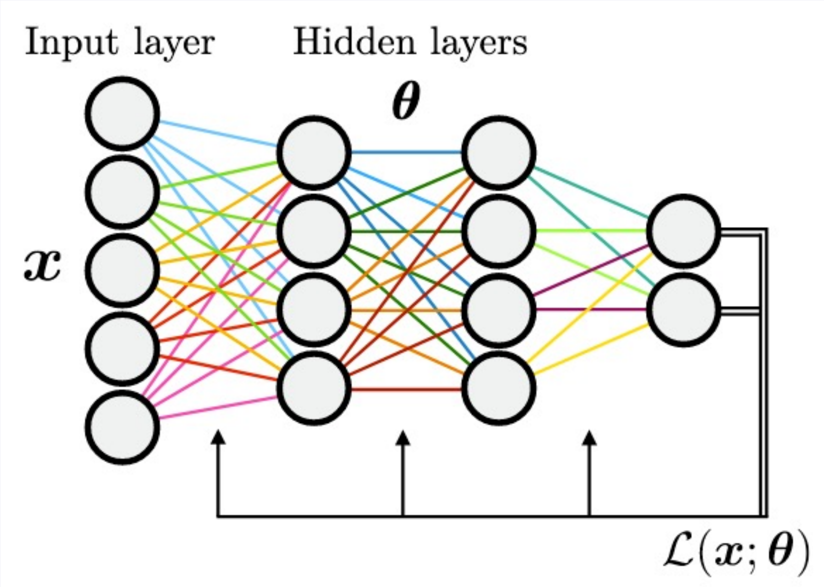
Introduction to quantum models and applications

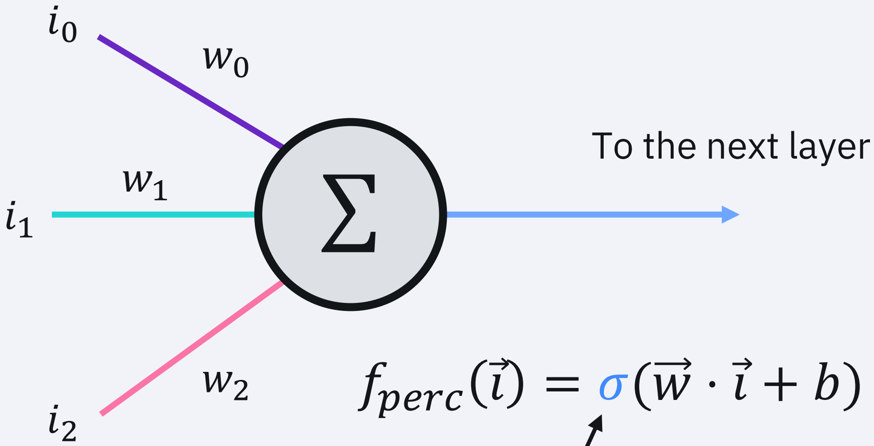
Francesco Tacchino
Quantum Applications Researcher
IBM Quantum, IBM Research – Zurich

Qiskit

Classical feed-forward neural networks



$$y_{NN}(x) = f_L \circ f_{L-1} \ldots \circ f_1(x)$$



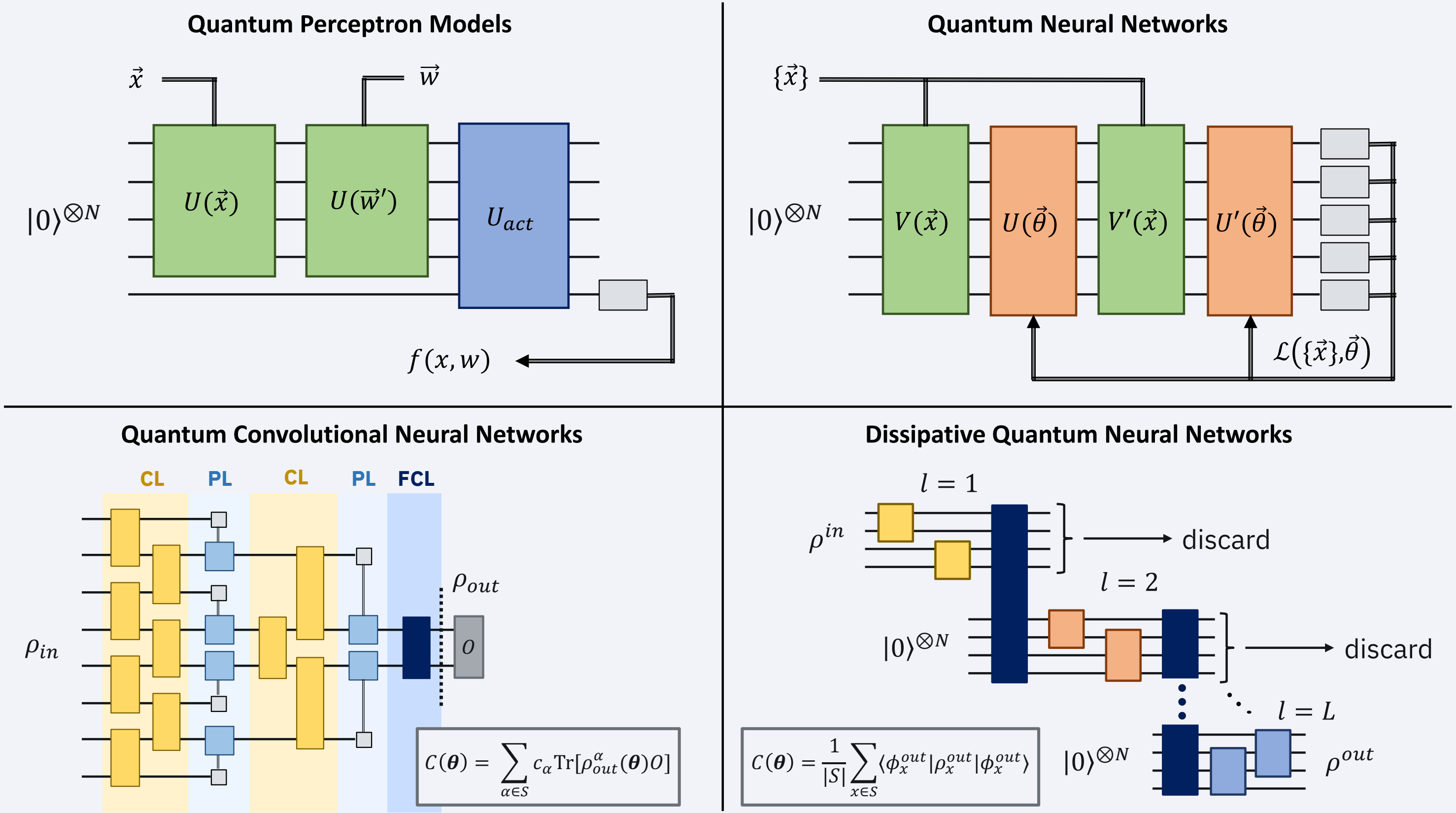
$$f_{perc}(\vec{\imath}) = \sigma(\vec{w} \cdot \vec{\imath} + b)$$

Non-linear activation function

Supervised learning: $\mathcal{T} = \{(x_i, y_i)\}$

$$\mathcal{L} = \sum_{\mathcal{T}} (y_{NN}(x_i) - y_i)^2$$

Quantum neural network models



S. Mangini *et al.*, EPL (Europhysics Letters) 134, 10002

Quantum perceptrons

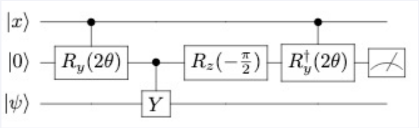
Strategy

Implement a quantum algorithm reproducing the functionalities of classical perceptrons, e.g. for classification

Key Features

- ➢ First approaches to the realization of non-linear transformations of the inputs from the charachteristic **linear** unitary evolution of quantum systems
- ➢ Transition between “first wave” and “second wave” QML
- ➢ Precursor of Dissipative QNNs

- **Repeat Until Success circuits**



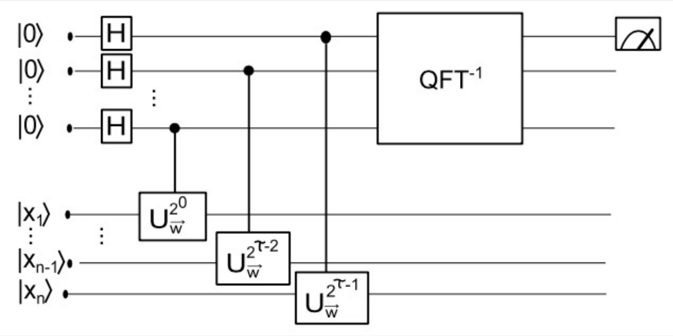
Y. Cao et al., arXiv:1711.11240

$$U_{perceptron} = \sum_{\alpha} |\alpha\rangle\langle\alpha| \otimes U(\alpha)$$

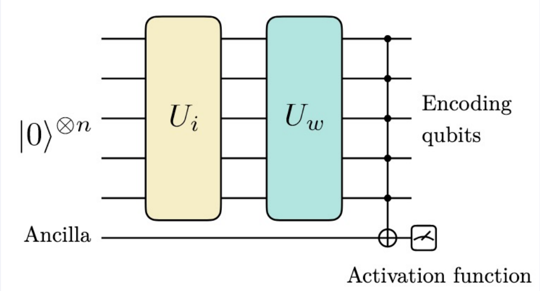
E. Torrontegui and J. J. Garcia-Ripoll, EPL **125**, 3004 (2019)

- **QFT-based perceptron**

M. Schuld et al., Phys. Lett. A **379**, 660 (2015)



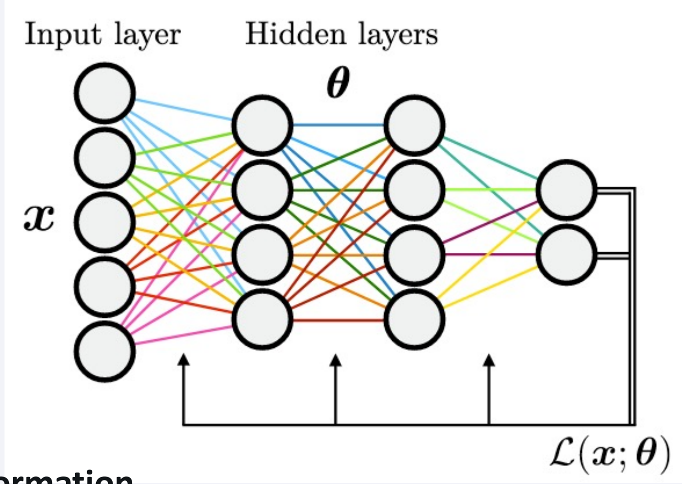
- **Non-linearity from measurement**



F. Tacchino et al., npj Quantum Inf. **5**, 26 (2019)

Quantum neural networks

Classical feed-forward design



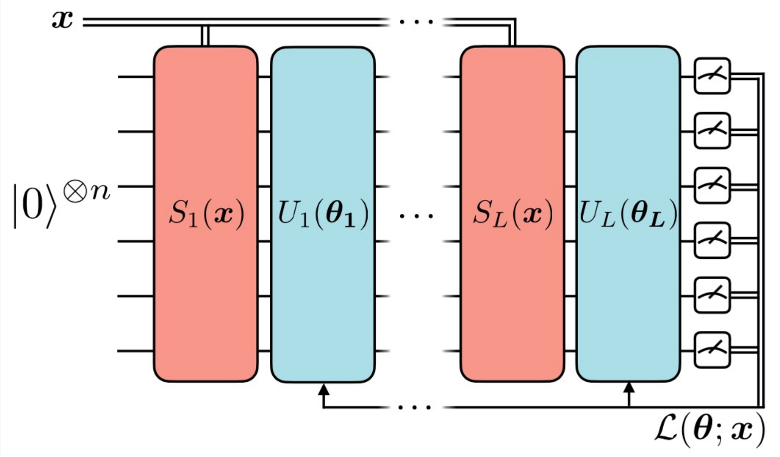
Information

Training

$$f_{\boldsymbol{\theta}}(\boldsymbol{x}) = \sigma(W_{\boldsymbol{\theta}}\boldsymbol{x} + b_{\boldsymbol{\theta}})$$

$$y_{NN}(\boldsymbol{x}) = f_{\boldsymbol{\theta}_L} \circ f_{\boldsymbol{\theta}_{L-1}} \dots \circ f_{\boldsymbol{\theta}_1}(\boldsymbol{x})$$

QNN ~ parametrized quantum circuits



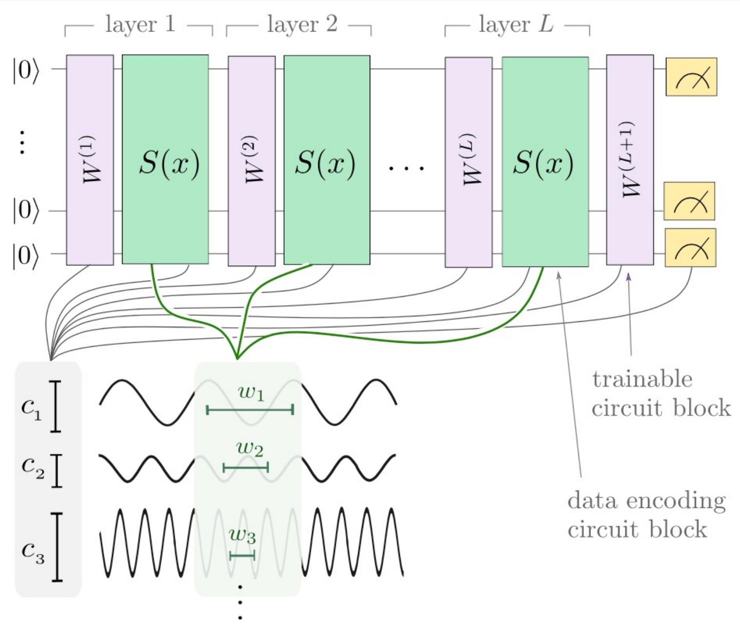
$$y_{QNN}(\boldsymbol{x}) = \langle 0|\mathcal{U}^\dagger(\boldsymbol{x};\boldsymbol{\theta})M\mathcal{U}(\boldsymbol{x};\boldsymbol{\theta})|0\rangle$$

$$\mathcal{U}(\boldsymbol{x};\boldsymbol{\theta}) = \prod_i U_i(\boldsymbol{\theta}_i)S_i(\boldsymbol{x})$$

- Similar layered structure, but different "information flow"
- Quantum layers: **(linear)** unitary operations

Universality of quantum neural networks

Qiskit



$$f_{\boldsymbol{\theta}}(\boldsymbol{x}) = \langle 0|U(\boldsymbol{x},\boldsymbol{\theta})MU(\boldsymbol{x},\boldsymbol{\theta})|0\rangle = \sum_{\omega\in\Omega} c_\omega e^{i\omega x}$$

Schuld *et al.*, Phys. Rev. A 103, 032430 (2021)

- **Available frequencies** are determined by the **encoding**

$S(x) \sim e^{-iHx}$ $\qquad$ $\Omega = \{(\lambda_{j_1} + \cdots + \lambda_{j_L}) - (\lambda_{k_1} + \cdots + \lambda_{k_L})\}$

$K(L,d) \le \dfrac{d^{2L}}{2} - 1$ $\qquad$ Spectrum size if H acts on dimension d
(The actual bound can be much tighter!)

- **Trainable unitaries and the observable** fix the coefficients

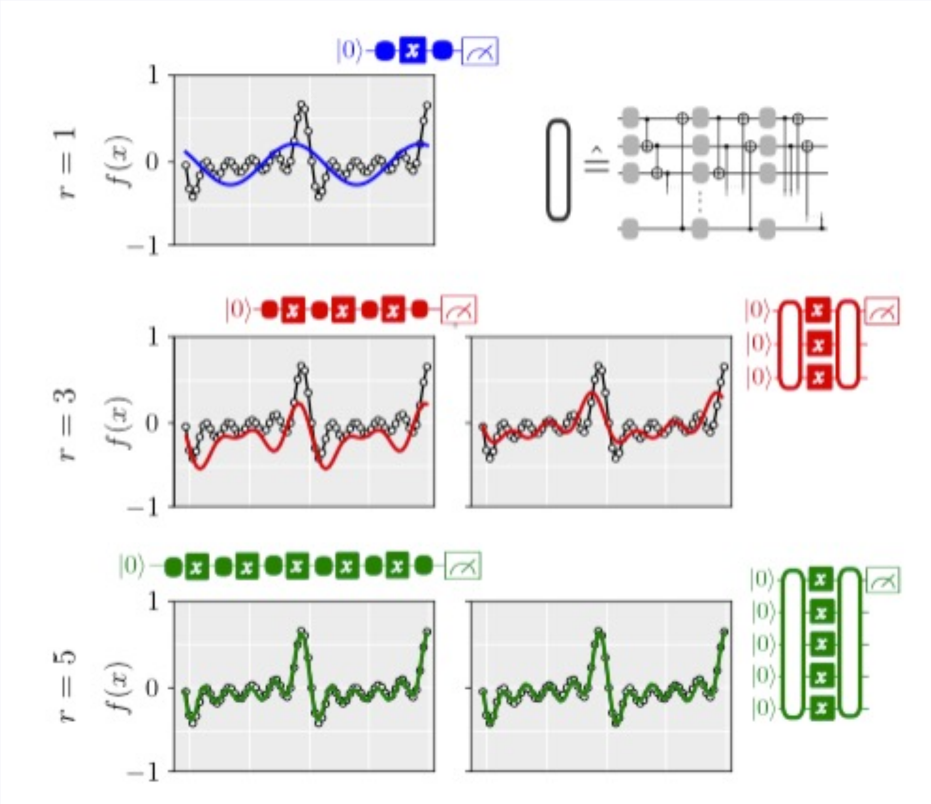
$$f_m(x) = \langle \Gamma|S_{H_m}^\dagger(x) M S_{H_m}(x)|\Gamma\rangle$$

Theorem (Universality). *Let $\{H_m\}$ be a universal Hamiltonian family and $\{f_m\}$ the associated quantum model family. For all functions $g \in L_2([0,2\pi]^N)$ and $\forall \epsilon > 0$ there exist $m' \in \mathbb{N}$, some initial state $|\Gamma\rangle$ and some observable M such that*

$$||f_{m'} - g||_2 \le \epsilon$$

Mitarai *et al.*, PRA **98**, 032309 (2018) Gil Vidal and Theis, Front. Phys. **8**, 297 (2020)
Pérez-Salinas *et al.*, Quantum **4**, 226 (2020) Pérez-Salinas *et al.*, arxiv:2102.04032

Universality of quantum neural networks



Example

With L repetitions of the same single-qubit encoding gate the spectrum size is

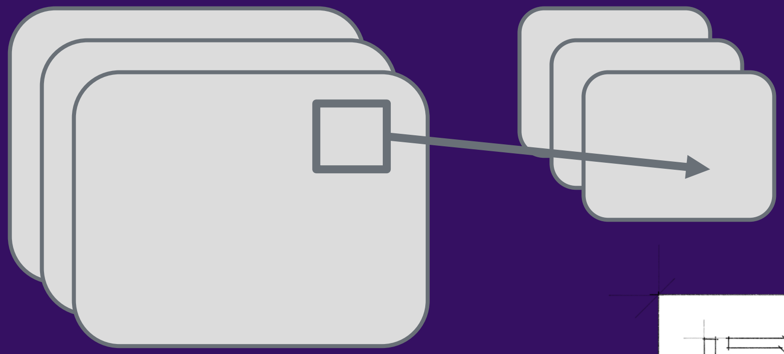
$$K = L$$

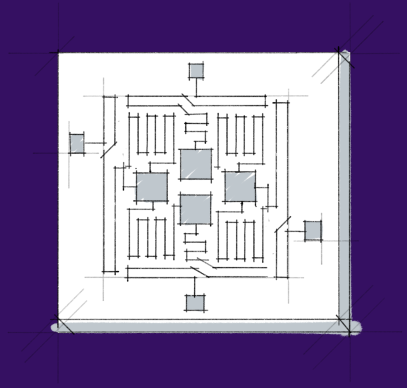
The frequency spectrum increases *linearly* with the number of repetitions of single qubit Pauli gates

Notice that ***series*** and ***parallel*** repeated encodings are equivalent for universality purposes.

Schuld *et al.*, Phys. Rev. A 103, 032430 (2021)

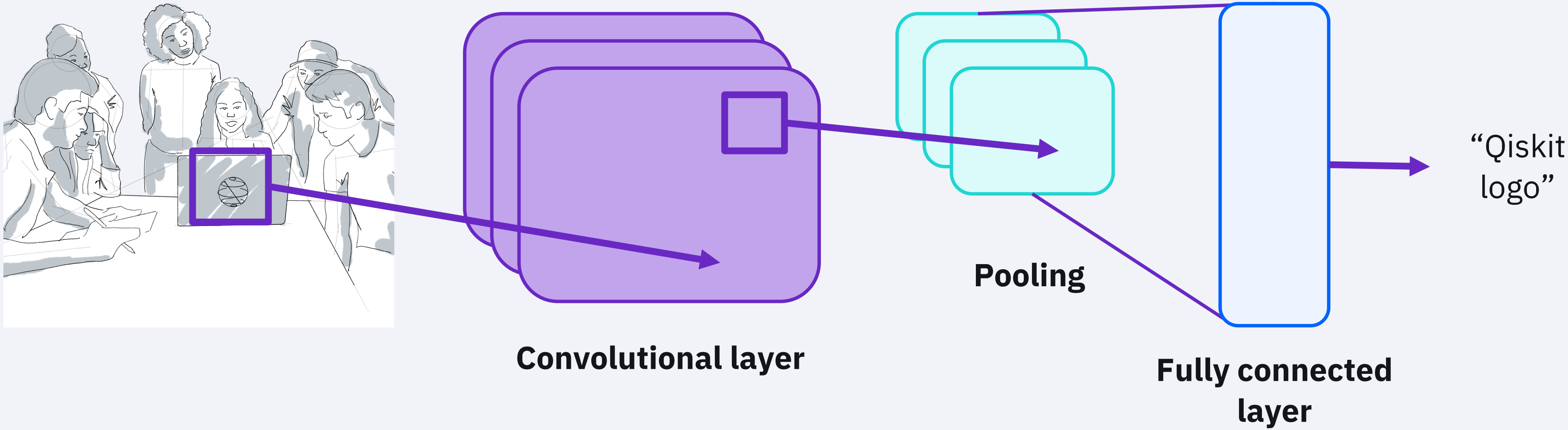
Quantization of a classical model: Convolutional Neural Networks





Convolutional Neural Networks

Qiskit

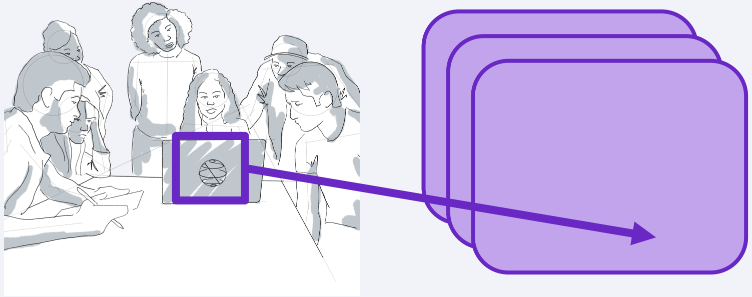


Famous examples:

“LeNet” Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," in *Proceedings of the IEEE*, **86**, 2278 (1998)

“AlexNet” A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks” *Commun. ACM* **60**, 84 (2017)

CNNs: Convolutional layers



Stride

0	1	0	1	0
0	1	0	0	0
0	0	1	1	1
0	1	1	0	1
0	0	1	0	1

Local portion of the image

(e.g. matrix of b/w pixels)

0	1	0
1	-4	1
0	1	0

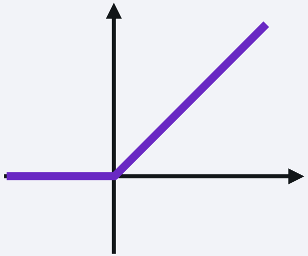
One (or more) filters

Dot product

+ ReLU

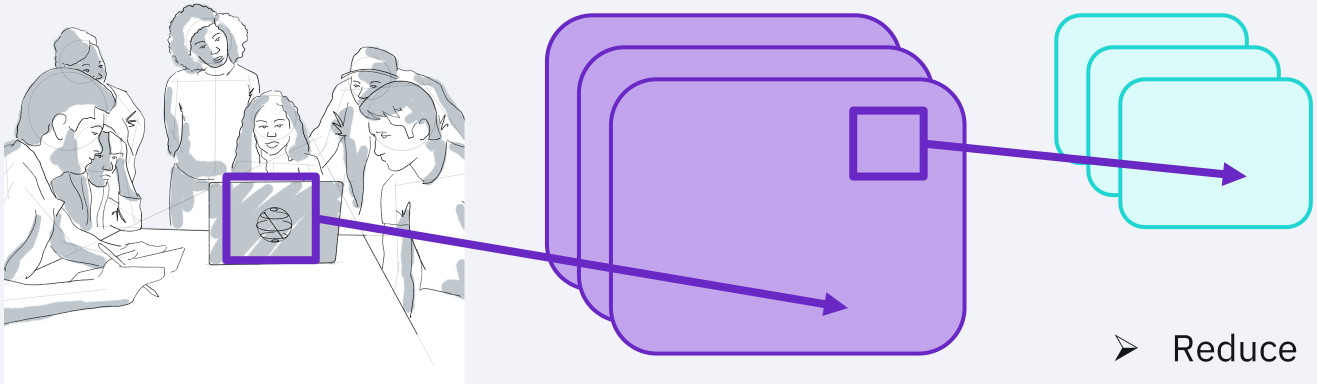
a	b	c
d	e	f
g	h	i

(Rectified) Convolved feature



Rectified Linear Unit

CNNs: Pooling layers



1	1	3	4
1	3	1	1
2	2	1	0
0	1	2	4

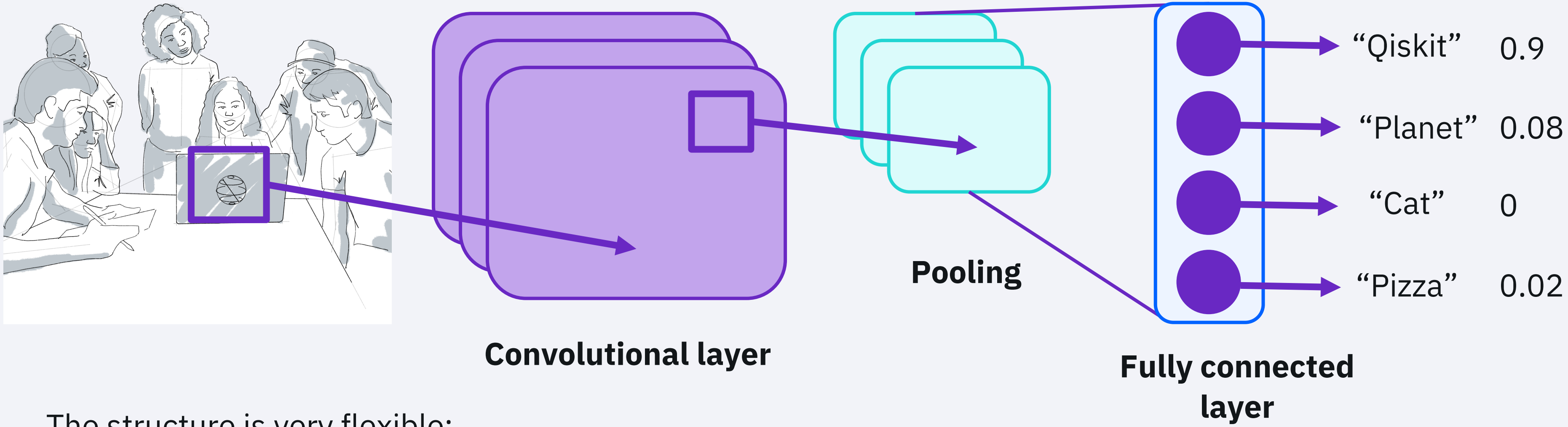
Max, Average, Sum ...

3	4
2	4

Convolved feature

- Reduce feature dimension, making it more manageable for successive layers
- Make the network resilient to small changes
- Avoid overfitting (less parameters involved)
- Reach scale invariance

CNNs: Classification



The structure is very flexible:

- Add more convolutional/pooling layers
- Add more complex FCL at the end
- **Change the number and nature of filters**

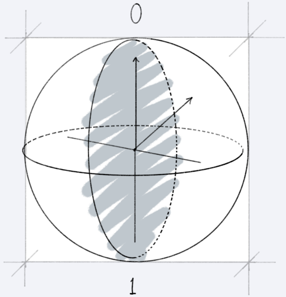
Training

Can be done with differential methods, most notably via Backpropagation

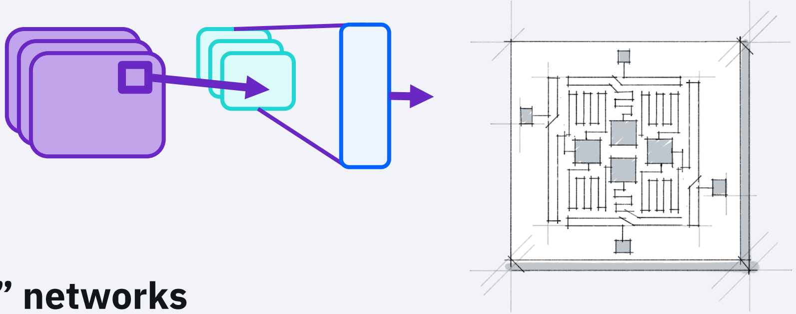
Quantum approaches to CNNs

The most complex and successful CNNs represent computationally intensive models

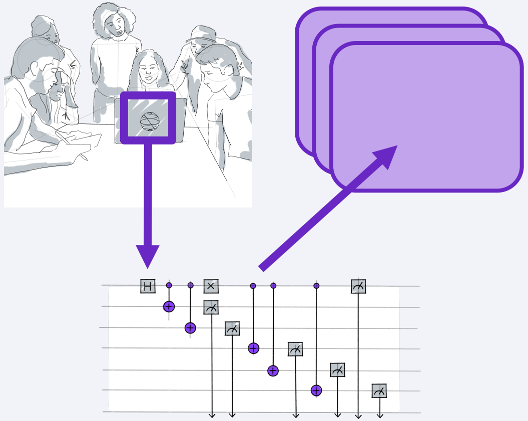
Can Quantum Computing become a resource?



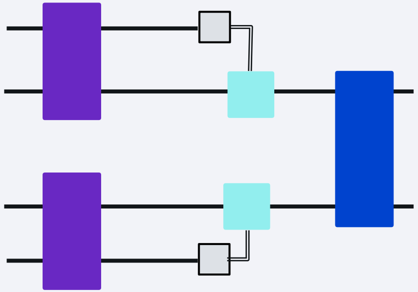
“Direct” approach



“Quanvolutional” networks



CNN-inspired Quantum Neural Networks



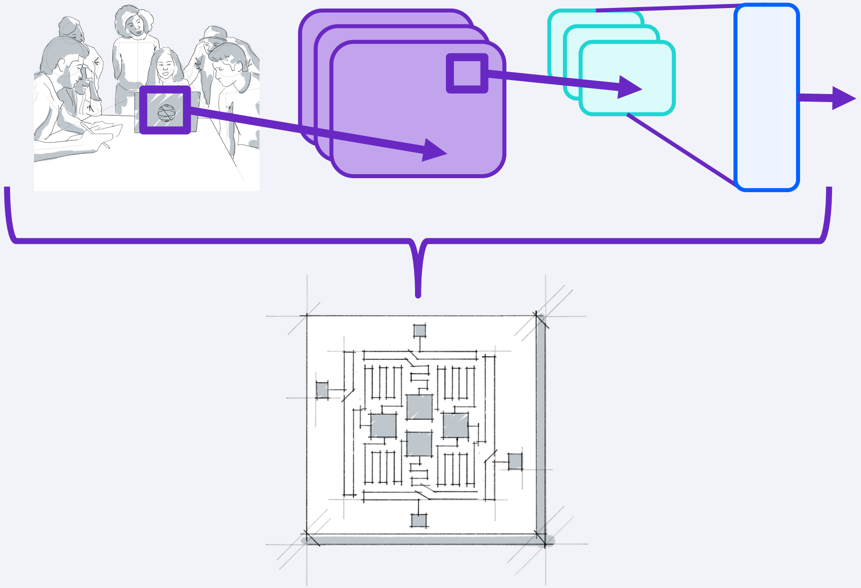
“Direct” quantum implementation of CNNs

Strategy

Design and implement a quantum circuit model performing all the CNN operations

Key Features

- (With some assumptions) provable speedups can in principle be achieved
- Leverage quantum superposition effects for quantitative advantages
- Classical data encoding bottleneck and/or QRAM needed
- The overall algorithm can require non-trivial subroutines (QFT, quantum phase estimation)



Examples

Li *et al.*, Quantum Sci. Technol. **5**, 044003 (2020)

Kerenidis *et al.*, arXiv 1911.01117 (2019)

"Quanvolutional" neural networks

Qiskit

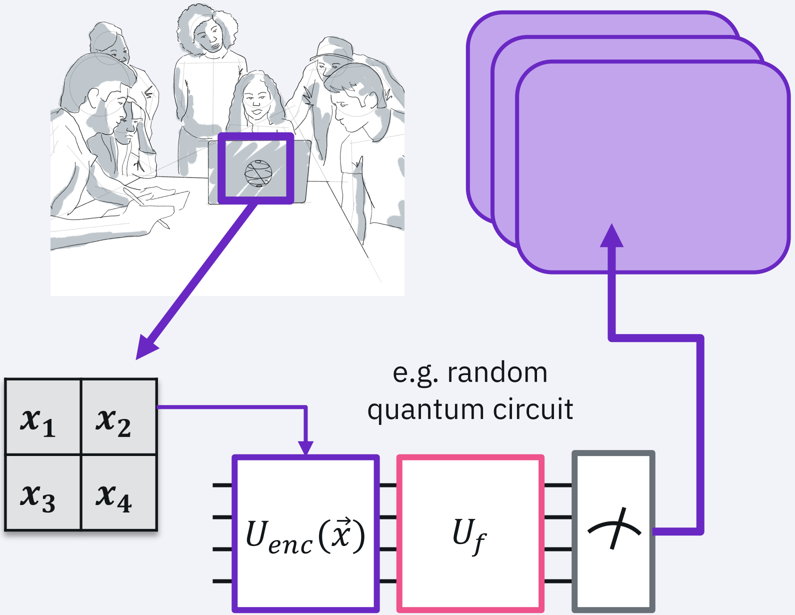
Strategy

Extend the capabilities of classical CNNs by introducing quantum kernels as a new kind of transformation/filter to be employed in convolutional layers.

Key Features

- Compatible and directly integrable with existing classical architectures
- Suitable for near term implementations: shallow and noise-resilient circuits can be used
- No QRAM required (but optimal interfaces with classical data remain an open problem in general)
- Targets qualitative advantage, only conjectured so far, assuming hard-to-simulate kernels

Example

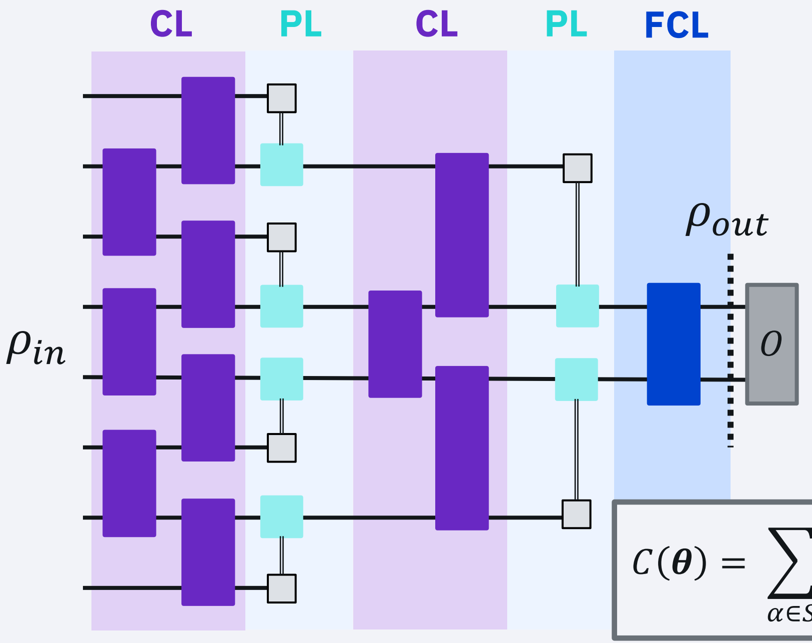


M. Henderson *et al.*, Quantum Machine Intelligence **2**, 2 (2020)

Quantum convolutional networks

Strategy

Build quantum machine learning models inspired by CNNs for the analysis of classical and quantum data



$$C(\boldsymbol{\theta}) = \sum_{\alpha \in S} c_\alpha \mathrm{Tr}[\rho_{out}^{\alpha}(\boldsymbol{\theta}) O]$$

Key Features

- Can have as few as $O(\log N)$ parameters
- Related to hierarchical quantum circuits like tree-tensor networks and multi-scale entanglement renormalization ansatz
- Can be used to analyze classical data or quantum states
  - Quantum phase recognition, quantum error correction, entanglement detection, ...
- **Recent proof**: absence of barren plateaus

Grant *et al.*, npj Quant. Inf. **4**, 65 (2018)
I. Cong *et al.*, Nature Physics **15**, 1273 (2019)
Pesah *et al.*, arXiv:2011.02966 (2020)

Dissipative quantum neural networks

Strategy

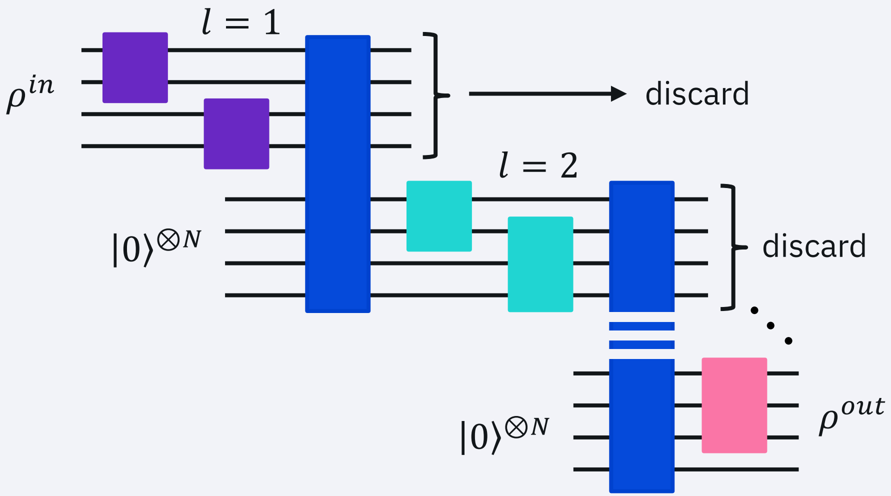
Bulid a feed-forward architecture with general quantum operations to achieve universal quantum computation capabilities

Key Features

- It formulates in QML terms the task of <u>learning an unknown quantum transformation</u>
- The feedforward architecture is represented by the layer-wise composition of quantum CP maps
- Backpropagation-like training can be performed by acting on one layer at a time (depth of the network is not a bottleneck)

K. Beer *et al.*, Nature Communications **11**, 808 (2020)

K. Sharma *et al.*, arXiv:2005.12458 (2020)



$$\rho^{l+1} = \mathrm{Tr}_l\left[\prod_j U_j^l \left(\rho^l \otimes |0 \dots 0\rangle_{l+1}\langle 0 \dots 0|\right) \prod_j U_j^{l,\dagger}\right]$$

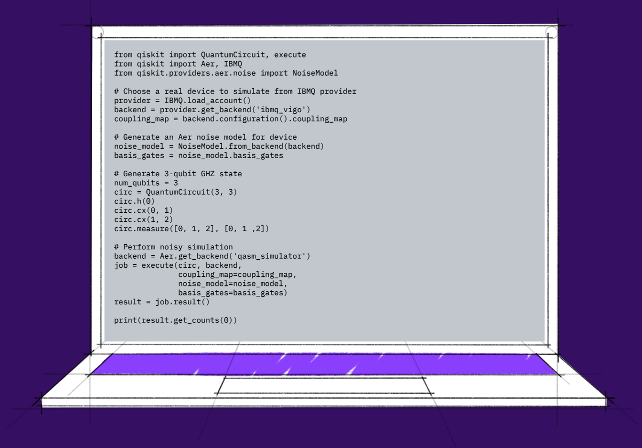
$$\rho^{out} = \mathcal{E}^{out}(\mathcal{E}^L(\dots\mathcal{E}^2(\mathcal{E}^1(\rho^{in}))$$

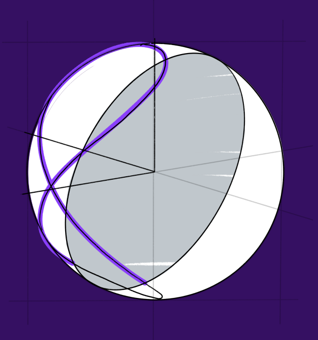
$$C = \frac{1}{|S|}\sum_{x \in S}\langle \phi_x^{out}|\rho_x^{out}|\phi_x^{out}\rangle$$

Target transformation

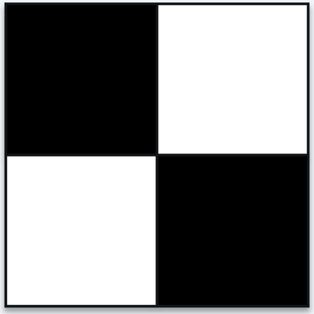
$$|\phi_x^{in}\rangle \to |\phi_x^{out}\rangle$$

Application: simple image classification on real quantum processors

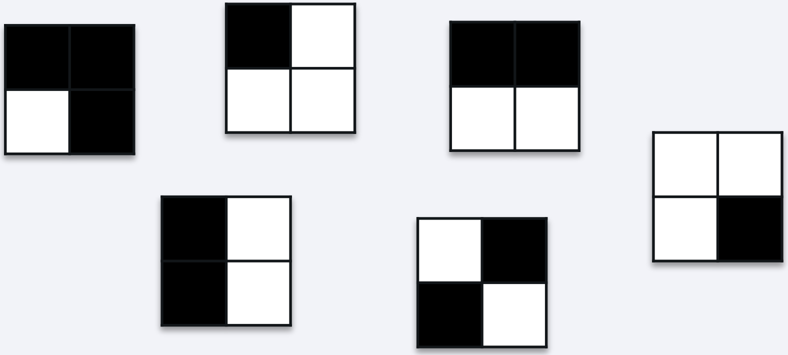




Quantum implementation of a feed-forward network

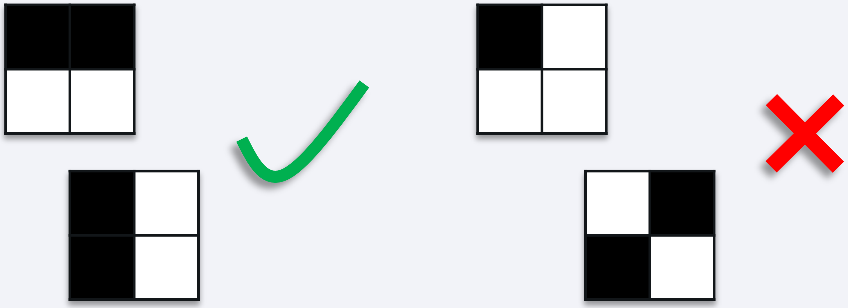


2x2 b/w pixels "images"



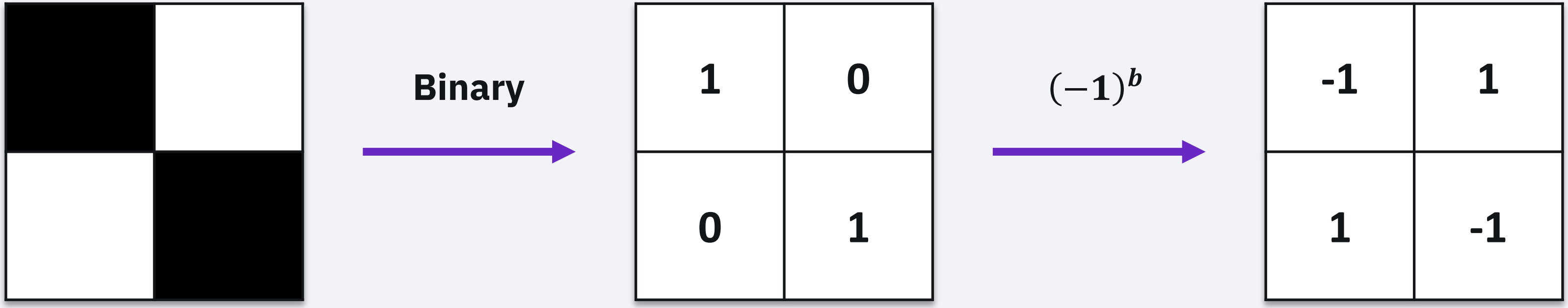
Task

Recognize, among all possible 2×2 b/w pictures, only those with a vertical *or* horizontal line



$2^4 =$ **16 different patterns**

Step 1: encoding



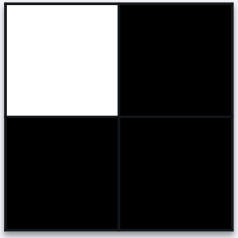
$$\vec{i} = \begin{pmatrix} i_0 \\ i_1 \\ \vdots \\ i_{2^N-1} \end{pmatrix} \qquad i_j = -1, +1$$

2^N-bit **inputs** can be encoded as ± 1 factors in a balanced superposition (**Real Equally Weighted**) of the computational basis states of N qubits.

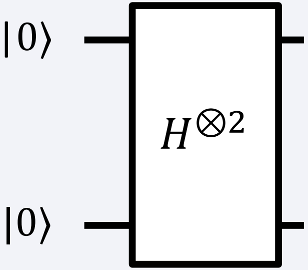
$$\boxed{\langle \psi_i | \psi_w \rangle \propto \sum_j i_j w_j}$$

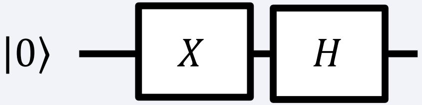
$$|\psi_i\rangle = \frac{1}{\sqrt{2^N}} \sum_{j=0}^{2^N-1} i_j |j\rangle$$

Preparing the encoding states: direct sign flips

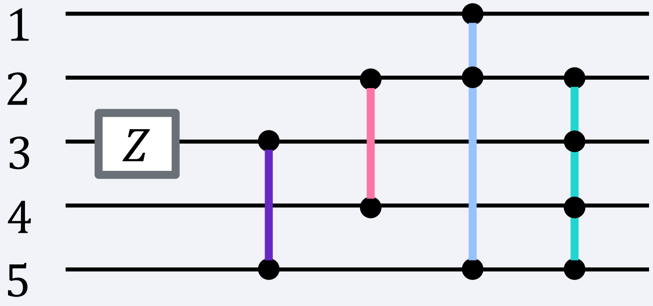


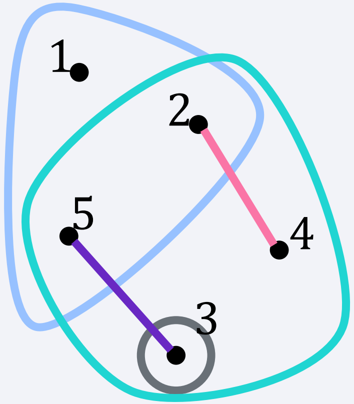
$$2|\psi_i\rangle = |00\rangle - |01\rangle - |10\rangle - |11\rangle$$





Quantum hypergraph states





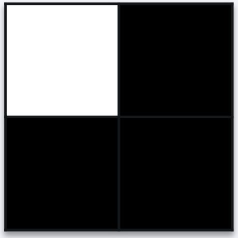
$$|g_{\leq N}\rangle = \prod_{k=1}^{N} \prod_{\{q_1,\dots,q_k\}\in E} C^k Z_{\{q_1,\dots,q_k\}} |+\rangle^{\otimes N}$$

M. Rossi *et al.*, New Journal of Physics **15**, 113022 (2013)

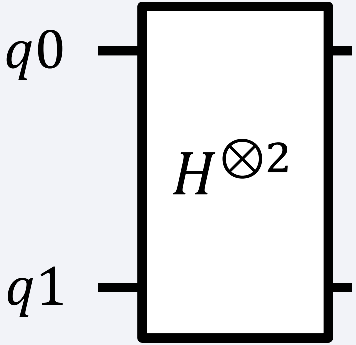
REW ⟷ **Hypergraph states**

The algorithm for the generation of hypergraph states can be adapted to design an optimal preparation procedure for encoding states using at most a single N-controlled operation and a number of p-controlled operations, with $p < N$.

Preparing the encoding states: hypergraphs



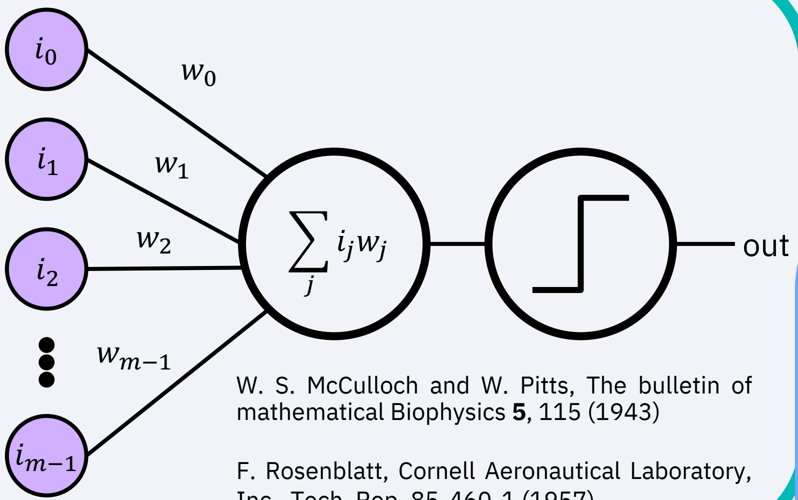
$$2|\psi_i\rangle = |00\rangle - |01\rangle - |10\rangle - |11\rangle$$



Hypergraph States Generation Algorithm

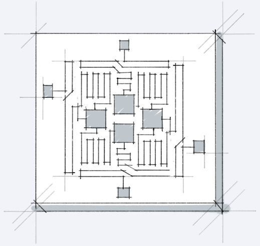
for $P = 1$ to N **do**
  for $j = 0$ to $m - 1$ **do**
    if $(|j\rangle$ has exactly P qubits in $|1\rangle$ **and** $i_j = -1)$ **then**
      Apply $C^P Z$ to those qubits
      Flip the sign of i_k in $\vec{i}\ \forall k$ such that $|k\rangle$ has the same P qubits in $|1\rangle$
    end if
  end for
end for

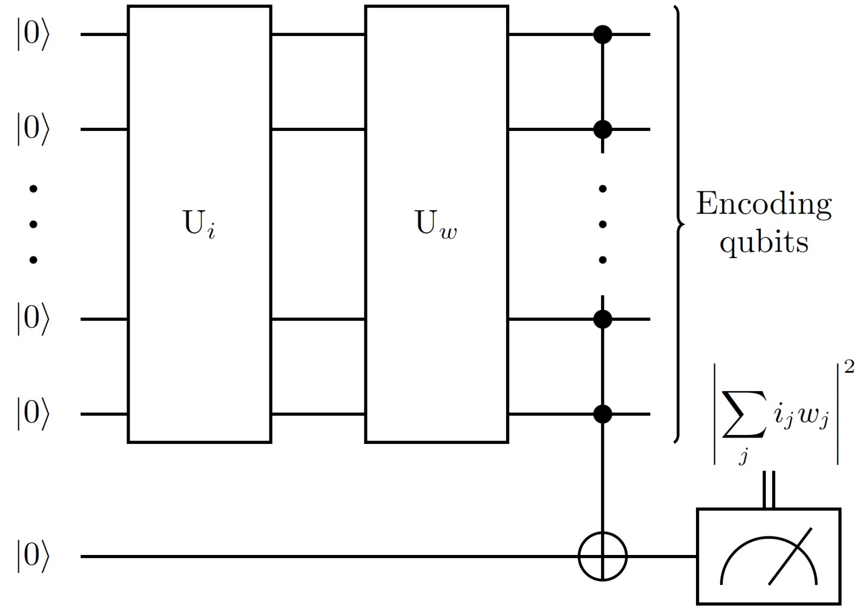
Step 2: single nodes (neurons)



W. S. McCulloch and W. Pitts, The bulletin of mathematical Biophysics **5**, 115 (1943)

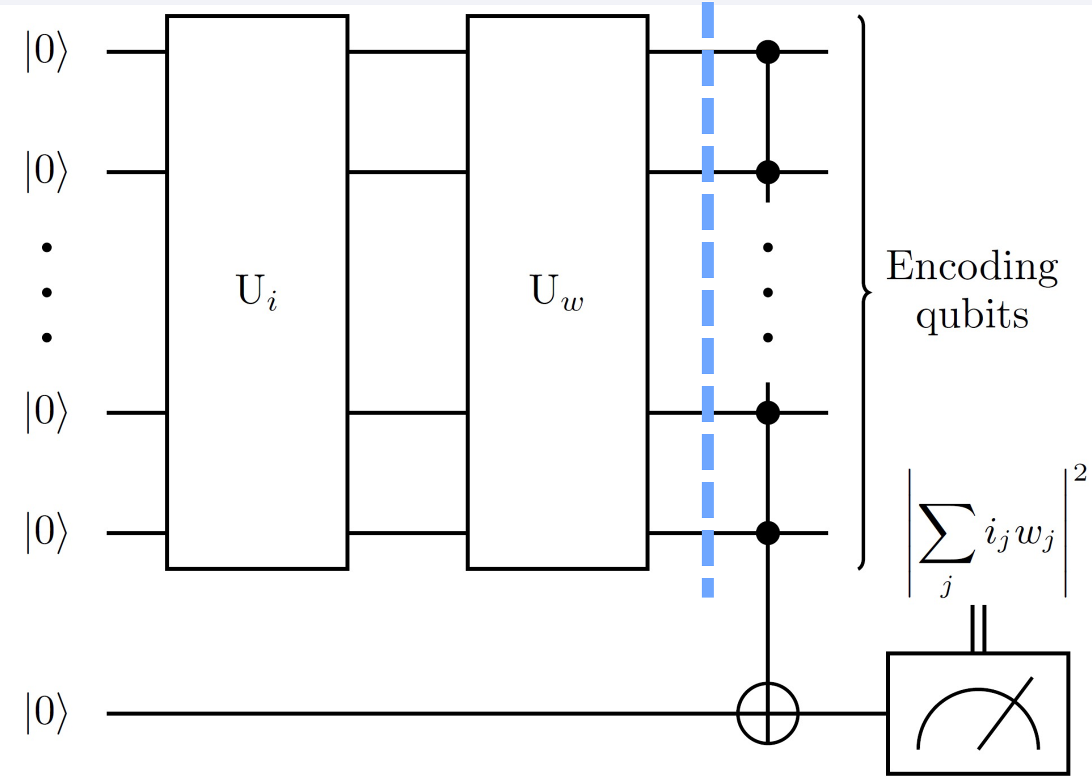
F. Rosenblatt, Cornell Aeronautical Laboratory, Inc., Tech. Rep. 85-460-1 (1957)





We will use these units as filters to recognize single 2x2 patterns

Quantum implementation of artificial neurons



F. Tacchino *et al.*, npj Quantum Information **5**, 26 (2019)

$$|\psi_i\rangle = U_i|0\rangle^{\otimes N}$$

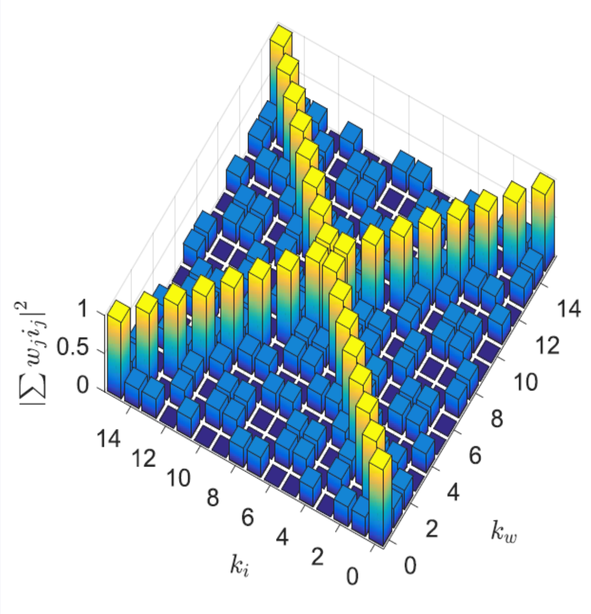
$$|1\rangle^{\otimes N} = U_w|\psi_w\rangle$$

U_w can be taken as essentially the inverse of the input transformation for the vector $\vec{w}$

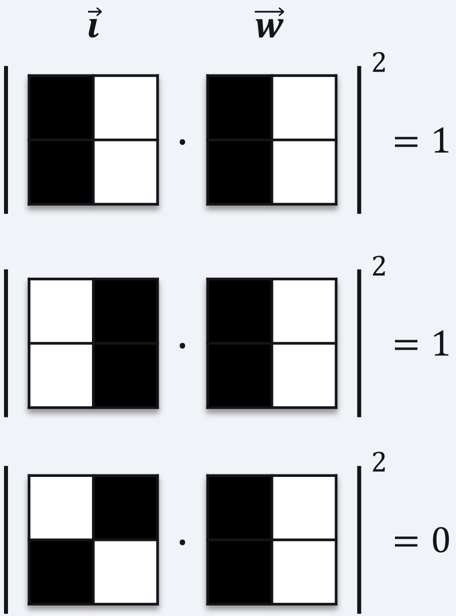
$$\sum_{j=0}^{2^N-2} c_j|j\rangle|0\rangle_a + c_{2^N-1}|2^N-1\rangle|1\rangle_a$$

with $c_{2^N-1} = \langle\psi_i|\psi_w\rangle$

Single neurons on IBM Quantum processors



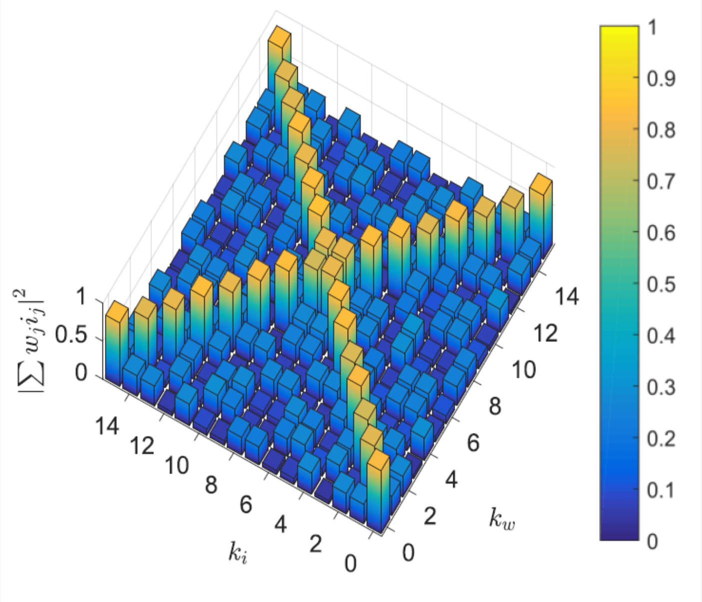
Ideal



Global phase symmetry

↓

An image and its negative are treated equally

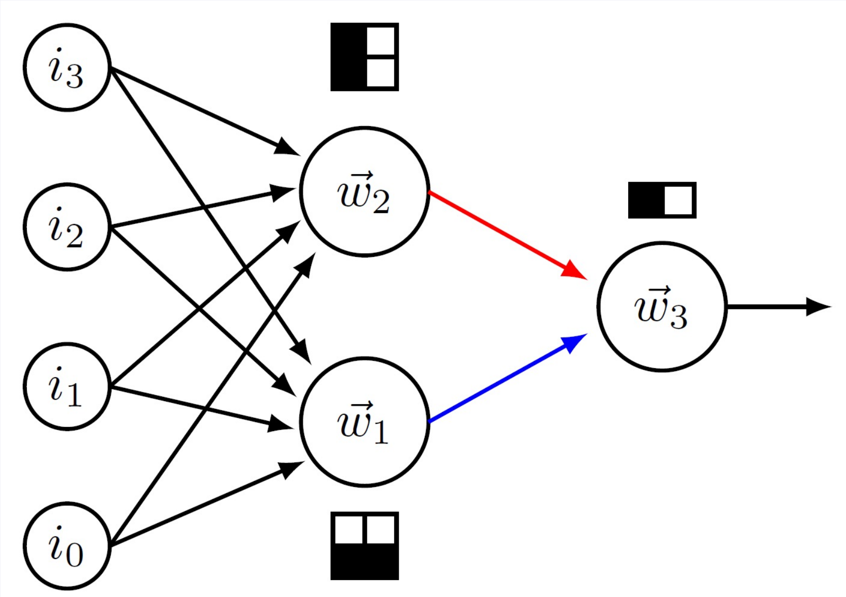


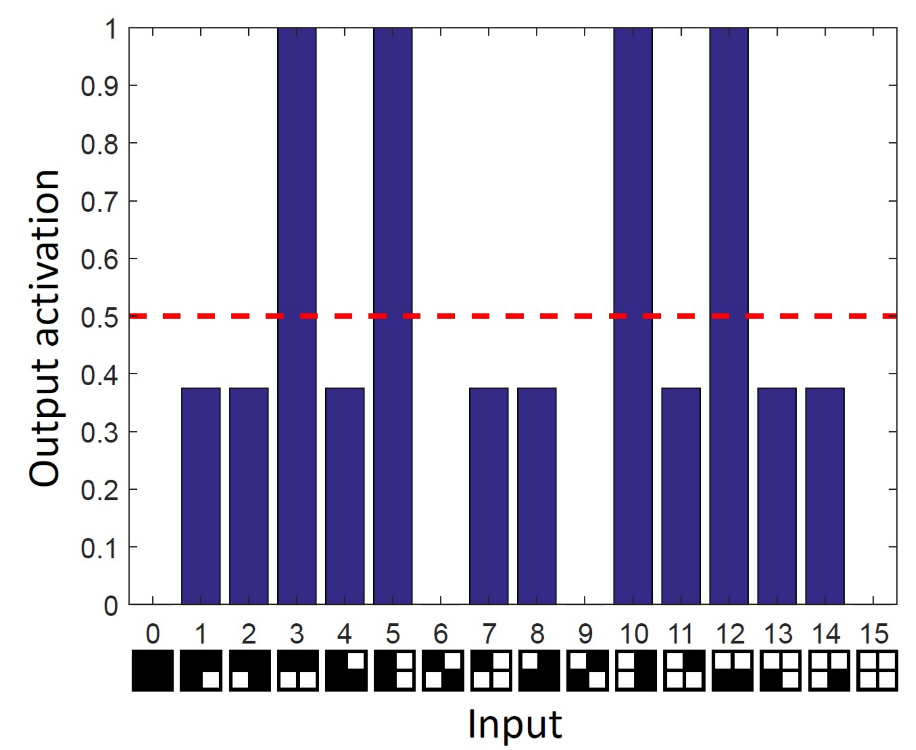
Results from ibmq_5_yorktown

(2 qubits, hypergraph version)

F. Tacchino *et al.*, npj Quantum Information **5**, 26 (2019)

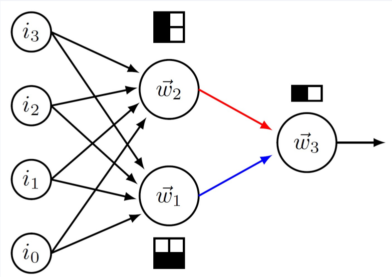
Step 3: build the feed-forward network

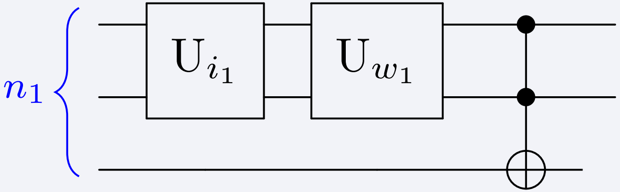


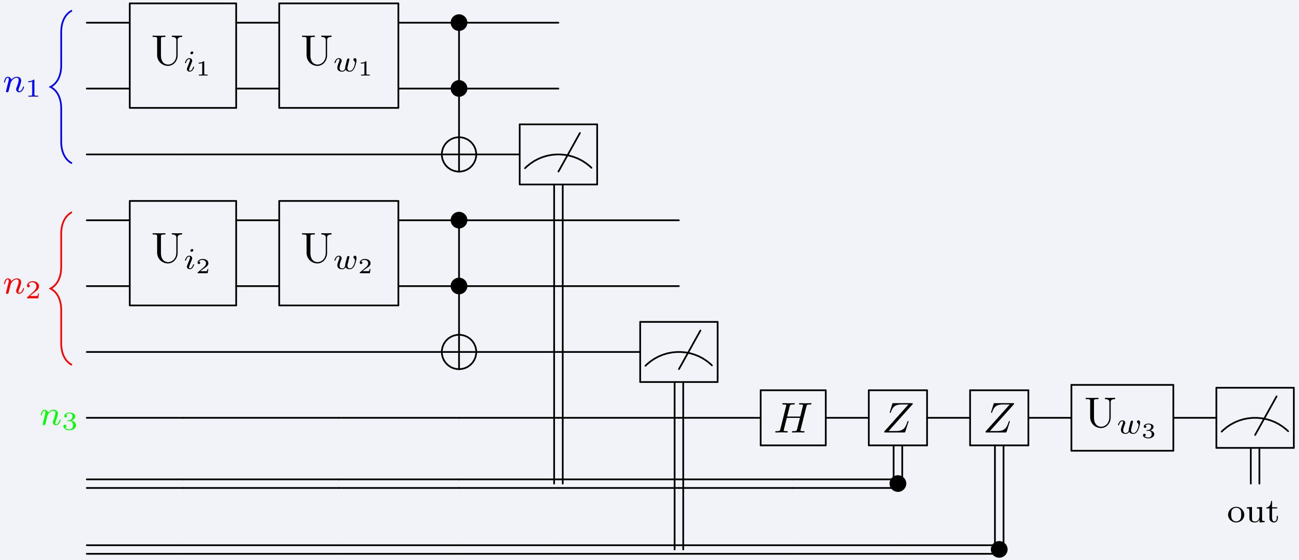


F. Tacchino *et al.*, Quantum Science and Technology **5**, 044010 (2020)

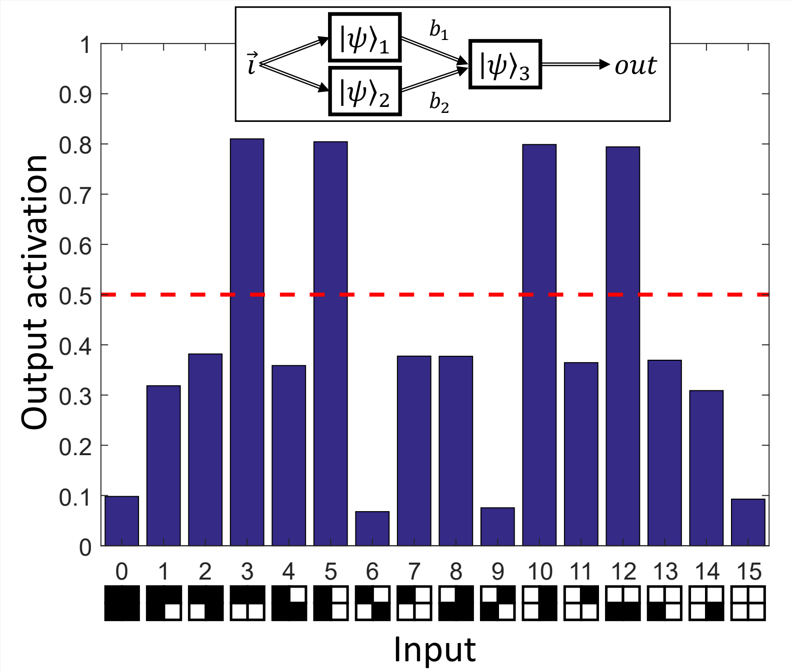
Step 3: build the feed-forward network

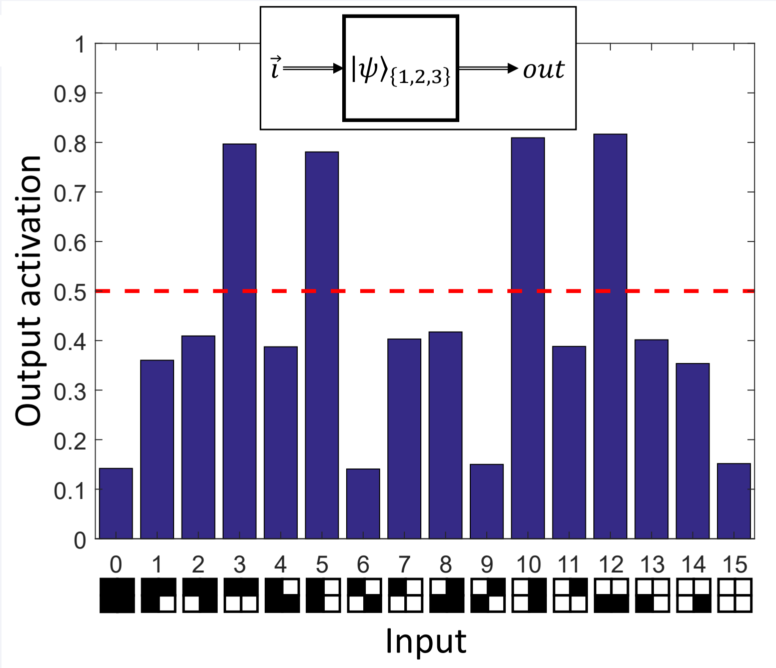






Results





Classical post-processing of single-node outputs

$$p^{out} = \sum_{[b_1,b_2]} p([b_1,b_2])p(a_3 = 1|[b_1,b_2])$$

No post-processing, requires 7 simultaneously coherent qubits

F. Tacchino *et al.*, Quantum Science and Technology **5**, 044010 (2020)