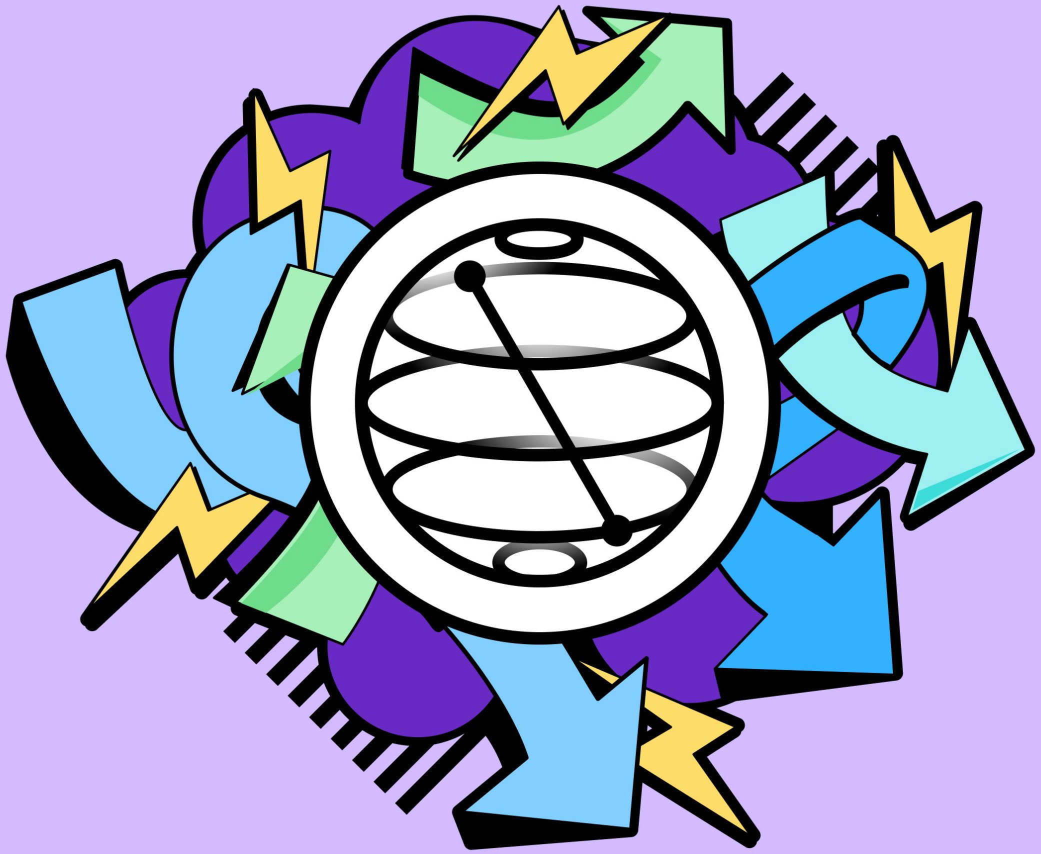
Qiskit | Fall Fest

The capacity and power of Quantum Machine Learning



Nouhaila Innan*, Walid EL Maouki

*PhD Student at Hassan II University of Casablanca

What Makes QML Quantum ?

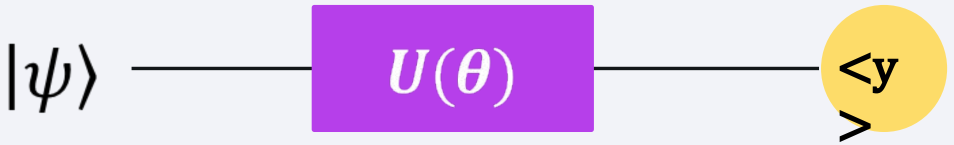
- The algorithm itself runs on a quantum computer.

- The data operates on quantum or come from a quantum state/process.

QML Models

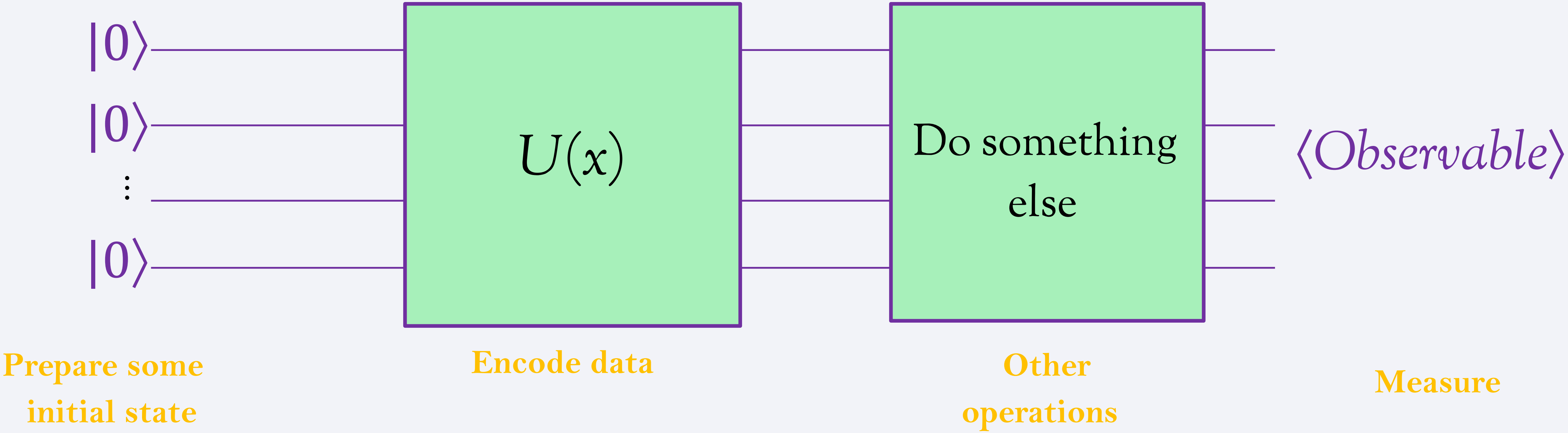
- **Deterministic quantum models**
  - ❑Deutsh-Jozsa algorithm

- **Variational quantum models :**
  - ❑Variational quantum eigensolver
  - ❑Variational classifier
  - ❑Quantum support vector machines
  - ❑Quantum neural networks
  - ❑...



QML Algorithms

In general...



What does capacity mean?

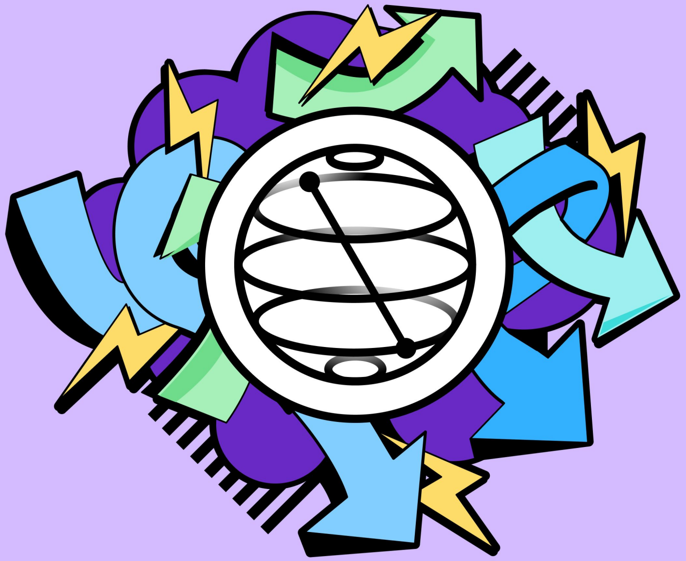
Statistical complexity

Expressivity

Power

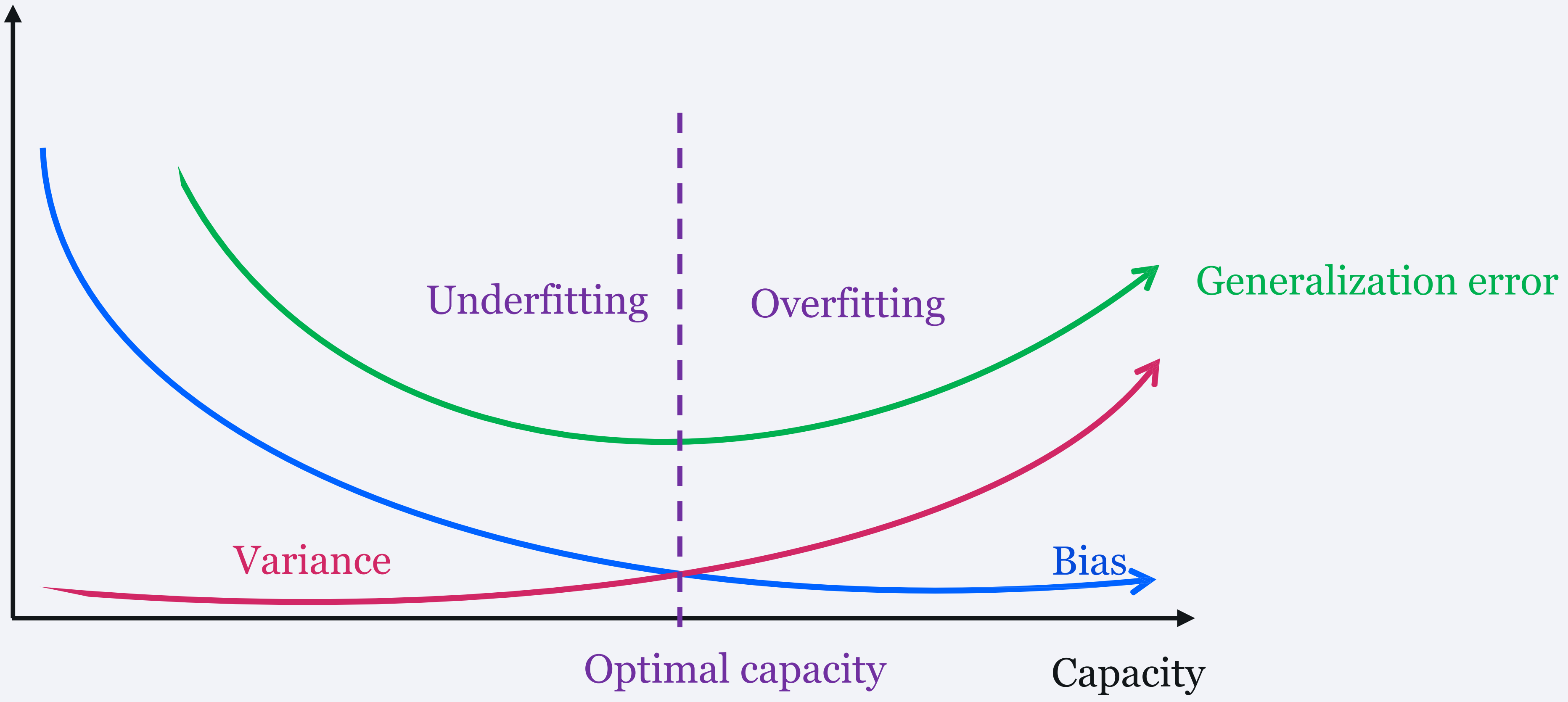
How many functions can my model approximate?

Is a higher capacity always better?!



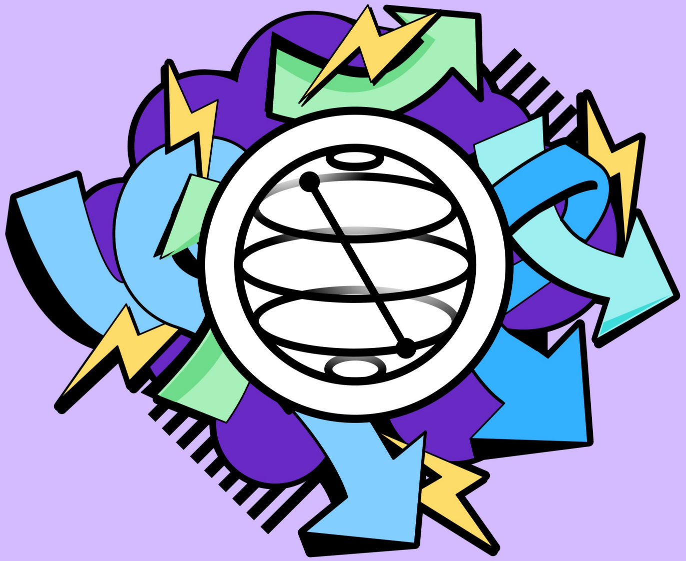
Generalization

We can understand capacity in the context of generalization performance



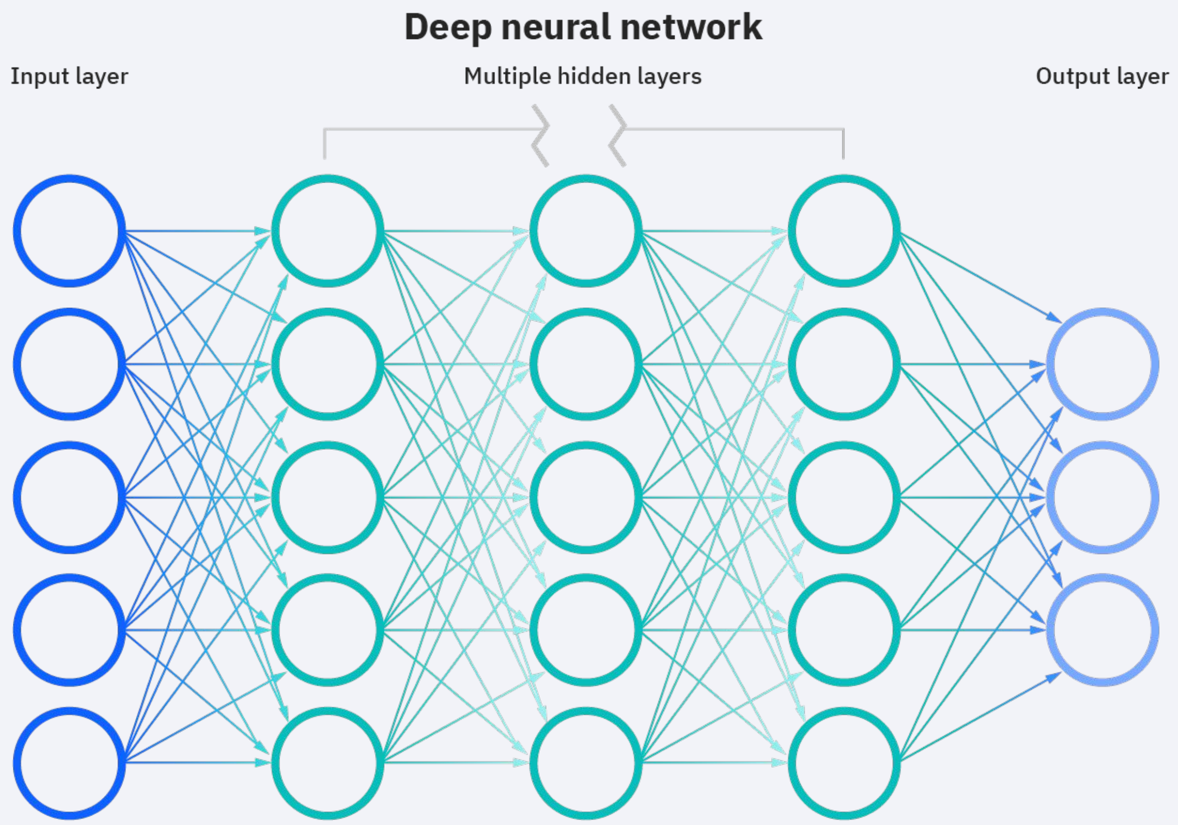
How do we measure capacity?

(Classically)



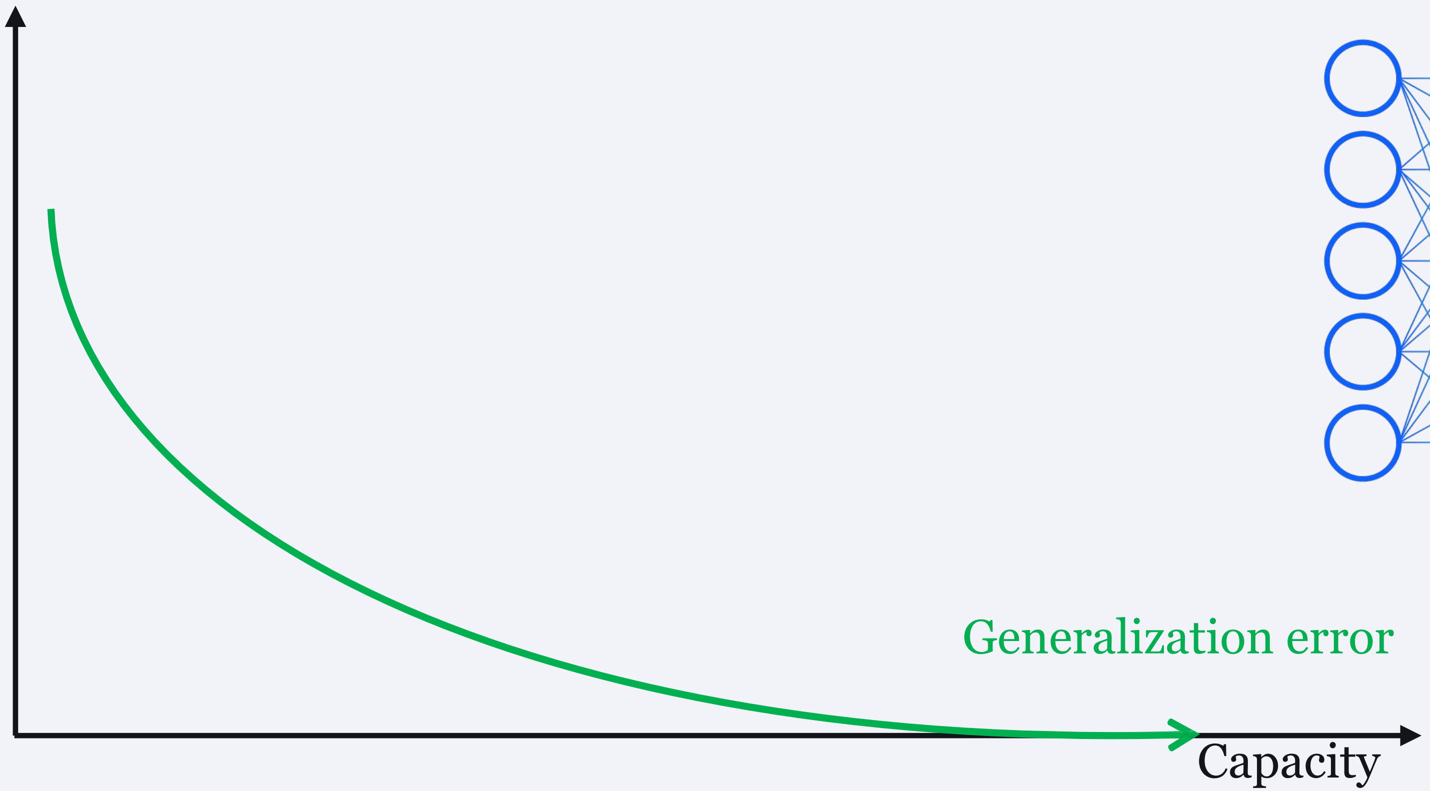
Capacity measures

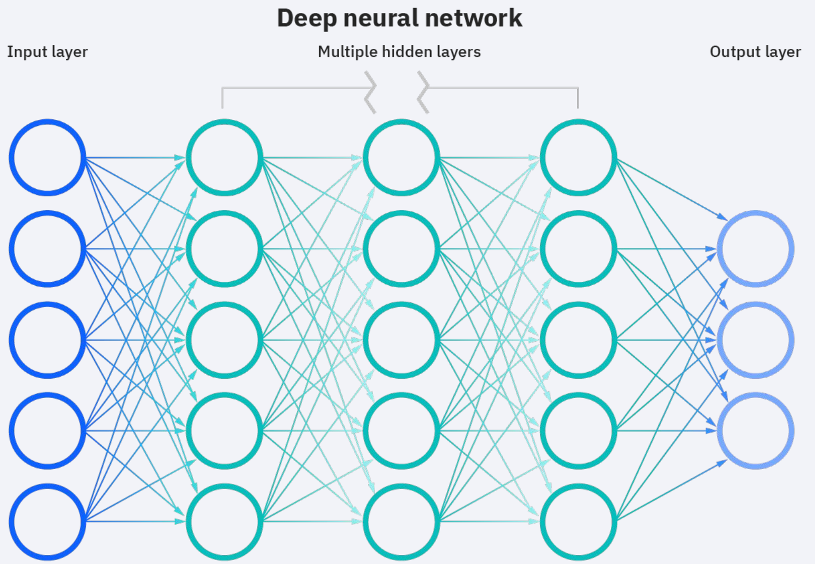
As a start, making your model larger (adding more parameters, increasing the depth) was assumed to add to capacity



Deep neural networks

A common misconception or seemingly paradoxical behavior was that DNNs do not overfit

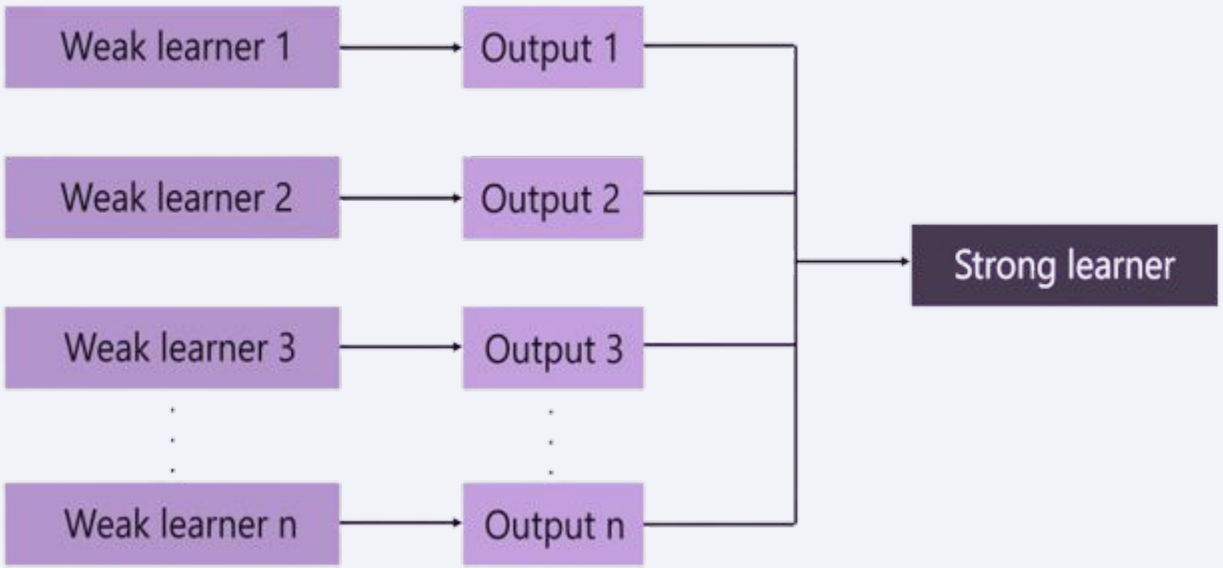




Capacity measures

Distribution of margins

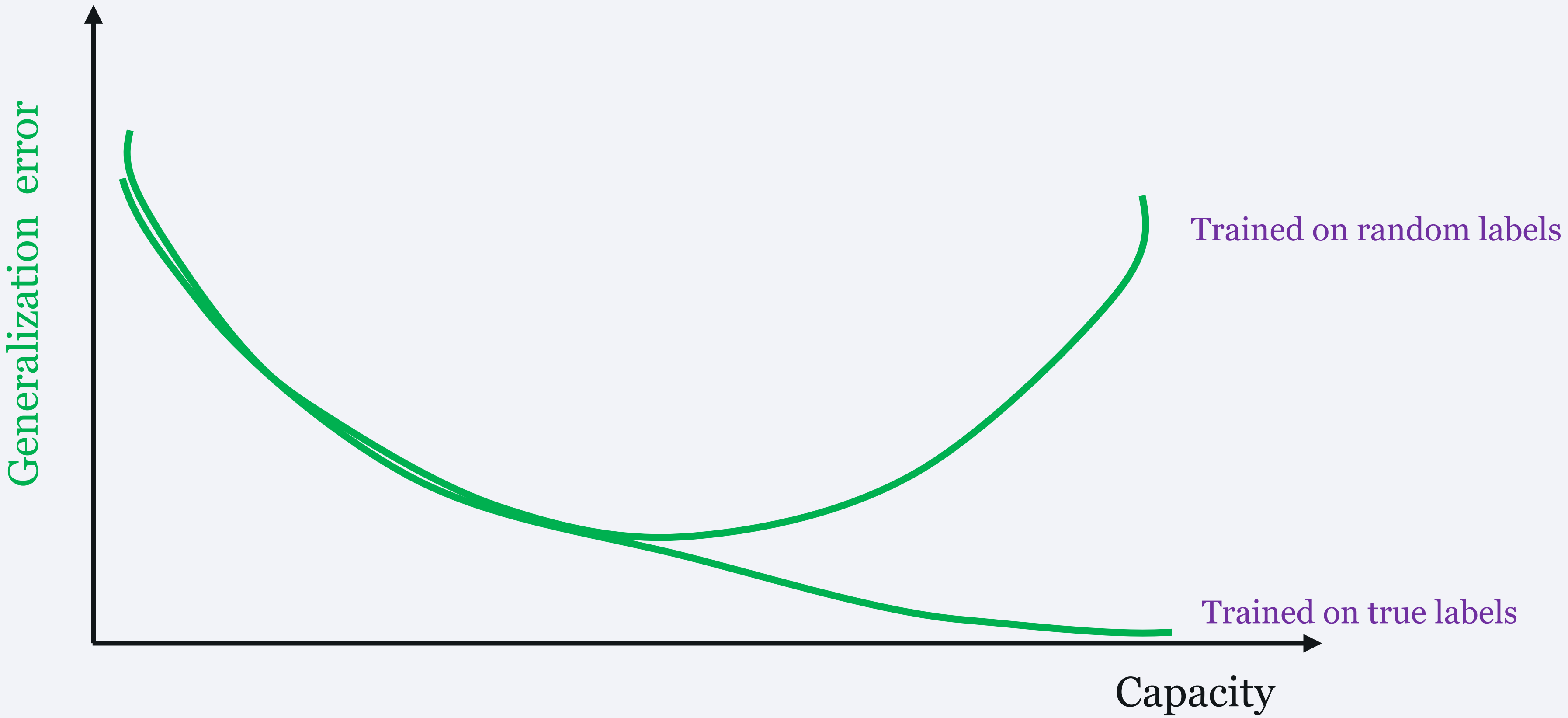
- Linear classifiers that produce large margins (Vapnik, V., & Chervonenkis, A. (1974). Theory of pattern recognition.)
- Explained the “paradox” observed in boosting methods (Bartlett, P., Freund, Y., Lee, W. S., & Schapire, R. E. (1998). Boosting the margin: A new explanation for the effectiveness of voting methods. The annals of statistics, 26(5), 1651- 1686.)



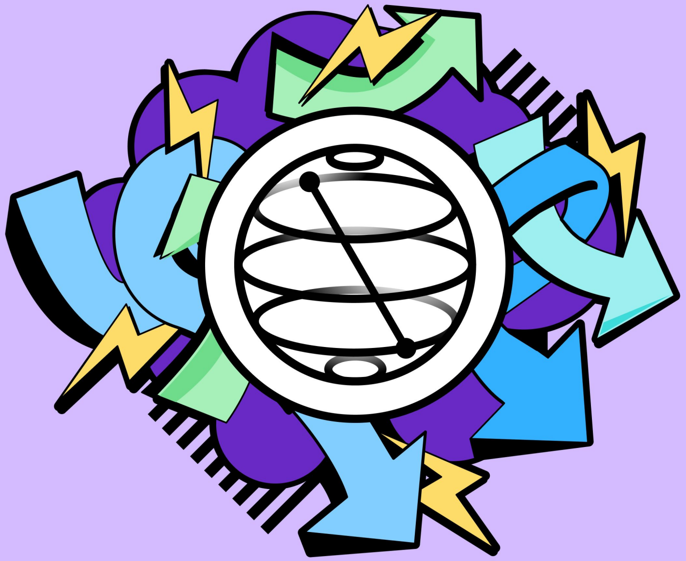
Capacity measures

Qiskit | Fall Fest

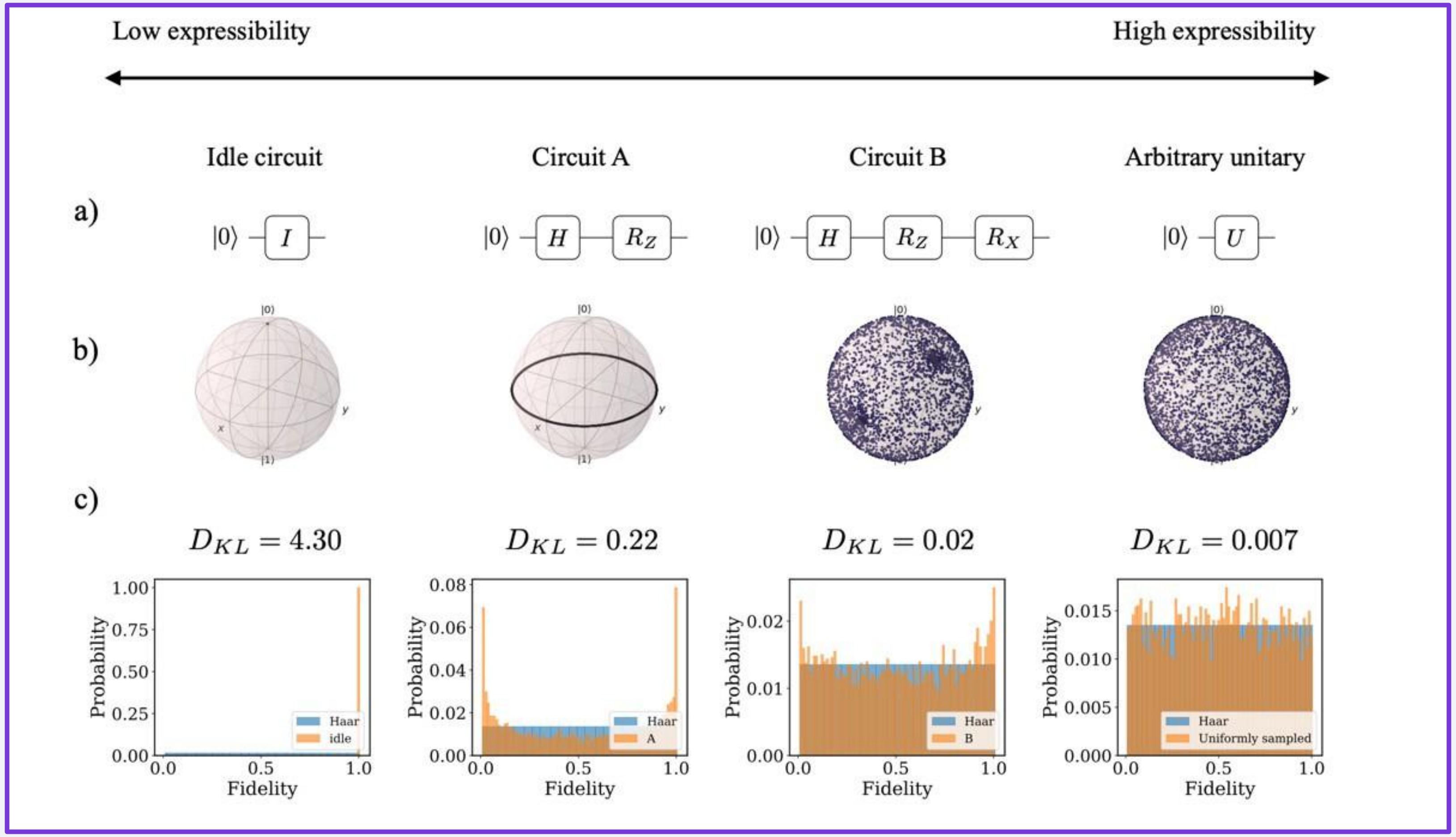
- Unfortunately, margins don't seem to work with neural networks
- The margin does not inform us about generalization behavior when we consider the randomized experiment example by Zhang et al.



Capacity of quantum models!

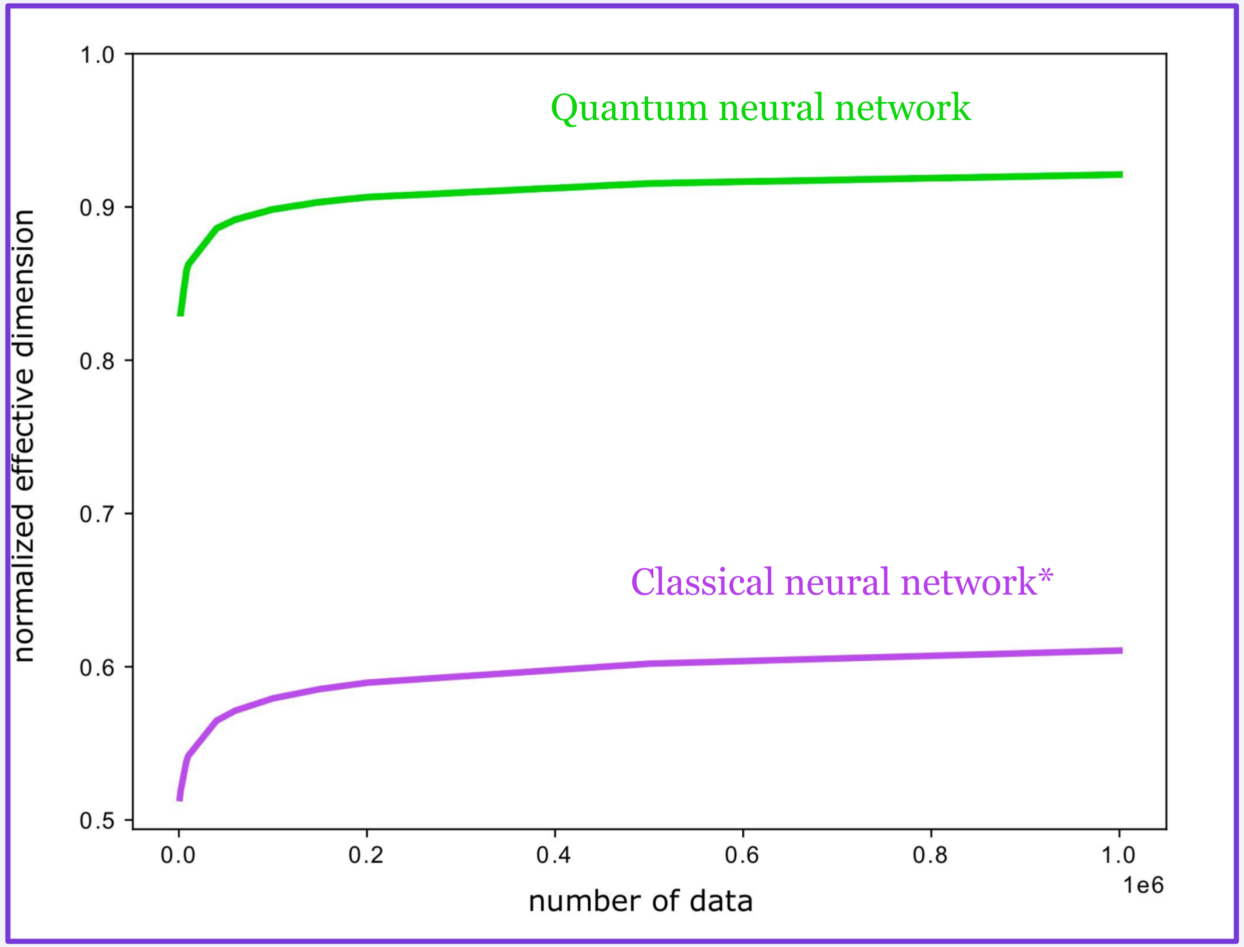


Variational circuits



Sim et al. "Expressibility and Entangling Capability of Parameterized Quantum Circuits for Hybrid Quantum-Classical Algorithms." Advanced Quantum Technologies 2.12 (2019): 1900070.

Effective Dimension



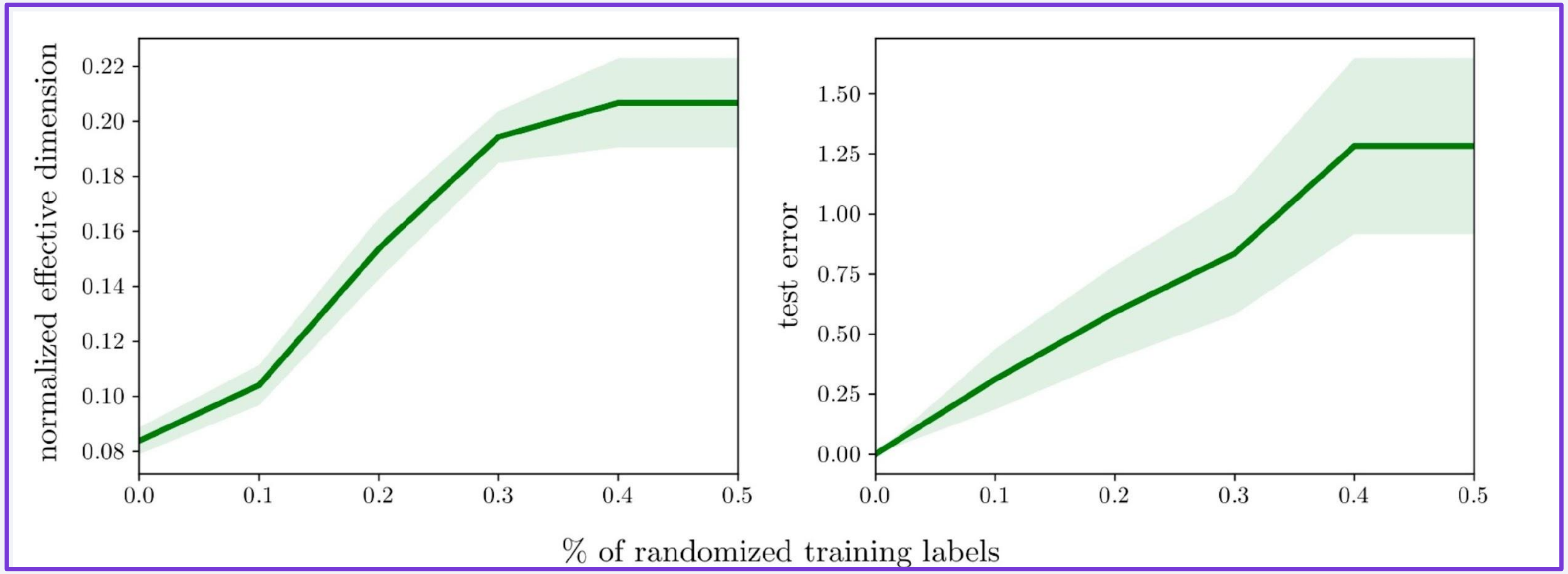
**Chose the best possible classical ffnn out of all possible configurations*

Abbas, A., Sutter, D., Zoufal, C., Lucchi, A., Figalli, A., & Woerner, S. (2020). The power of quantum neural networks. arXiv preprint arXiv:2011.00027.

Effective dimension and generalization

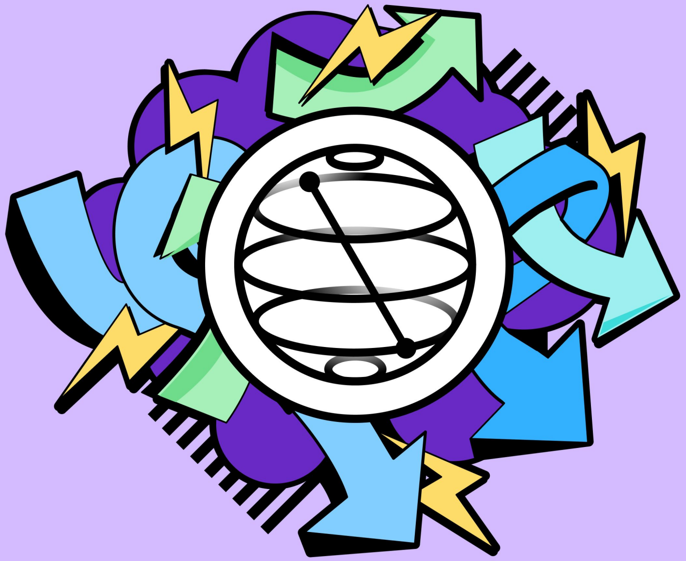
Qiskit | Fall Fest

Found that the effective dimension for a model trained on confusion sets with increasing label corruption accurately captures generalisation behaviour



Abbas, A., Sutter, D., Zoufal, C., Lucchi, A., Figalli, A., & Woerner, S. (2020). The power of quantum neural networks. arXiv preprint arXiv:2011.00027.

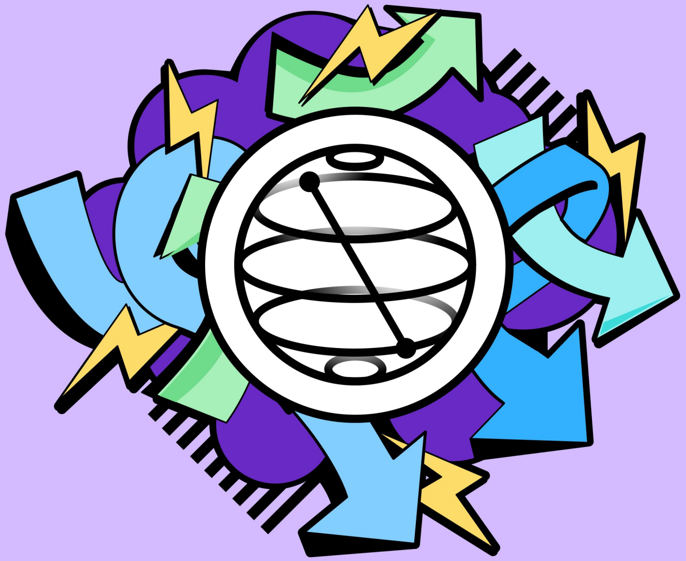
How will QML models generalize?



Where can we hope for an advantage?

Capacity?	Generalization?	Data?
Computational?	Quantum kernels?	Training?
Statistical?	Variational circuits?	Applications?

QML challenges



Variational circuits steps

Task: Train a quantum circuit on labelled samples in order to predict labels for new data

Step 1: Encode the classical data into a quantum state
Step 2: Apply a parameterized model/ Ansatz
Step 3: Measure the circuit to extract labels
Step 4: Use optimization techniques (like gradient descent) to update model parameters

Step1: Encoding data into quantum state

Encoding classical data into the quantum state is an open problem:

- There is **no generic encoding method** addressing all types of datasets.
- It **depends very much on your problem**. Hence, multiple data encodings are possible.

Each encoding is essentially a trade-off between two major factors:

1. **The number of required qubits.**
2. **The number of operations to prepare the quantum state.**

Basis encoding

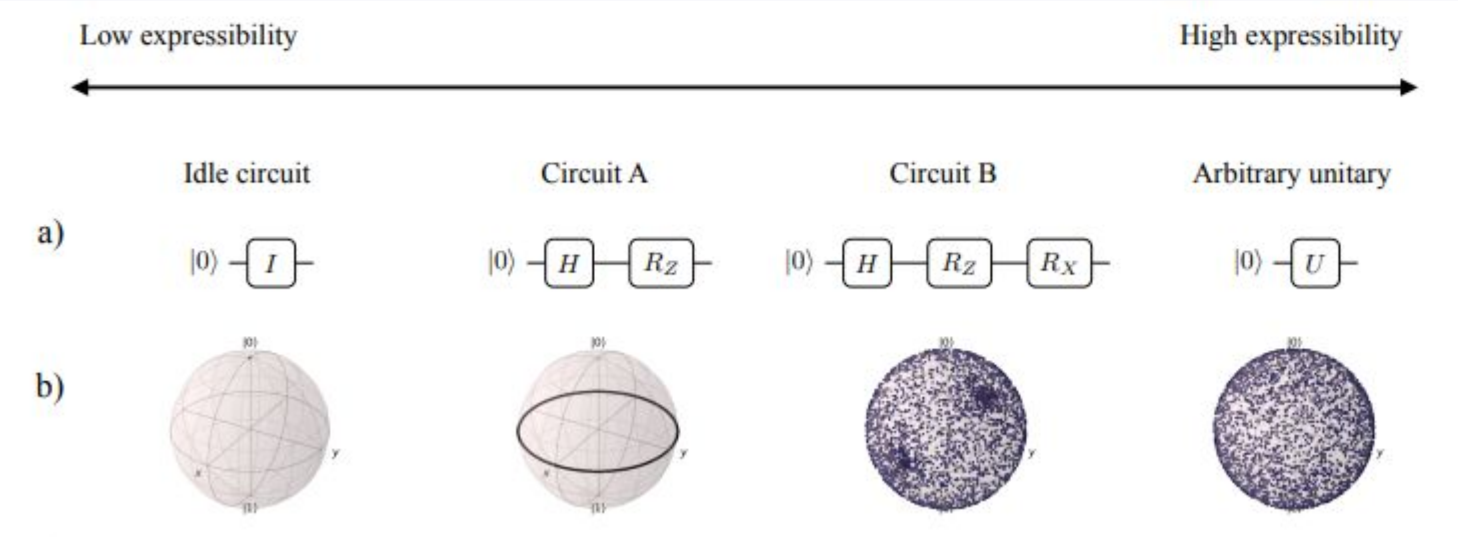
- Pros: suitable for arithmetic computations.
- Cons: **requires a lot of qubits**. A real number is approximated by $(n + k + 2)$ bits, and thus prepared as an $(n + k + 2)$-dimensional quantum state.

Amplitude encoding

- Pros: it requires only $\log_2(n)$ qubits to represent an n-dimensional data point.
- Cons: **Requires a deep circuit to implement** (in terms of no. operations/gates).

Step 2-1: Parametrized model or Ansatz

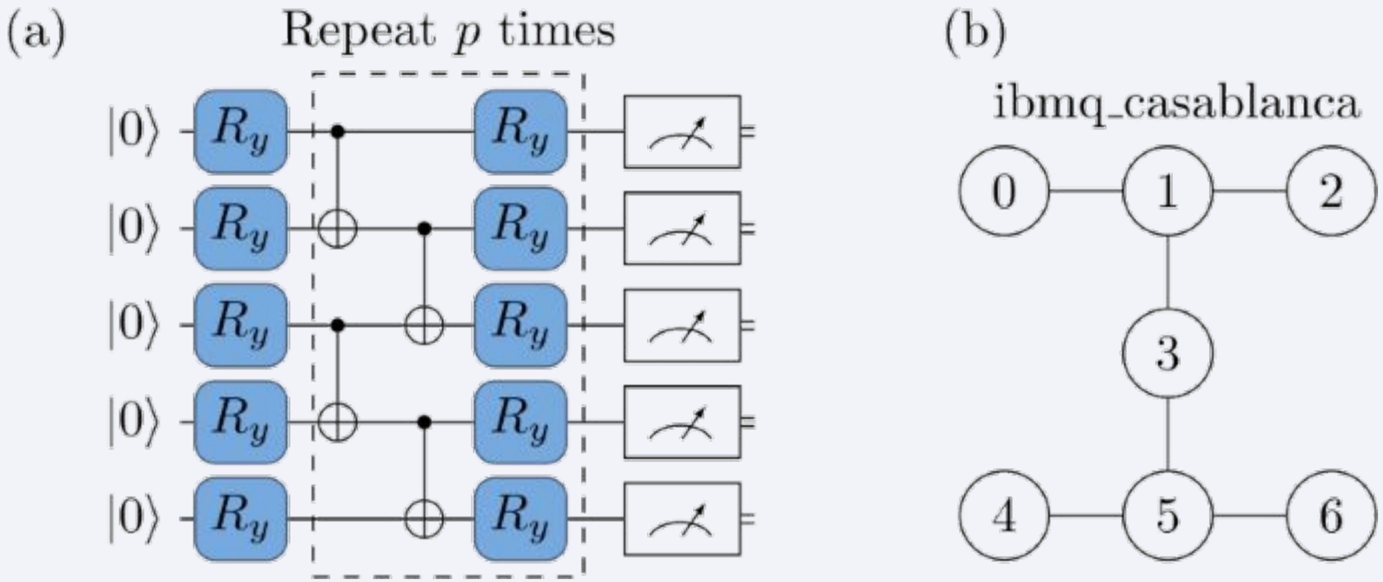
- **Expressibility and entanglement capability**: Expressive ansatz to span enough Hilbert space where our solution state reside



Sim, Sukin et *al*. "Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms." *Advanced Quantum Technologies* 2.12 (2019): 1900070.

Step 2-2: Parametrized model or Ansatz

- heuristic/hardware efficient Ansatz



2-qubit gates are applied only to adjacent qubits.

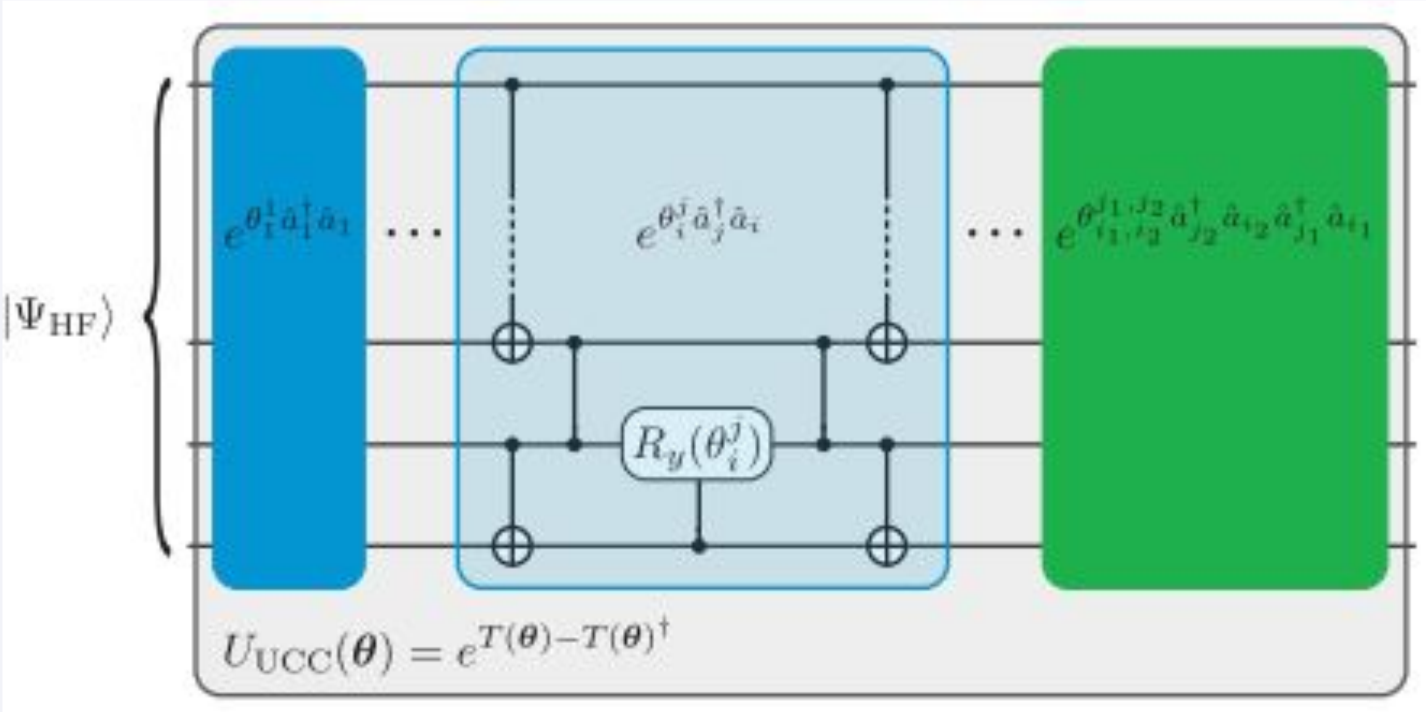
Amaro, *et al.* A case study of variational quantum algorithms for a job shop scheduling problem. *EPJ Quantum Technol.* **9**, 5 (2022). https://doi.org/10.1140/epjqt/s40507-022-00123-4

Kandala, et al. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. Nature 549, 242–246 (2017). https://doi.org/10.1038/nature23879

Step 2-3: Parametrized mode or Ansatz

❖ Problem inspired ansatz

[Unitary Coupled Cluster Ansatz]



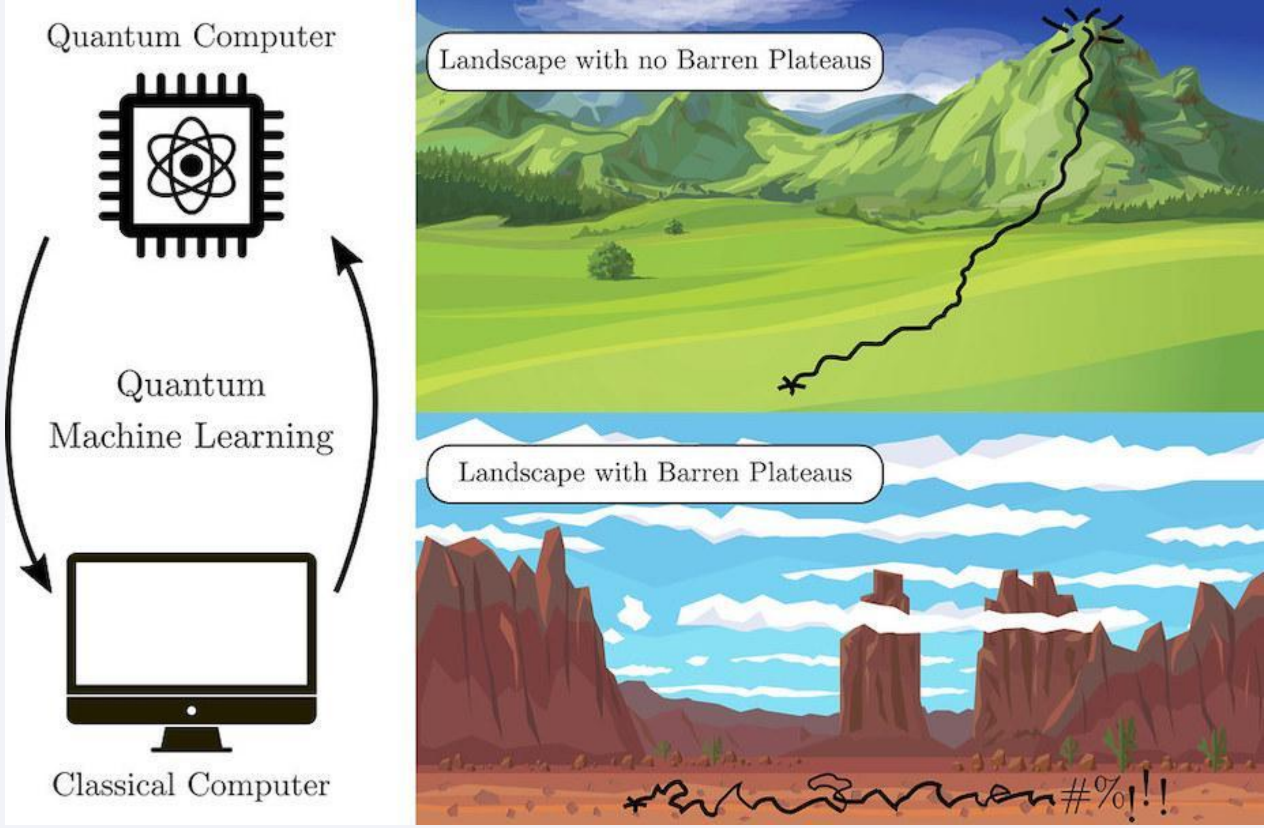
Jonathan Romero *et al* 2019 *Quantum Sci. Technol.* **4** 014008

Step 2-4: Ansatz

❖ The Barren plateau problem

Reasons

1. **Deep** parametrized circuit and no. qubits
2. **Random** initialization of parameters
3. hardware **noise** induced barren plateaus
4. **Entanglement**-induced barren plateaus



Consequence

1. The expected value of the gradient is zero! $\langle \partial_k C \rangle = 0$
2. Gradients vanish exponentially with the number of qubits! $Var[\partial_k C] \approx 2^{-n}$

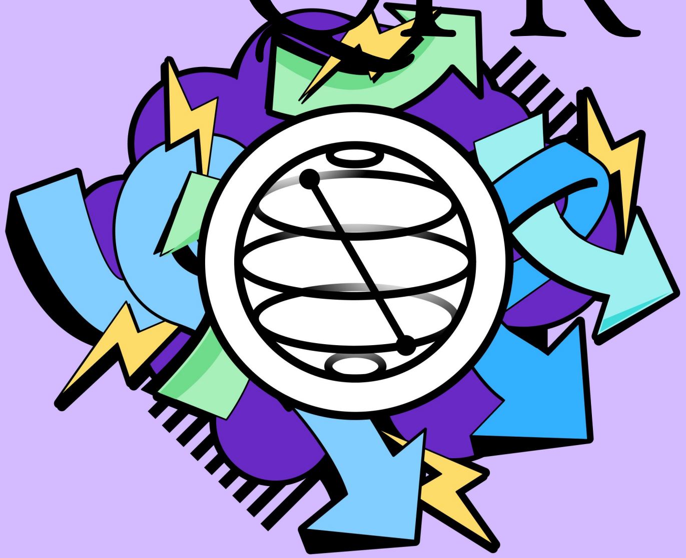
Solutions

1. Use parameters **close to the solution**. [1]
2. Use **local cost functions** instead of global ones. [2]
3. Introduce **correlations between parameters**. [3][4]

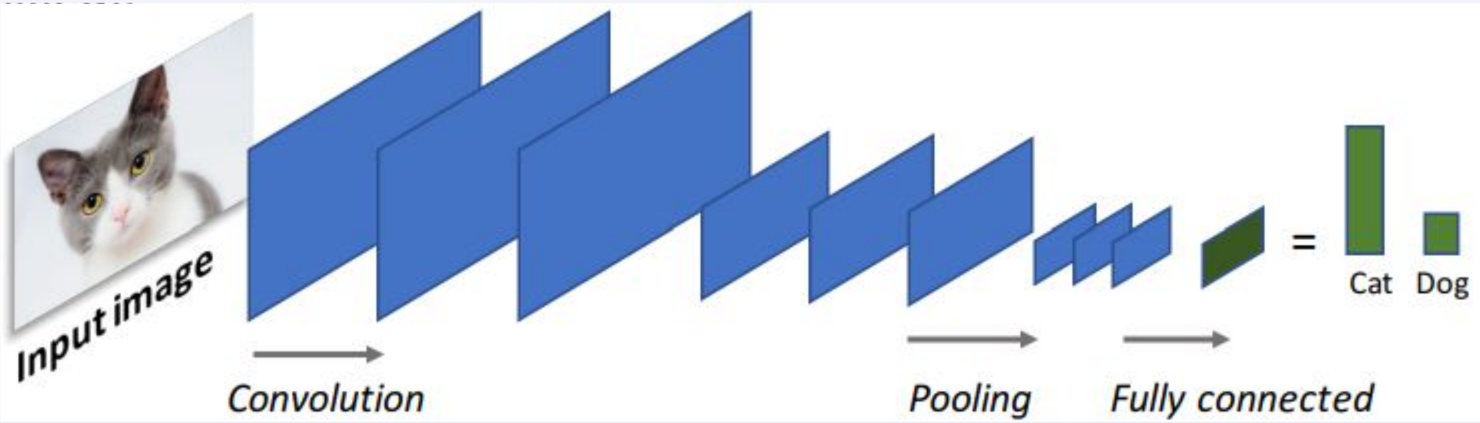
[1] Grant et al Quantum 3, 214 (2019) [2] Cost Function Dependent Barren Plateaus in Shallow Parameterized Quantum Circuits," M. Cerezo, et al, arXiv:2001.00550 [3] Z. Holmes et al., arXiv:2101.02138 (2021) [4] T. Volkoff and P. J. Coles, arXiv:2005.12200 (2020)

Qiskit | Fall Fest

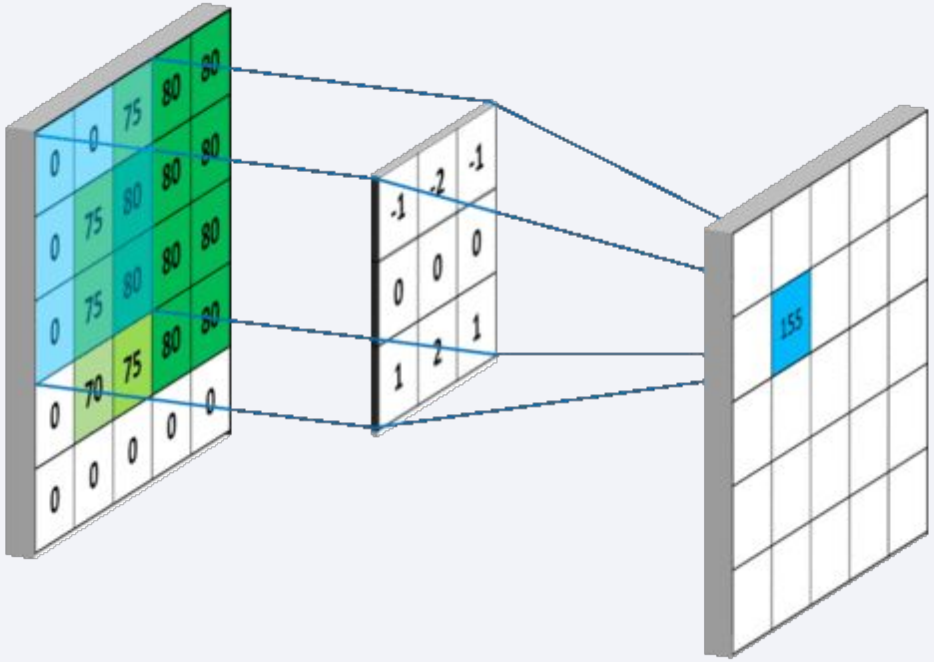
Application: QCNN for QPR



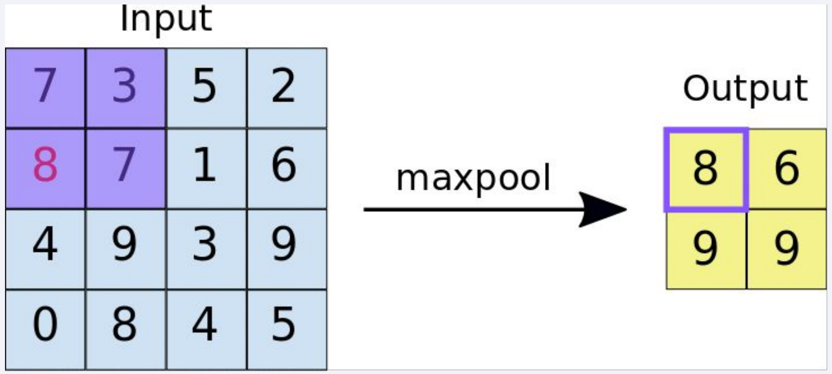
Classical convolutional neural network



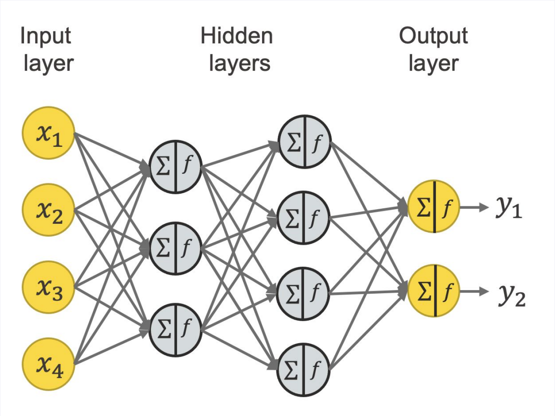
Convolution operation



Pooling operation



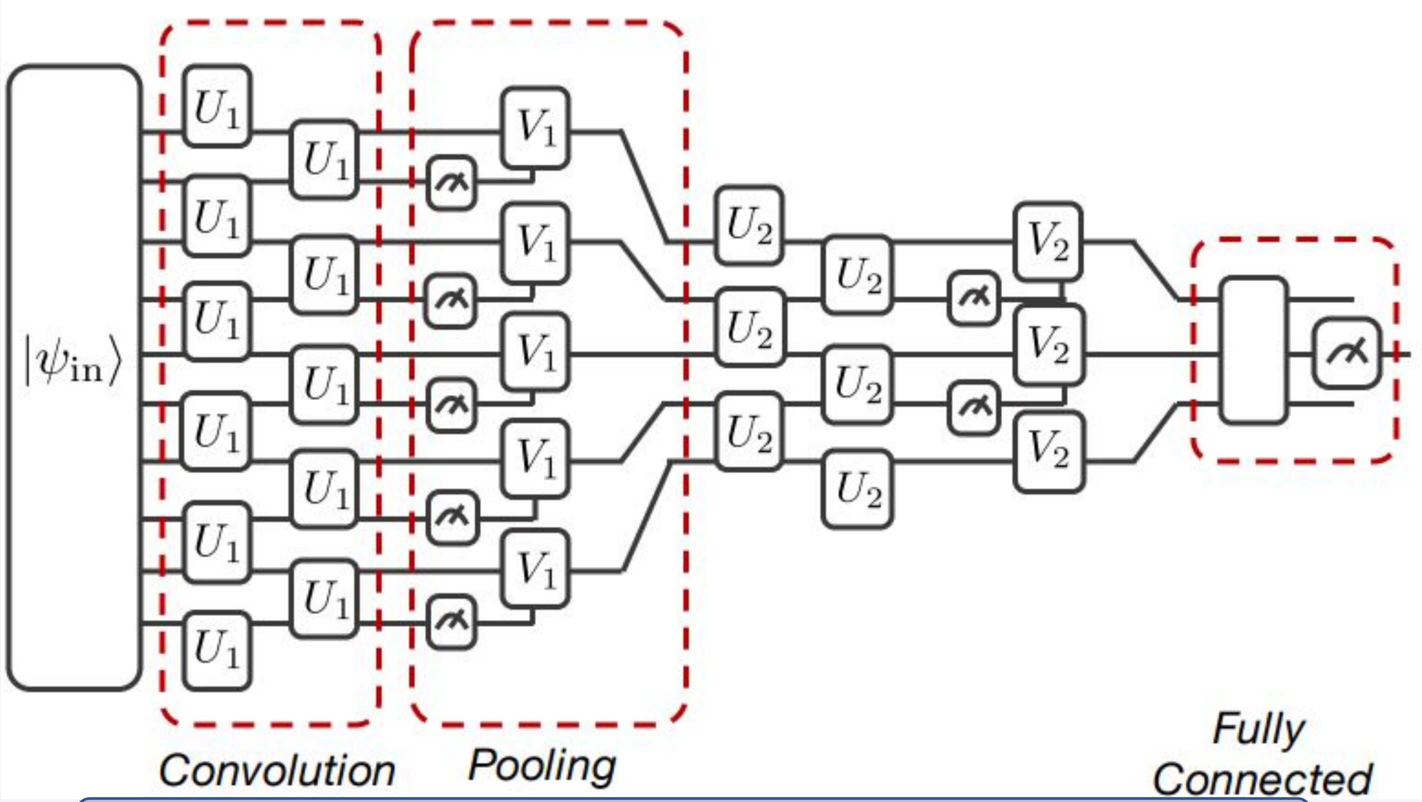
Fully connected or DNN



Yamashita, R., Nishio, M., Do, R.K.G. *et al.* Convolutional neural networks: an overview and application in radiology. *Insights Imaging* **9**, 611–629 (2018).

Quantum convolutional neural network

Input: Classical or quantum state



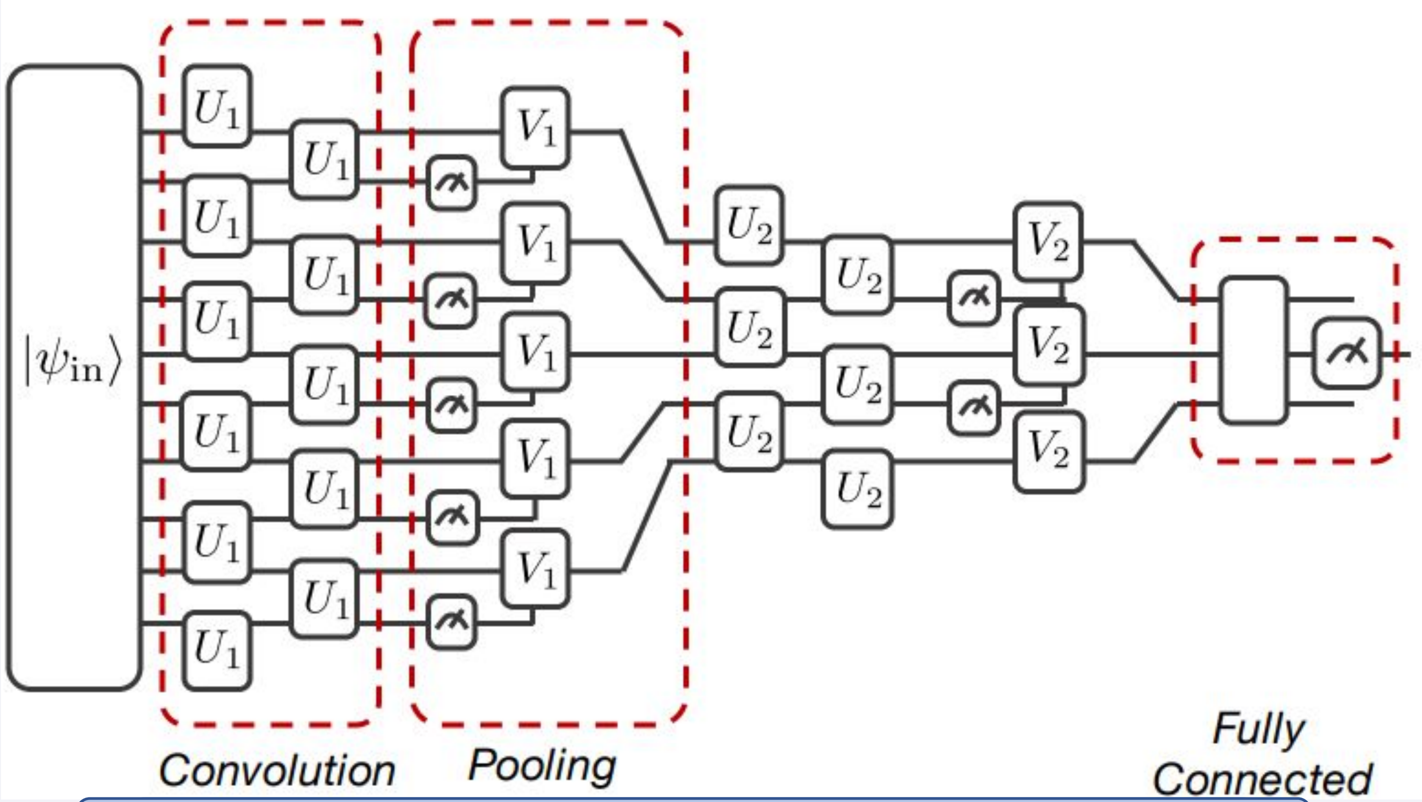
Convolution

- Combination of entangling and parametrized gates applied between **neighboring pairs of qubits**

Cong, I., et al. Quantum convolutional neural networks. *Nat. Phys.* **15**, 1273–1278 (2019). https://doi.org/10.1038/s41567-019-0648-8

Quantum convolutional neural network

Input: Classical or quantum state



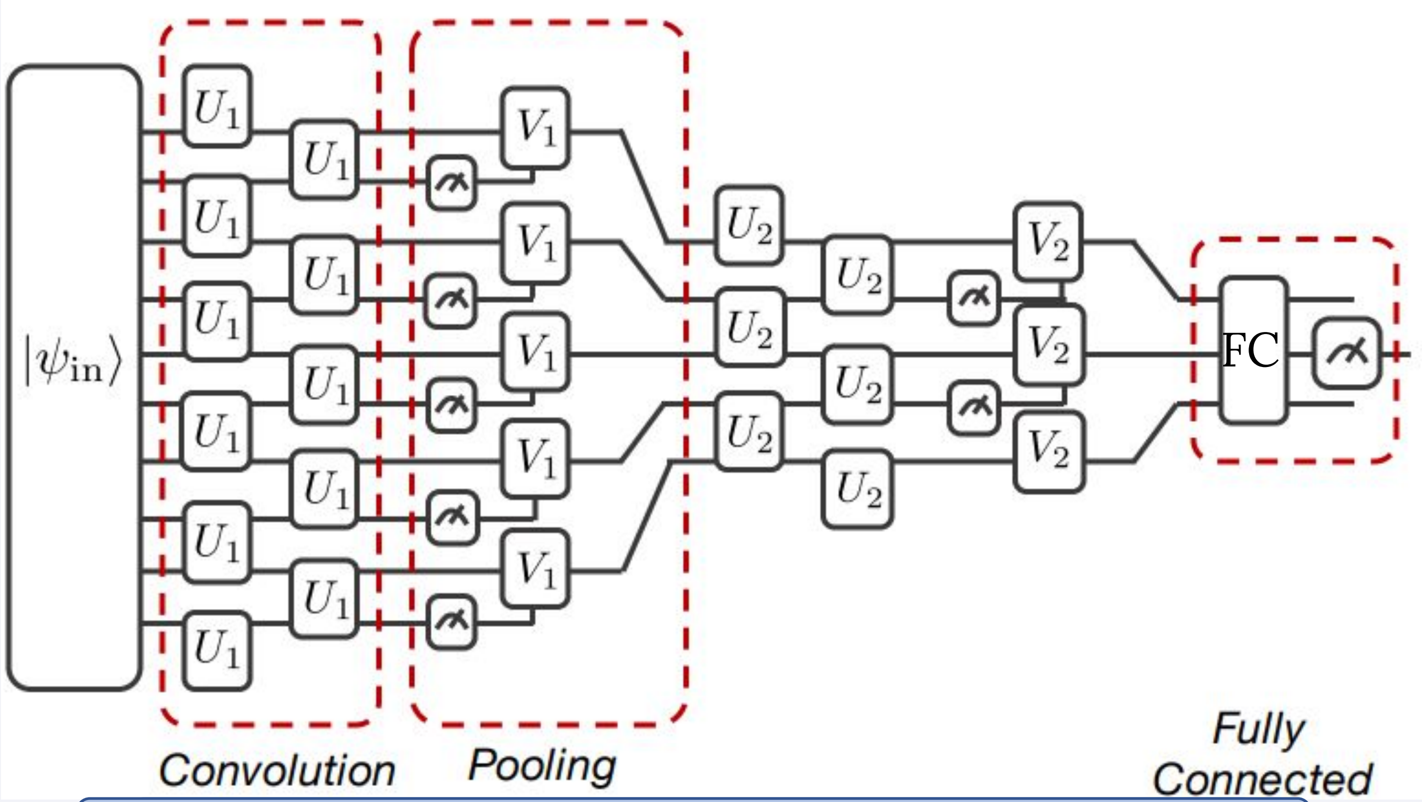
Pooling

- Measure a subset of the qubits and use the **measurement results to control other operations** on the remaining qubits.
- **Reduces the number of qubits** while retaining characteristic features of the input state vector

Cong, I., et al. Quantum convolutional neural networks. *Nat. Phys.* **15**, 1273–1278 (2019). https://doi.org/10.1038/s41567-019-0648-8

Quantum convolutional neural network

Input: Classical or quantum state



Fully connected or PQC

- Applied on the remaining qubits
- **Perform classification** by mapping the result onto a single output qubit

Cong, I., et al. Quantum convolutional neural networks. *Nat. Phys.* **15**, 1273–1278 (2019). https://doi.org/10.1038/s41567-019-0648-8

Quantum Phase recognition

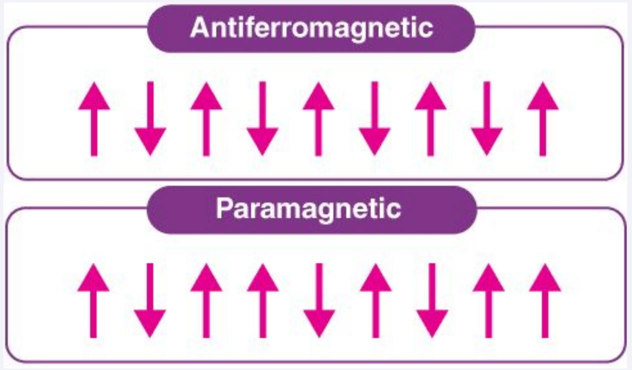
Model: a family of Hamiltonians on a **spin-1/2 chain** with open boundary conditions:

1D input chain of spins: 

$$H = -J\sum_{i=1}^{N-2} Z_i X_{i+1} Z_{i+2} - h_1 \sum_{i=1}^{N} X_i - h_2 \sum_{i=1}^{N-1} X_i X_{i+1}$$

- J: coupling constant (let's take $J = 1$)
- $\boldsymbol{h_1}$: parametrize **the strength of an external field**.
- $\boldsymbol{h_2}$: parametrize the **nearest neighbor Ising-type coupling**.
- $\{X_i, Y_i, Z_i\}$ are the **Pauli operators** acting on the spin at site i or qubit i.
- N: is the **number of spins** in our model or number of qubits

Phase of matter of spin system



Symmetry-protected topological order - Wikipedia

Quantum Phase recognition

By varying the model parameters h_1 and h_2 in H, **we obtain different systems, each with its own ground state**. This latter **allows us to identify which phase of matter our system is in.**

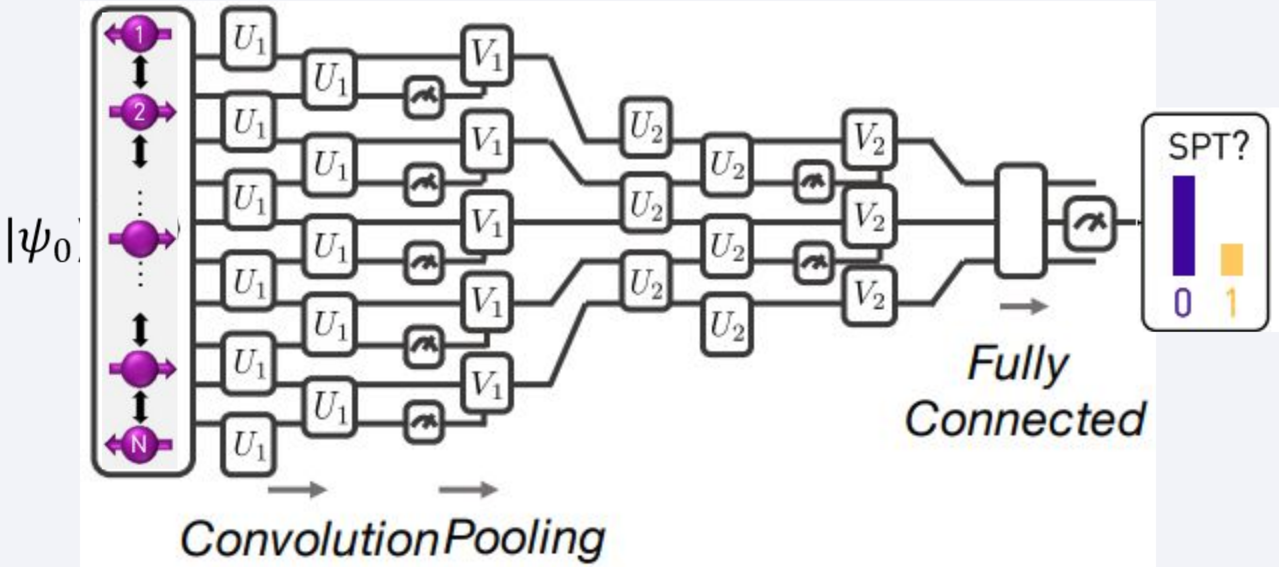
$$H = -J \sum_{i=1}^{N-2} Z_i X_{i+1} Z_{i+2} - h_1 \sum_{i=1}^{N} X_i - h_2 \sum_{i=1}^{N-1} X_i X_{i+1}$$

→ paramagnetic phase
→ symmetry-protected topological (SPT) phases
→ anti-ferromagnetic phase

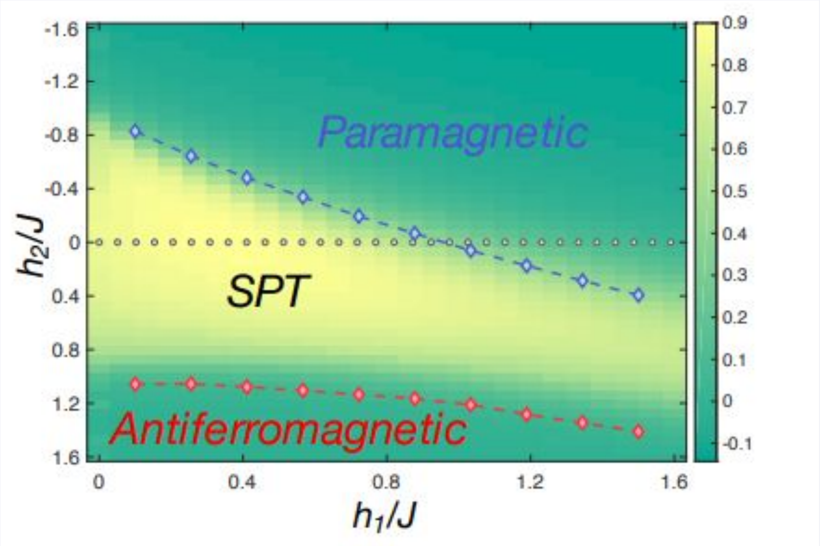
Problem: We have three phase of matter, and we are trying to recognize the ground states that belongs to the **SPT phase.**

QCNN is very efficient in quantum phase recognition

<u>input states are the ground states of a family of Hamiltonians</u>



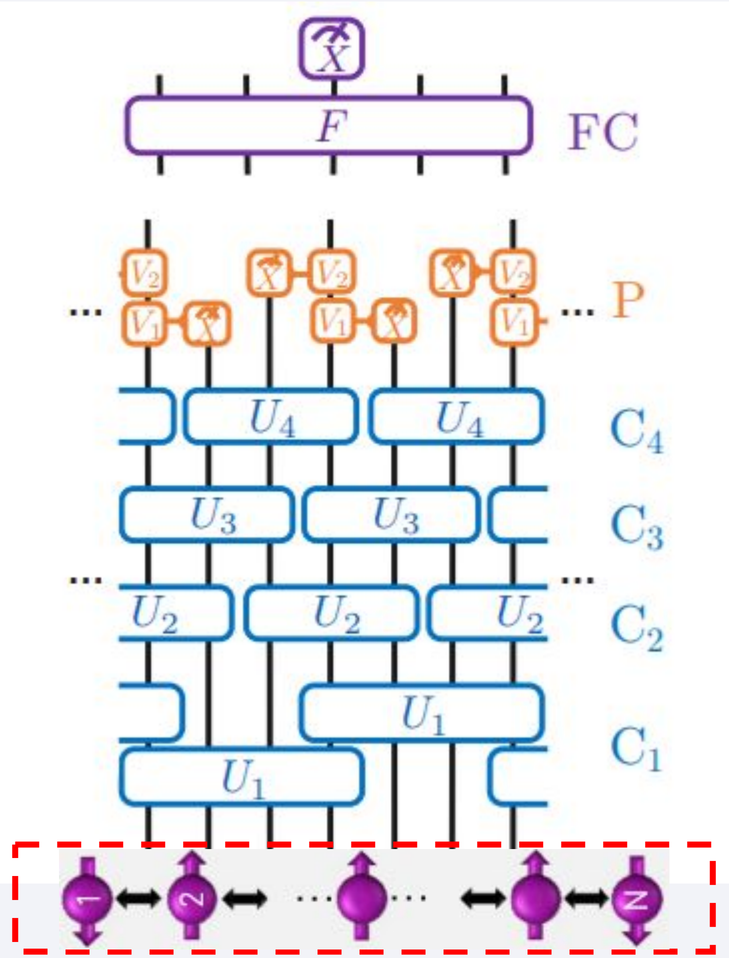
2D heat map



Task: train a QCNN on labelled samples in order to predict labels for new data

Cong, I., et al. Quantum convolutional neural networks. *Nat. Phys.* **15**, 1273–1278 (2019). https://doi.org/10.1038/s41567-019-0648-8

Training and testing data

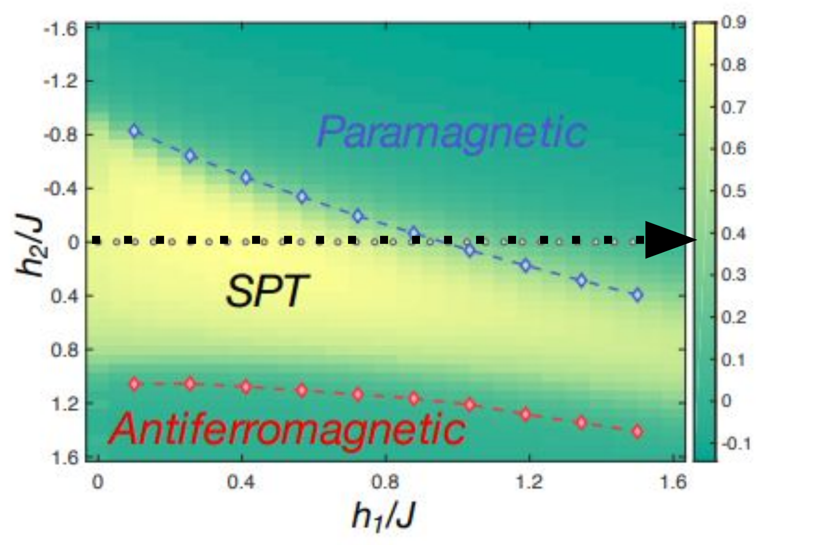


$|\psi\rangle_\alpha$ is input state=ground state

Training Data:

- 40 ground state wavefunctions
- Train along $h_2 = 0, h_1 =$ sampled 40 times between 0 and 1.6, J=1
- Classified training data: $\{|\psi\rangle_\alpha\ ,\ \ y_\alpha\}$, $\alpha = 1 \ldots M$, $y_\alpha = 0\ or\ 1$ corresponding binary classification outputs.
- 0 corresponds to paramagnetic or anti-ferromagnetic phase while 1 corresponds to SPT phase

Case where $h_2 = 0 \rightarrow h_1 \leq\ 1$ are in the SPT phase and thus assigned the label 1 while $h_1\ >\ 1$ are in the paramagnetic phase

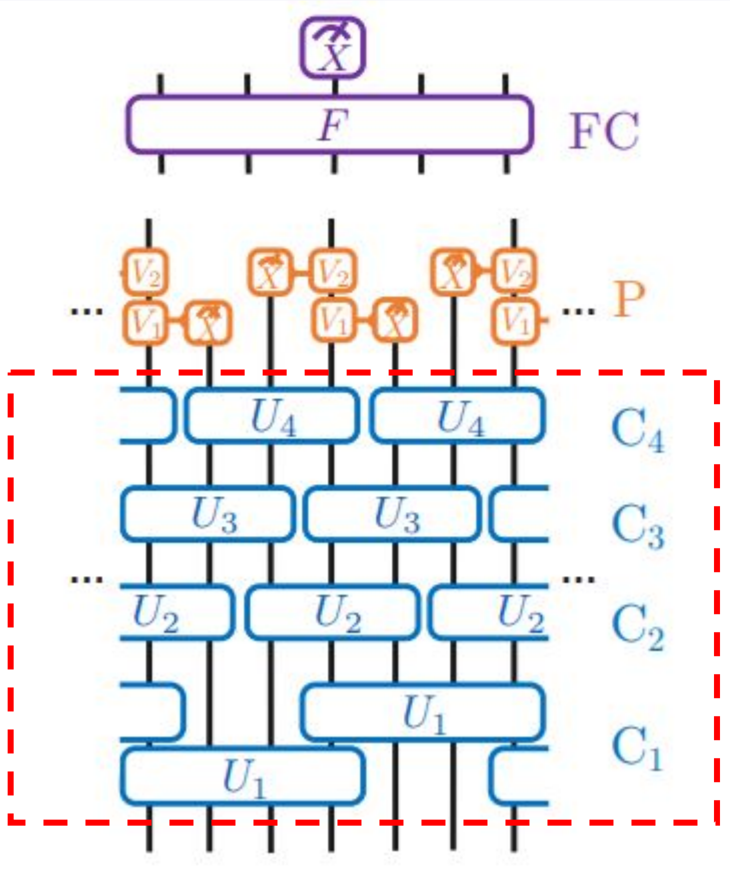


Test Data:

Combination of (h1/J, h2/J) in the range (0, 1.6) and (-1.6, 1.6) respectively

These wavefunctions would be fed into the QCNN and the final measurement would determine the predicted label/phase of matter that the wavefunction belongs to.

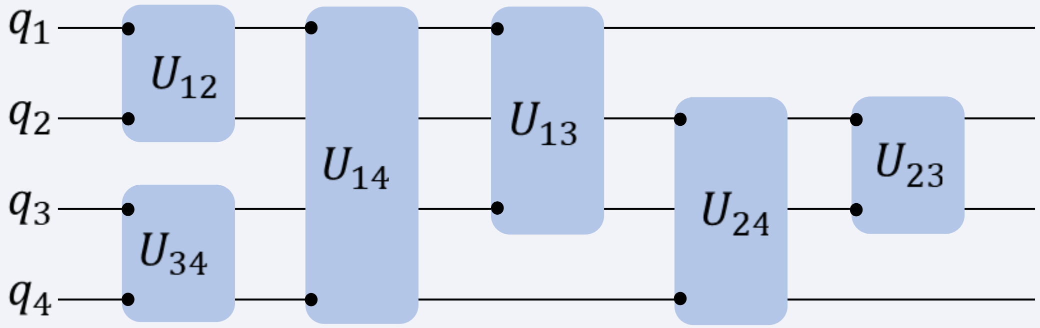
Convolution layers



We parameterize unitaries as exponentials of generalized $a \times a$ Gell-Mann matrices $\{\Lambda_j\}$, where $a = 2^w$ and w is the number of qubits involved in the unitary: $U = \exp(-i\sum_j \theta_j \Lambda_j)$; $i^2 = -1$
There are $a^2 - 1$ Gell-Mann matrices, hence $a^2 - 1$ trainable parameters in U

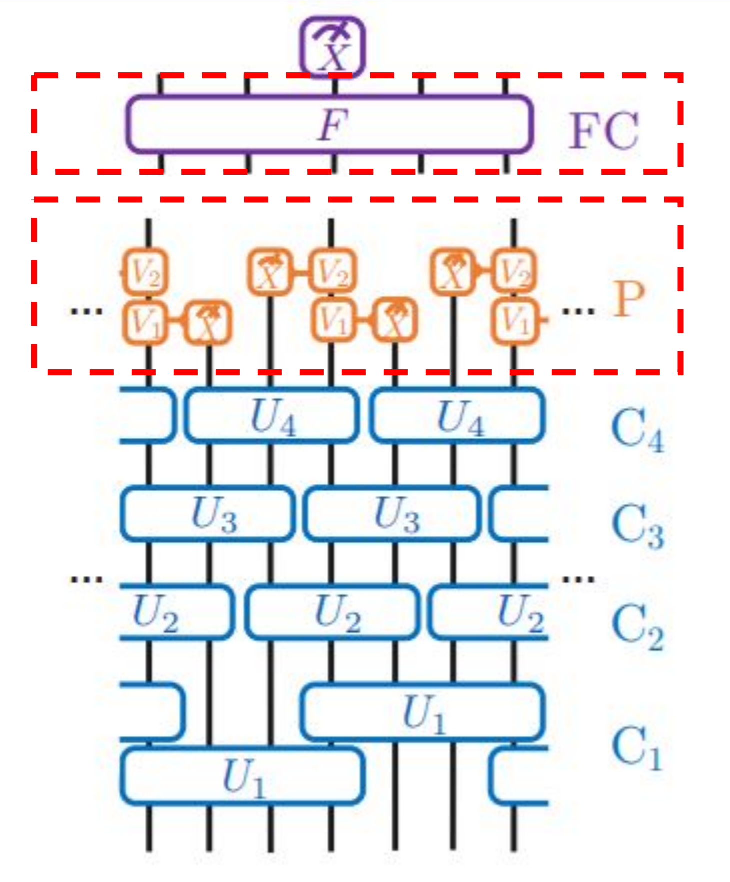
1. C_1: we apply U_1, a 4-qubits convolution. U_1 is a product of six two-qubit unitaries between each possible pair of qubits:

$$U_1 = U_{23}U_{24}U_{13}U_{14}U_{12}U_{34}\,,\ a = 2^2$$

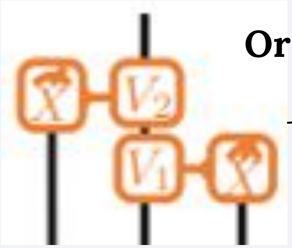


2. $C_{2,3,4}$: $U_{2,3,4}$ are three qubits unitary, $a = 2^3$

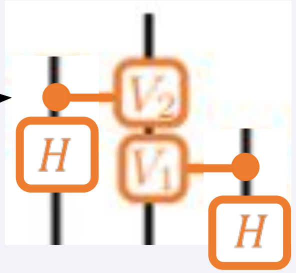
Pooling and fully connected layers



3. P: reduces the total number of "active" qubits in the circuit by a factor of 3. The pooling unitaries are controlled by a measurement in the X-basis.



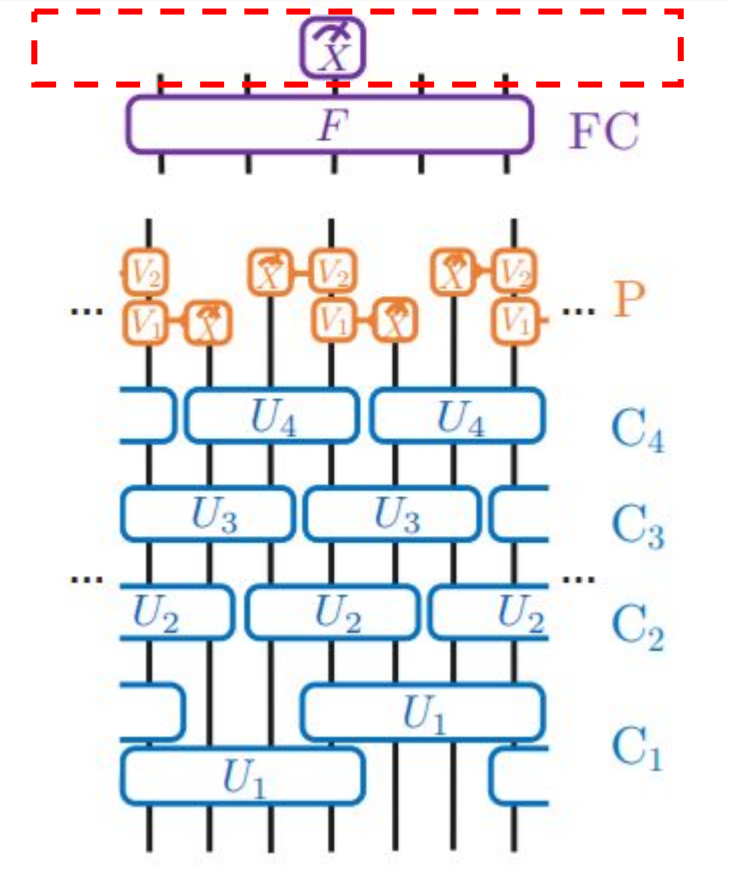
Or controlled-gate instead of measurement



$V_1 = \exp(-i\sum_j \lambda_j \Lambda_j)$ $V_2 = \exp(-i\sum_j \beta_j \Lambda_j)$ $a = 2^1$

4. F: is a U with $2^w \times 2^w$ Gell-Mann matrices where w is the remaining qubits

Measurement and learning



5. $\langle X \rangle = f_{\{U_i, V_j, F\}}(|\psi_\alpha\rangle)$ this result help us compute the cost function via the mean error square formula:

$$MSE = \frac{1}{2M} \sum_{\alpha=1}^{M} \left(y_i - f_{\{U_i, V_j, F\}}(|\psi_\alpha\rangle) \right)^2,$$

y_i is the predifined label of $|\psi_\alpha\rangle$, M: # of training samples (in our case= 40)

6. Update parameters θ_μ:

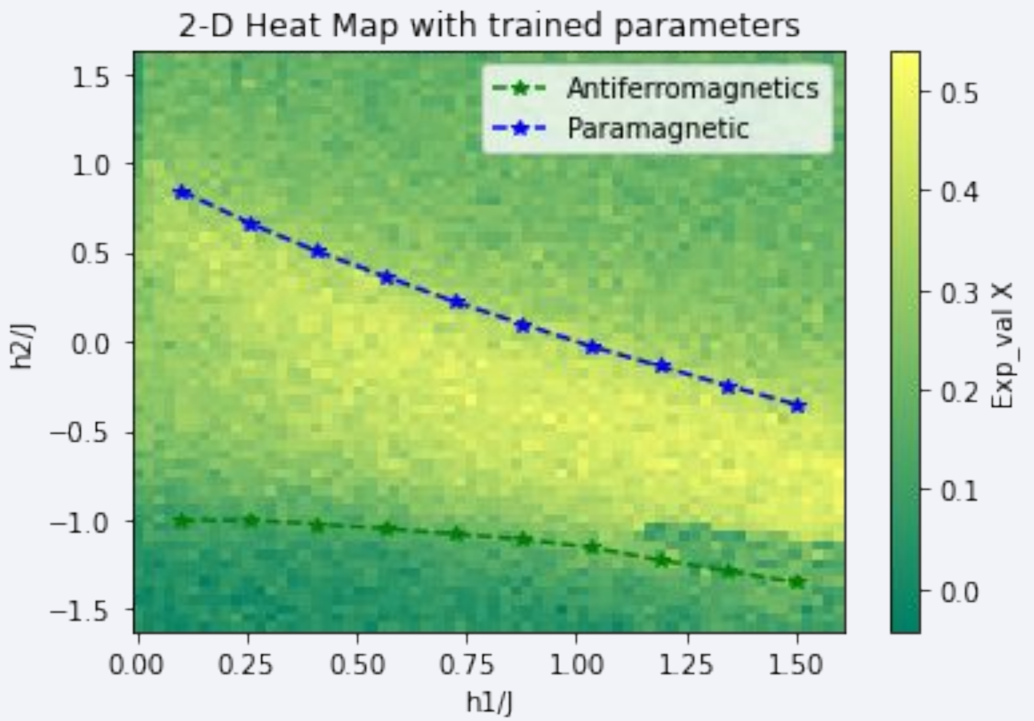
$$\theta_\mu \rightarrow \theta_\mu - \lambda \frac{\partial MSE}{\partial \theta_\mu}$$

Parameter shif-rule:

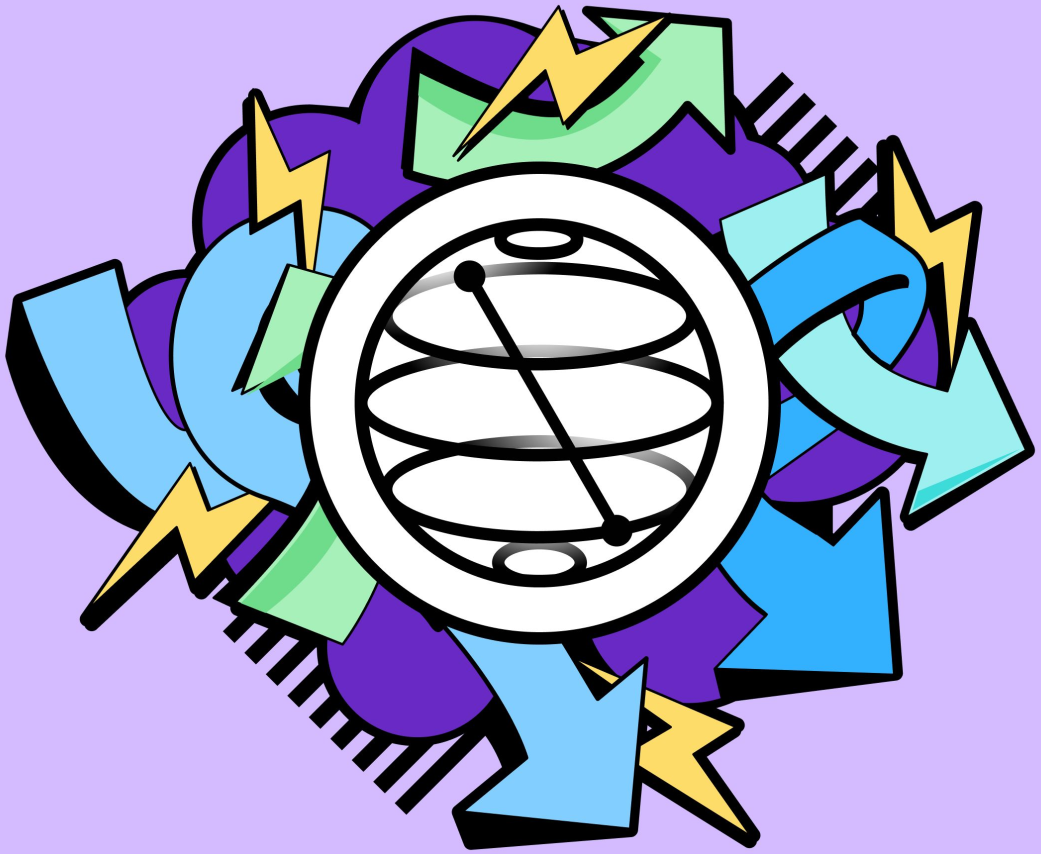
$$\frac{\partial MSE}{\partial \theta_\mu} = \frac{1}{2}\left(\mathrm{MSE}\left(\theta_\mu + \frac{\pi}{2}\right) - \mathrm{MSE}\left(\theta_\mu - \frac{\pi}{2}\right) \right)$$

Test results

Used Gradient free optimizer: Cobyla with 500 iterations

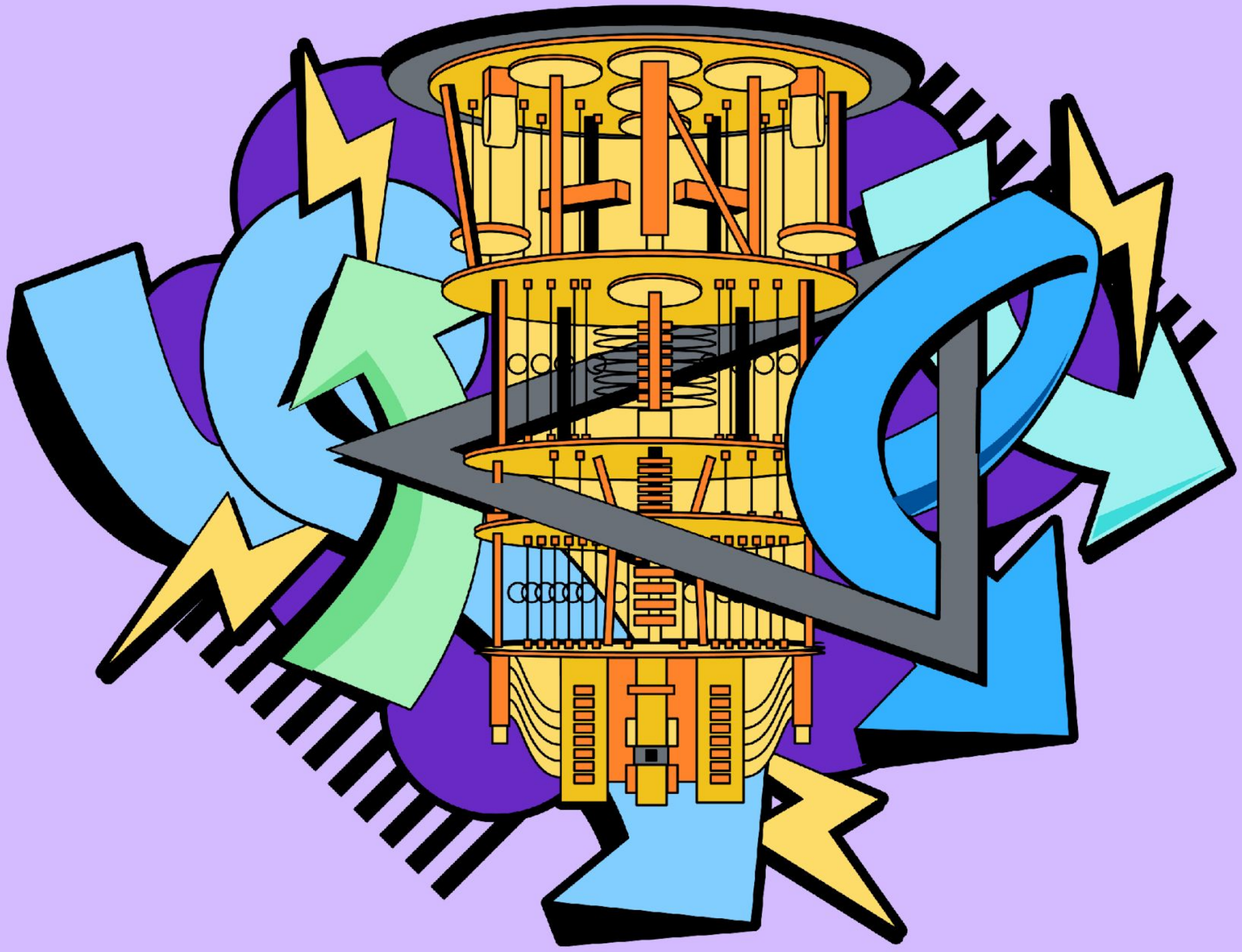


Qiskit | Fall Fest



Thank you!

Nouhaila Innan & Walid EL Maouki
nouhailainnan@gmail.com , elmaouaki.walid@gmail.com



Qiskit | Fall Fest 2022