Quantum Machine Learning

State of the art, drawbacks and future possibilities

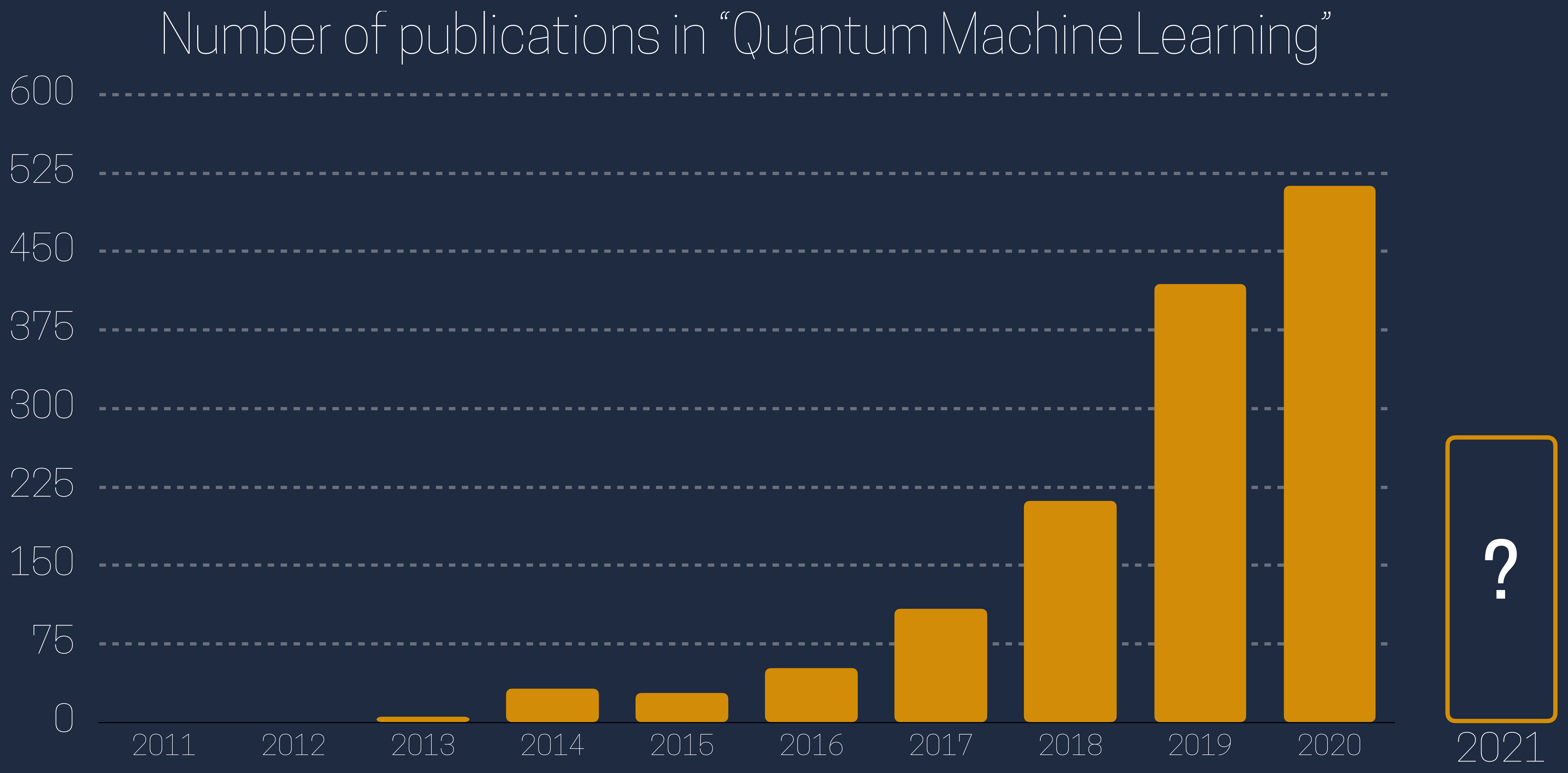
PhD End-of-Year Seminars
University of Pavia, Department of Physics

Stefano Mangini XXXV° cycle
Supervisor Prof. Chiara Macchiavello
Quantum Information Theory Group (QUIT)

1\October\ 2020

UNIVERSITÀ DI PAVIA

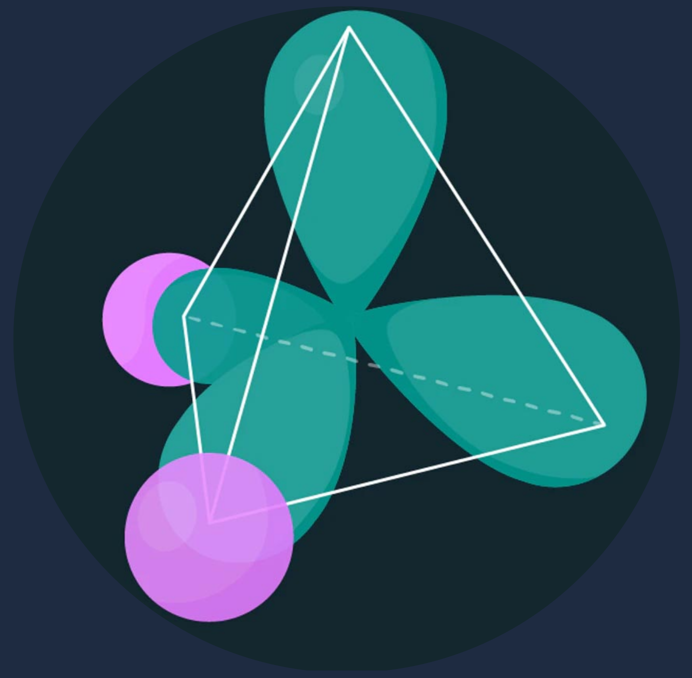
Hype behind QML



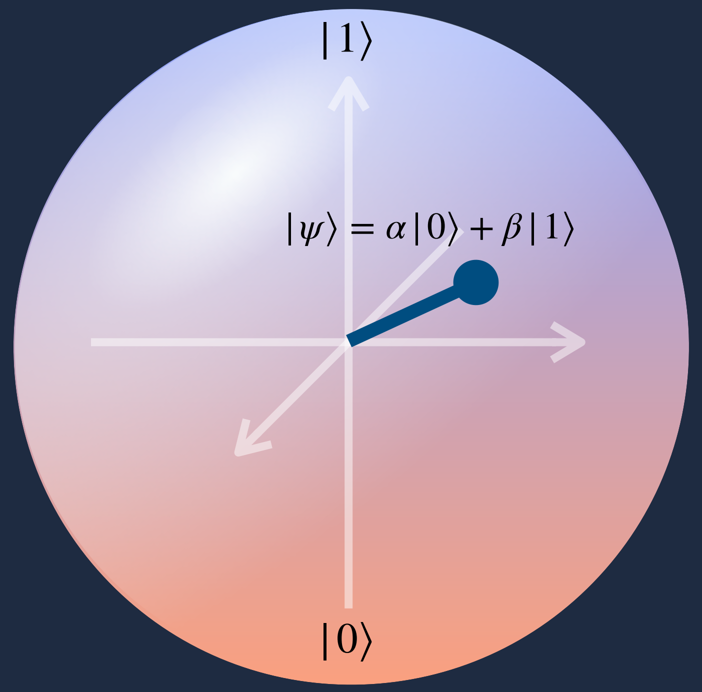
Source: Dimension.ai using the keyword "quantum machine learning"

Why?

Range of possible applications:



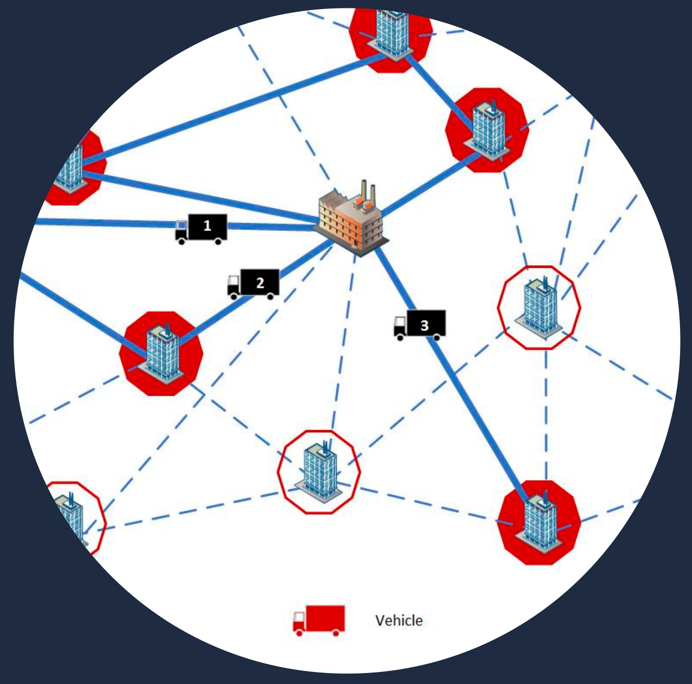
Quantum Chemistry
Drug Discovery
Condensed matter



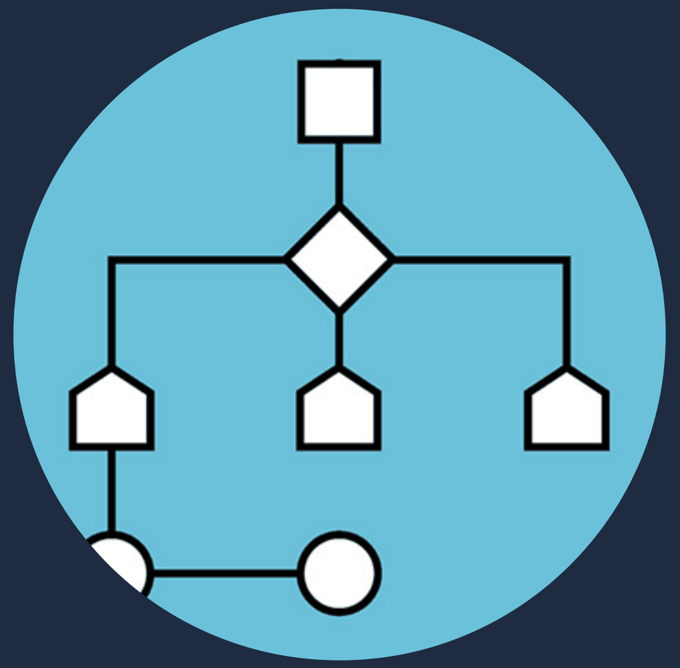
Optimize quantum computers

Portfolio optimization

Self driving cars



Logistic problems
like vehicle routing



New algorithms
Understand older ones

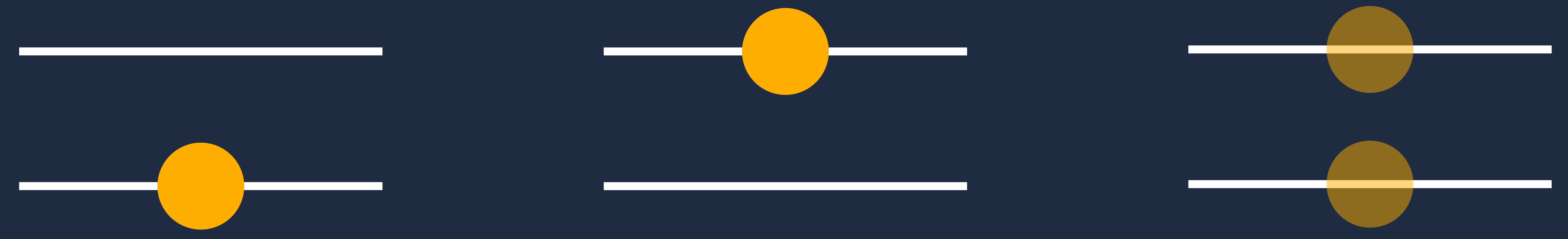
Quantum Computers

Quantum computers are physical systems capable of implementing quantum computations.

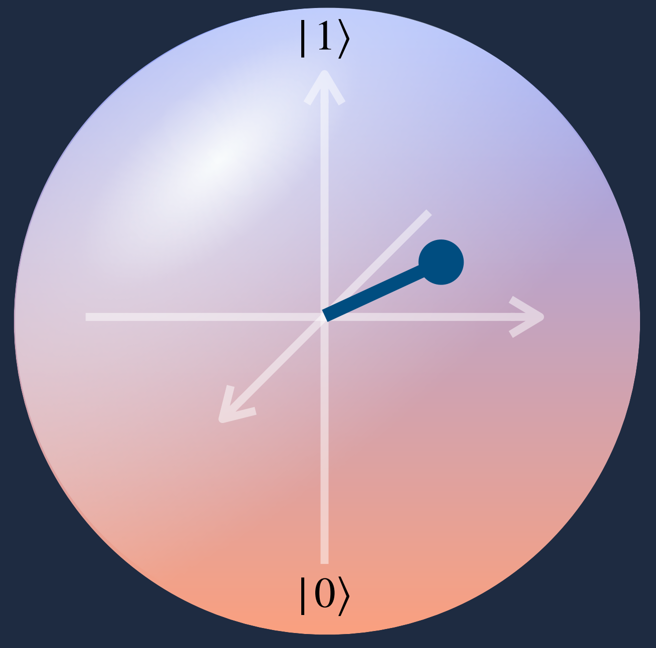
Qubit

$|1\rangle$

$|0\rangle$



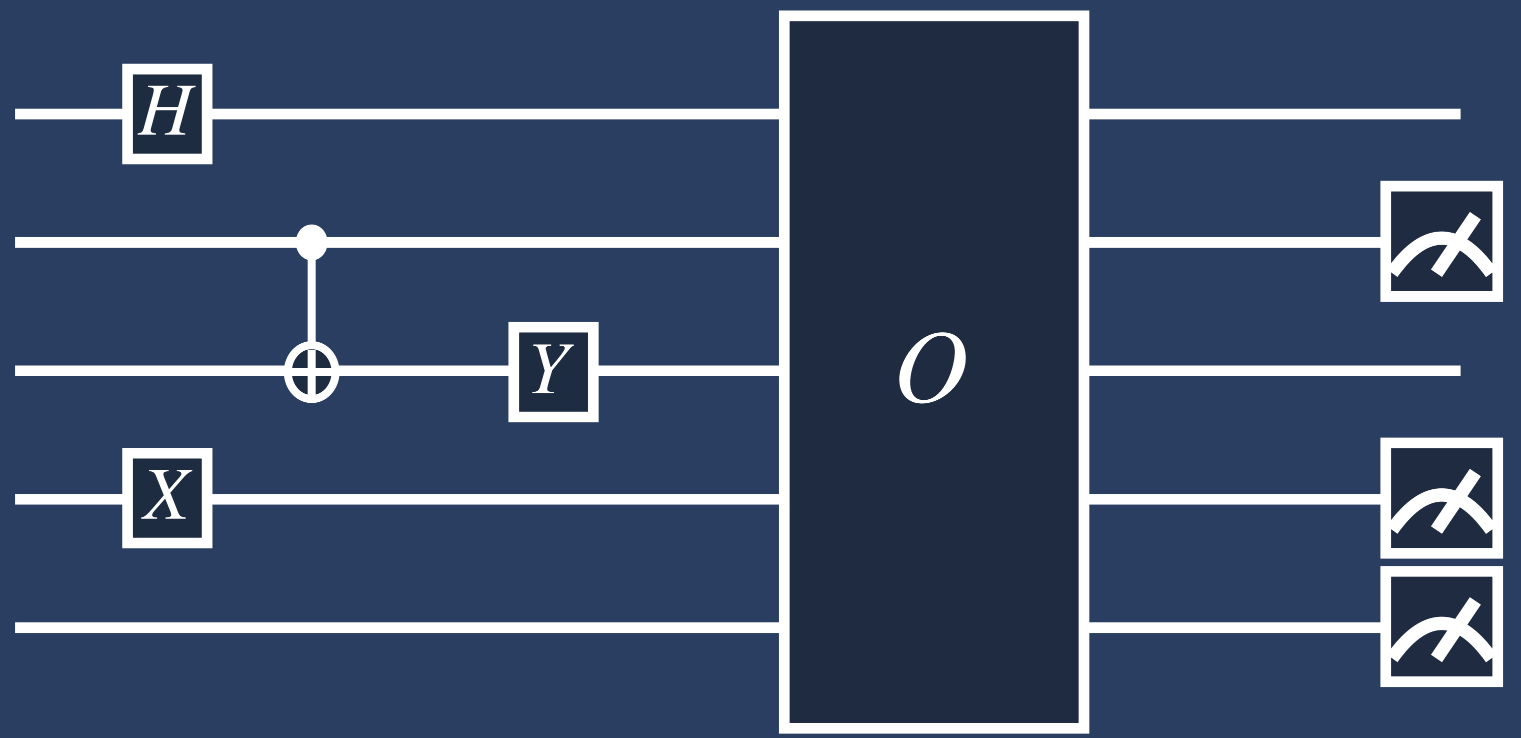
$\dim \mathcal{H} = 2$ $\quad |\psi\rangle = |0\rangle \quad |\psi\rangle = |1\rangle \quad |\psi\rangle = \dfrac{|0\rangle + |1\rangle}{\sqrt{2}}$



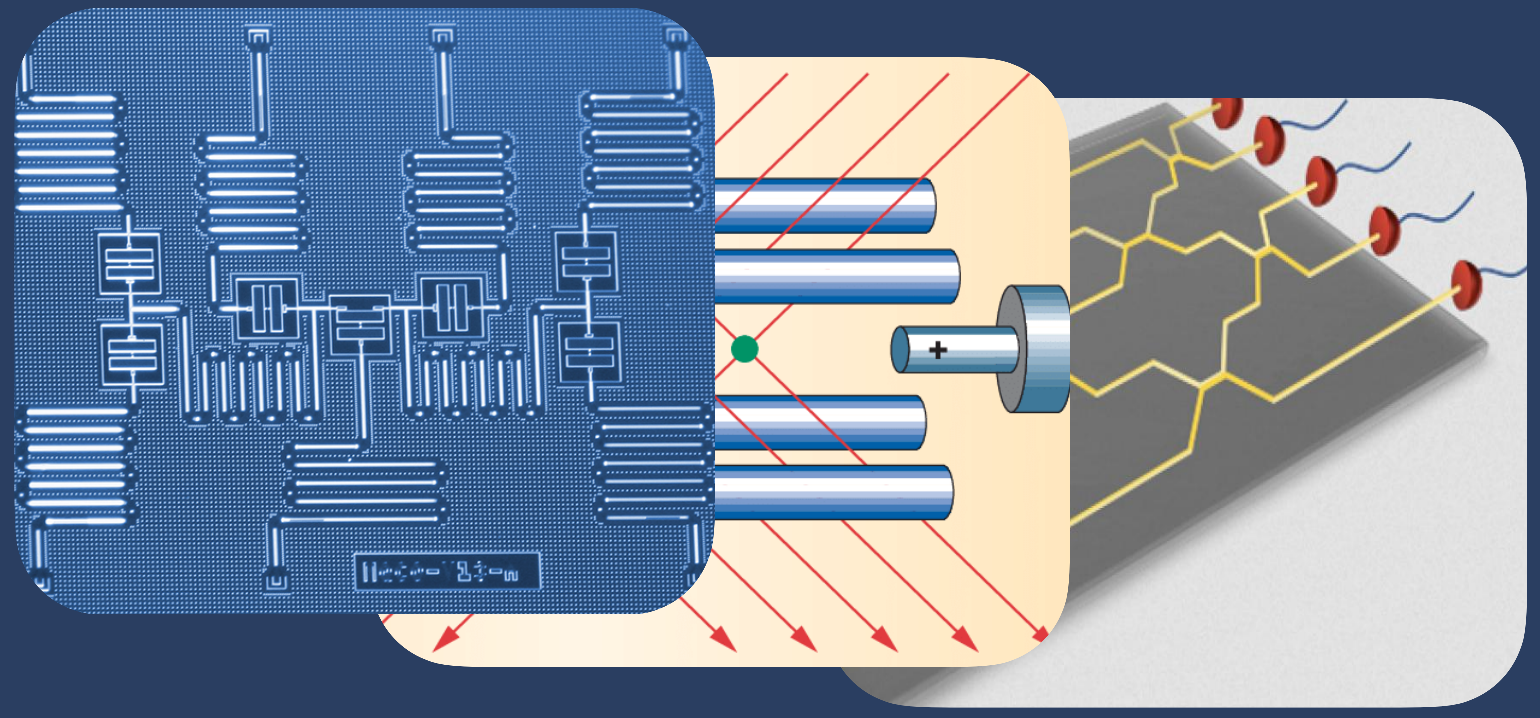
$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

Multiple qubits $\quad \mathcal{H} = \mathcal{H}_0 \otimes \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n \quad \dim \mathcal{H} = 2^n$ Exponential!

Quantum circuit model



- Superconducting circuits
- Ion Traps
- Photonics



Quantum Advantage

Shor's Factoring

$$N = p \times q$$

Quantum: $\exp\left(O\left((\log N)^{1/3}(\log\log N)^{2/3}\right)\right)$

Classical: $O\left((\log N)^3\right)$

Exponential!

Hidden subgroup problem: Discrete Logarithm, Order Finding, …

Quantum Fourier Transform

Grover's search

Target

Quantum: $O\left(\sqrt{N}\right)$

Classical: $O(N)$

Polynomial!

HHL for Linear Equations (aka matrix inversion)

$$\boldsymbol{Ax} = \boldsymbol{b}$$

Quantum: $O(\log N)$ *

Classical: $O(N)$

Exponential!

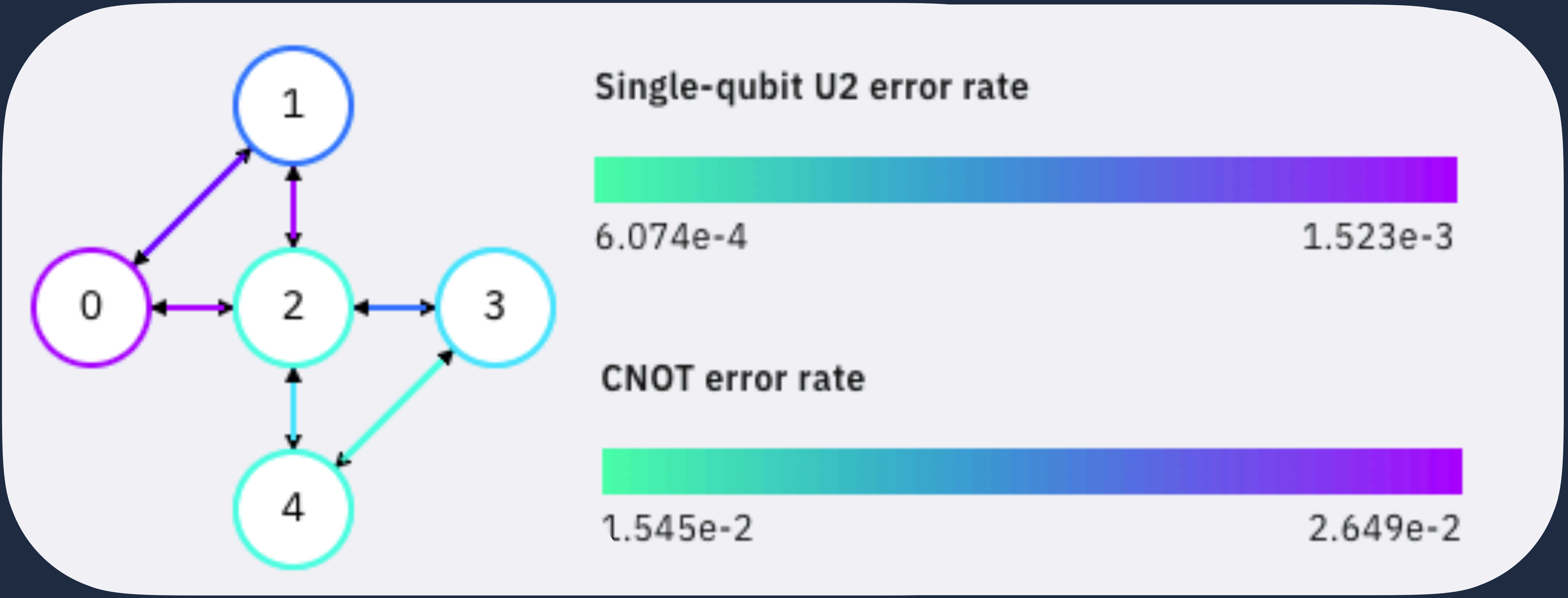
*given constraints on A

NISQ

Noisy **I**ntermediate **S**cale **Q**uantum (NISQ) devices:

- 10-10^2 qubits
- Gate Errors
- Low connectivity

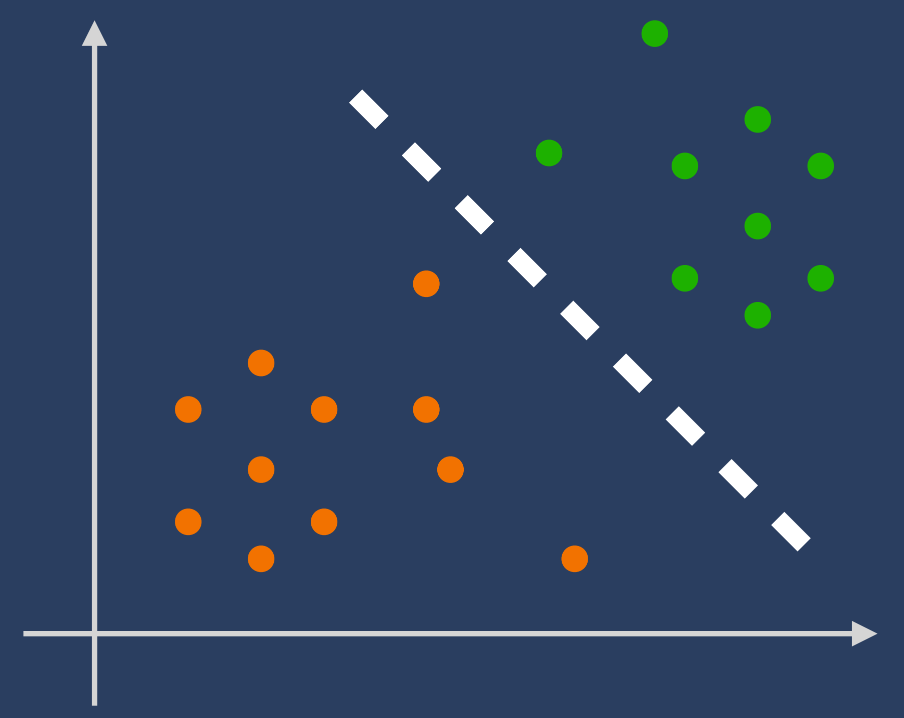
IBMQ's Roadmap:
1121 (physical) qubits by 2023



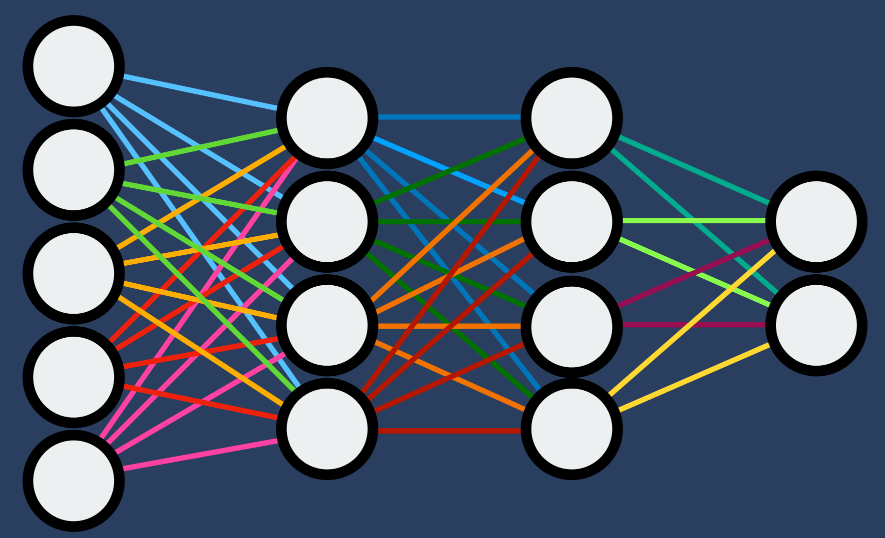
IBM Quantum Experience : ibmqx2-yorktown quantum processor

A primer on classical AI

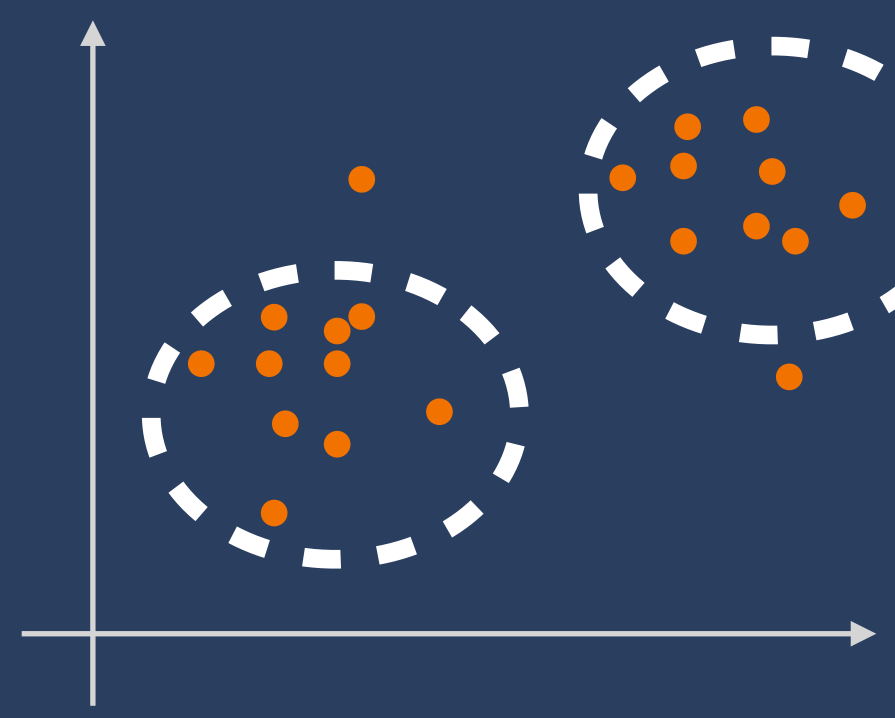
Supervised Learning



Perceptrons,
SVM,
NN,
...

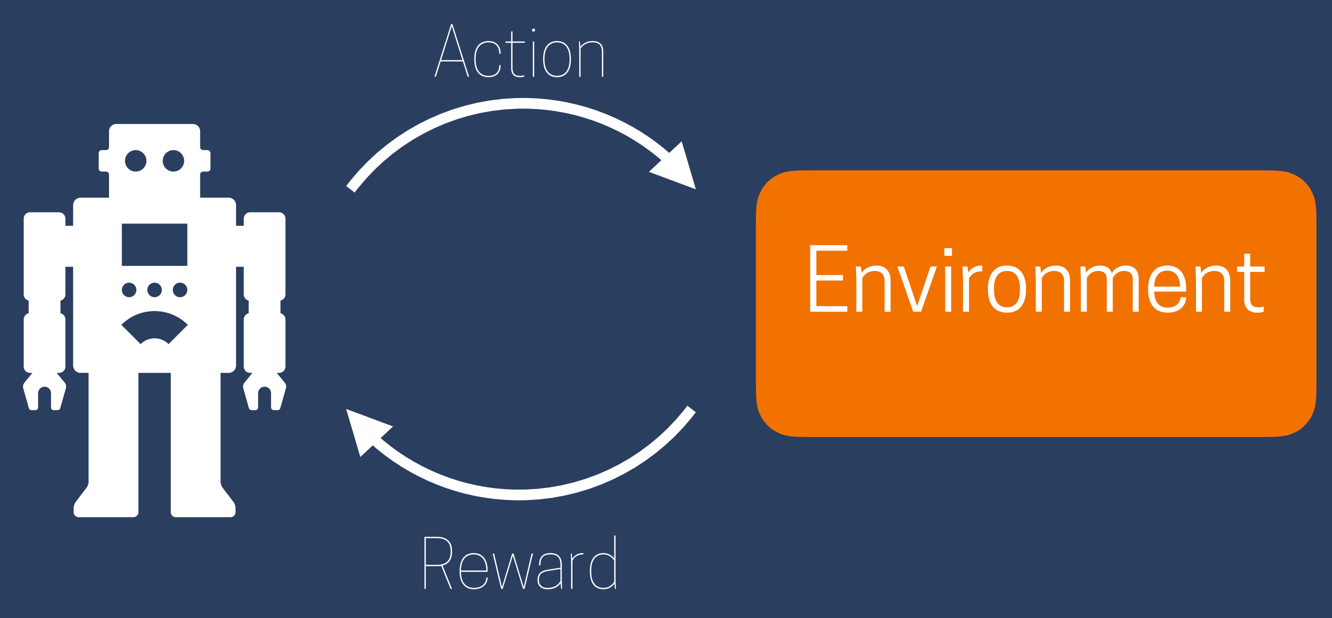


Unsupervised Learning



PCA,
k-means,
...

Reinforcement Learning



...but Quantum.

The four-fold way

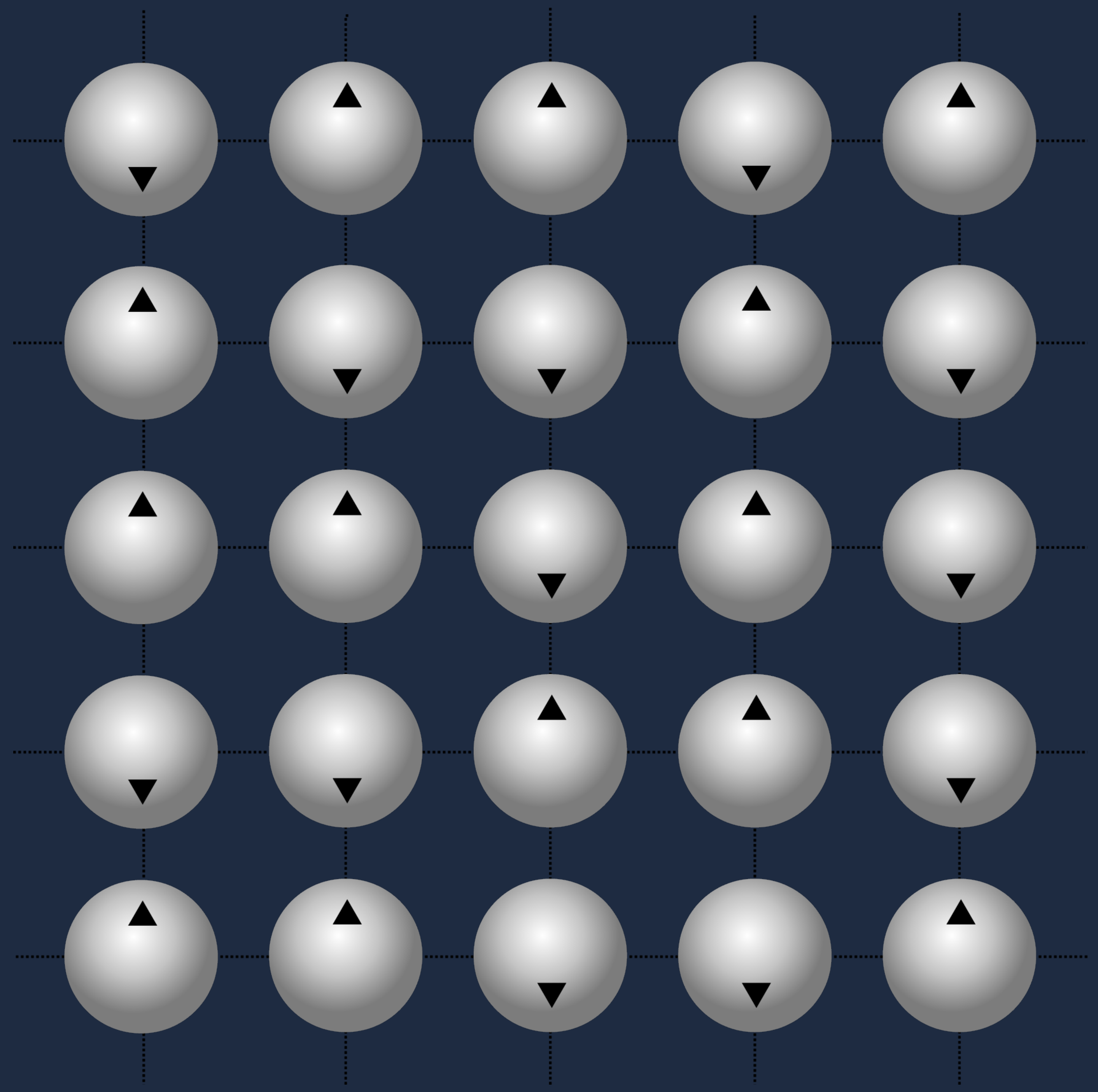
CC *Classical for classical*	CQ *Classical for Quantum*
QC *Quantum for classical*	QQ *Quantum for Quantum*

Quantum-Applied
Machine Learning

Quantum-Enhanced
Machine Learning

ML for Quantum Physics

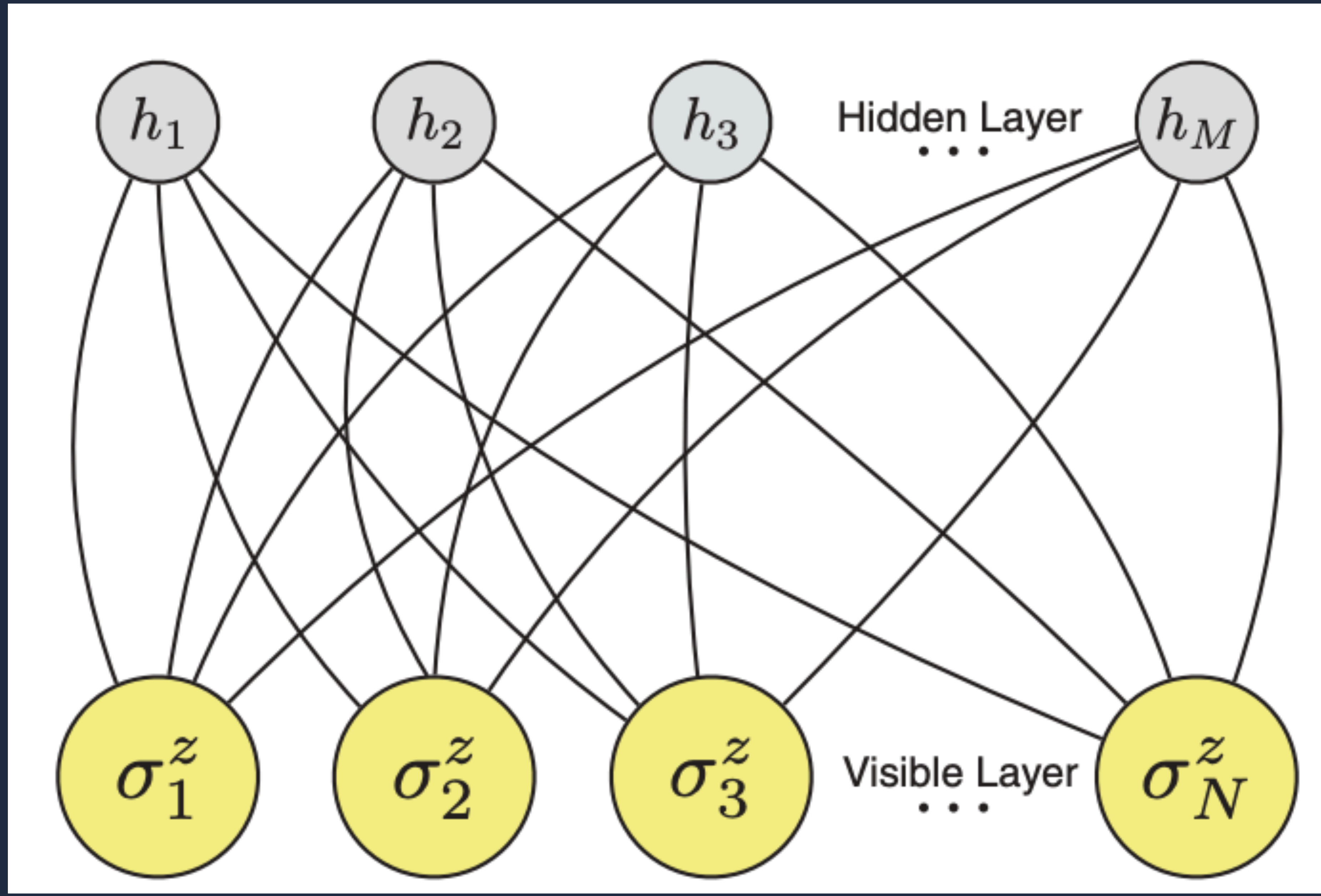
Phase transitions



Unsupervised: PCA, Clustering
Supervised: NN, CNN

Representing quantum states

Boltzmann Machines

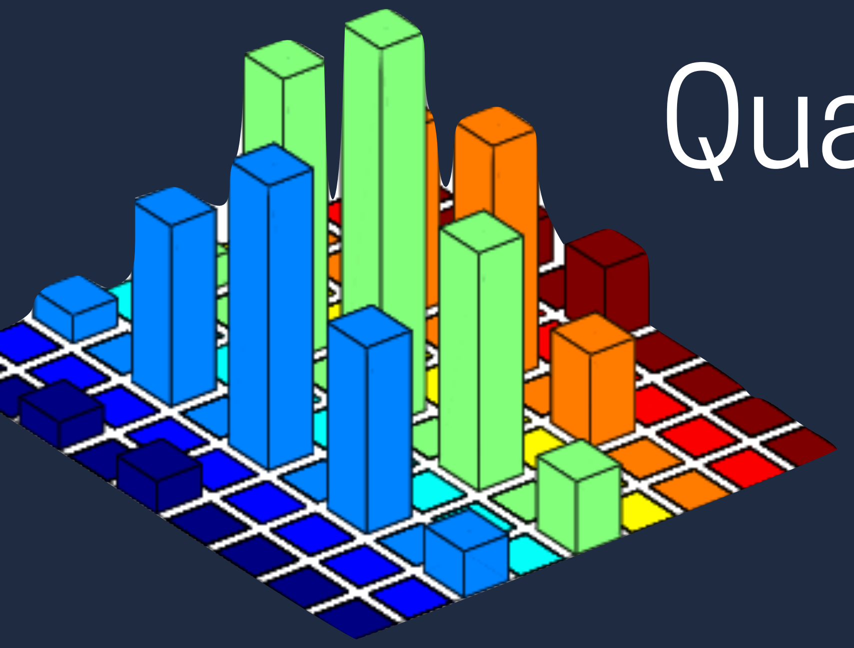


Neural Network Quantum States (NQS):

$$\psi = \sum_{\{h\}} \exp\left(\sum_j a_j \sigma_j^z + \sum j b_j h_j + \sum_{ij} W_{ij} h_i \sigma_j^z \right)$$

Carleo, G., et al. Machine learning and the physical sciences. *Reviews of Modern Physics* **91**.4 (2019): 045002

ML for Quantum Control



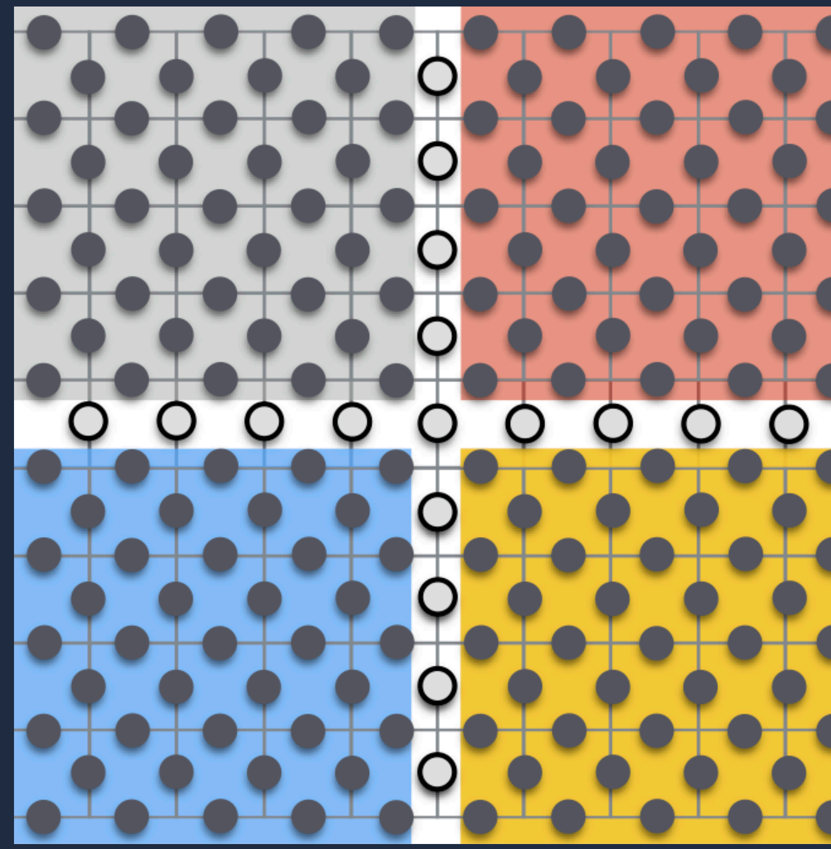
Quantum State Tomography (QST)

Reconstruct density matrix ρ from measurements
Exponential in number of qubits

→

Recurrent Neural Networks optimizing gates
RBMs using parametrization of the state

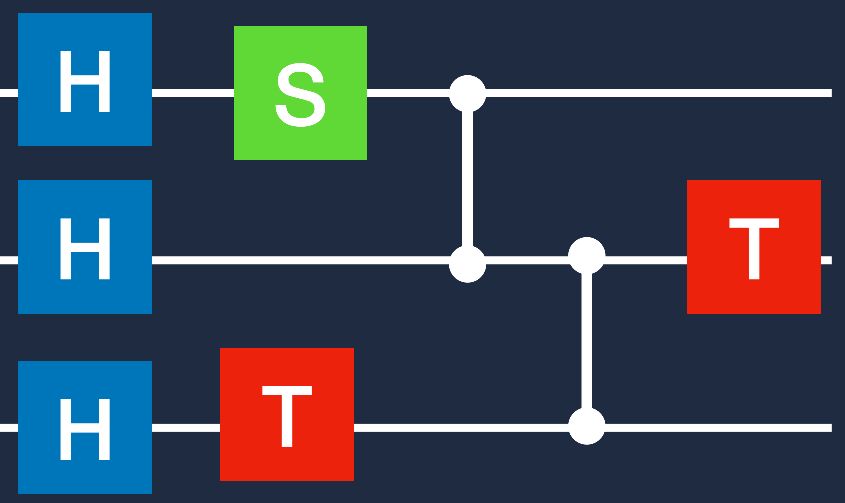
Quantum Error Correction (QEC)



Find strategies to protect quantum computation against noise and errors

←

Reinforcement learning for $\boldsymbol{S}$ and $\boldsymbol{e}$
BMs learning $p_\lambda(\boldsymbol{S}, \boldsymbol{e})$



Quantum Algorithms

Develop new quantum algorithms for specialized tasks

→

Reinforcement learning for new experiments
Optimization techniques

Carleo, G., et al. Machine learning and the physical sciences. *Reviews of Modern Physics* **91**.4 (2019): 045002

IBM Quantum Challenge

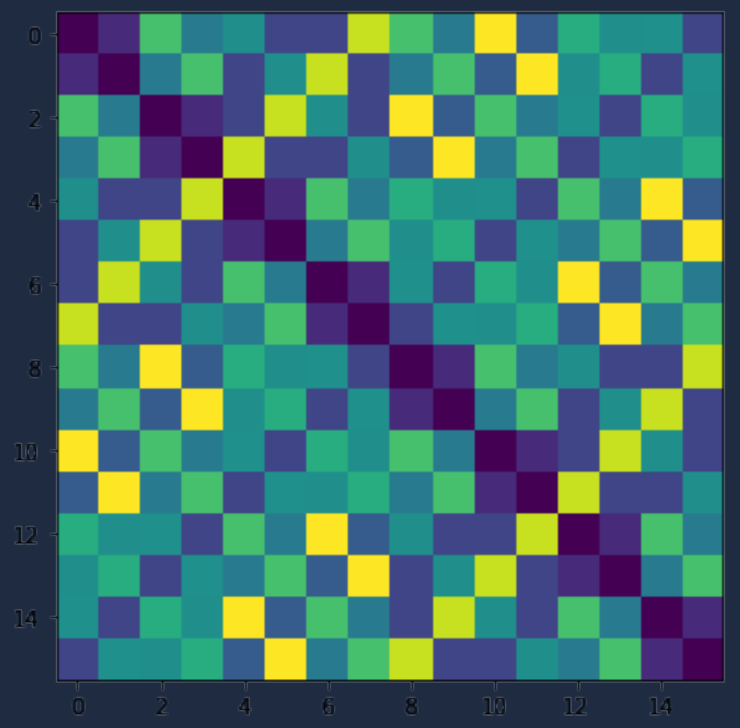
UNIVERSITÀ DI PAVIA

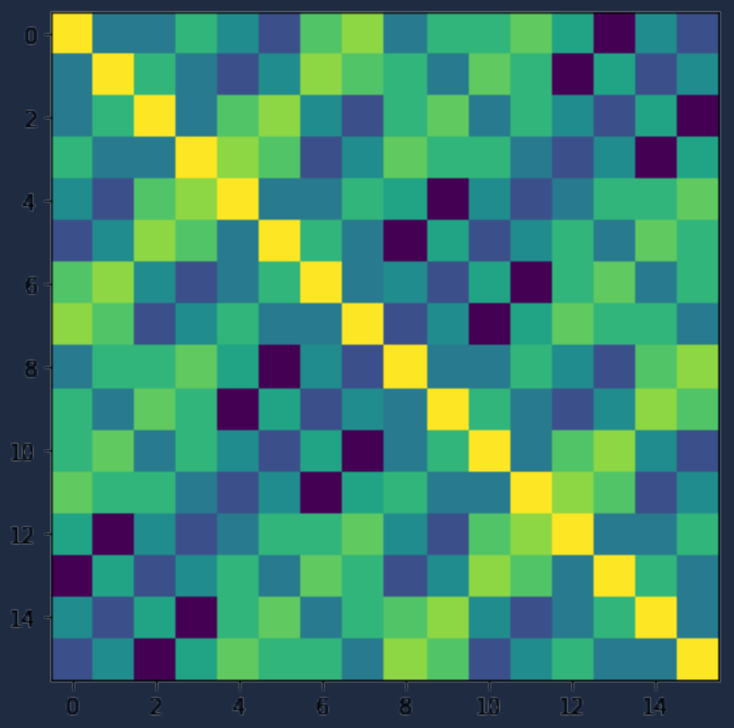
Given unitary U, find an approximation V, such that

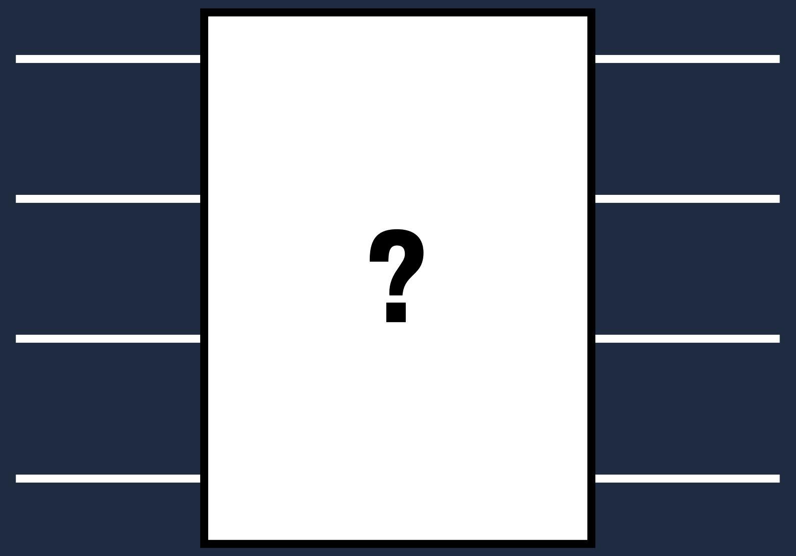
$$\|U - V\|_2 < \varepsilon, \qquad \varepsilon = 0.01 \qquad \|A\|_2 = \max_{|\psi\rangle} \|A\,|\psi\rangle\|_2$$

Using only single qubit gates and CNOT, minimizing the cost = $10n_{cx} + n_{u3}$

$U =$

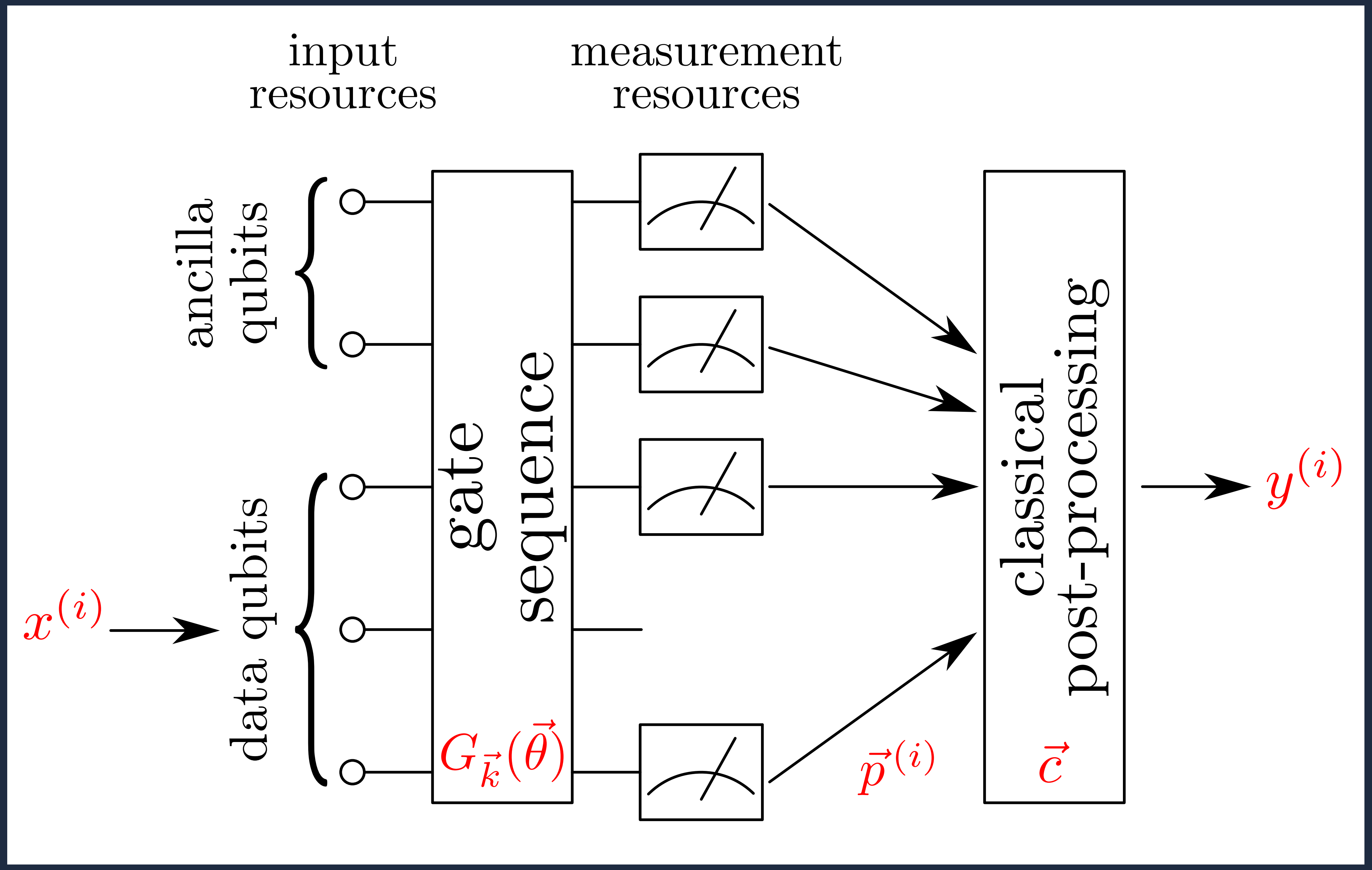






Best cost

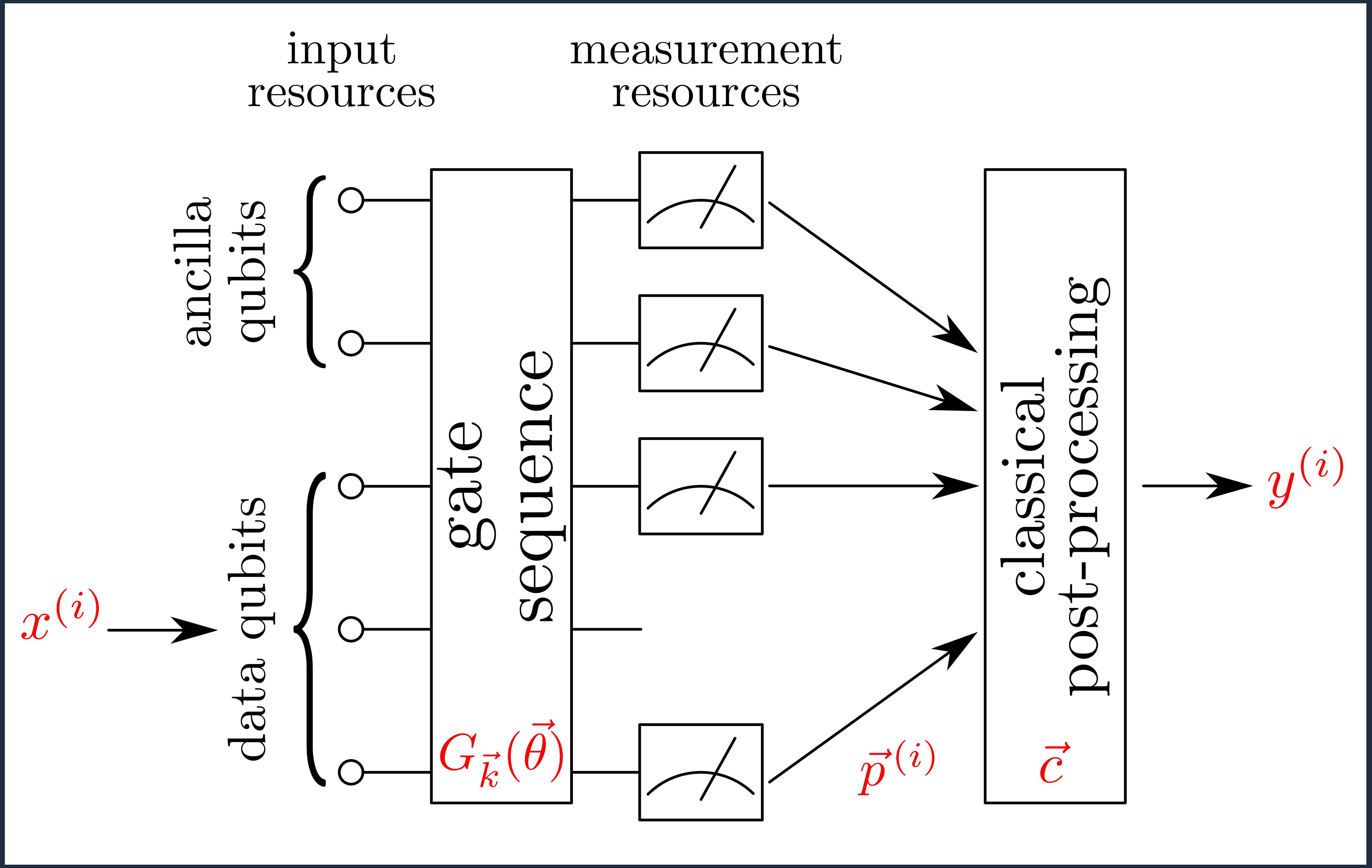
46

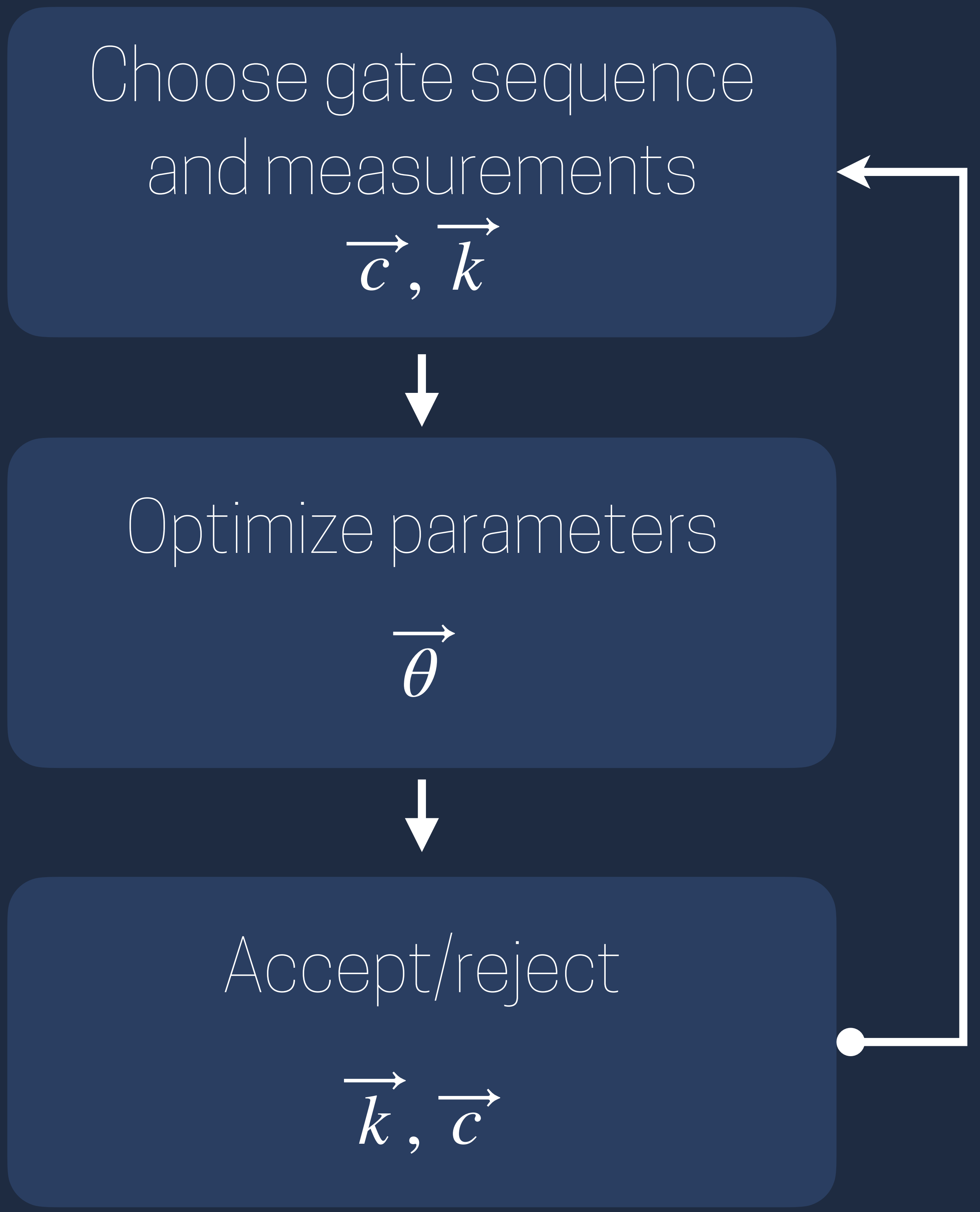


45!

IBM Quantum Challenge

Using a Machine Learning approach someone got 45!





Cincio, L., et al. Learning the quantum algorithm for state overlap, *New J. Phys* **20** 113022

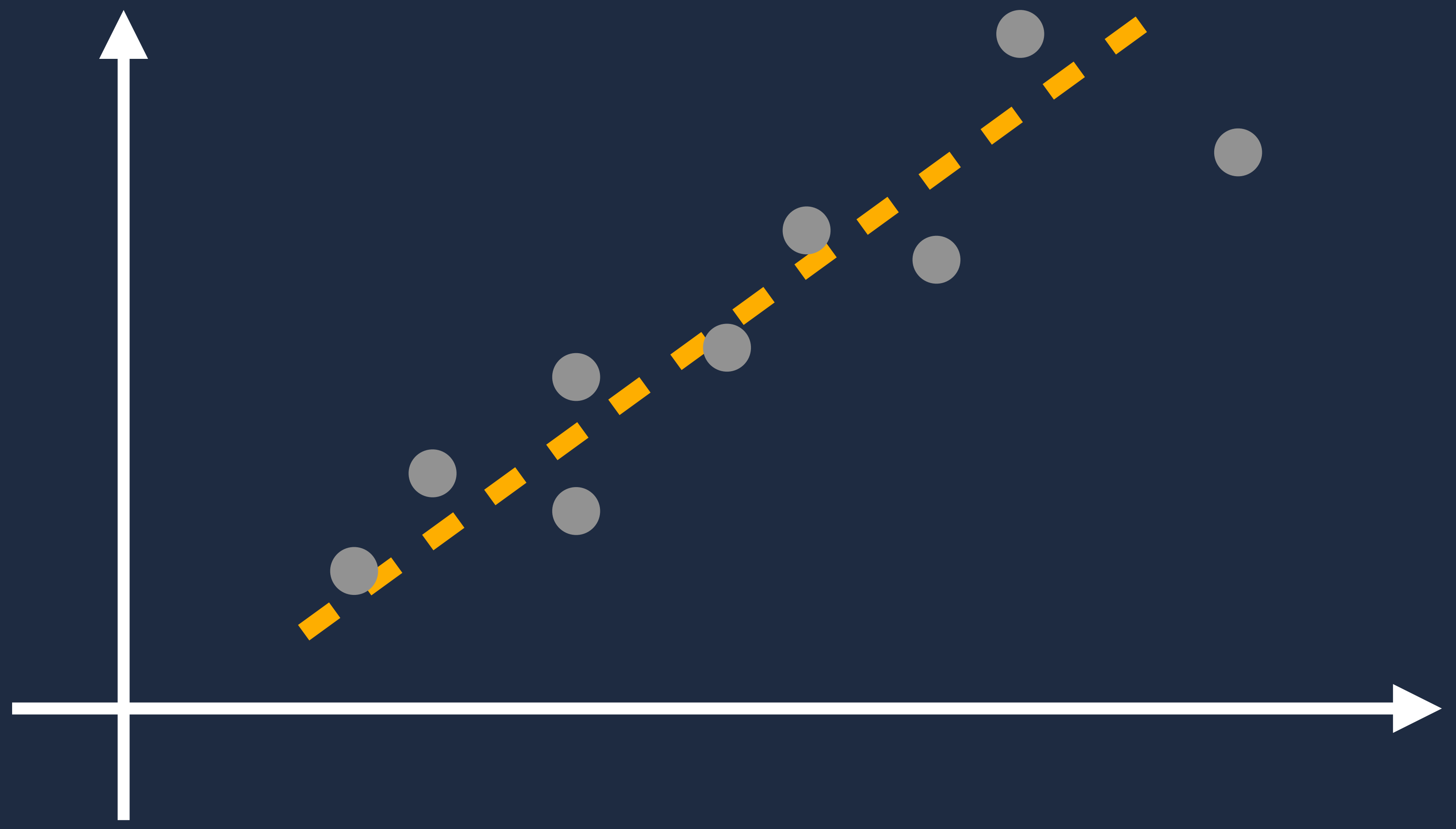
Quantum Linear Algebra

Linear regression problems

Unknown function $y = f(x)$

Linear approximation $\tilde{y} = \vec{w} \cdot \vec{x} + b$

Define a loss function $\mathscr{L}(\vec{w}, b) = \sum_{i=1}^{M} (\tilde{y}_i - y_i)^2$



Matrix form $\mathscr{L}(\vec{\theta}) = (X\vec{\theta} - \vec{y})^2$

$$\vec{\theta} = \begin{bmatrix} w_1 \\ \vdots \\ w_d \\ b \end{bmatrix} \qquad X = \begin{bmatrix} x_1^{(1)} & \cdots & x_d^{(1)} & 1 \\ \vdots & \ddots & \vdots & \vdots \\ x_1^{(M)} & \cdots & x_d^{(M)} & 1 \end{bmatrix} \qquad \vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_M \end{bmatrix}$$

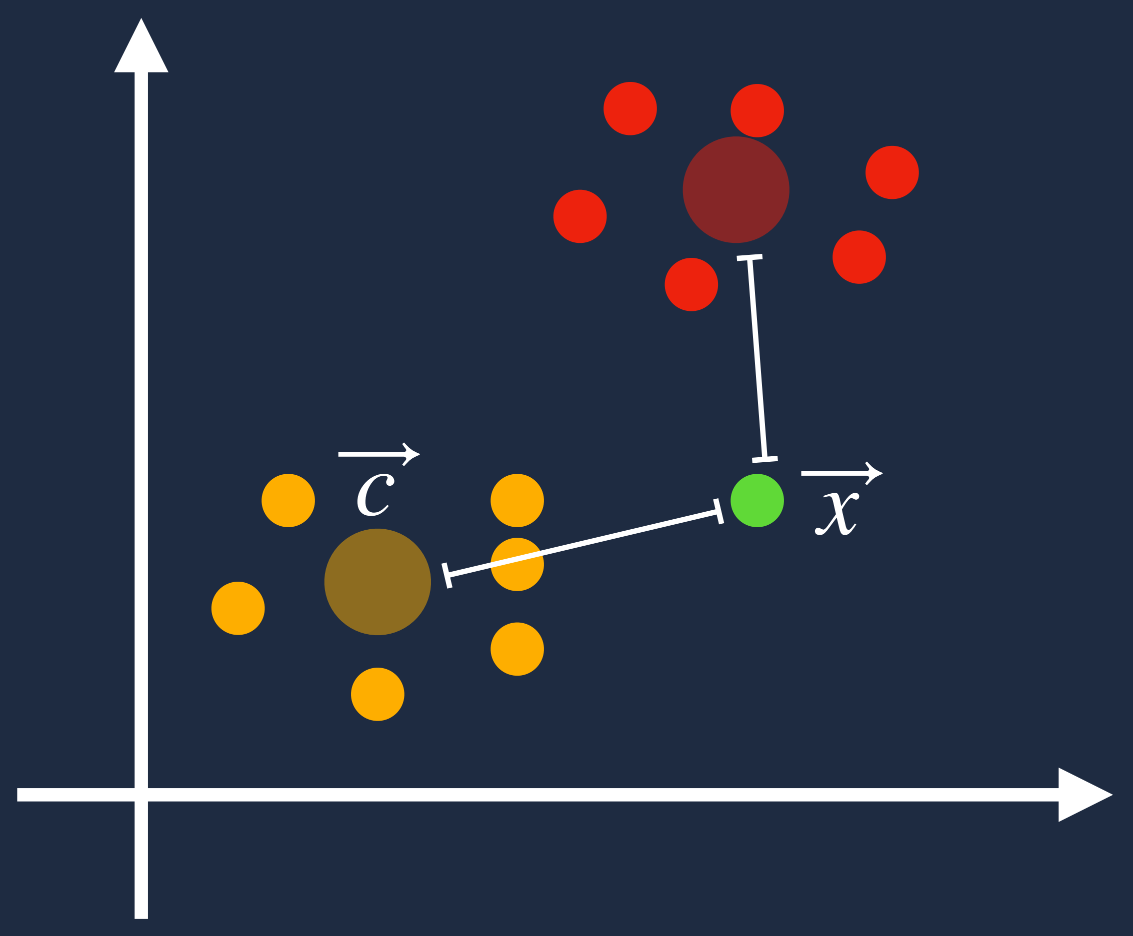
Optimization $\dfrac{\partial \mathscr{L}(\vec{\theta})}{\partial \vec{\theta}} = 0$ $\qquad \vec{\theta} = \boxed{(X^\dagger X)^{-1}} X^\dagger \vec{y}$

HHL algorithm for matrix inversion!

Biamonte, J., Wittek, P., Pancotti, N. *et al.* Quantum machine learning. *Nature* **549,** 195–202 (2017).

Quantum Linear Algebra

Nearest neighbors



$$\vec{x} \in \mathbb{R}^N \to |x\rangle = \sum_{j=0}^{n} \frac{x_j}{|\vec{x}|} |j\rangle$$

$$\vec{c} = \frac{1}{M} \sum_{i=1}^{M} \vec{v}_i \to |c\rangle = \sum_{j=0}^{n} \frac{c_j}{|\vec{c}|} |j\rangle$$

$$n = \log N$$

Amplitude encoding

(with normalization)

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0, x\rangle + |1, c\rangle) \qquad |\phi\rangle = \frac{1}{\sqrt{|\vec{x}|^2 + |\vec{c}|^2}}(|\vec{x}||0\rangle - |\vec{c}||1\rangle)$$

SWAP test

$$\sqrt{|\vec{x}|^2 + |\vec{c}|^2}\,\left|\langle\psi|\phi\rangle\right|^2 = |\vec{x} - \vec{c}|^2$$

Classical $O(\text{poly} MN)$

Quantum $O(\log MN)$

Fast Scalar product!

QA based on Fast Linear Algebra:

Quantum PCA

Quantum SVM

Quantum clustering

Quantum data fitting

...

Drawbacks:

Not suited for NISQ

Requires high resources

Strong limits of applicability

Biamonte, J., Wittek, P., Pancotti, N. *et al.* Quantum machine learning. *Nature* **549,** 195–202 (2017).

Dequantization

Quantum algorithms giving birth to quantum-inspired classical algorithms

Quantum PCA
Quantum SVM
Quantum Supervised Clustering
Quantum Recommendation system
...

Dequantization →

Classical random procedure doing as well up to polynomial overhead

$\vec{x} \in \mathbb{R}^N$

Quantum RAM

$$|x\rangle = \sum_{j=0}^{n} \frac{x_j}{|\vec{x}|} |j\rangle$$

Requires only $n = \log N$ resources

$\vec{x} \in \mathbb{R}^N$

Classical Data Structure
("Sample and Query access")

$$\mathscr{D}_{x_i} = \frac{x_i^2}{\|\vec{x}\|^2}$$

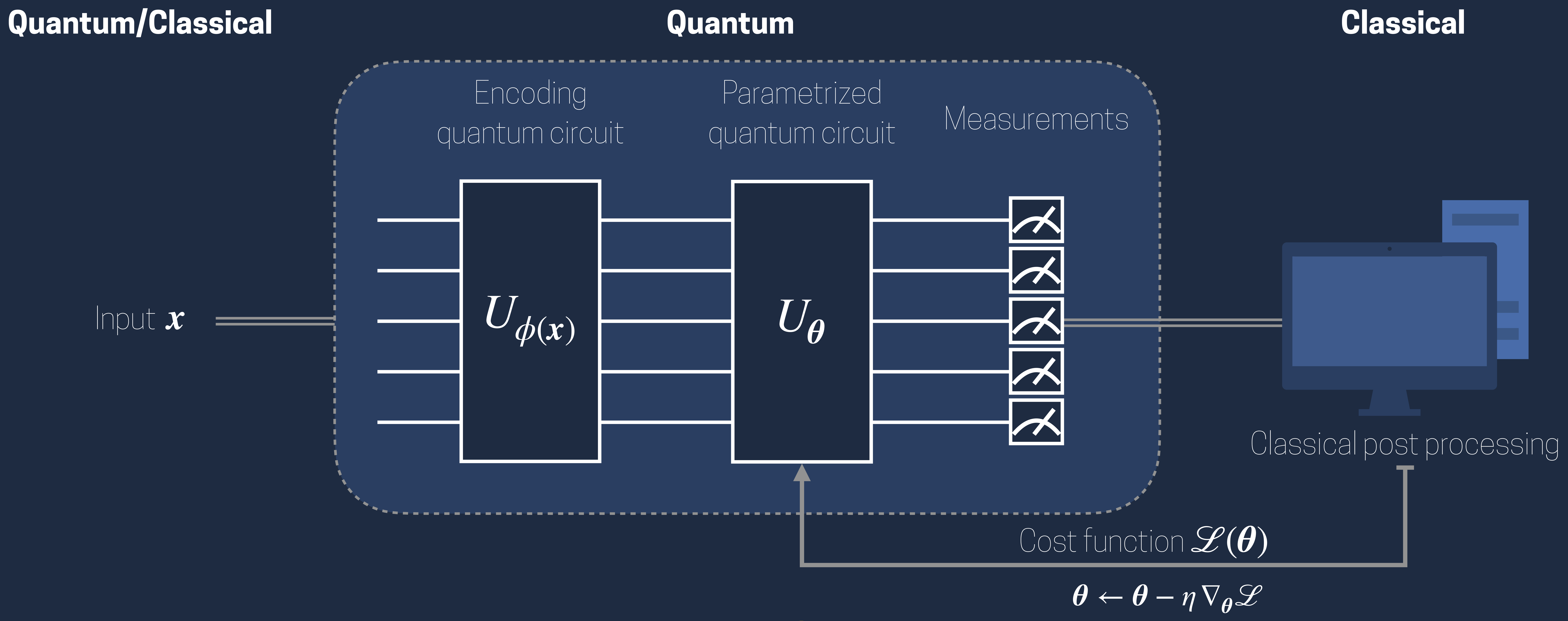
Replaced by a classical sampling procedure
(if conditions are met)

...polynomial speedups still matters.

Tang, E. A quantum-inspired Classical Algorithm for Recommendation Systems, Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing (2019)

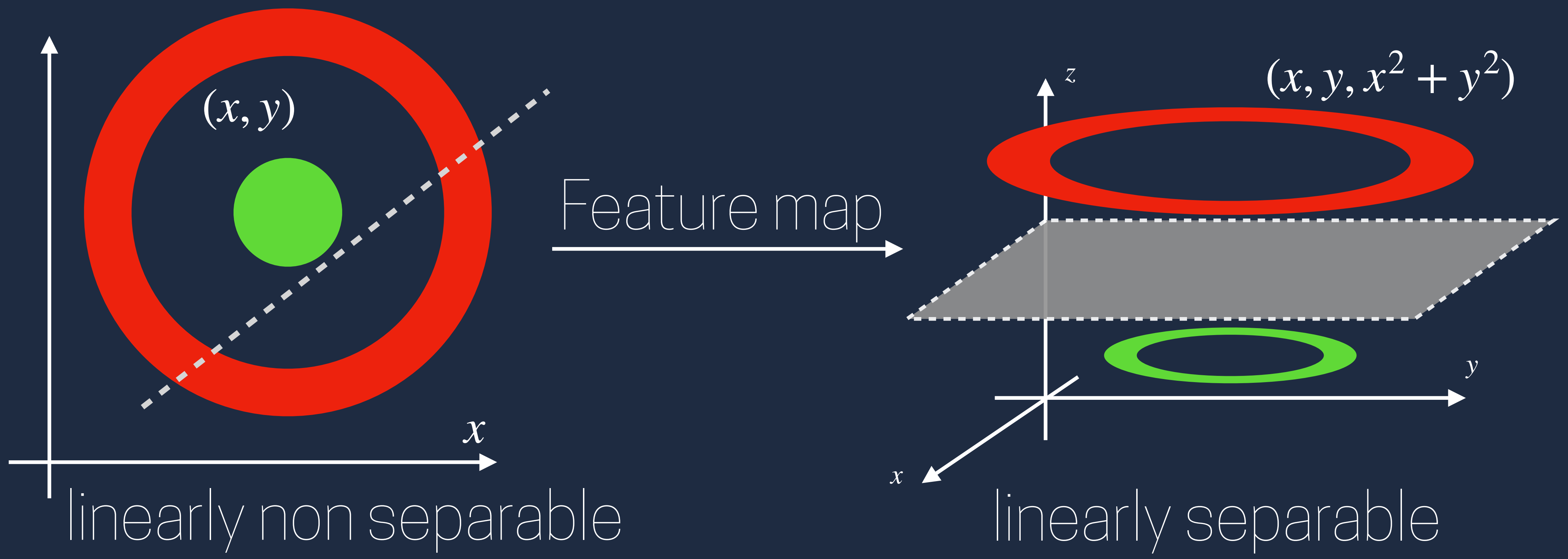
Hybrid models

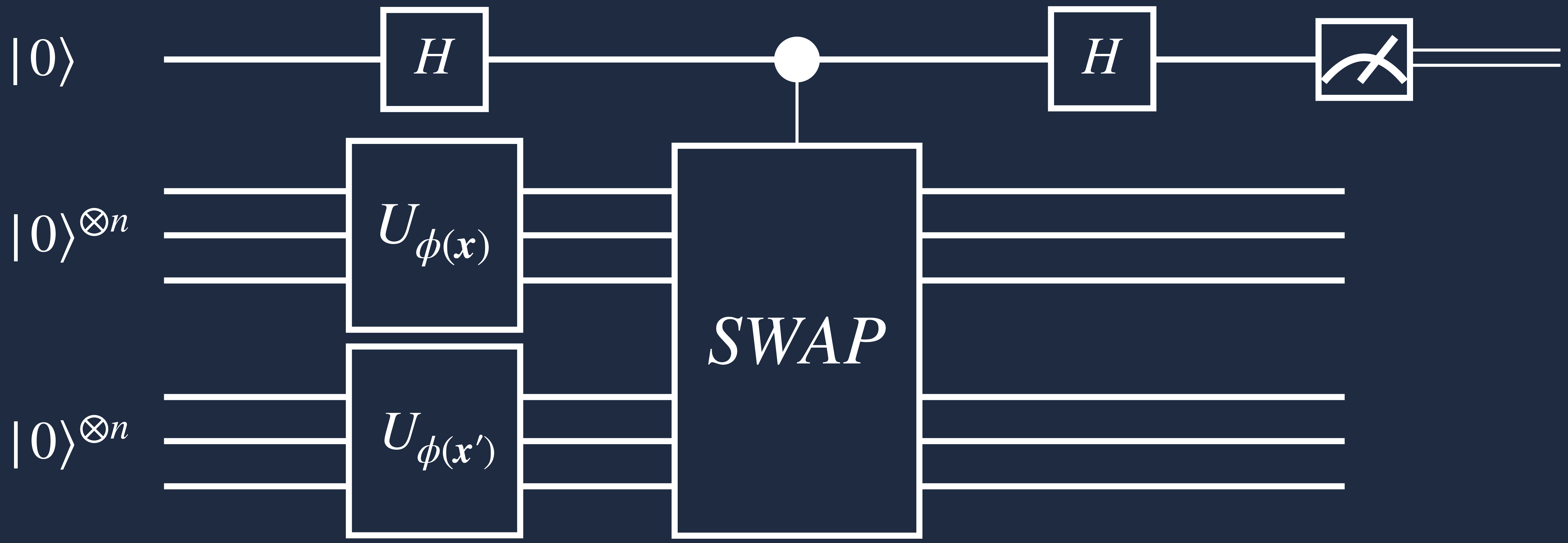
In the NISQ era, a promising way is to use hybrid quantum-classical learning models



Benedeti, M., *et al.* Parametrized quantum circuits as learning models. *Quantum Sci. Technol.* **5** 019601 (2019)

Kernel methods





$$P(0) = \frac{1}{2} + \frac{1}{2}\left|\langle \phi(\boldsymbol{x}') \,|\, \phi(\boldsymbol{x}) \rangle\right|^2$$

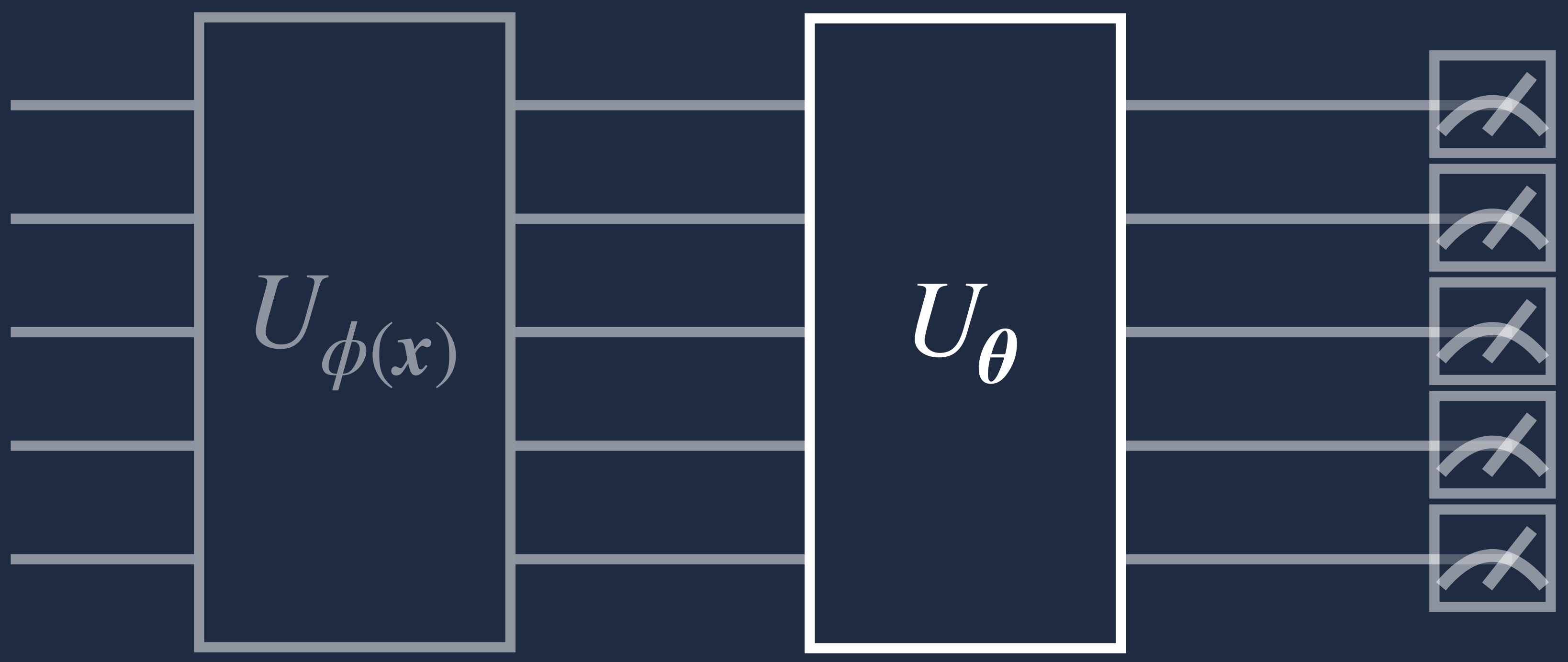
$$= \frac{1}{2} + \frac{1}{2} \underbrace{\left|\langle \boldsymbol{0} \,|\, U^{\dagger}_{\phi(\boldsymbol{x})} U_{\phi(\boldsymbol{x})} \,|\, \boldsymbol{0} \rangle\right|^2}_{\text{Quantum Kernel function } \mathscr{K}(\boldsymbol{x}, \boldsymbol{x}')}$$

Quantum advantage

kernels which are difficult to simulate classically

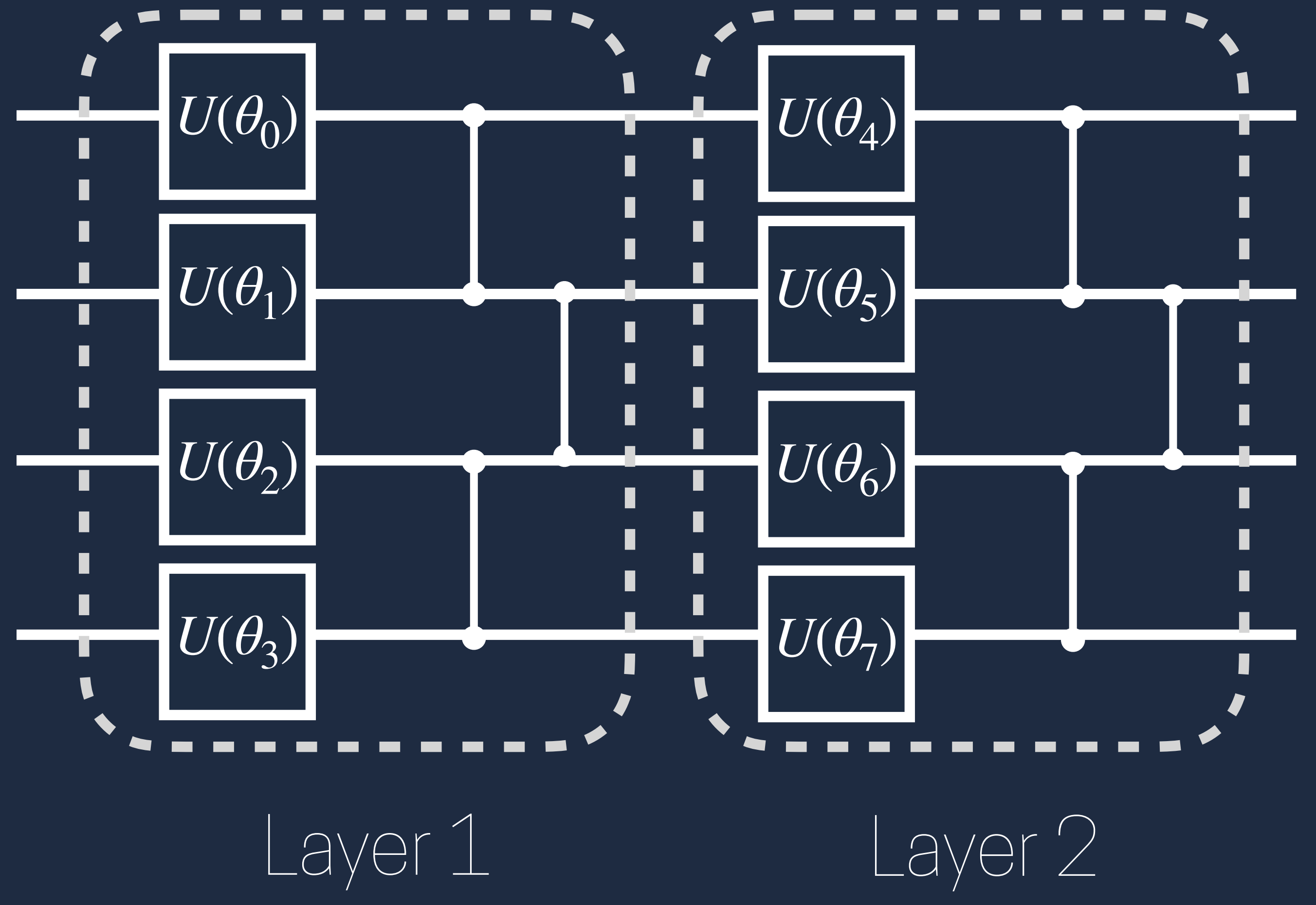
Variational Quantum Models

Classification performed by the parametrized quantum circuit



Variational Ansatz

$U(\boldsymbol{\theta}) =$



Until condition is met, repeat:

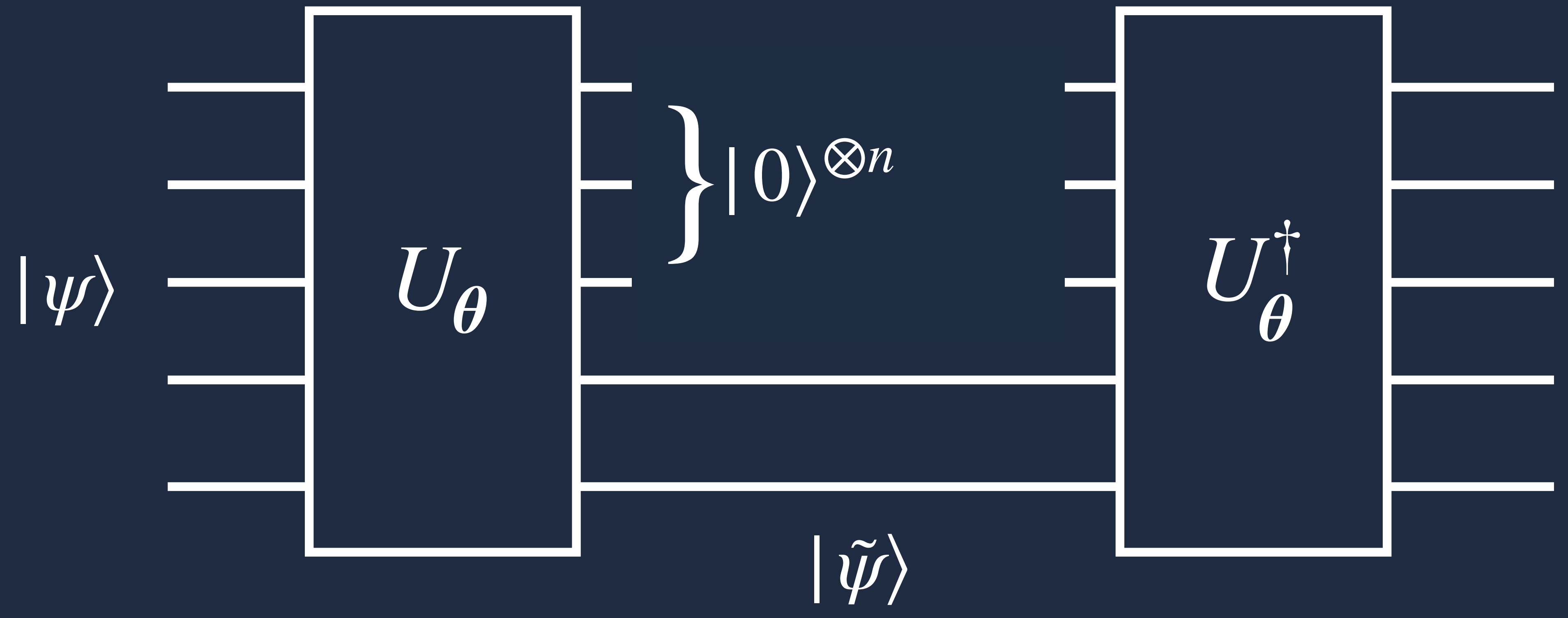
Measurement outcomes $\{\langle M_k \rangle_{\boldsymbol{x},\boldsymbol{\theta}}\}_k$

Evaluate cost function $\mathscr{L}(\langle M_k \rangle_{\boldsymbol{x},\boldsymbol{\theta}})$

Update variational parameters $\boldsymbol{\theta} \leftarrow \boldsymbol{\theta} - \eta \nabla_{\boldsymbol{\theta}} \mathscr{L}$

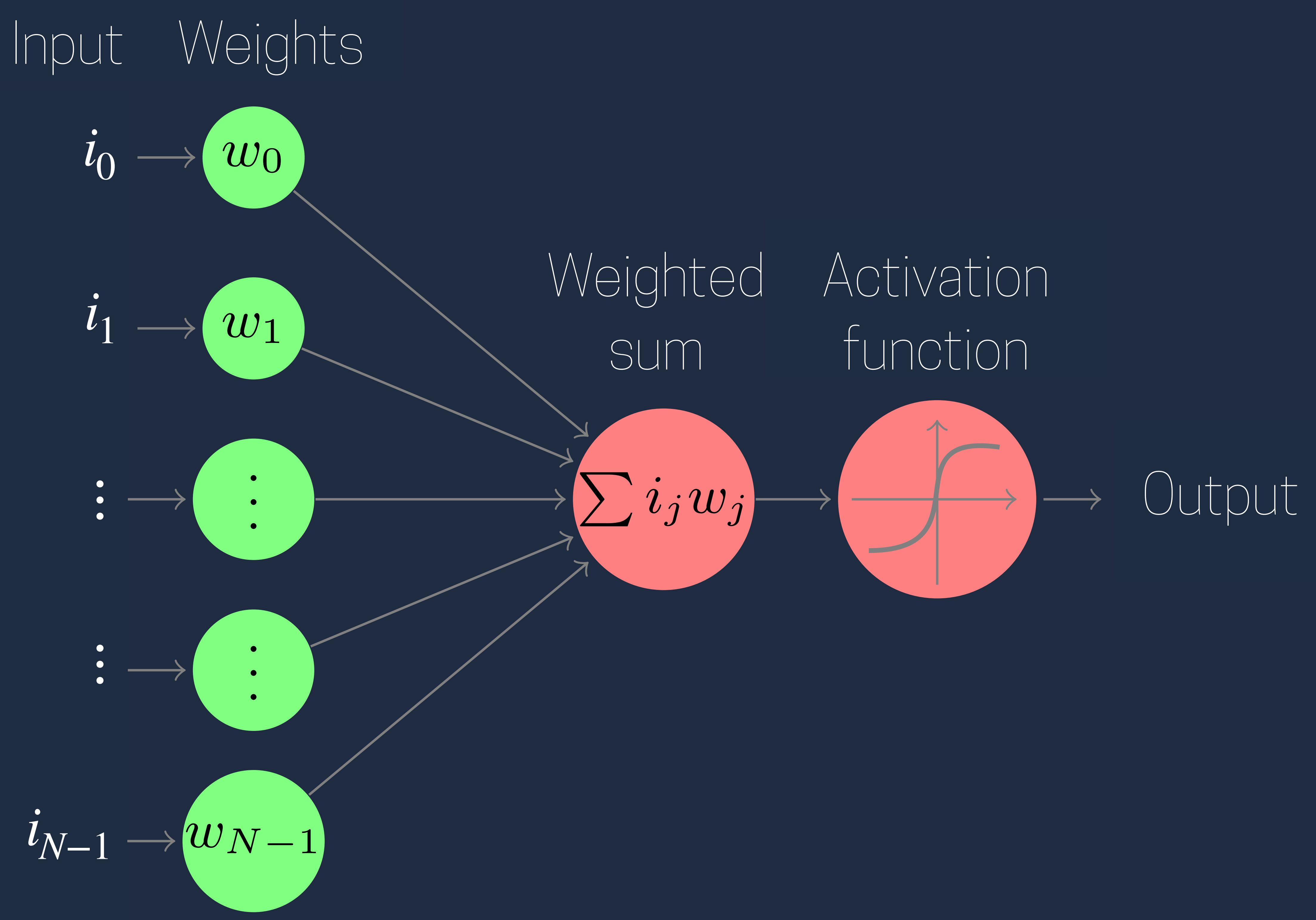
Note: Gradients by numerical methods (SPSA),
Parameter Shift rules, "Barren Plateaus"

Quantum Variational Autoencoder

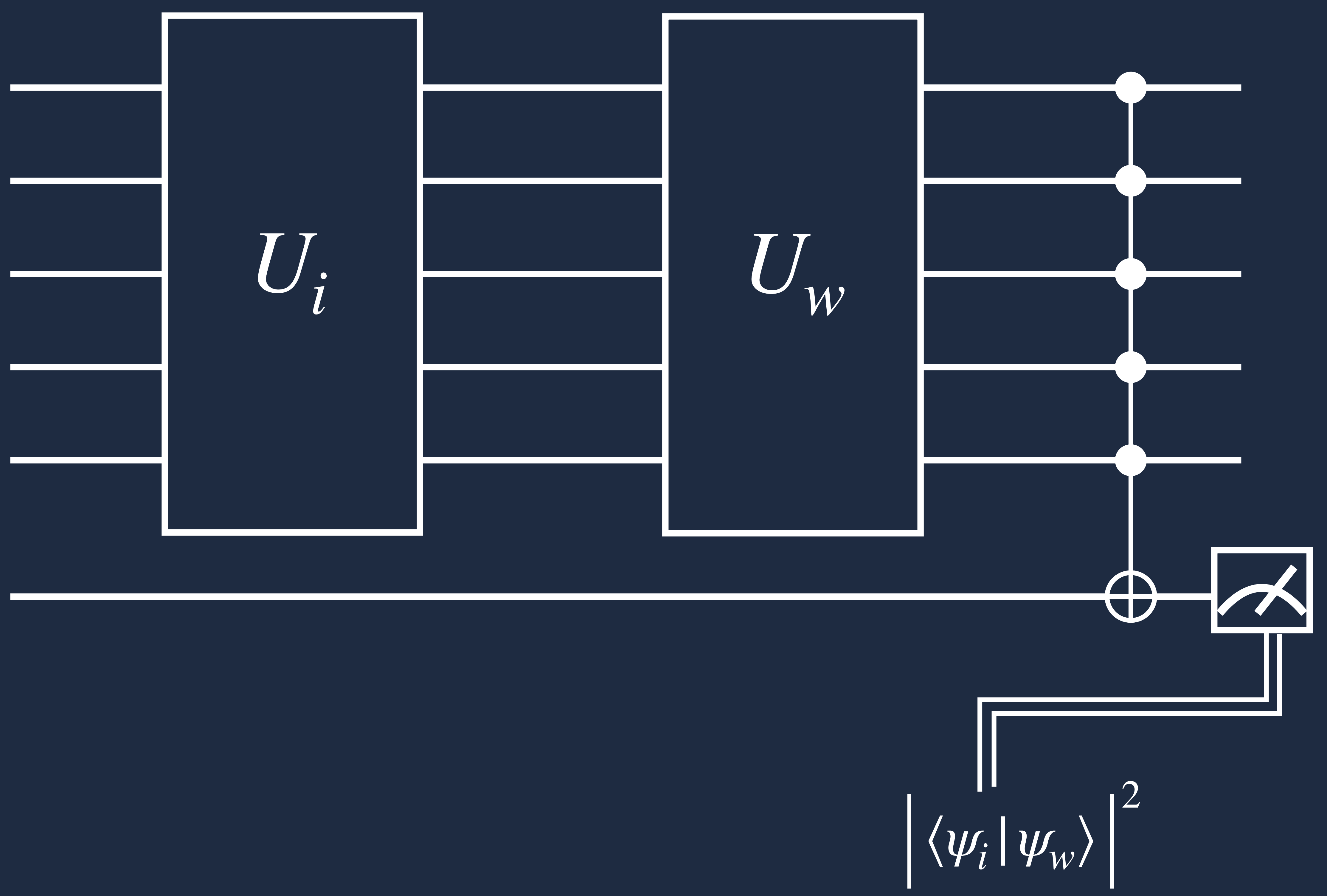


Here in Pavia

Quantum model of neurons



LME states $|\psi_i\rangle = \frac{1}{\sqrt{2^n}} \sum_{j=0}^{N-1} e^{i\alpha_j} |j\rangle$



Take Home Message

Quantum Machine Learning, as well as Quantum Computing, promise to greatly enhance computational tasks.

Quantum-enhanced machine learning

Faster linear algebra,
Parametrized quantum circuits,
Creation of quantum-inspired algorithms, ...

Quantum tomography,
Quantum simulation,
Quantum control, ...

Quantum-applied machine learning

However

Yet, no real actual real speed up
Much still do be done

Extra 1\ *Dequantization*

Quantum algorithms giving birth to quantum-inspired classical algorithms

Recommendation systems

$T \in \mathbb{R}^{n \times m}$ $\quad T = \begin{bmatrix} T_1^{(1)} & T_1^{(2)} & \cdots & T_1^{(m)} \\ \vdots & \ddots & & \vdots \\ T_n^{(1)} & T_n^{(2)} & \cdots & T_n^{(m)} \end{bmatrix} \longleftarrow$ User

Preferences

Quantum Recommendation System

$O(\text{poly}(k)\ \text{polylog}(mn))$

$\vec{x} \in \mathbb{R}^N$

Classical input data

$$|x\rangle = \sum_{j=0}^{n} \frac{x_j}{|\vec{x}|} |j\rangle$$

Requires only $n = \log N$ resources

$$\mathcal{D}_{x_i} = \frac{x_i^2}{\|\vec{x}\|^2} \longrightarrow$$

Replaced by a classical sampling procedure

(if conditions are met)

Dequantization!

$O(\text{poly}(k)\ \text{polylog}(mn))$

Extended to:

Supervised clustering

Quantum PCA

...

...polynomial speedups still matters.

Tang, E. A quantum-inspired Classical Algorithm for Recommendation Systems, Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing (2019)

Extra 2\ GPT-3

GPT-3 is a model for **N**atural **L**anguage Processing (NLP) capable of interpreting and forming sentences

Key facts:

175B	355y	\$4.6 Milion
Parameters	Training time	Training cost

The Guardian

A robot wrote this entire article. Are you scared yet, human?

We asked GPT-3, OpenAI's powerful new language generator, to write an essay for us from scratch. The assignment? To convince us robots come in peace

- For more about GPT-3 and how this essay was written and edited, please read our editor's note below

Quantum systems produce atypical patterns that classical systems are thought not to produce efficiently, so it is reasonable to postulate that quantum computers may outperform classical computers on machine learning tasks.

The field of quantum machine learning explores how to devise and implement quantum software that could enable machine learning that is faster than that of classical computers.

$175 \cdot 4 \cdot 10^9 = 700GB$

$\dim \mathscr{H} = 2^n$

only $n = 43$ qubits!

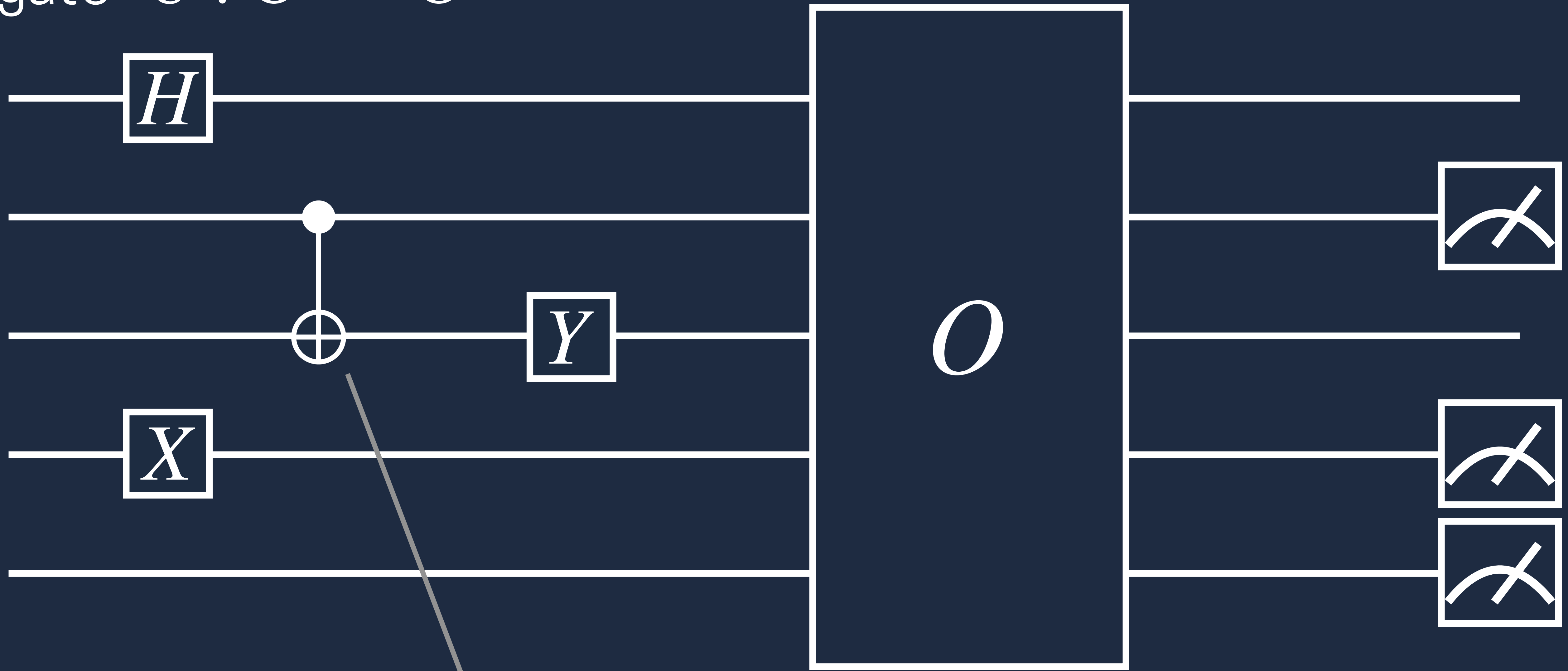
Extra3\ Quantum circuit model

A box represent a unitary **quantum gate** $U : \mathbb{C}^2 \to \mathbb{C}^2$

Quantum gate acting on all qubits in the circuit

Each line represent a **qubit**

$|\psi\rangle \in \mathbb{C}^2$



Measurement in the computational basis $\{|0\rangle, |1\rangle\}$

$$CNOT = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$|00\rangle \to |00\rangle \quad |10\rangle \to |11\rangle$

$|01\rangle \to |01\rangle \quad |11\rangle \to |10\rangle$

Superposition

$|0\rangle$

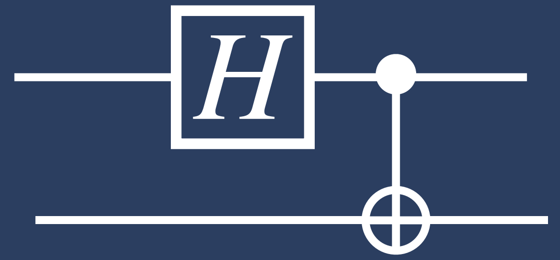
$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$

$$H|0\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Entanglement (Bell state)

$|0\rangle$

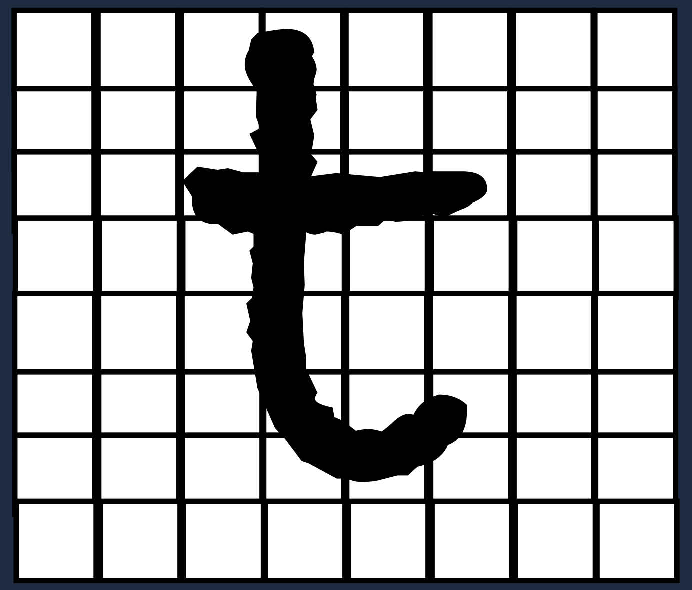
$|0\rangle$



$|\psi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$

Extra 4\ Quantum Learning Theory

Study the theoretical aspects of quantum machine learning, and results are framed in the language of **Co**mputational **L**earning **T**heory (COLT).



Learner $\mathscr{A}$

Concept $c : \{0,1\}^n \rightarrow \{0,1\}$ ⟵ recognize letter "t"

Concept class $\mathscr{C} = \{c \,|\, c : \{0,1\}^n \rightarrow \{0,1\}\}$ ⟵ recognize all letters

Probably Approximately Correct (PAC) Learning:

Learner $\mathscr{A}$ Oracle $P(c, D) \longrightarrow$ Example $(x, c(x))$

with probability $1 - \delta$ $\Pr_{x \sim D}[h(x) \neq c(x)] < \epsilon$

Quantum PAC $\sum_x \sqrt{D(x)}\,|x, c(x)\rangle$

Results:

- Disjunctive Normal Forms (DNF) are efficientlyQuantum PAC-learnable faster than classically
- Concept classes built upon factorization, are learnable exponentially faster with quantum resources (Shor)

Arunachalam, S. & de Wolf, R. A survey on Quantum Learning Theory, arXiv:1701.06806 (2017)