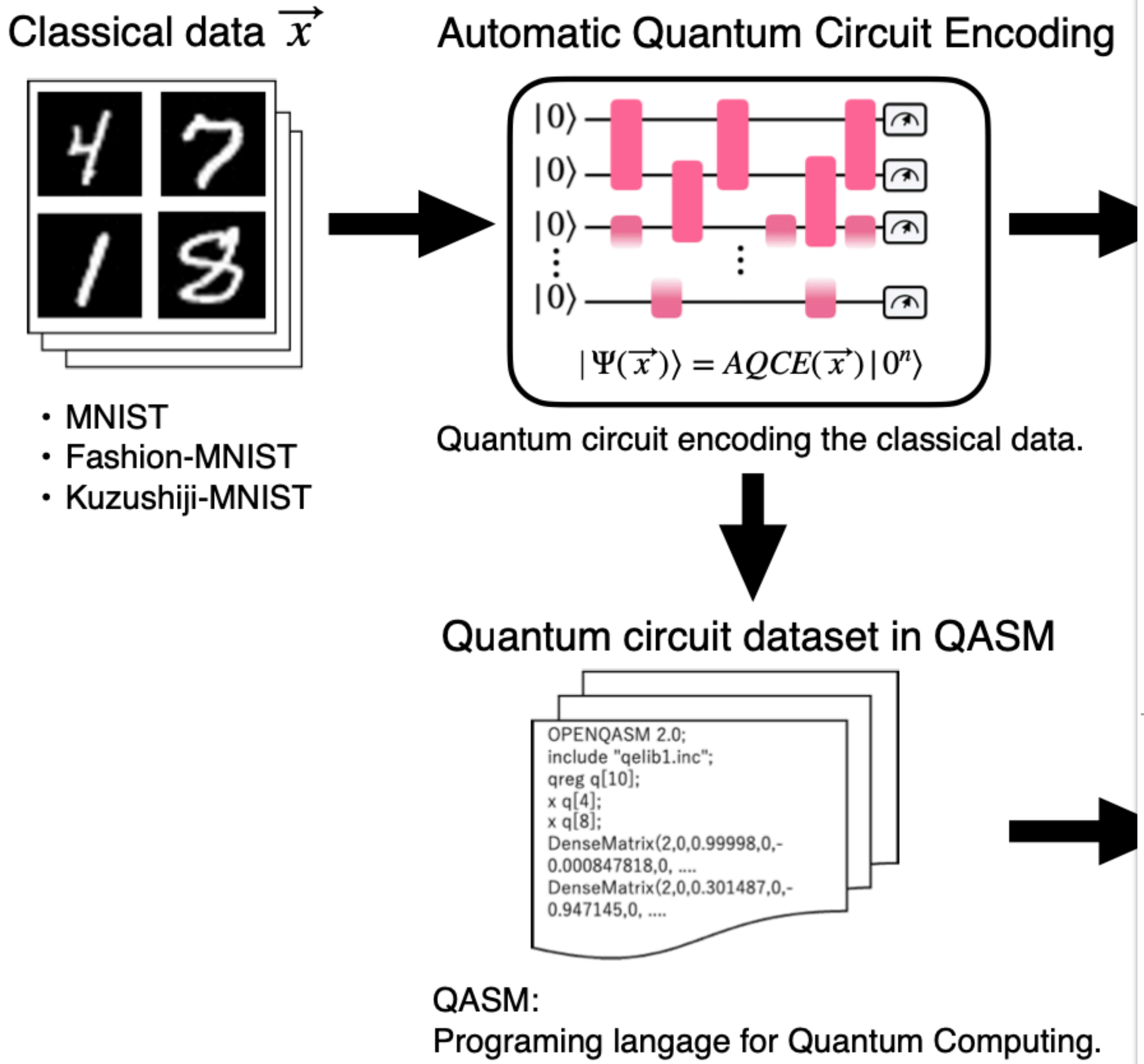


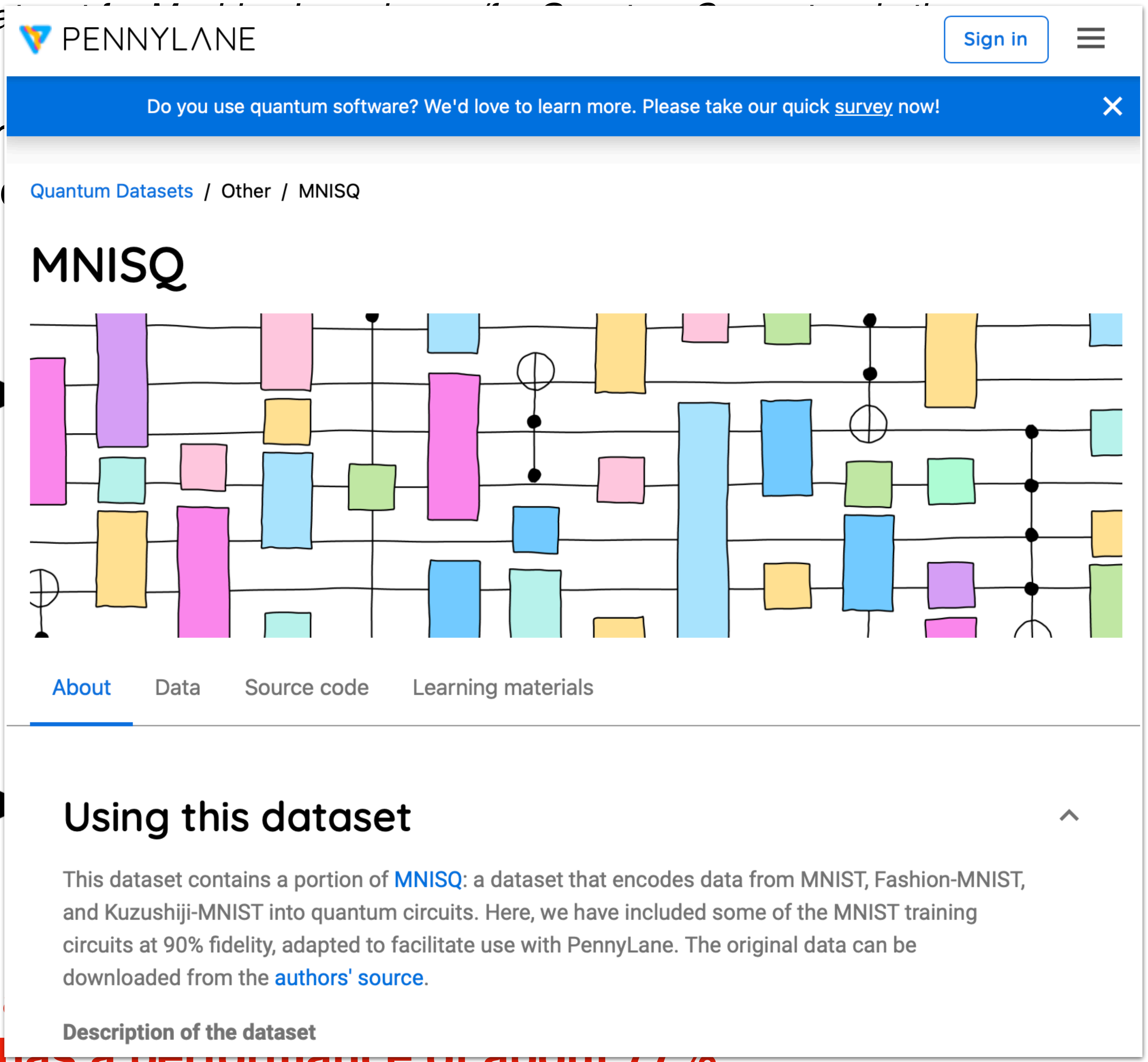
While QML(quantum kernel method) has a classification performance >97%, the classical ML language model (S4, transformer) has a performance of about 77%.

MNISQ dataset: World Largest Quantum “Circuit” Dataset

Placidi *et al.*, *“MNISQ: A Large-Scale Quantum Circuit Da*
NISQ era.” arXiv:2306.16627.

Embedding classical dataset (MNIST, Fash
generates a large (4,950,000) quantum cir





While QML(quantum kernel method) h
ML language model (S4, transformer) has a performance of about 77%.

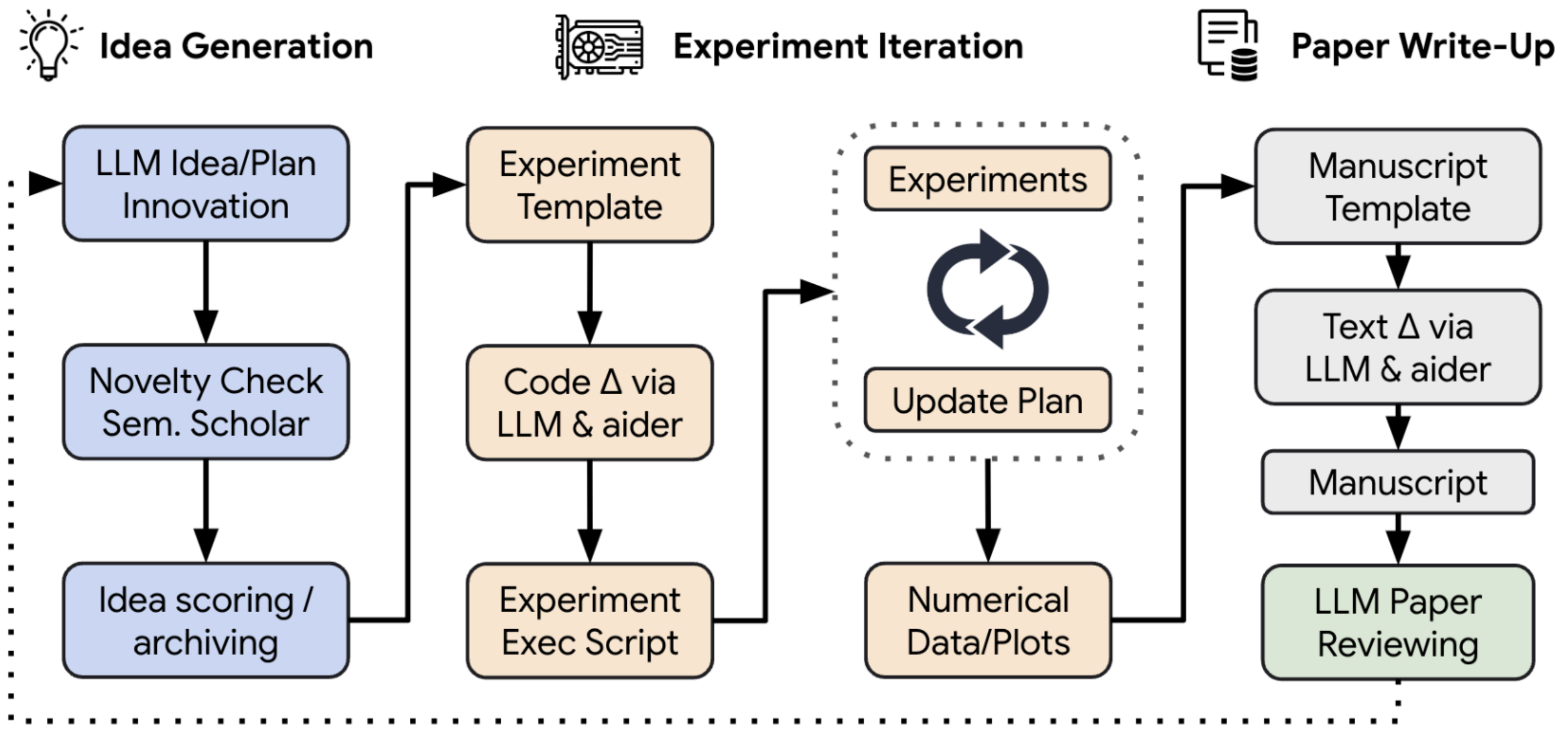
Why not use LLMs for QML research?

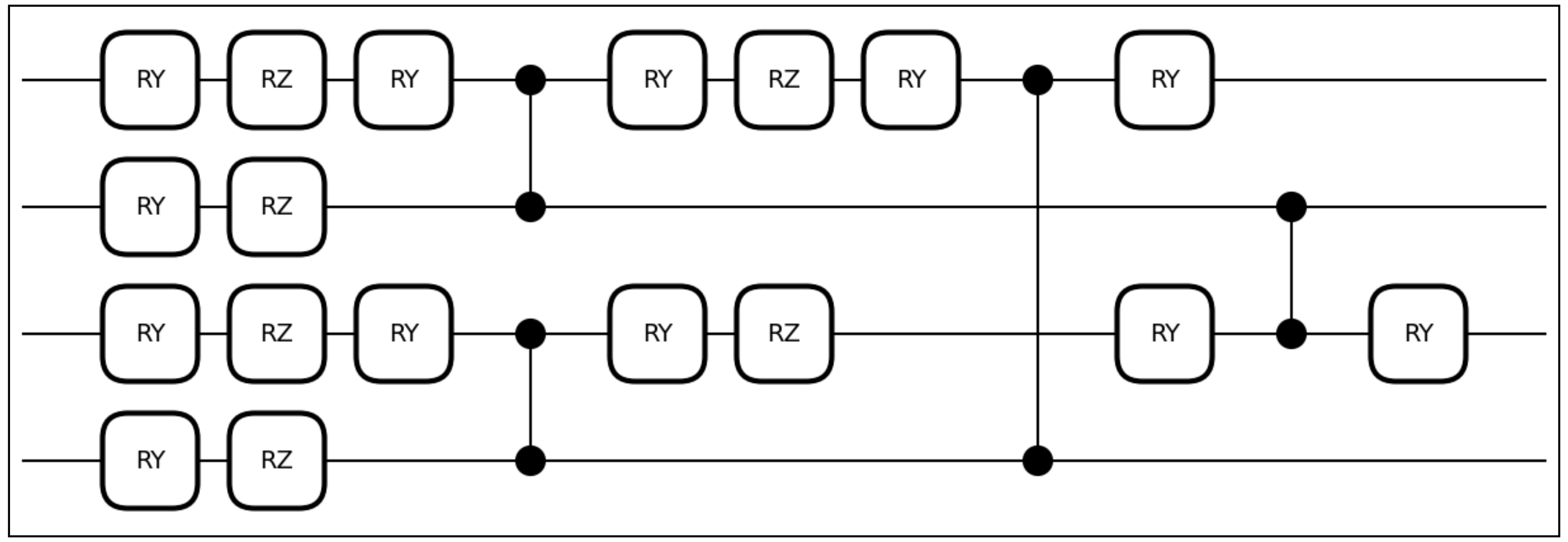
Automating Quantum Feature Map Design Using LLM

K Sakka, K Mitarai, K Fujii , “Automating quantum feature map design via large language models”
arXiv:2504.07396

- AI—especially LLMs—enables research automation across fields
- Designing effective **quantum feature maps** still relies on human intuition
- **Goal:** Automate quantum feature map design using LLMs

Sakana AI: The AI Scientist

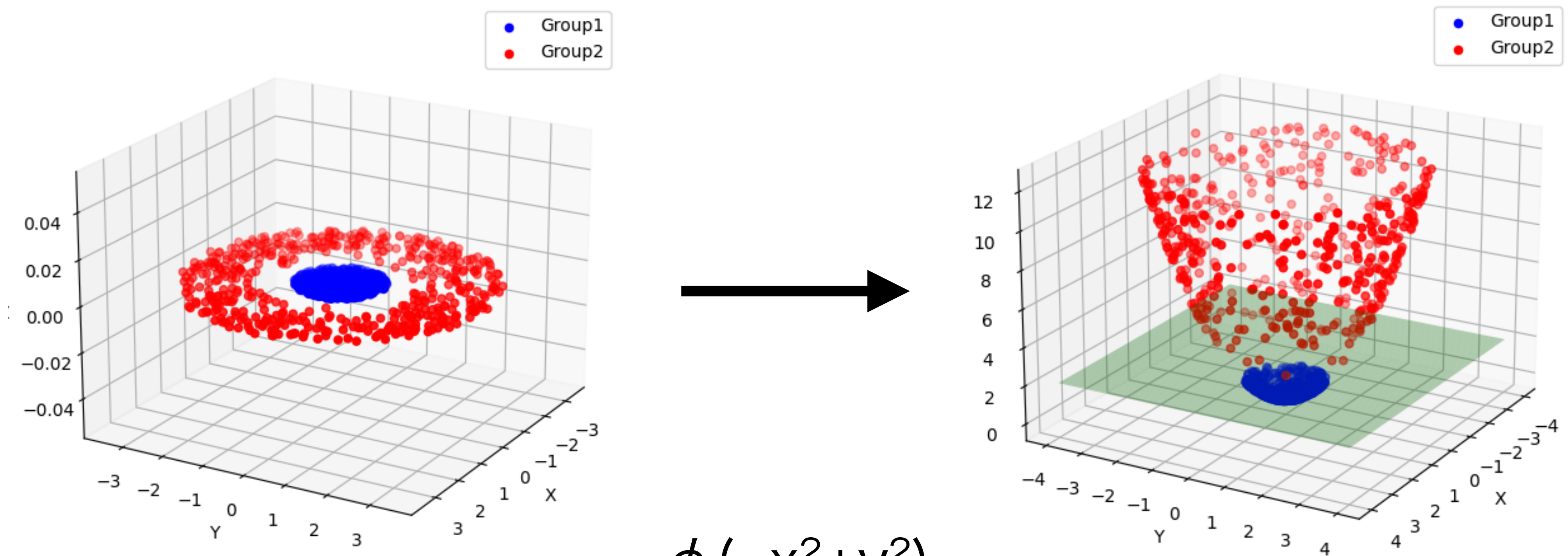




Quantum Feature Map : Quantum circuit that maps an input data to a quantum state.

input data

quantum feature space



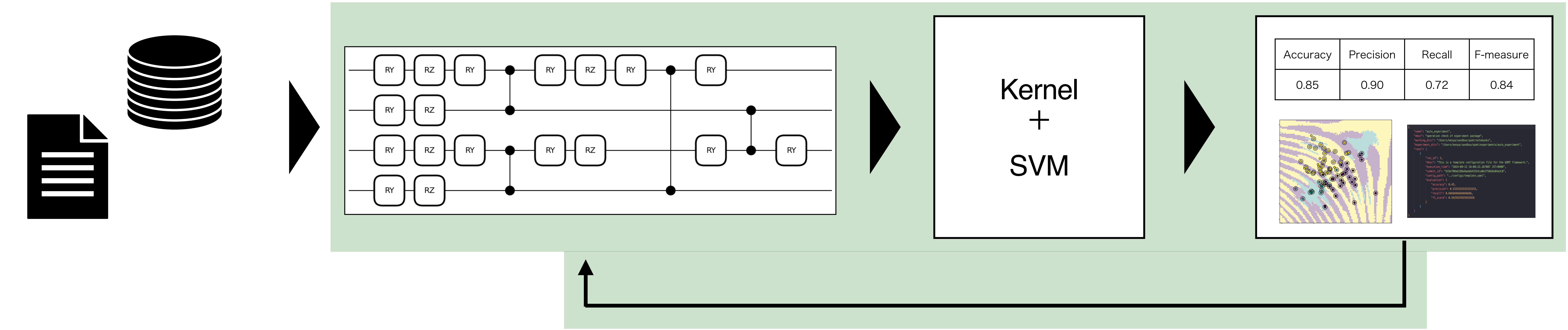
ϕ (=x^2+y^2)

2dim→3dim space

Automating Quantum Feature Map Design Using LLM

K Sakka, K Mitarai, K Fujii , “Automating quantum feature map design via large language models” arXiv:2504.07396

Dataset | **Quantum Feature Map** | **QML model** | **Performance analysis**



LLM Feedback Loop
Automating Trial-and-Error Design of Quantum Feature Maps

Organizing Information for LLM Utilization

- Use up-to-date research information, including papers and reports
- Extract papers via arXiv API and store them in a searchable database

Prompt Engineering

- Encode expert reasoning processes into prompts
- Guide LLM toward feasible solutions leveraging quantum properties

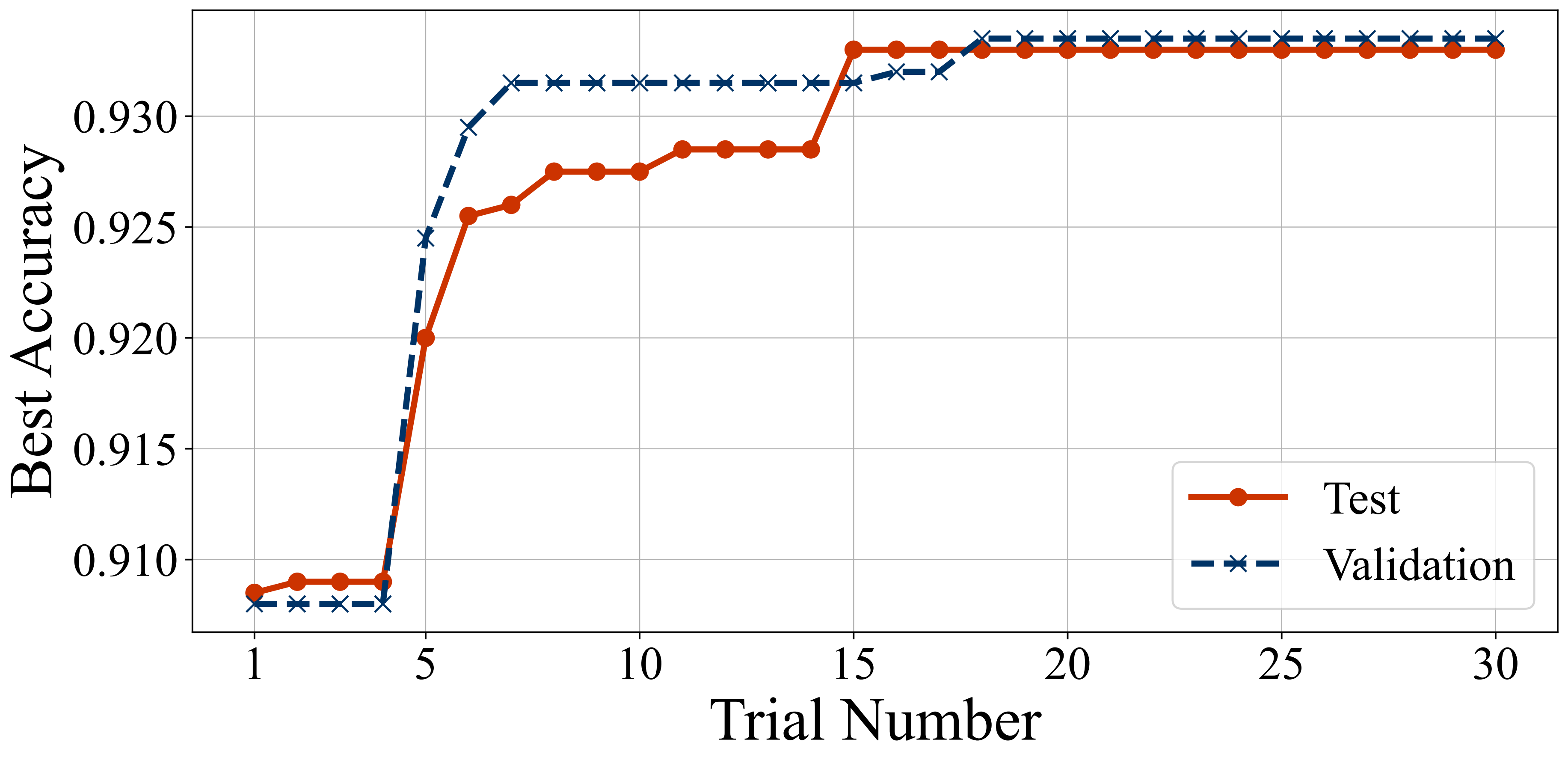
Experiment Support & Management

- Automate all processes except feature map generation using tools
- Store trial results as feedback logs for iterative improvement

Automating Quantum Feature Map Design Using LLM

K Sakka, K Mitarai, K Fujii , “Automating quantum feature map design via large language models”
arXiv:2504.07396

Performance Improves Through Iterative Trials



Type	Method	Accuracy		
		MNIST	Fashion-MNIST	CIFAR-10
Classical	Linear Kernel	0.9385 ± 0.0002	0.8437 ± 0.0009	0.4087 ± 0.0011
	Polynomial Kernel	0.9667 ± 0.0058	0.8702 ± 0.0030	0.5375 ± 0.0014
	Sigmoid Kernel	0.9343 ± 0.0002	0.8189 ± 0.0120	0.4079 ± 0.0006
	RBF Kernel	0.9765 ± 0.0005	0.8864 ± 0.0014	0.5669 ± 0.0085
Quantum	ZZ FeatureMap [5]	0.9255 ± 0.0009	0.8252 ± 0.0023	0.3907 ± 0.0016
	NPQC FeatureMap [20]	0.9644 ± 0.0028	0.8749 ± 0.0026	0.4903 ± 0.0188
	YZCX FeatureMap [20]	0.9727 ± 0.0006	0.8778 ± 0.0049	0.4753 ± 0.0341
	Astronaut V1 (Previous, Section D)	0.9731 ± 0.0008	0.8835 ± 0.0021	0.5290 ± 0.0030
	Astronaut V2 (Ours, 10 qubits)	$\mathbf{0.9772 \pm 0.0002}$	0.8880 ± 0.0014	0.5450 ± 0.0057
	Astronaut V2 (Ours, 14 qubits)	0.9767 ± 0.0002	$\mathbf{0.8888 \pm 0.0007}$	$\mathbf{0.5734 \pm 0.0006}$

- The generated feature maps outperform existing quantum feature maps across all datasets.
- Compared with classical methods, they outperform linear, polynomial, sigmoid, RBF kernels.

Released as OSS QXMT
Quantum Experiment Management Tool

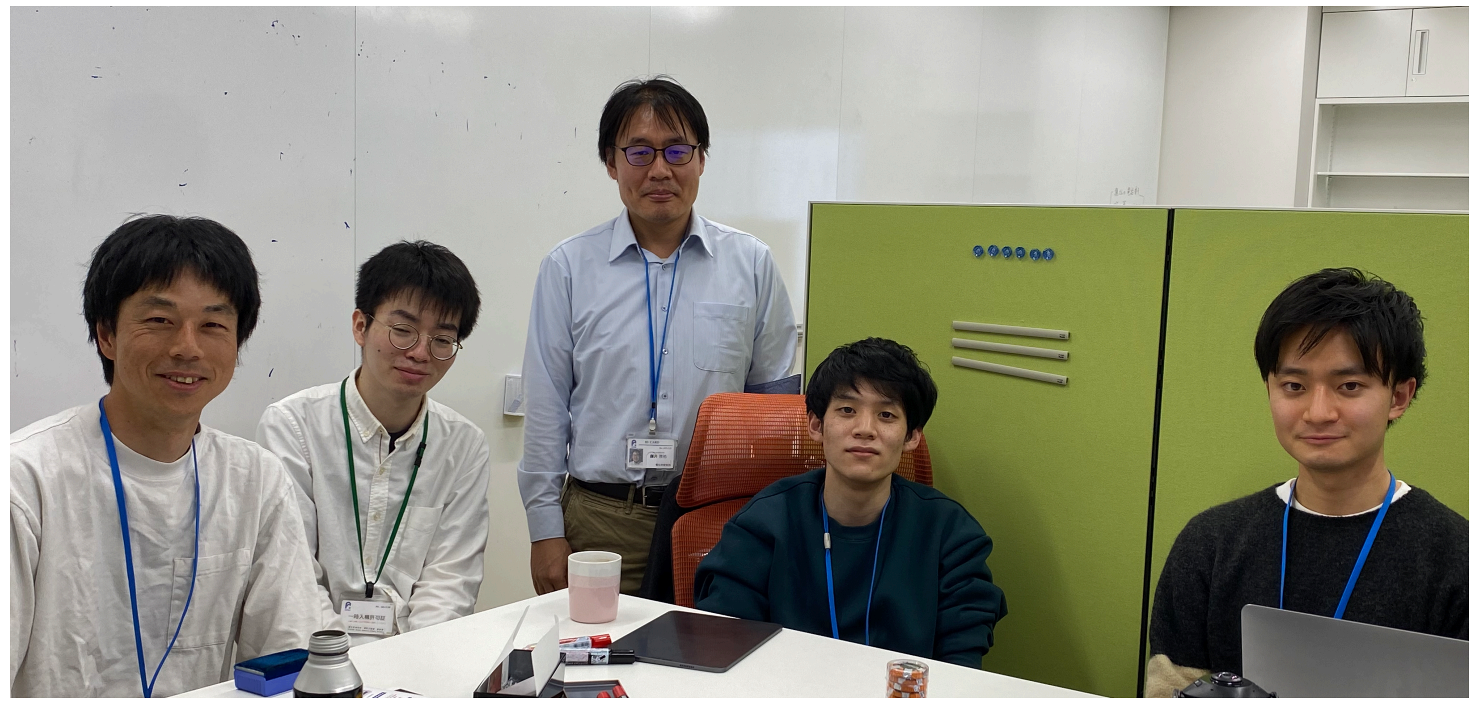
Astronaut
Implementation of automated quantum feature map design (as demonstrated in the paper)

This approach can be extended beyond quantum machine learning to any research domain that requires iterative optimization — such as algorithm design or noise mitigation.

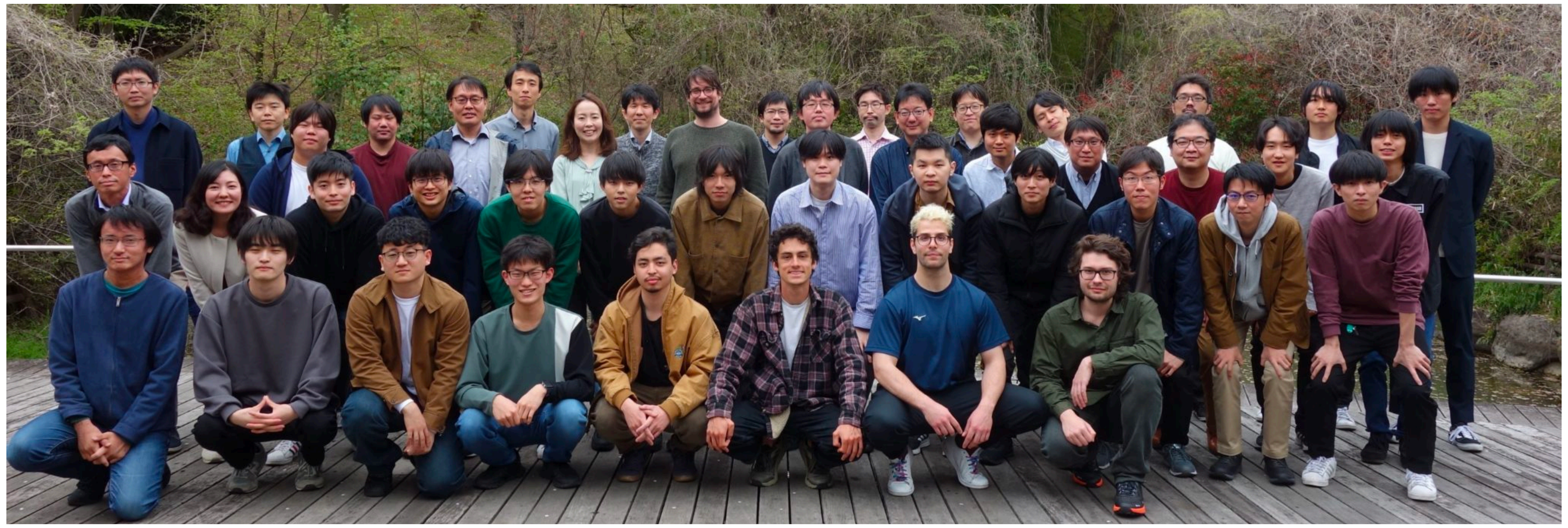
Summary

- Quantum Machine Learning leverages high-dimensional Hilbert spaces by quantum feature maps.
- Evidence of verifiable quantum advantage emerges from quantum-chaos/OTOC experiments.
- Quantum data (e.g., MNISQ) is crucial for benchmarking and training QML.
- AI/LLMs can automate quantum-feature-map design and accelerate quantum-computing research.

RIKEN QUANTUM COMPUTING



Quantum Computing Theory Research Team at RQC, RIKEN



大阪大学 OSAKA UNIVERSITY

Quantum Computing (theory) Group at Osaka Univ.

Funding acknowledgement:
MEXT Q-LEAP Quantum AI flagship
JST COI-NEXT Quantum Software Research Hub
JST Moonshot Goal6