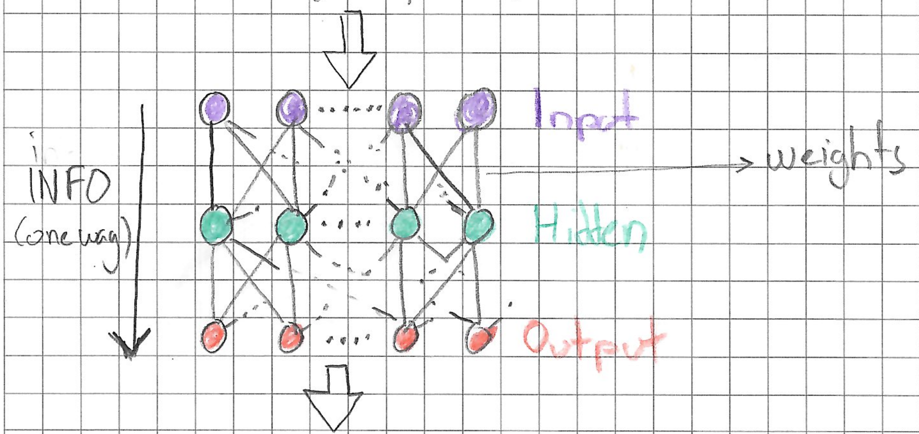
Neural Networks

Universal approximation theorem:

For neural networks every continious function that maps intervals of real numbers to some output interval of real numbers can be approximated arbitrarily closely by a multi-layer perceptron with just one hidden layer (for a wide range of activation functions)

Multi layer perceptron / feed forward network



The sum of the weights is fed to a function called "activation function". When this function is above a threshold, the node becomes "activated" otherwise, it is in its "inactive" state.

Mathematically a single layer $L: \mathbb{R}^n \to \mathbb{R}^m$ performs

$$L(\vec{x}) = f(W\vec{x} + \vec{b})$$

where $W \in \mathbb{R}^{m \times n}$ is the weight matrix $\vec{x} \in \mathbb{R}^n$ is the input vector and $\vec{b} \in \mathbb{R}^m$ is the bias vector

activation function f typically has no free parameters

"Deep" in deep learning comes from stacking multiple such layers, using the output of one layer as input of the next

An N-layer neural network with model parameter set Θ is denoted as

$$\vec{y} = f_{\Theta}(\vec{x}) = \mathcal{L}_N \circ \cdots \circ \mathcal{L}_1(\vec{x}) \qquad (1)$$

Quantum Neural Networks

A full "Quantum Neural Network" title has not been claimed yet. Generally, the term is used for quantum and hybrid algorithms which are optimized/trained by classical co-processors.

Continious Variable Model (CV model)

Instead of using qubits, the information can be coded into quantum effects in continuum. This strategy is arguably better suited for the continious-variable function $f_{\Theta}(\vec{x})$ in (1).

At least, it is realizable with much lower cost and less error-prone at the moment.

CV formalism can be physically realised using optical, microwave, and in ion traps.

In the CV model information is carried in the quantum states of bosonic modes called "qumodes". These form the "wires" of a quantum circuit

Continious-variable quantum information can be encoded using "wave-function representation" or "phase-space representation"

Wave-Function representation

A continious variable (i.e. x) is specified The state of the qumode is then represented by complex-valued wave-function $\Psi(x)$

x can equivalently be p, or real and imaginary part of a quantum field

Phase-Space representation

In the phase-space picture, the conjugate variables x and p are treated on equal footing (i.e. similar to Hamiltonian mechanics), State of the qumode is represented by $(x,p) \in \mathbb{R}^2$, for N qumodes the phase space has 2N variables

Qumode states are represented as real valued functions $F(\vec{x}, \vec{p})$ in phase space called "quasiprobability distributions" Quasi refers to the fact that these functions share some but not all properties of classical probability distributions. Specifically, they can be negative.

Normalisation constraint in CV formalism is that $F(\vec{x}, \vec{p})$ has unit integral over phase space. This is much looser than the qubit systems

Qumode states can also be represented as vectors or density matrices in the countably infinite Hilbert space spanned by the Fock states $\{|n\rangle\}_{n=0}^{\infty}$

Fock states in photonic case are the eigenstates of photon number operator $\hat{n} = (\hat{x}^2 + \hat{p}^2 - 1)/2$ where $\hat{x}$ is the position operator and $\hat{p}$ is the momentum operator

The phase-space and Hilbert space formulations give equivalent predictions. The CV quantum systems can be approached in wavelike or particle like perspective

Gaussian operations

Gaussian gates are "easy" gates in CV quantum computer

For a system of N modes, the most general Gaussian transformation has the effect

$$\begin{bmatrix} \vec{x} \\ \vec{p} \end{bmatrix} \longmapsto M \begin{bmatrix} \vec{x} \\ \vec{p} \end{bmatrix} + \begin{bmatrix} \vec{\alpha}_r \\ \vec{\alpha}_i \end{bmatrix}$$

i.e. they are linear, corresponding to affine transformations

M is a real valued symplectic matrix and $\alpha \in \mathbb{C}^N \cong \mathbb{R}^{2N}$ is a complex vector with real $(\vec{\alpha}_r)$ and imaginary $(\vec{\alpha}_i)$ parts

M by definition obeys

$$M^T \Omega M = \Omega$$

$$\Omega = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

The gates that obey the above criteria are

$$\hat{R}(\phi) : \begin{bmatrix} \vec{x} \\ \vec{p} \end{bmatrix} \longmapsto \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix} \quad \text{(rotation)}$$

GÜLPAŞ

$$\hat{D}(\alpha) : \begin{bmatrix} \vec{x} \\ \vec{p} \end{bmatrix} \mapsto \begin{bmatrix} x + \sqrt{2}\, Re(\alpha) \\ p + \sqrt{2}\, Im(\alpha) \end{bmatrix} \qquad \text{(displacement)}$$

$$\hat{S}(r) : \begin{bmatrix} \vec{x} \\ \vec{p} \end{bmatrix} \longmapsto \begin{bmatrix} e^{-r} & 0 \\ 0 & e^{r} \end{bmatrix} \begin{bmatrix} \vec{x} \\ \vec{p} \end{bmatrix} \qquad \text{(squeezing)}$$

$$\hat{BS}(\theta) : \begin{bmatrix} x_1 \\ x_2 \\ p_1 \\ p_2 \end{bmatrix} \rightarrow \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & \cos\theta & -\sin\theta \\ 0 & 0 & \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ p_1 \\ p_2 \end{bmatrix} \qquad \text{(Phaseless beam splitter)}$$

Fully Connected Quantum Layers

Each layer in CV NN consists of

$$\mathcal{L} := \hat{\phi} \circ \hat{D} \circ \hat{U}_2 \circ \hat{S} \circ \hat{U}_1$$

where $\hat{U}_i = \hat{U}_i(\theta_i, \phi_i)$ are general N-port linear optical interferometers containing BS and R gates. $\hat{D} = \otimes_{i=1}^{N} \hat{D}(\alpha_i)$ and $\hat{S} = \otimes_{i=1}^{N} \hat{S}(r_i)$ are collective displacement and squeezing operators. $\hat{\phi} = \hat{\phi}(\lambda)$ is some non-gaussian gate (cubic phase, kerr gate, etc.)