Quantum GANs
A Physics Perspective

Roberto Cappuccio

Dipartimento di Scienze Fisiche, della Terra e dell’Ambiente

Uiversità di Siena

Index

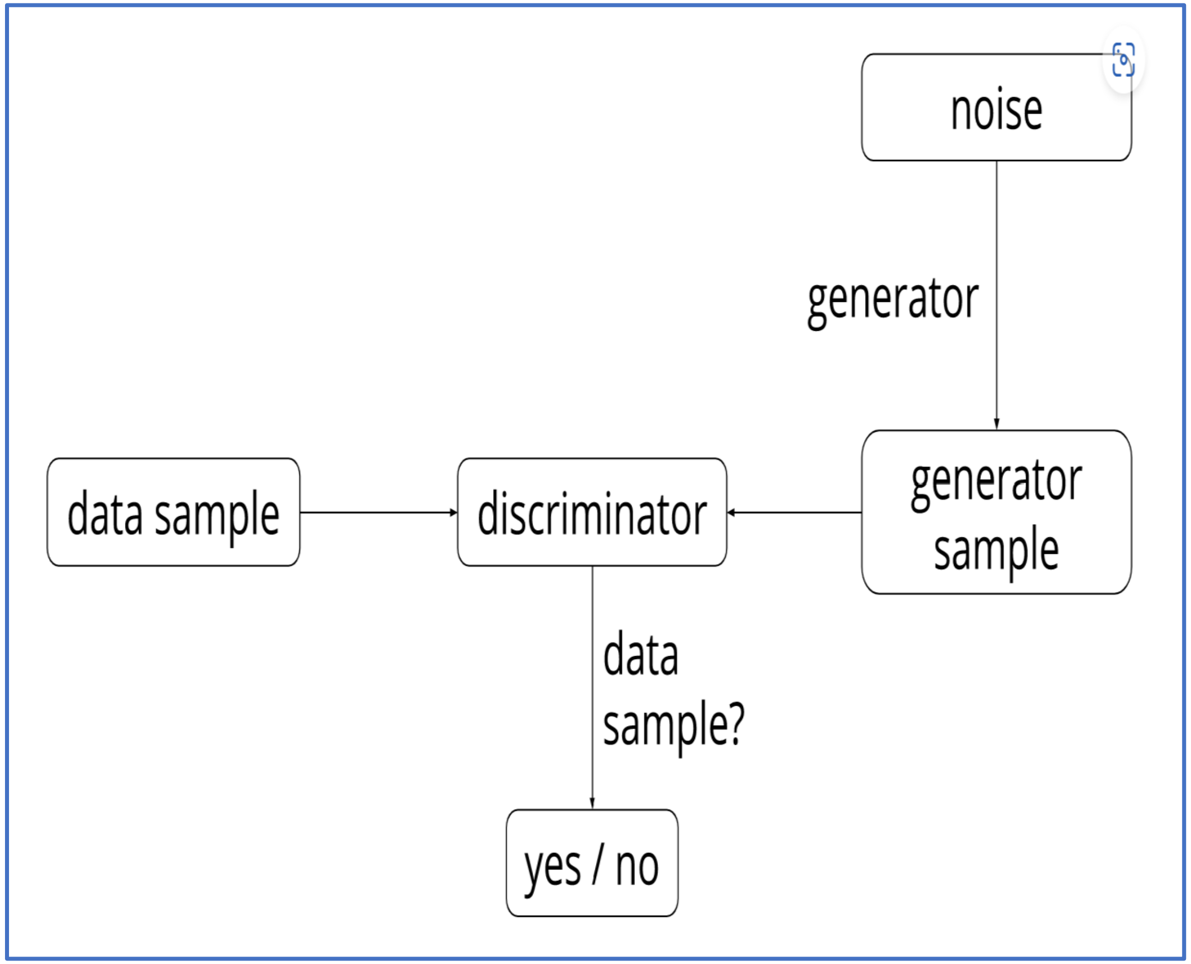
What's a GAN

GAN stands for Generative Adversarial Network

- The specific goal of GANs is to generate data from scratch
- Typical domain is Images but not only (Videos, Music, Texts, Functions, Simulation Data etc)
- GANs are made up by 2 main blocks:
  - **G**ENERATOR (Leo): it generates fake data from noise distribution, by pretending they are true
  - **D**ISCRIMINATOR (Tom): it wants to prove the generated date are fake

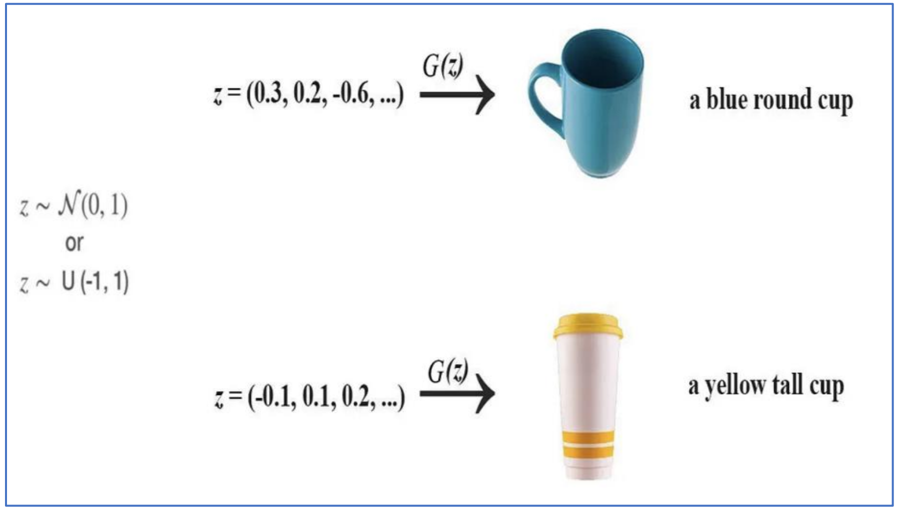
What's a GAN

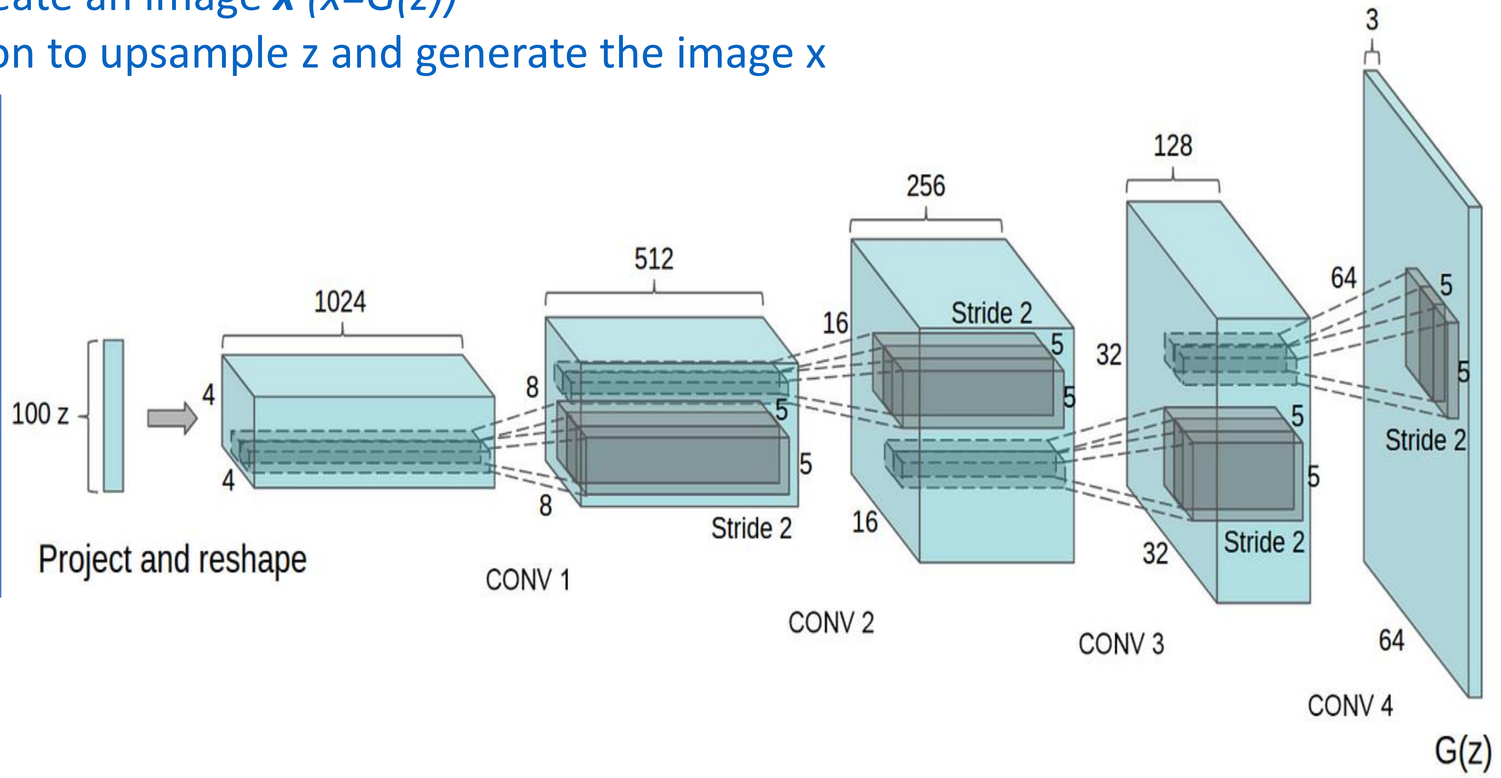
- GAN pits two neural networks against each other:
  - a generator network G(z)
  - a discriminator network D(x).
- The generator tries to mimic examples from a training dataset, which is sampled from the true data distribution q(x)). It does so by transforming a random source of noise received as input into a synthetic sample.
- The discriminator receives a sample, but it is not told where the sample comes from. Its job is to predict whether it is a real data sample or a synthetic sample.
- The discriminator is trained to make accurate predictions, and the generator is trained to output samples that fool the discriminator into thinking they came from the data distribution.



What's G?

- The Generator samples noise ***z*** using a normal or uniform distribution.
- With ***z*** as an input, it uses a generator ***G*** to create an image ***x*** *(x=G(z))*
- It performs multiple transposed de-convolution to upsample z and generate the image x





- The Generator alone will just create random noise.
- The Discriminator "conceptually" provides guidance to the Generator on what images to create.

Information Theory Background

- The information *I* of an event *x* with probability distribution *P(x)* is defined as: $I(x) = -\log_2 P(x)$
- The entropy *H* measures the amount of information and is defined as: $H(x) = E_{x\sim P}[\,I(x)]$
- Then we have: $H(x) = -\,E_{x\sim P}[\log_2 P(x)] = \sum_x P(x)\log_2 P(x))$
- Example: entropy of a coin launch: $H(x) = -p(head)\log_2 p(head) - p(tail)\log_2 p(tail) = -\frac{1}{2}\log_2\left(\frac{1}{2}\right) = 1$
- The Cross-Entropy *H* of an event *x* with probability ditribution *P(x)*, but relative to a differente sample *z* with probability distribution *Q(z)*, is defined as: $H(x) = \sum_x P(x)\log_2 Q(z)$

- *D* is a classifier $X \rightarrow \{0, 1\}$ that tries to distinguish between:
  - a sample from the data distribution *(D(x) = 1, for x* $\sim p_{data}$*)*
  - a sample from the model distribution *(D(G(z)) = 0, for z* $\sim p_{noise}$*)*

- *G* is a generator $Z \rightarrow X$ trained to produce samples *G(z)* for $z \sim p_{noise}$ that are difficult for *D* to distinguish from data.

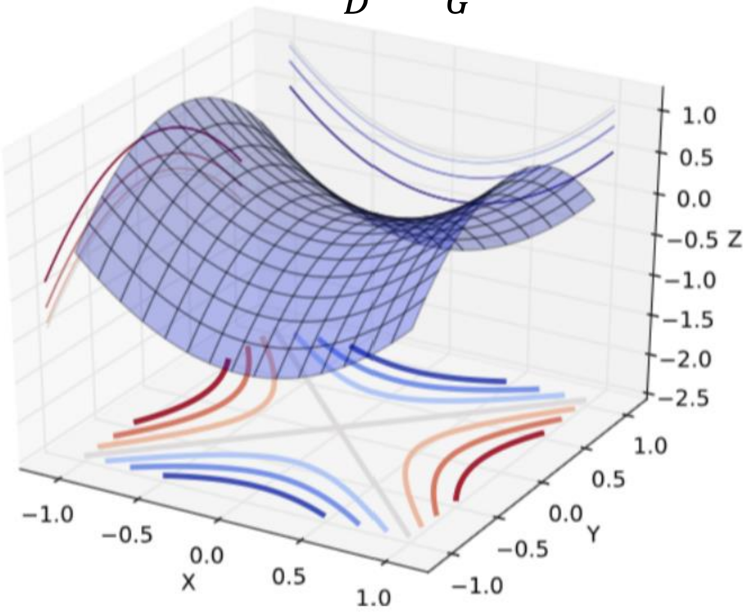
GAN Value Function (1)

- Consider the value function: $V(D,G) = E_{x \sim p_{data}}[\log_2(D(x))] + E_{z \sim p_{noise}}[\log_2(1 - D(G(z)))]$

- We want to:
  - for fixed *G*, find *D* which maximizes *V(D, G)*
  - for fixed *D*, find *G* which minimizes *V(D, G)*

$$(D^*, G^*) = \max_D \min_G V(D,G)$$

- In other words, we are looking for the saddle point:



- An "optimal" discriminator is :

$$D_G^*(x) = \frac{p_{data}(x)}{p_{data}(x) + p_G(x)}$$

GAN Value Function (2)

- The D_{KL} Kullback-Leibler divergence is defined as: $D_{KL}\,(P||Q) = E_x \log_2 \frac{P(x)}{Q(x)}$

- Using the optimal discriminator $D_G^*\,(x)$ we have:

$$D_G^*\,(x) = \frac{p_{data}(x)}{p_{data}(x)\,+\,p_G\,(x)}$$

$$V(D,G) = \underbrace{D_{KL}\left[p_{data}, \frac{p_{data}(x)+p_G\,(x)}{2}\right] + D_{KL}\left[p_G, \frac{p_{data}(x)+p_G\,(x)}{2}\right]}_{\text{notice the difference}} = \underbrace{D_{JSD}[p_{data}(x), p_G\,(x)]}_{2\times\text{Jensen-Shannon Divergence (JSD)}} - \log_2(4)$$

The GAN Training Algorithm

- Sample minibatch of *m* training points *x (1) , x (2) , . . . , x (m)* from *D*

- Sample minibatch of *m* noise vectors *z (1) , z (2) , . . . , z (m)* from p_z

- Update the discriminator parameters ϕ by stochastic gradient **ascent**:

$$\nabla_\phi V(G_\theta, D_\phi) = \frac{1}{m} \nabla_\phi \sum_{i=1}^{m} [\log D_\phi\, (\boldsymbol{x}^{(i)}) \;+\; \log(1 \;-\; D_\phi\, (G_\theta\, (\boldsymbol{z}^{(i)}\,)))\,]$$

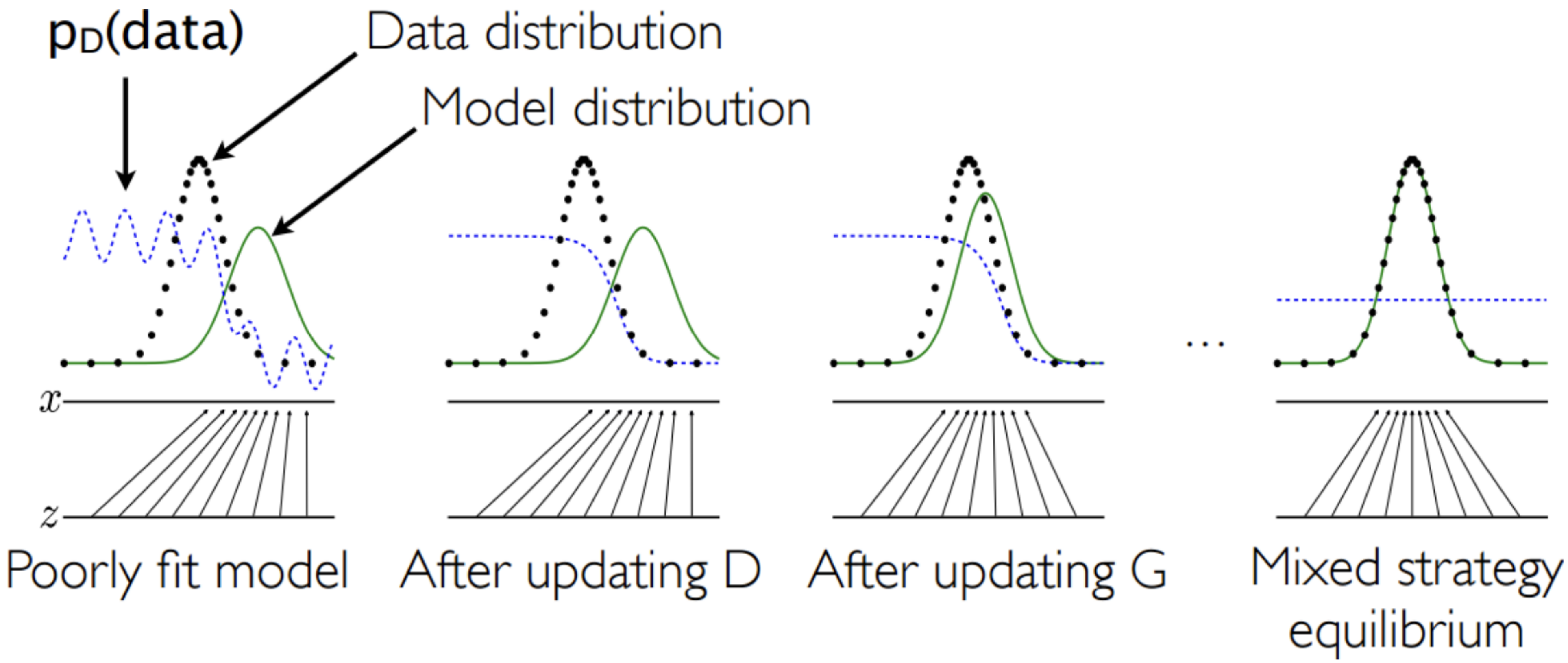
- Update the generator parameters ϑ by stochastic gradient **descent**:

$$\nabla_\theta V(G_\theta, D_\phi) = \frac{1}{m} \nabla_\theta \textstyle\sum_{i=1}^{m} \log(1 \;-\; D_\phi\, (G_\theta\, (\boldsymbol{z}^{(i)}\,)))$$

- Repeat for fixed number of epochs

GAN Gradient Calculation

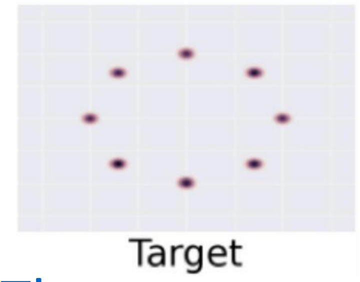
- Assuming *D* and *G* are neural networks parameterized by ϑ_D and ϑ_G , backpropagation can be used to optimize *D*'s and *G*'s objectives alternatively until convergence.



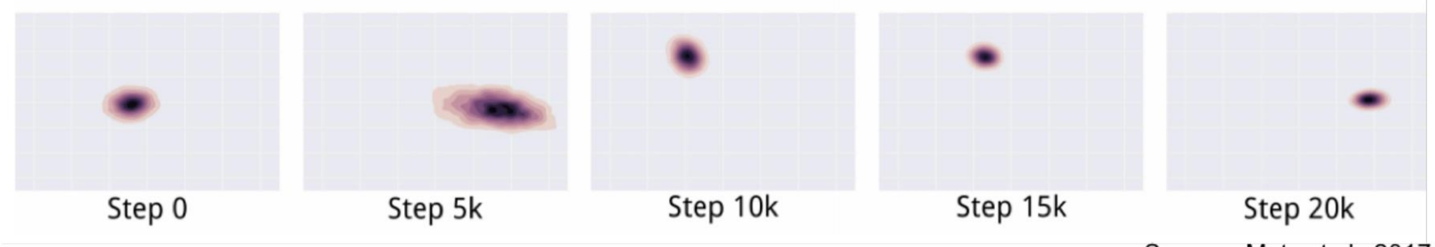
GANs Collapse

- GANs are notorious for suffering from mode collapse
- Intuitively, this refers to the phenomena where the generator of a GAN collapses to one or few samples (dubbed as “modes”)
- There is not a general solution but just empirical-driven methods (*Q-GAN may have it*)

- Example: the target distribution is a mixture of Gaussians



- The generator distribution keeps oscillating between different modes



Source: Metz et al., 2017

Demo Classical GAN

Demo Classical GAN

Quantum Computing Short Introduction

- Quantum Logic, Gates and Circuits
- Quantum Algorithms
- Demo Pennylane Quantum GAN
- Demo Qiskit Quantum GAN

Go to Q.C. Intro

Quantum GAN

- Quantum GANs are the transposition of the GAN model on quantum computing

- Generalizing their framework to the quantum settings:

- classical data → density matrix $\sigma = \sum_i p_i |\psi_i\rangle\langle\psi_i|$

- Generator → unitary operator on a quantum circuit U that outputs the quantum state ρ:
  $\rho = U(\theta_g)\rho_0 U^\dagger(\theta_g)$ from the initial state ρ_0

- Discriminator → positive operator T that takes both true data σ and fake data ρ in input and outputs the probability the data being true: $D(\theta_d, \rho_{\text{in}}) = \text{Tr}[T\rho_{\text{in}}]$

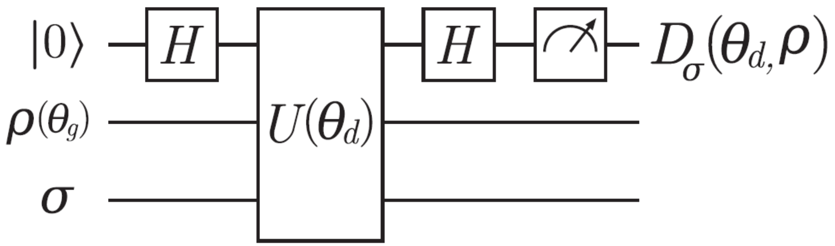
- The Q-GAN must solve the "minimax game": $\min_{\theta_g} \max_T \left(\text{Tr}[T\sigma] - \text{Tr}[T\rho(\theta_g)]\right)$

Convergence of Quantum GAN

- To evaluate the fidelity of the ρ density matrix respect to σ the usage of Helstrom measurement operator $T = P^{+}(\sigma - \rho)$ is not a good choice since it may lead to oscillation

- Niu et alii proposed to leverage a specific feature of Quantum Computing that is the entanglement: rather than evaluating either fake or true data individually the propsed discriminator perform a measurement on the **joint system** of the true data σ and the generated data $\rho(\Theta_g)$

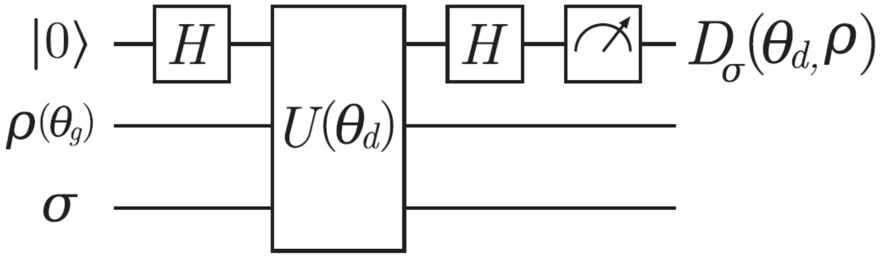
- The «entangled» fidelity is: $D_\sigma^{\text{fid}}[\rho(\theta_g)] = \left[\text{Tr}\sqrt{\sigma^{1/2}\rho(\theta_g)\sigma^{1/2}}\right]^2$



EQ-GAN architecture:

- the generator varies Θ_g to fool the discriminator;
- the discriminator varies Θ_d to distinguish the state;
- the global optimum of the EQ-GAN occurs when $\rho(\Theta_g) = \sigma$

Entangling Q-GAN



Given this architecture the unitary *U* operator must be implemented in order to evaluate :

$$\min_{\theta_g} \max_{\theta_d} V(\theta_g, \theta_d) = \min_{\theta_g} \max_{\theta_d} \{1 - D_\sigma[\theta_d, \rho(\theta_g)]\}$$

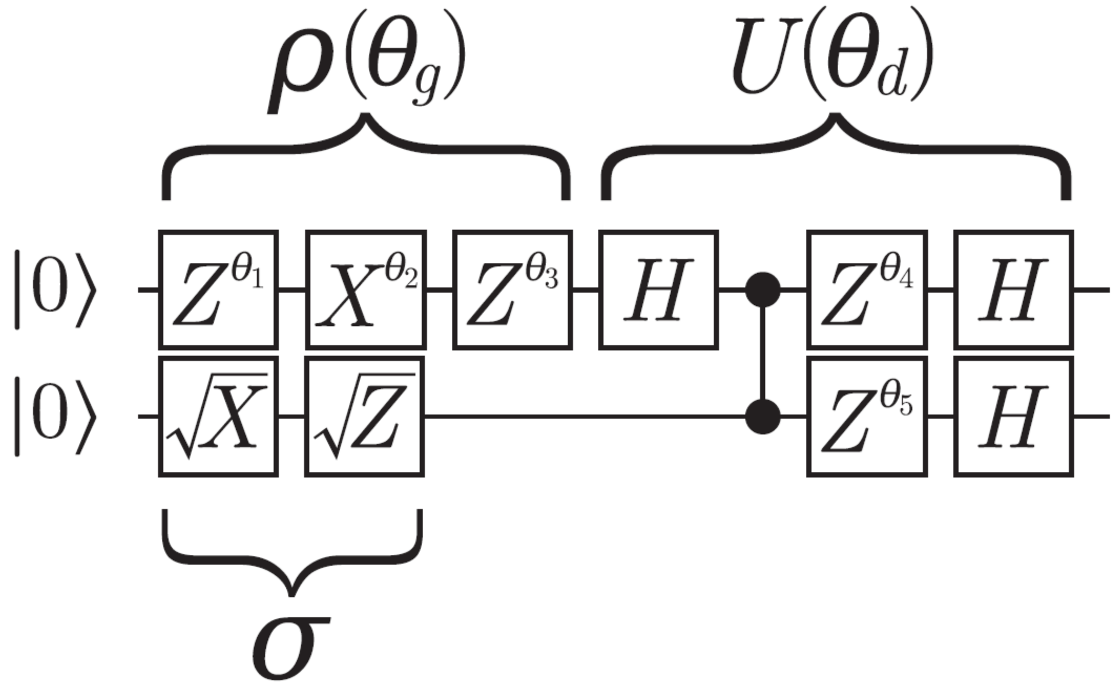
Instead simply swapping *ρ* and *σ*, let's take *U* such that: $U(\theta_d) = \exp[-i\theta_d \mathrm{CSWAP}]$

So we have:

$$\begin{aligned} HU(\theta_d)H|0\rangle_{\mathrm{a}}|\psi\rangle|\zeta\rangle &= \frac{i \sin\theta_d}{2} |1\rangle_{\mathrm{a}}[|\zeta\rangle|\psi\rangle - |\psi\rangle|\zeta\rangle] \\ &+ \frac{1}{2}|0\rangle_{\mathrm{a}}[(e^{-i\theta_d} + \cos\theta_d)|\psi\rangle|\zeta\rangle \\ &- i\sin\theta_d|\zeta\rangle|\psi\rangle] \end{aligned}$$

$$D_\sigma[\theta_d, \rho(\theta_g)] = \frac{1}{2}\{1 + \cos^2\theta_d + \sin^2\theta_d D_\sigma^{\mathrm{fid}}[\rho(\theta_g)]\}.$$

Entangling Q-GAN Circuit



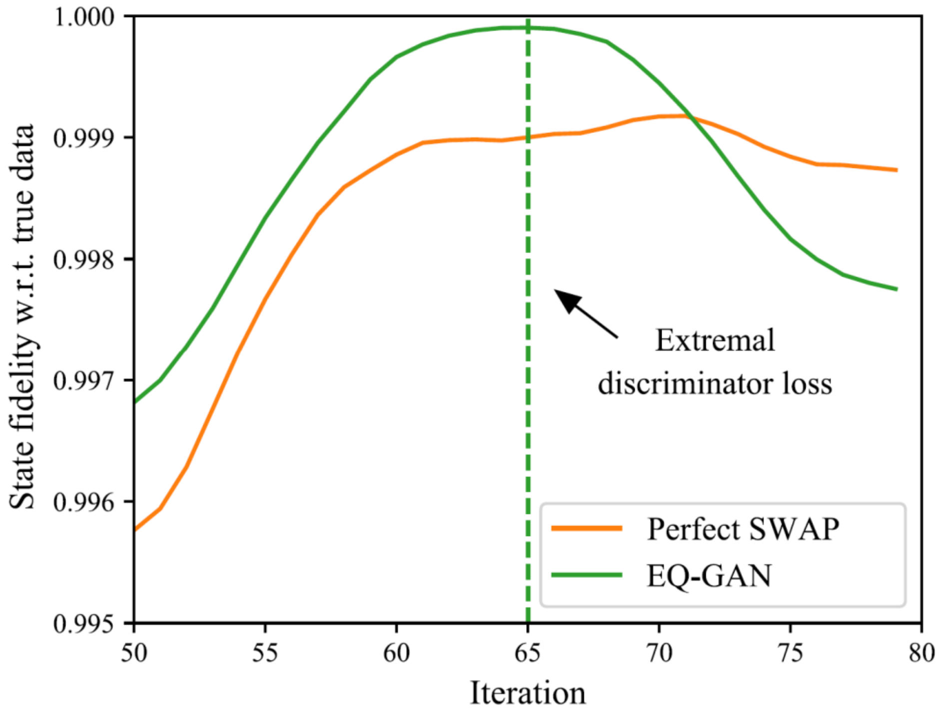
View the circuit in the IBM composer

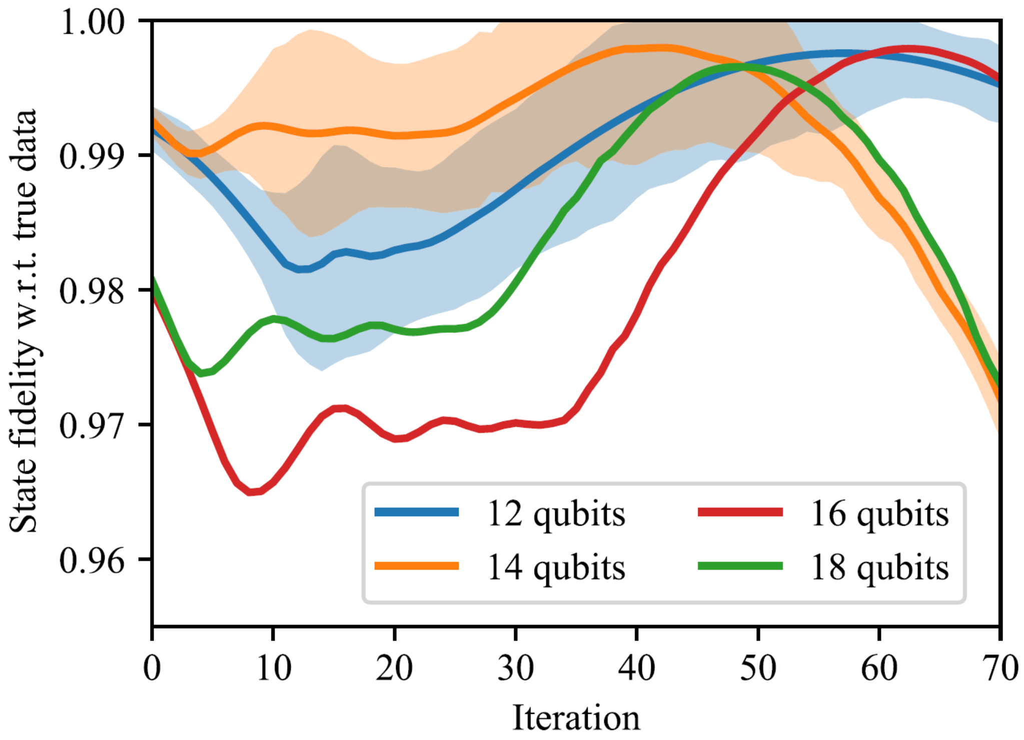
θ_1, θ_2, …, θ_5 are the rotation parameters that are randomly set and then modified by the learning algorithm

Entangling Q-GAN Results

Minimum error in state fidelity → (1-state fidelity)

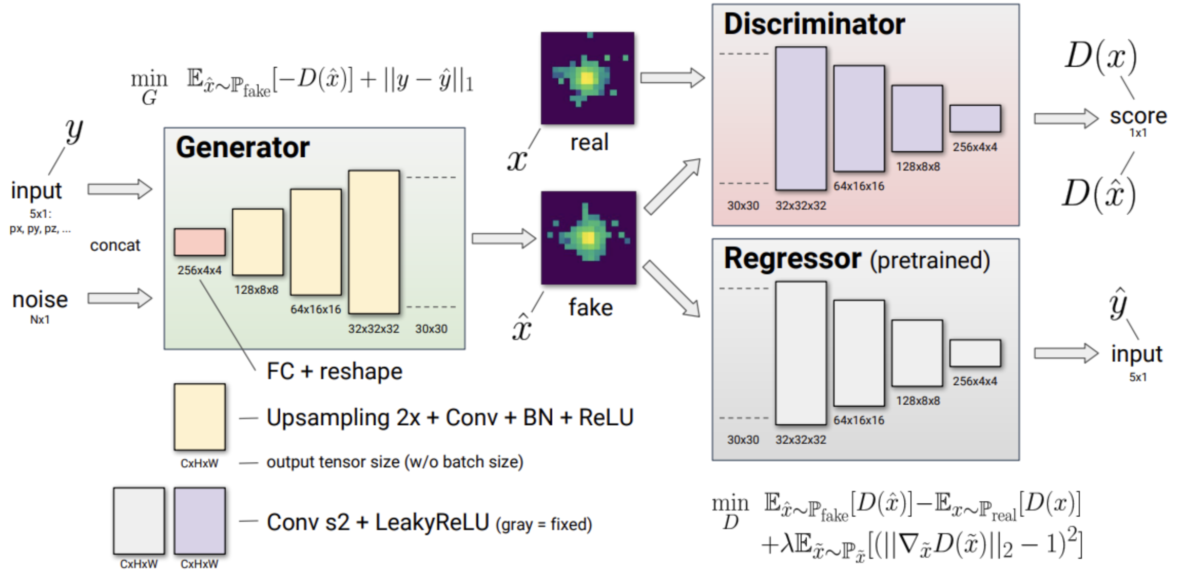
QML model	Minimum error in state fidelity
Perfect swap	$(2.4 \pm 0.5) \times 10^{-4}$
EQ-GAN	$(\mathbf{0.6 \pm 0.2}) \times \mathbf{10^{-4}}$

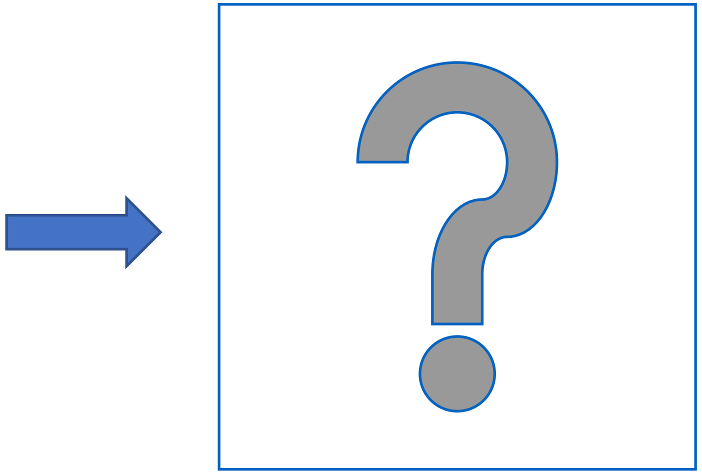




Q-GAN in Physics

- **Dual-Parameterized Quantum Circuit GAN Model in High Energy Physics**
  Su Yeon Chang, Steven Herbert, Sofia Vallecorsa, Elías F. Combarro, Ross Duncan
- **Quantum Generative Adversarial Networks in a Continuous-Variable Architecture to Simulate High Energy Physics Detectors**
  Su Yeon Chang, Sofia Vallecorsa, Elías F. Combarro, Federico Carminati
- **Generative Adversarial Networks for LHCb Fast Simulation**
  Fedor Ratnikov





Q-GAN Version

Demo Classical GAN

<u>Demo Quantum GAN</u>

References

- Quantum Computation and Quantum Information
  Michael A.Nielsen – Isaac L.Chuang – Cambridge Press

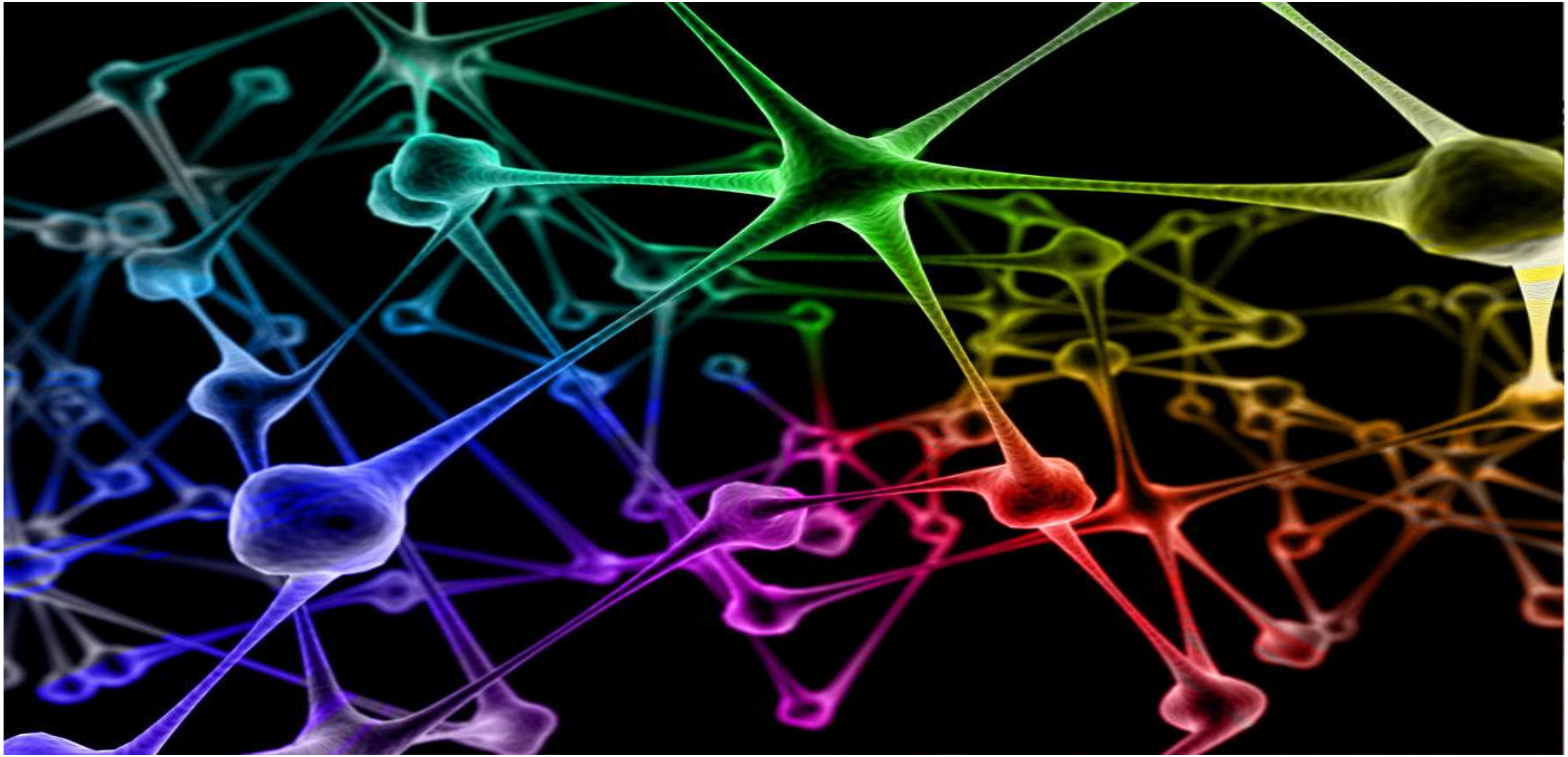
- An Introduction to Quantum Computing
  Philippe Kaye – Raymond Laflamme – Michele Mosca – Oxford University Press

- Entangling Quantum Generative Adversarial Networks
  Murphy Yuezhen Niu, Alexander Zlokapa, Michael Broughton, Sergio Boixo, Masoud Mohseni, Vadim Smelyanskyi, Hartmut Neven

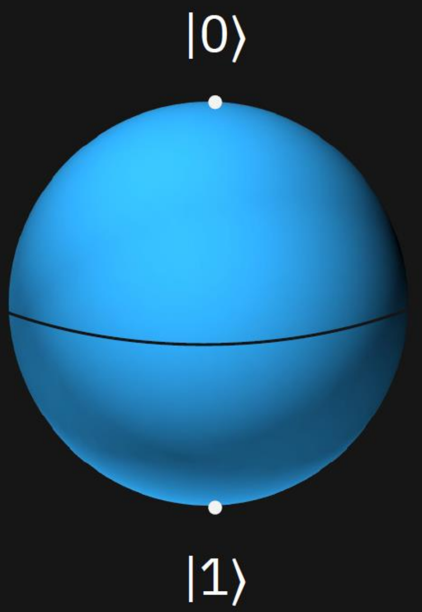
Questions & Answers

Q&A

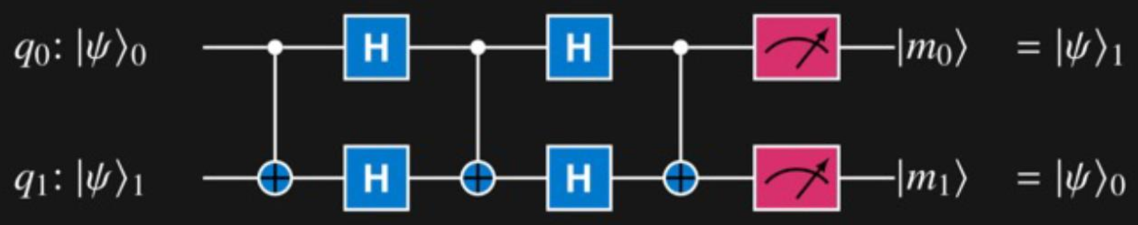
Thanks for your attention!



Classical vs Quantum Logic

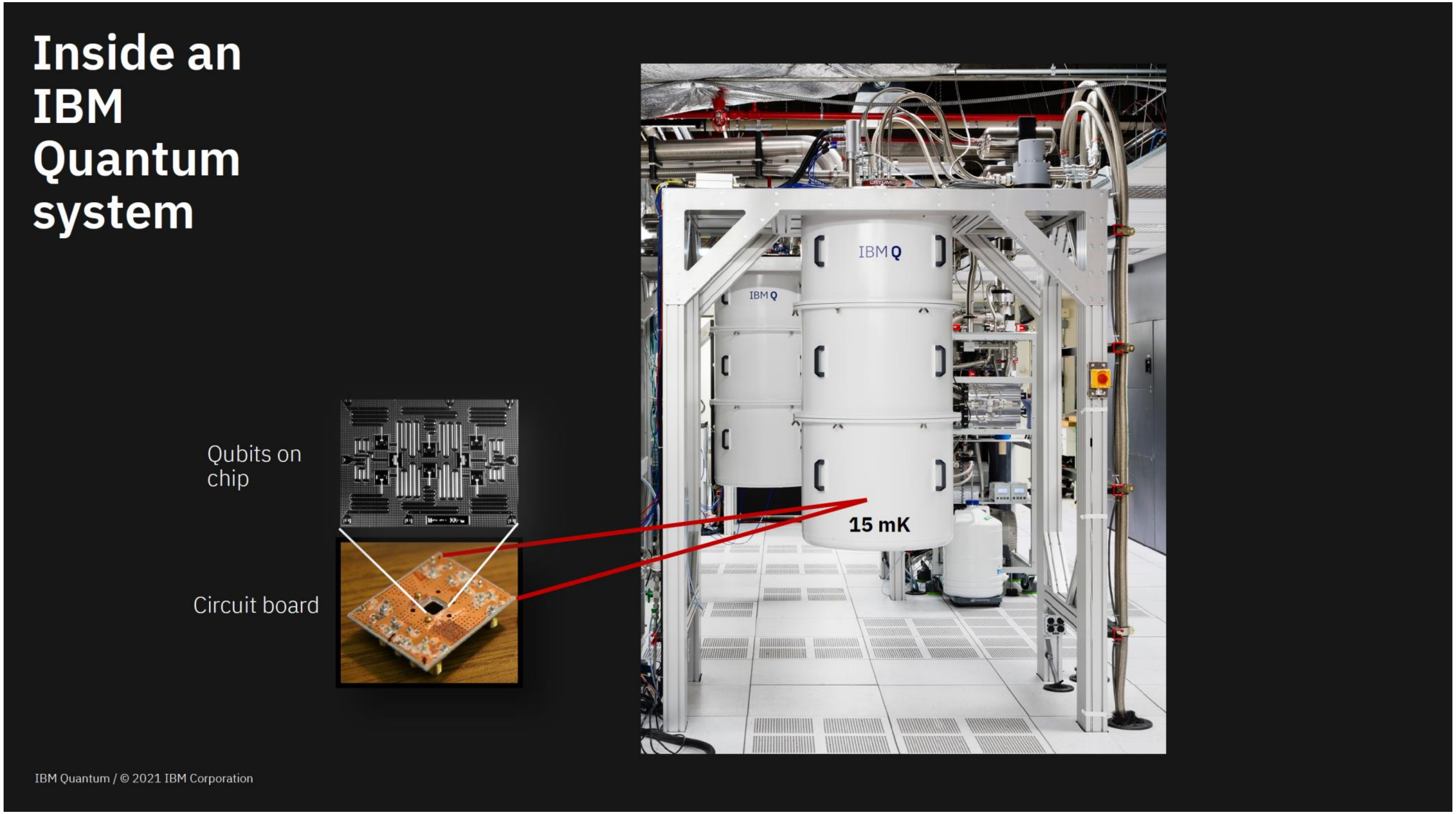


A quantum bit or qubit is a controllable quantum object that is the unit of information



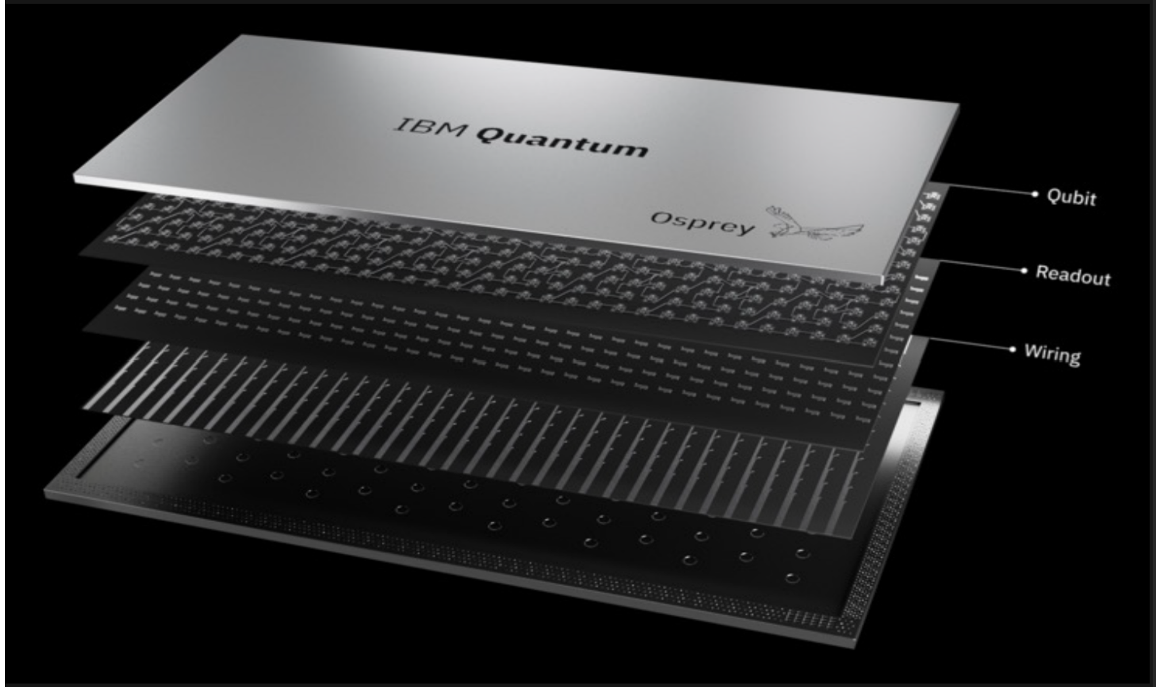
A quantum circuit is a set of quantum gate operations on qubits and is the unit of computation

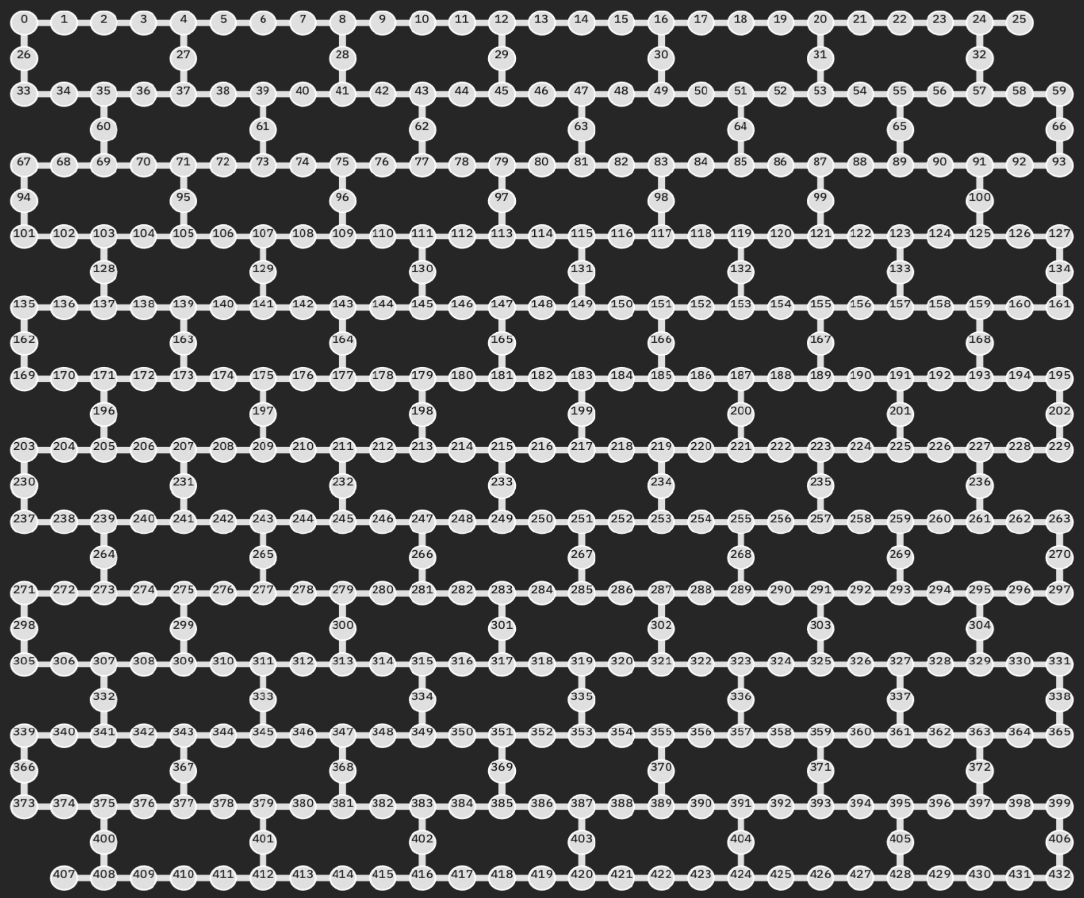
Quantum Computer Hardware?



IBM Osprey Chip

Osprey – 433 Qubits

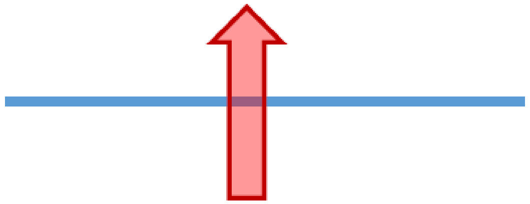




Osprey connectivity map

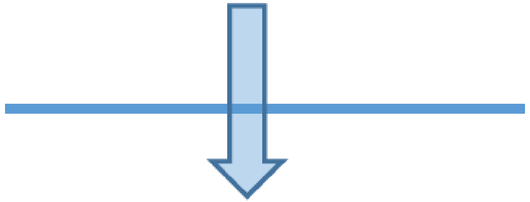
The basis of Quantum Computing

The basis of Quantum Computation is the Qubit:
a 2-level Quantum Mechanics system



$$|1\rangle = |\uparrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$|\psi\rangle = a|0\rangle + b|1\rangle$$

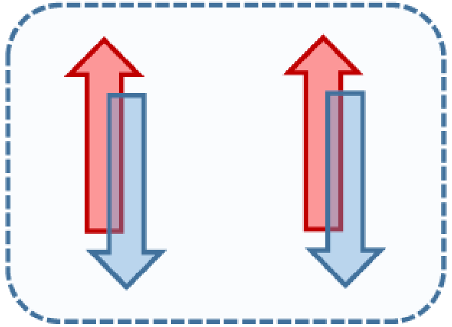


$$|0\rangle = |\downarrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

What does Quantum Computing rely on?

Qubits can be stacked together:

$$\mathcal{H}_A \otimes \mathcal{H}_B$$



$|\psi\rangle_A \otimes |\psi\rangle_B$ basis: $\begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}1\\0\end{pmatrix} = \begin{pmatrix}1\\0\\0\\0\end{pmatrix}$; $\begin{pmatrix}1\\0\end{pmatrix} \otimes \begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\1\\0\\0\end{pmatrix}$ etc.

general state: $|\psi\rangle_{AB} = a_{00}|00\rangle + a_{01}|01\rangle + a_{10}|10\rangle + a_{11}|11\rangle$

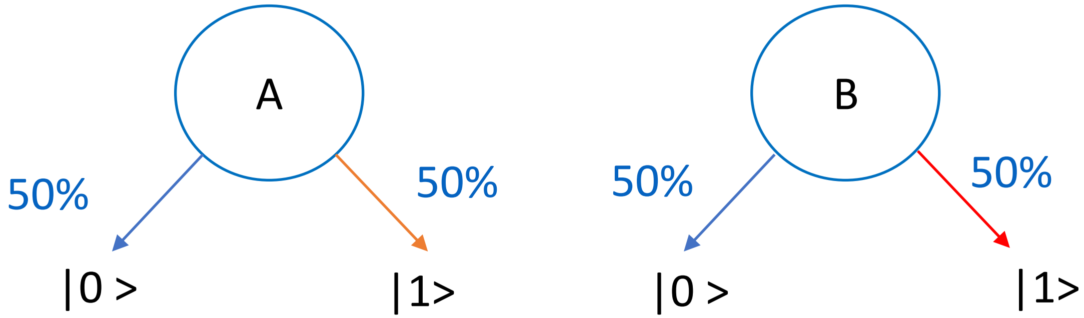
non-separable state - entangled: $|\psi\rangle_{Bell} = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \neq |\psi\rangle_A \otimes |\psi\rangle_B$

Entanglement

Entangled State

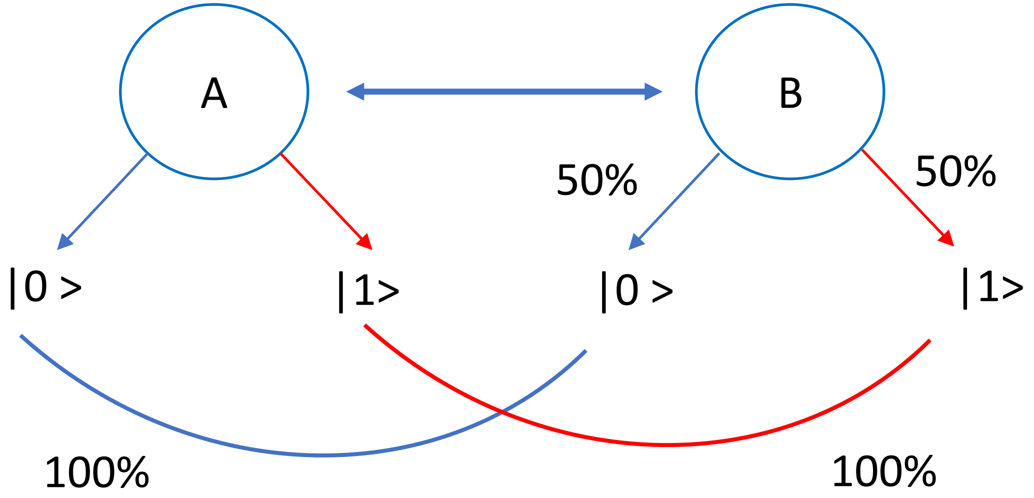
$$|\Phi_+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$$

Non-entangled states:



Two entangled states (particles, photons, macroscopic states …) show a deep correlation even if separated.

Entangled states «spooky» action at a distance:



Important Qubit Gates

$$-\boxed{X}- = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$ "bit flip"

$$-\boxed{Y}- = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$ not a straightforward meaning

$$-\boxed{Z}- = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$ "phase flip"

$$|0\rangle \; -\boxed{X}- \; |1\rangle$$
$$|1\rangle \; -\boxed{X}- \; |0\rangle$$

"bit flip" is just the classical not gate

Pauli's matrices

$$\sigma_0 = I$$
$$\sigma_1 = X$$
$$\sigma_2 = Y$$
$$\sigma_3 = Z$$

$$\sigma_j^2 = I$$

Gate composition

$$-\boxed{X}-\boxed{Y}- = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} = -\boxed{-iZ}-$$

$$-\boxed{Y}-\boxed{X}- = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} = -\boxed{iZ}-$$

Important Qubit Unitaries

Hadamard gate: $-\boxed{H}- = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$

$$H^2 = I$$

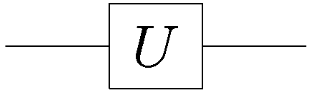
Jacques Hadamard 1865-1963

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

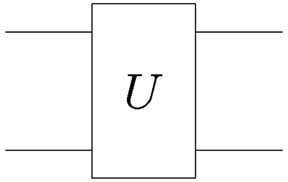
$$= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = -\boxed{X}-$$

$$-\boxed{H}-\boxed{Z}-\boxed{H}- = -\boxed{X}-$$

Quantum Circuit Elements



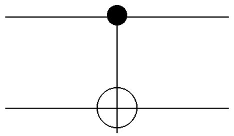
single qubit rotations



two qubit rotations

control

target



controlled-NOT

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

$|00\rangle \rightarrow |00\rangle$

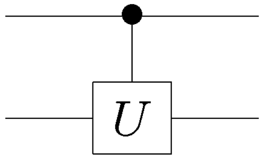
$|01\rangle \rightarrow |01\rangle$

$|10\rangle \rightarrow |11\rangle$

$|11\rangle \rightarrow |10\rangle$

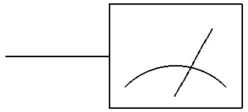
control

target



controlled-U

$$= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & U_{00} & U_{01} \\ 0 & 0 & U_{10} & U_{11} \end{bmatrix}$$



measurement in the $|0\rangle, |1\rangle$ basis

Quantum Composer

Demo Composer

Quantum GAN

Back to Quantum GAN

Appendix EQ-GAN

Helstrom Measurament

The Helstrom measurement is the measurement that has the minimum error probability when trying to distinguish between two states.

For example, let's imagine you have two pure states $|\psi\rangle$ and $|\phi\rangle$, and you wish to know which it is that you have. If $\langle\psi|\phi\rangle = 0$, then you can specify a measurement with three projectors

$$P_\psi = |\psi\rangle\langle\psi| \qquad P_\phi = |\phi\rangle\langle\phi| \qquad \bar{P} = \mathbb{I} - P_\psi - P_\phi.$$

(For a two-dimensional Hilbert space, $\bar{P} = 0$.)

The question is what measurement should you perform in the case that $\langle\psi|\phi\rangle \neq 0$? Specifically, let's assume that $\langle\psi|\phi\rangle = \cos(2\theta)$, and I'll concentrate just on projective measurements (IIRC, this is optimal). In that case, there is always a unitary U such that

$$U|\psi\rangle = \cos\theta|0\rangle + \sin\theta|1\rangle \qquad U|\phi\rangle = \cos\theta|0\rangle - \sin\theta|1\rangle.$$

Now, those states are optimally distinguished by $|+\rangle\langle+|$ and $|-\rangle\langle-|$ (you get $|+\rangle$, and you assume you had $U|\psi\rangle$). Hence, the optimal measurement is

$$P_\psi = U^\dagger|+\rangle\langle+|U \qquad P_\phi = U^\dagger|-\rangle\langle-|U \qquad \bar{P} = \mathbb{I} - P_\psi - P_\phi.$$

The success probability is

$$\left(\frac{\cos\theta + \sin\theta}{\sqrt{2}}\right)^2 = \frac{1 + \sin(2\theta)}{2}.$$

More generally, how do you distinguish between two density matrices ρ_1 and ρ_2? Start by calculating

$$\delta\rho = \rho_1 - \rho_2,$$

and finding the eigenvalues $\{\lambda_i\}$ and corresponding eigenvectors $|\lambda_i\rangle$ of $\delta\rho$. You construct 3 measurement operators

$$P_1 = \sum_{i:\lambda_i>0} |\lambda_i\rangle\langle\lambda_i| \qquad P_2 = \sum_{i:\lambda_i<0} |\lambda_i\rangle\langle\lambda_i| \qquad P_0 = \mathbb{I} - P_1 - P_2.$$

If you get answer P_1, you assume you had ρ_1. If you get P_2, you had ρ_2, while if you get P_0 you simply guess which you had. You can verify that this reproduces the pure state strategy described above. What's the success probability of this strategy?

$$\frac{1}{2}\text{Tr}((P_1 + P_0/2)\rho_1) + \frac{1}{2}\text{Tr}((P_2 + P_0/2)\rho_2)$$

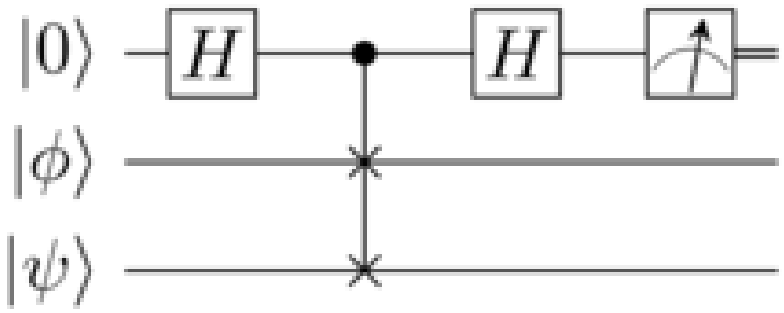
We can expand this as

$$\frac{1}{4}\text{Tr}((P_1 + P_2 + P_0)(\rho_1 + \rho_2)) + \frac{1}{4}\text{Tr}((P_1 - P_2)(\rho_1 - \rho_2))$$

Since $P_1 + P_2 + P_0 = \mathbb{I}$ and $\text{Tr}(\rho_1) = \text{Tr}(\rho_2) = 1$, this is just

$$\frac{1}{2} + \frac{1}{4}\text{Tr}((P_1 - P_2)(\rho_1 - \rho_2)) = \frac{1}{2} + \frac{1}{4}\text{Tr}|\rho_1 - \rho_2|.$$

Swap Test



- The swap test is a procedure in quantum computation that is used to check how much two quantum states differ, appearing first in the work of Barenco et al.

- It takes two input states and outputs 1 with probability $\frac{1}{2} - \frac{1}{2} |< \psi|\phi >|^2$

Swap Test

Consider two states: $|\phi\rangle$ and $|\psi\rangle$. The state of the system at the beginning of the protocol is $|0,\phi,\psi\rangle$. After the Hadamard gate, the state of the system is $\frac{1}{\sqrt{2}}(|0,\phi,\psi\rangle + |1,\phi,\psi\rangle)$. The controlled SWAP gate transforms the state into $\frac{1}{\sqrt{2}}(|0,\phi,\psi\rangle + |1,\psi,\phi\rangle)$. The second Hadamard gate results in

$$\frac{1}{2}(|0,\phi,\psi\rangle + |1,\phi,\psi\rangle + |0,\psi,\phi\rangle - |1,\psi,\phi\rangle) = \frac{1}{2}|0\rangle(|\phi,\psi\rangle + |\psi,\phi\rangle) + \frac{1}{2}|1\rangle(|\phi,\psi\rangle - |\psi,\phi\rangle)$$

The measurement gate on the first qubit ensures that it's 0 with a probability of

$$P(\text{First qubit} = 0) = \frac{1}{2}\Big(\langle\phi|\langle\psi| + \langle\psi|\langle\phi|\Big)\frac{1}{2}\Big(|\phi\rangle|\psi\rangle + |\psi\rangle|\phi\rangle\Big) = \frac{1}{2} + \frac{1}{2}|\langle\psi|\phi\rangle|^2$$

when measured. If ψ and ϕ are orthogonal $(|\langle\psi|\phi\rangle|^2 = 0)$, then the probability that 0 is measured is $\frac{1}{2}$. If the states are equal $(|\langle\psi|\phi\rangle|^2 = 1)$, then the probability that 0 is measured is 1.

In general, for P trials of the swap test using P copies of $|\phi\rangle$ and P copies of $|\psi\rangle$, the fraction of measurements that are zero is $1 - \frac{1}{P}\sum_{i=1}^{P} M_i$, so by taking $P \to \infty$, one can get arbitrary precision of this value.

Appendix Quantum Computing

Appendix Q.C.

Main Algorithms as «Lego Bricks»

• Teleportation • Superdende Coding	→	Quantum Cryptography
• Deutsch • Deutsch-Jozsa	→	«Oracle» Problems (guess the function)
• Grover	→	Search Speed-up
• Quantum Phase Estimation (QPE)	→	Linear Algebra / Shor

Quantum Teleportation

Alice wants to send her qubit to Bob

She does not know the wave function of her qubit

Alice → classical communication → Bob

- Alice and Bob share an entangled pair of qubits
- Alice makes a measure on her state and send him classical bits of information about the outcome
- Then Bob would have information about $|\psi\rangle$ as well

This would be a procedure for extracting information from without effecting the state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$

Quantum Teleportation

Alice shares with Bob an entangled state $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$

Alice does not know the wave function

ALICE

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

BOB

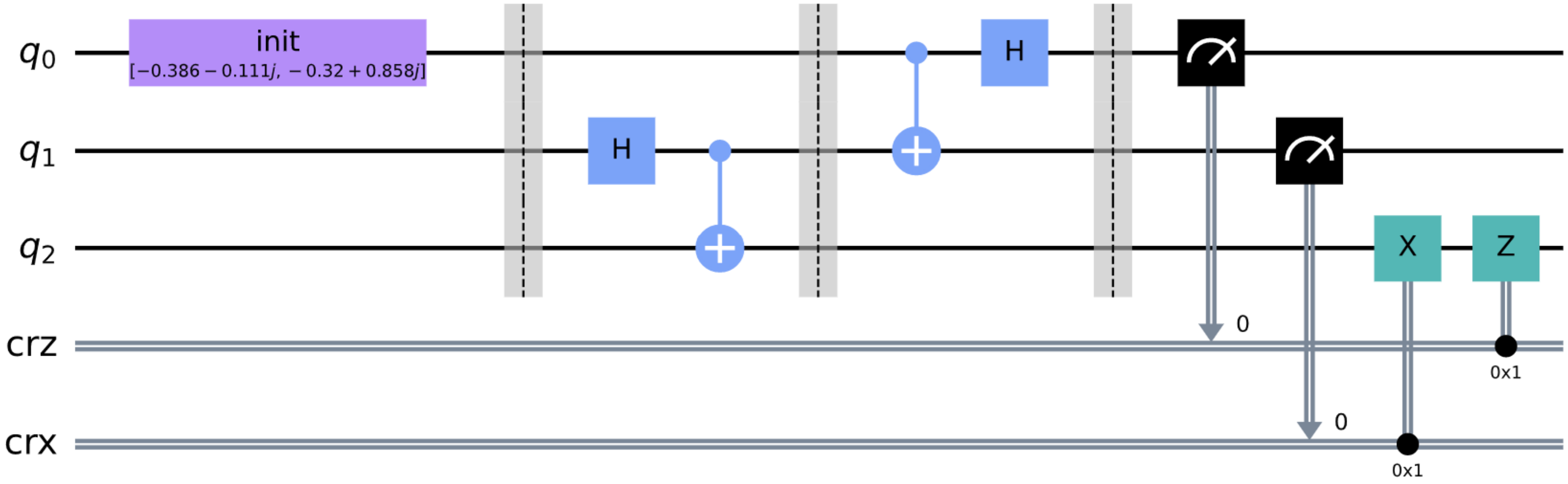
So we have three qubits whose wave function is:

$$\underbrace{|\psi\rangle}_{\text{qubit 1}} \otimes \underbrace{\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)}_{\text{qubit 2 and qubit 3}}$$

qubit 1 qubit 2 and qubit 3

Quantum Teleportation

Quantum Circuit used to simulate Teleportation:



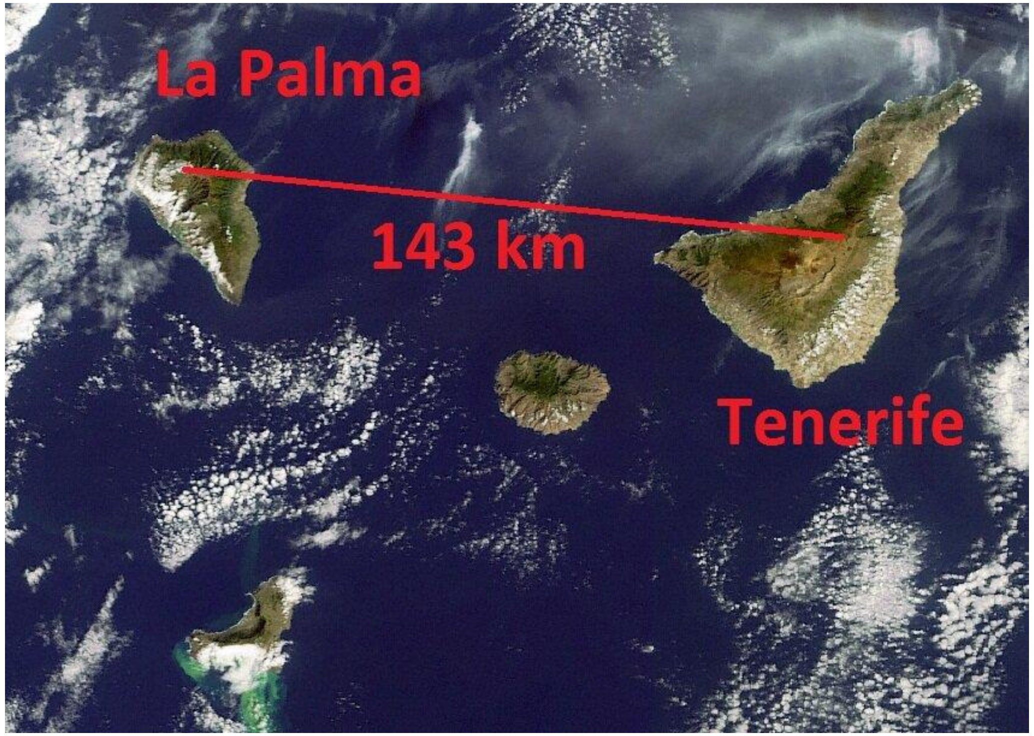
Alice makes a measure on q_0 and q_1 and send the bits to Bob

Bob reads them and if { 00 → does nothing; 01 → applies X gate; 10 → applies Z gate; 11 → applies ZX gate } then q_2 ⟺ q_0

Quantum Teleportation

ESA observatory breaks world quantum teleportation record

An international research team using ESA's Optical Ground Station in the Canary Islands has set a new distance world record in 'quantum teleportation' by reproducing the characteristics of a light particle across 143 km of open air.



Quantum Teleportation

Demo Teleportation

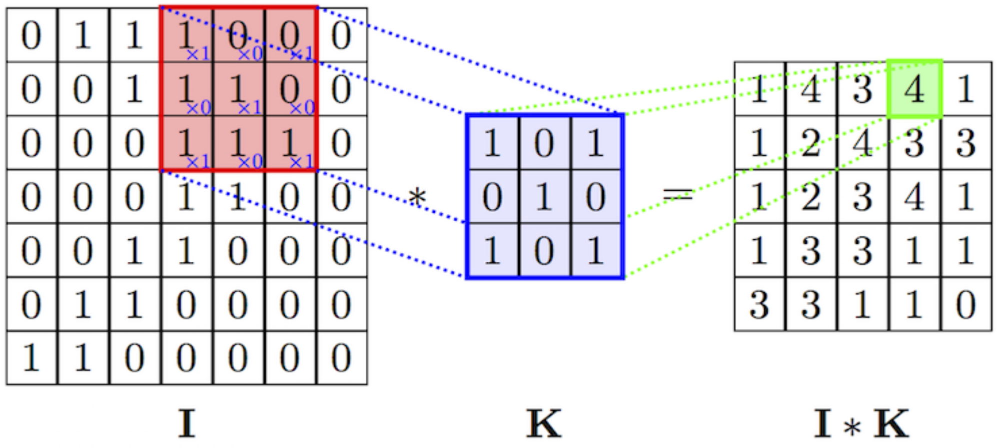
The Frontier: CNN vs Quantum Machine Learning

- Convolutional neural networks (ConvNets or CNNs) are a category of neural networks[1] that have proven to be very effective in tasks such as image recognition and classification
- Convolution works by extracting a feature map from a source image:

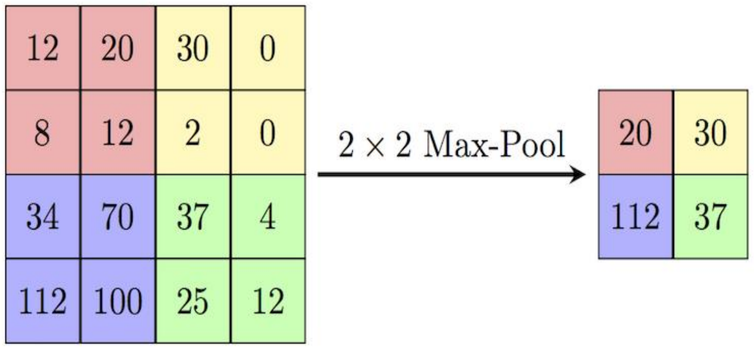
[1]https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/

CNN Schema

Filtering



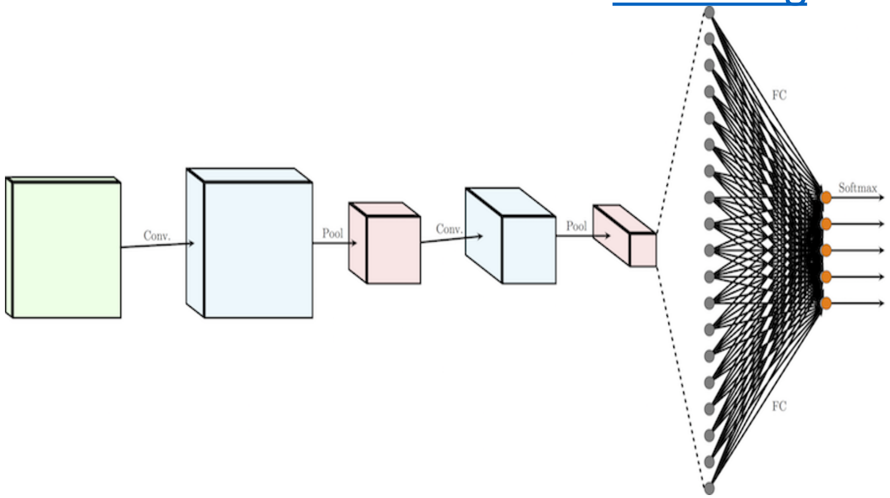
Pooling



- Convolutional Blocks:
  - Filtering
  - Pooling
- Alternating Convolutional Blocks
- Flattening
- Multi Perceptron Layers (MPLs)

Flattening

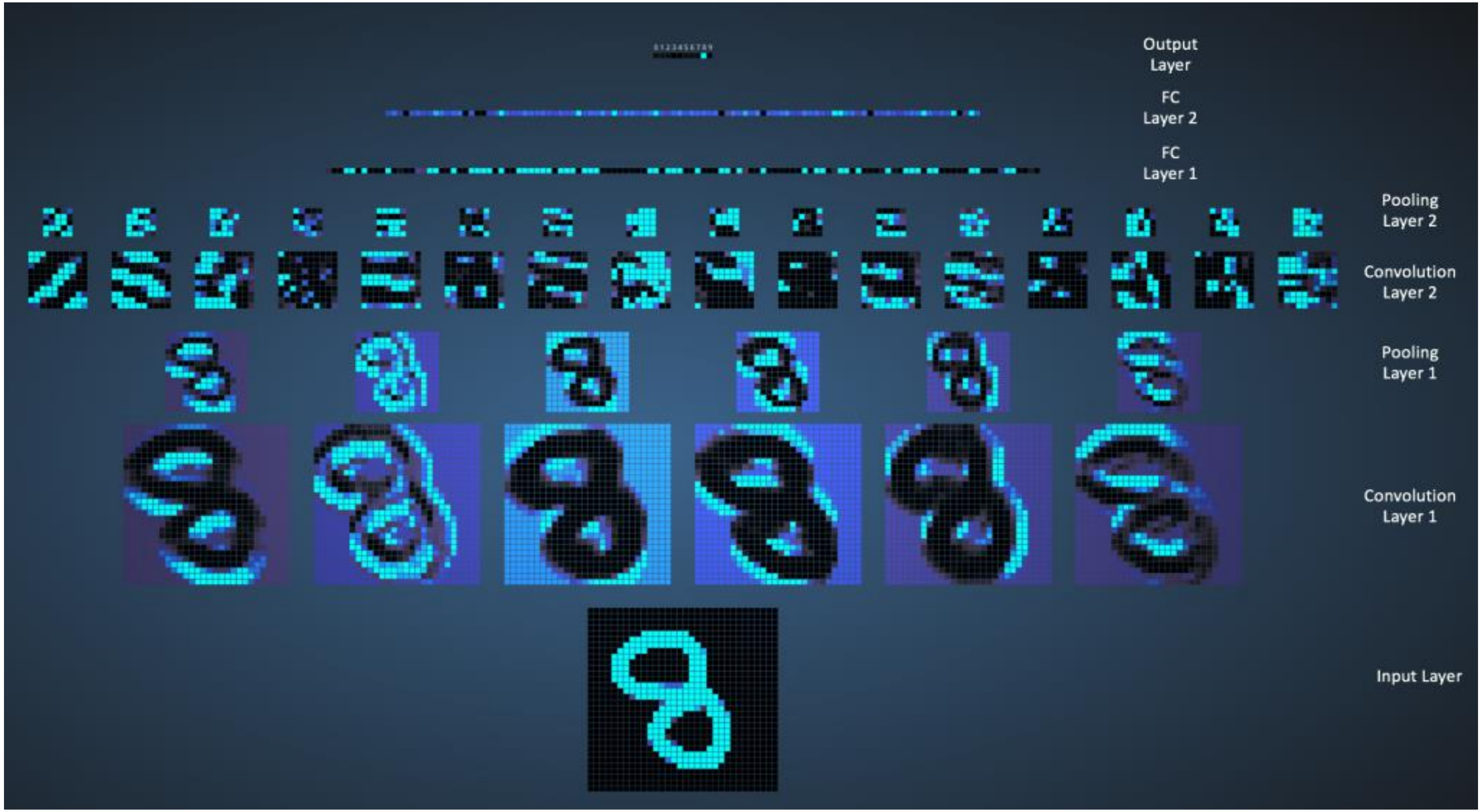
Alternating Blocks



MPLs

CNN ON MNIST

How a convolutional network classifies handwritten numbers:



CNN AT WORK

Demo CNN

The Frontier: Quantum Machine Learning

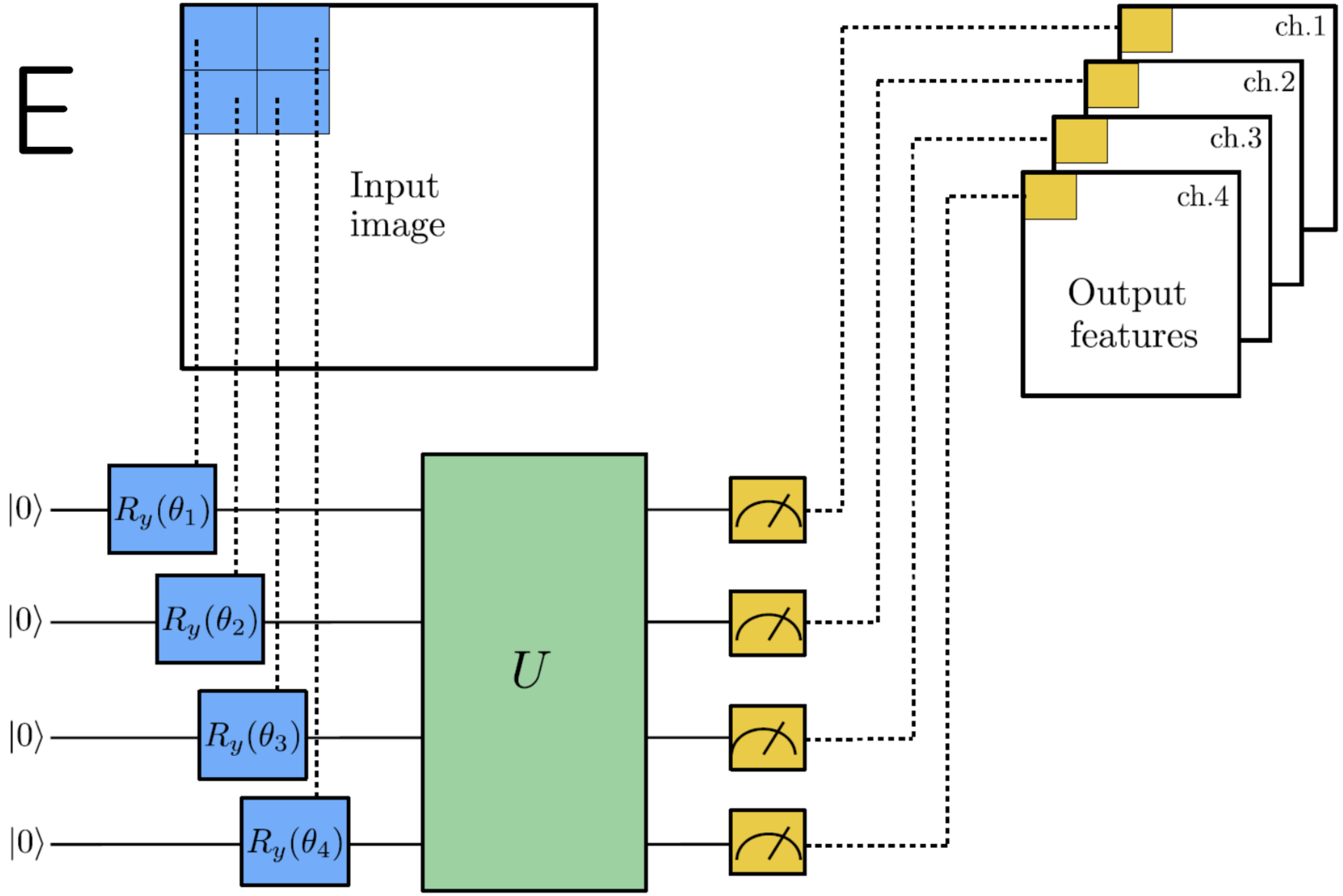
PENNYLANE

[Pennylane Quanvolution]

Quanvolutional Neural Networks

Author: Andrea Mari — Posted: 24 March 2020. Last updated: 15 January 2021

Originally introduced by Henderson et alii in 2019



Pennylane Quanvolution

Demo Quanvolution

Last but not Least…

Let’s ask ChatGPT to generate qiskit code to demonstrate teleportation protocol…

ChatGPT –> Qiskit

Demo ChatGPT

More Algorithms

«Oracle» Problems: Deutsch Algorithm

Given one bit function: $f : \{0,1\} \to \{0,1\}$

Four such function, it may be:

$$f(0) = f(1) = 0$$
$$f(0) = f(1) = 1$$

"constant"

$$f(0) = 0, f(1) = 1$$
$$f(0) = 1, f(1) = 0$$

"balanced"

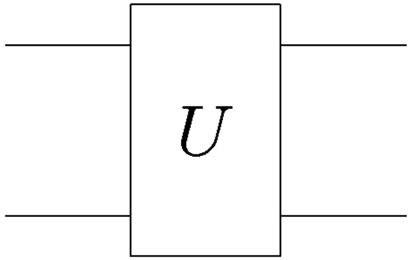
David Deutsch 1985

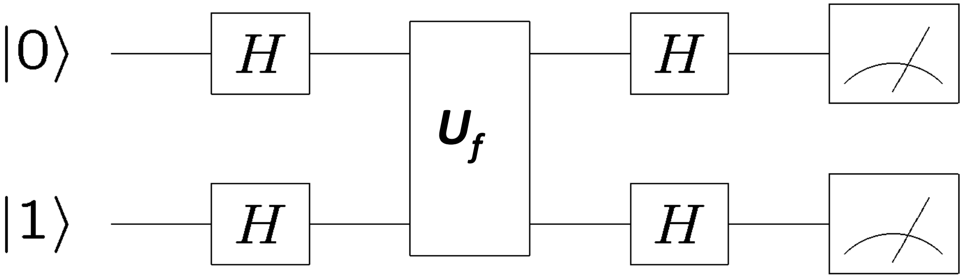
Deutsch's Problem

instance: function f is unknown

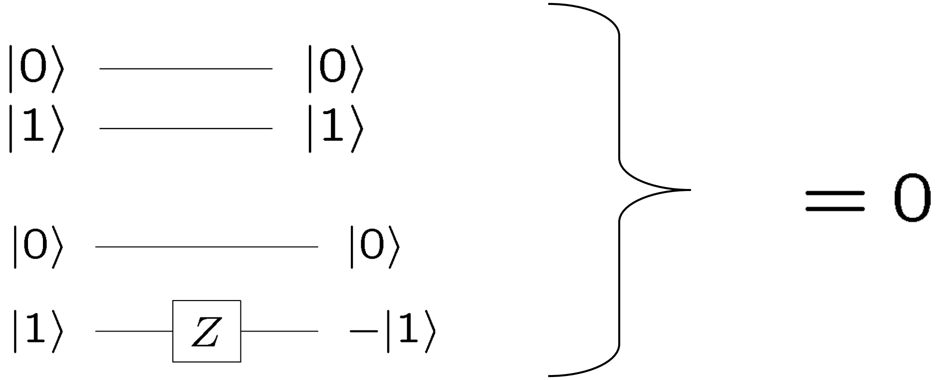
problem: determine whether function is constant or balanced

Quantum Deutsch

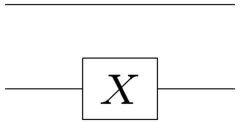




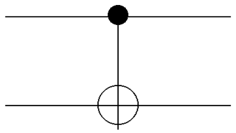
f constant f=0

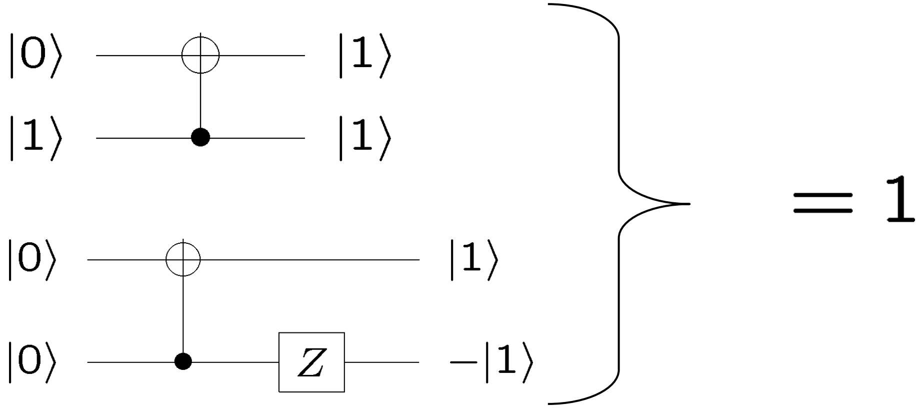


f constant f=1

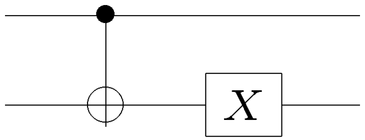


f balanced f={0,1}





f balanced f={1,0}

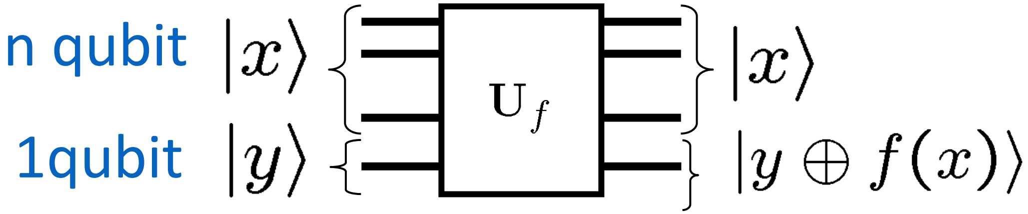


Deutsch Algorithm

Demo Deutsch

Searching speed-up: Grover Algorithm

Suppose we have a black box:



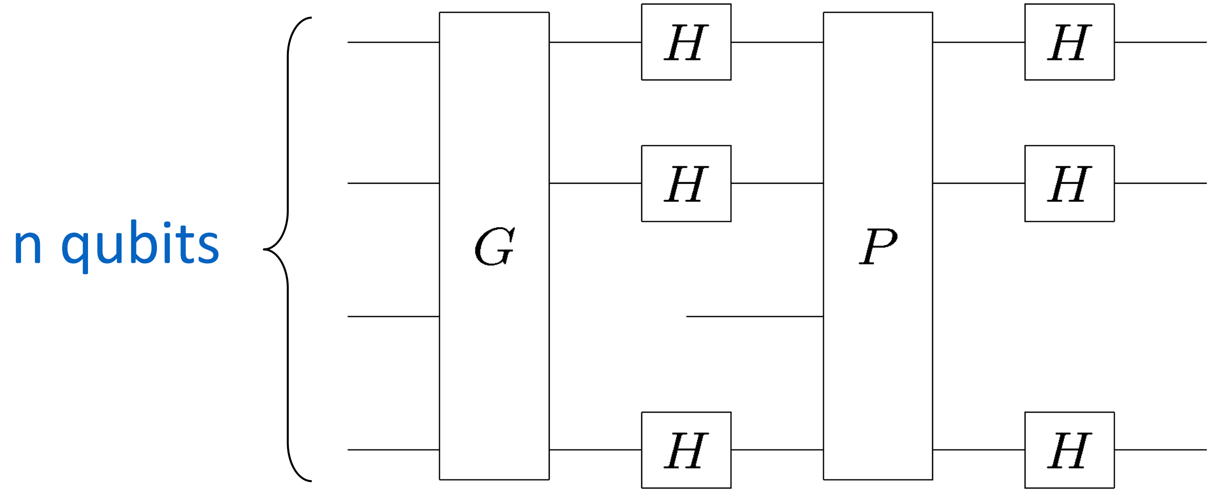
with the property:

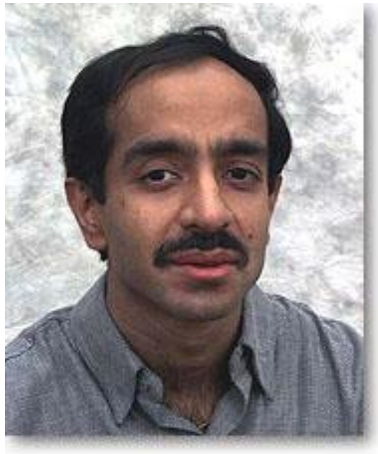
$$f(x_0) = 1$$
$$\forall x \neq x_0, f(x) = 0$$

Problem: find x_0 with as few queries as possible

Solution: identify the item in “only” $O(\sqrt{2^n})$ queries!

Grover Iterate Block





Lov Grover 1996

$$P = \sum_{x=0}^{2^n} -(-1)^{\delta_{x,0}} |x\rangle\langle x| = 2|0\rangle\langle 0| - I$$

$$H^{\otimes n} P H^{\otimes n} = H^{\otimes n}|0\rangle\langle 0|H^{\otimes n} - I = 2|\psi\rangle\langle\psi| - I$$

$$|\psi\rangle = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle$$

Grover’s iterate $R = (2|\psi\rangle\langle\psi| - I)G$

Grover Algorithm

Demo Grover

Quantum Phase Estimation Algorithm

We wanto to estimate a the phase ω given the quantum state $\sum_{y=0}^{2^n-1} e^{2\pi i \omega y}|y\rangle$

Note that in binary we can express $\omega = 0.x_1x_2x_3\ldots$ as:

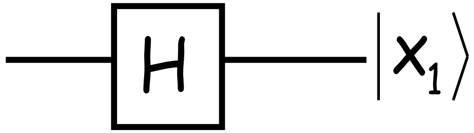
$$2\omega = x_1.x_2x_3\ldots$$

$$2^{n-1}\omega = x_1x_2x_3\ldots x_{n-1}.x_nx_{n+1}\ldots$$

If $\omega=0.x_1$ then we can do the following:

$$\frac{|0\rangle + e^{2\pi i(0.x_1)}|1\rangle}{\sqrt{2}}$$

$$= \frac{|0\rangle + (-1)^{x_1}|1\rangle}{\sqrt{2}}$$



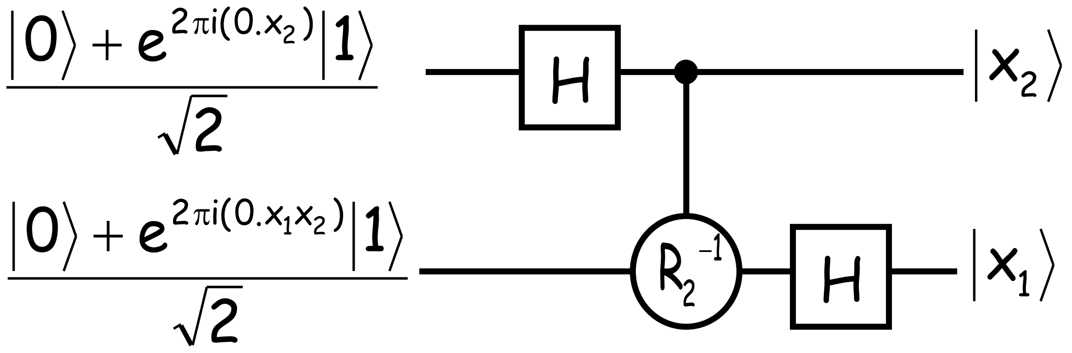
Quantum Phase Estimation Algorithm

It can be shown that:

$$\sum_{y=0}^{2^n-1} e^{2\pi\omega y}\,|y\rangle = \Big(|0\rangle + e^{2\pi i(2^{n-1}\omega)}|1\rangle\Big) \otimes \Big(|0\rangle + e^{2\pi i(2^{n-2}\omega)}|1\rangle\Big) \otimes \cdots \otimes \Big(|0\rangle + e^{2\pi i(\omega)}|1\rangle\Big)$$

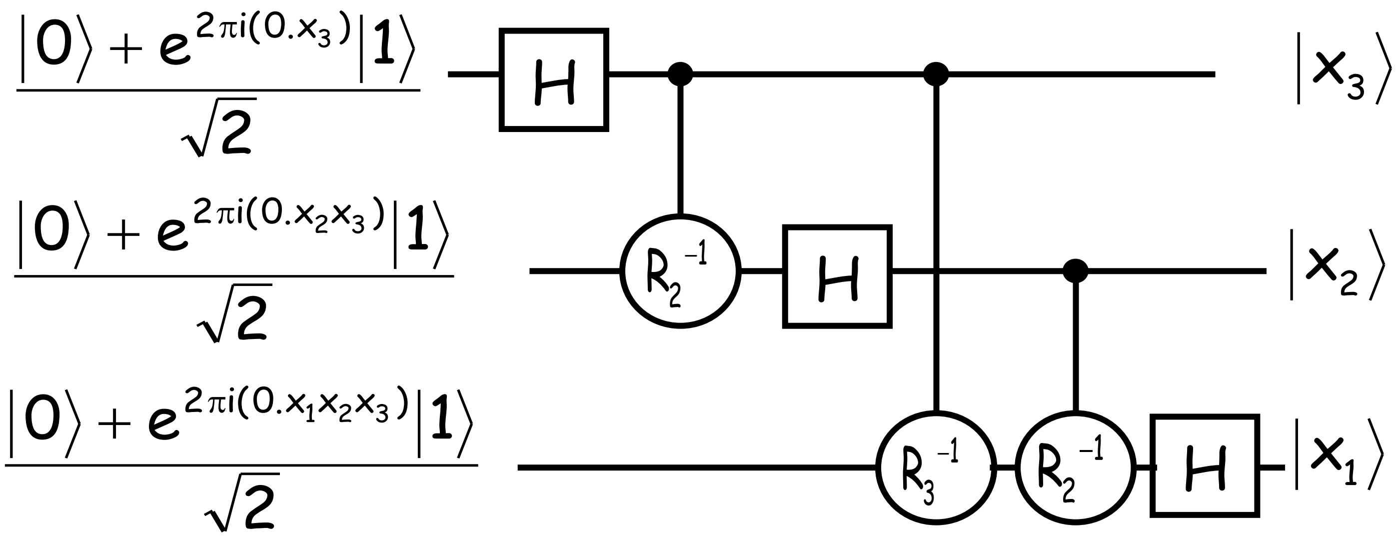
$$= \Big(|0\rangle + e^{2pii(0.x_n x_{n+1}\ldots)}|1\rangle\Big) \otimes \Big(|0\rangle + e^{2\pi i(0.x_{n-1}x_n x_{n+1}\ldots)}|1\rangle\Big) \otimes \cdots \otimes \Big(|0\rangle + e^{2\pi i(0.x_1 x_2\ldots)}|1\rangle\Big)$$

So if $\omega = 0.x_1x_2$ then we can do the following:



Quantum Phase Estimation Algorithm

If $\omega = 0.x_1 x_2 x_3$ then we can do the following:



Quantum Phase Estimation Algorithm

Demo QPE

Shor's Algorithm

A lecture on Shor's quantum factoring algorithm

Samuel J. Lomonaco Jr.

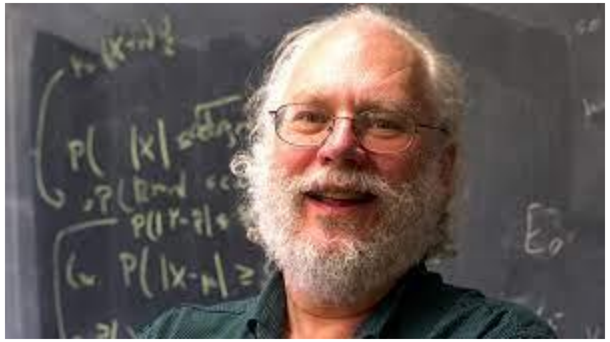
Step 1. Choose a random positive integer m. Use the polynomial time Euclidean algorithm to compute the greatest common divisor gcd (m, N) of m and N. If the greatest common divisor gcd (m, N) = 1, then we have found a non-trivial factor of N , and we are done. If, on the other hand, gcd (m, N) ≠ 1, then proceed to STEP 2.

Step 2. Use QUANTUM PHASE ESTIMATION to determine the unknown period P of the function

$$\begin{array}{ccc} \mathbb{N} & \xrightarrow{f_N} & \mathbb{N} \\ a & \longmapsto & m^a \bmod N \end{array}$$

Step 3. If P is an odd integer then goto Step 1. If P is even then goto Step 4.

Step 4. Since P is even, we have: $(m^{P/2} + 1)(m^{P/2} - 1) = m^P - 1 = 0 \bmod N$
If $m^{P/2} + 1 = 0 \bmod N$ then goto Step 1. else goto Step 5.



Peter Shor 1994

Step 5. Use the Euclidean algorithm to compute $d = \gcd(m^{P/2} - 1, N)$. Since $m^{P/2} + 1 \neq 0 \bmod N$ it can easily be shown that d is a non-trivial factor of N. Exit with the answer d.

Shor's Algorithm

Demo Shor

The Very End!