Day 3

- Comment on use of quantum algorithm for DM detection
- Grover's algorithm
- Distance based classifier
- Data reuploading

- PennyLane example for combinatorial problem
  - QAOA
  - FALQON
  - https://drive.google.com/file/d/1SopdB1k7GUQydhrMZPBFfkaPJsT9XcRT/view?usp=sharing

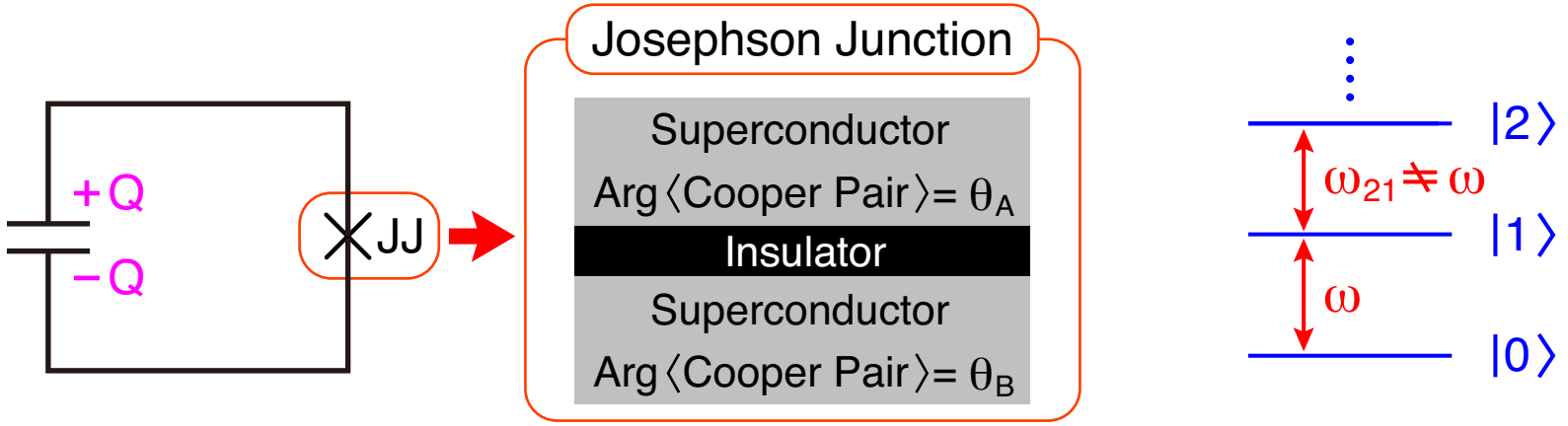
Application in Dark Matter Physics

- Detection of hidden photon dark matter using the direct excitation of transmon qubits
  - https://arxiv.org/pdf/2212.03884
- Quantum Enhancement in Dark Matter Detection with Quantum Computation
  - https://arxiv.org/pdf/2311.10413
- Search for QCD axion dark matter with transmon qubits and quantum circuit
  - https://arxiv.org/pdf/2407.19755
- Quantum entanglement of ions for light dark matter detection
  - https://arxiv.org/pdf/2311.11632

- Detection of hidden photon dark matter using the direct excitation of transmon qubits
  - https://arxiv.org/pdf/2212.03884

Slides taken from T. Moroi's talk (March 2024)

Transmon qubit: Capacitor + Josephson junction (JJ)



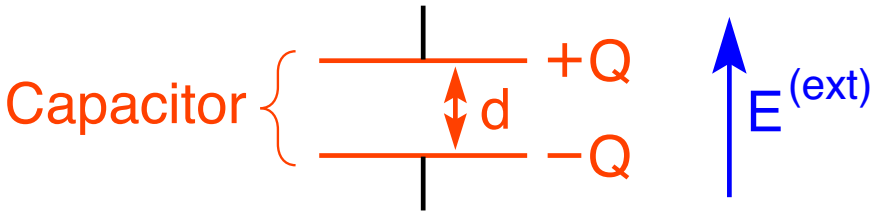
$$H_0 = \frac{1}{2C}Q^2 - J\cos\theta \simeq \frac{1}{2}\frac{C}{(2e)^2}\dot{\theta}^2 - J\cos\theta \quad \text{with } \theta = \theta_B - \theta_A$$

Transmon qubit has discrete energy levels

$\Rightarrow$ $|0\rangle$ and $|1\rangle$ can be used as $|g\rangle$ and $|e\rangle$, respectively

$\Rightarrow$ Transmon qubits are used in today's quantum computers

Transmon qubit couples to external electric field



$$\Leftrightarrow \; H_{\rm int} = QdE^{(\rm ext)}$$

Charge operator in the transmon limit: $CJ \gg (2e)^2$

$$Q \simeq \frac{C}{2e}\dot{\theta} \simeq \sqrt{\frac{C\omega}{2}}\left(|g\rangle\langle e| + |e\rangle\langle g| \right)$$

$|g\rangle \leftrightarrow |e\rangle$ transition occurs if DM field generates electric field

- Hidden photon
- Axion (if external magnetic field exists)
- ...

Case of hidden photon X_μ (in mass-eigenstate basis)

$$\mathcal{L} \ni -\frac{1}{4} X_{\mu\nu} X^{\mu\nu} + \frac{1}{2} m_X^2 X_\mu X^\mu + e\, \bar{\psi}_{\rm e} \gamma^\mu \psi_{\rm e} \left(A_\mu^{(\rm EM)} + \epsilon X_\mu \right)$$

ϵ: kinetic-mixing parameter

Oscillating hidden photon can play the role of DM

$$\vec{X} \simeq \bar{X} \vec{n}_X \sin(m_X t + \alpha) \quad \text{with } \rho_{\rm DM} = \frac{1}{2} m_X^2 \bar{X}^2$$

$\vec{X}$ = hidden photon field around the Earth

Hidden photon DM induces effective electric field

$$\vec{E}^{(X)} = -\epsilon \dot{\vec{X}} = -\bar{E}^{(X)} \vec{n}_X \cos(m_X t + \alpha)$$

$$\bar{E}^{(X)} = \epsilon \sqrt{2 \rho_{\rm DM}}$$

Hamiltonian for transmon qubit + hidden photon system

$$H = \omega|e\rangle\langle e| - 2\eta\cos(m_X t + \alpha)\big(\,|e\rangle\langle g| + |g\rangle\langle e|\,\big)$$

$$\eta \simeq \frac{1}{2\sqrt{2}} d\bar{E}^{(X)}\sqrt{C\omega} = \frac{1}{2}\epsilon d\sqrt{C\omega\rho_{\rm DM}}$$

$$\mathcal{H} = \mathcal{H}_0 + \Delta\mathcal{H}$$

$$\langle e|\Delta\mathcal{H}|g\rangle$$

Schrödinger equation

$$i\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle \quad \Rightarrow \quad |\psi(t)\rangle = U_{\rm DM}(t)\,|\psi(0)\rangle$$

Resonance limit $\omega = m_X$ (for $\eta t \ll 1$)

$$\begin{pmatrix}\psi_g(t)\\ \psi_e(t)\end{pmatrix} = U_{\rm DM}(t)\begin{pmatrix}\psi_g(0)\\ \psi_e(0)\end{pmatrix} \simeq \begin{pmatrix}1 & ie^{-i\alpha}\eta t\\ ie^{i\alpha}\eta t & 1\end{pmatrix}\begin{pmatrix}\psi_g(0)\\ \psi_e(0)\end{pmatrix}$$

$$|\psi(t)\rangle \equiv \psi_g(t)|g\rangle + e^{-i\omega t}\,\psi_e(t)|e\rangle$$

$|g\rangle \to |e\rangle$ transition probability (assuming $|\psi(0)\rangle = |g\rangle$)

$$|\psi_e(t)|^2 \simeq \begin{cases} \eta^2 t^2 & : \omega = m_X \text{ (on-resonance)} \\ \sim \eta^2(\omega - m_X)^{-2} & : \omega \neq m_X \text{ (off-resonance)} \end{cases}$$

Excitation probability for $\omega = m_X$:

$$P_{ge} \simeq 0.3 \times \left(\frac{\epsilon}{10^{-11}}\right)^2 \left(\frac{m_X}{10\,\mu\text{eV}}\right) \left(\frac{C}{0.1\,\text{pF}}\right) \left(\frac{d}{100\,\mu\text{m}}\right)^2 \left(\frac{\tau}{100\,\mu\text{s}}\right)^2$$

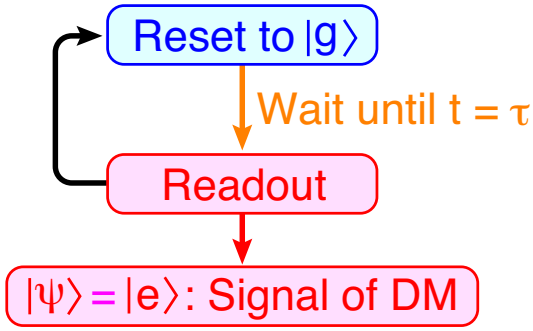
$\tau =$ coherence time

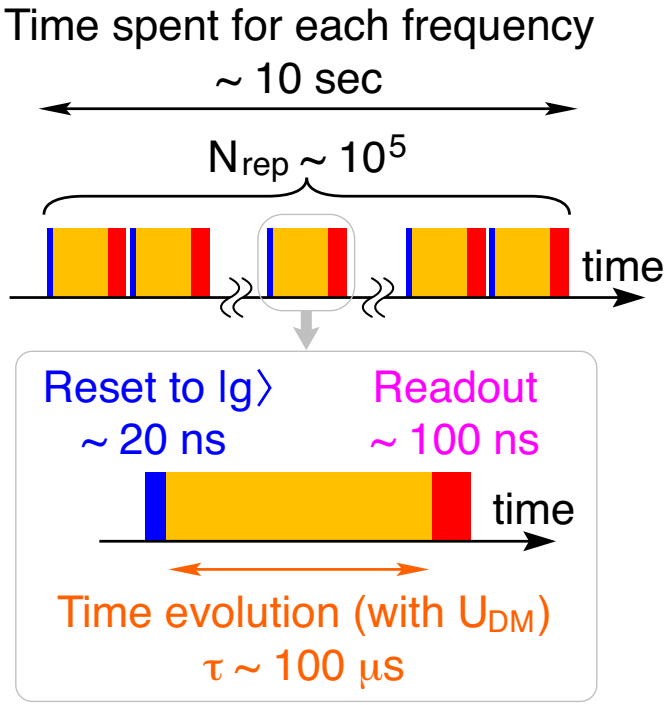
Excitation probability can be sizable

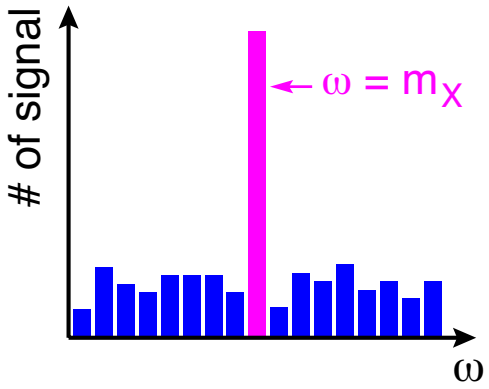
$\Rightarrow$ Transmon qubit as a DM detector

Search strategy

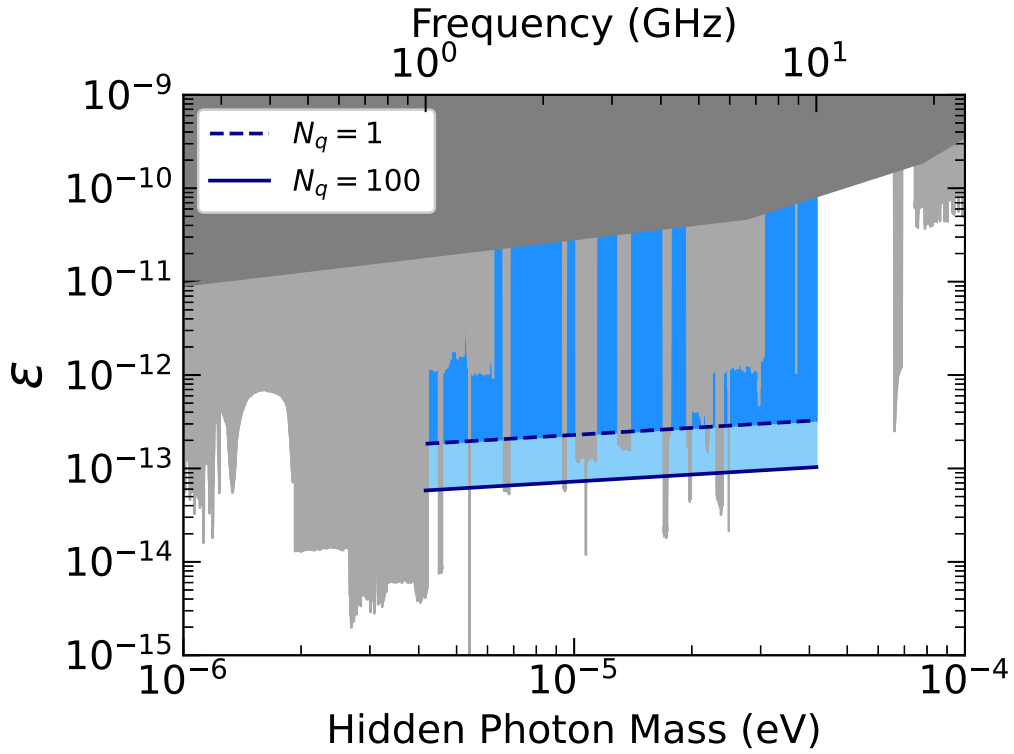
- For fixed ω, repeat the measurement cycle (reset, wait, and readout) as many time as possible
- Scan the qubit frequency ω







Discovery reach with 1 year frequency scan ($1 \leq f \leq 10$ GHz)



- $d = 100\ \mu$m
- $C = 0.1$ pF
- $\mathcal{Q} = 10^6$
- Error rate / qubit $= 0.1\ \%$

$\Rightarrow$ Using qubit, we may probe parameter region unexplored

$\Rightarrow$ We hope to use qubit for the detection of other DMs

- Quantum Enhancement in Dark Matter Detection with Quantum Computation
  - https://arxiv.org/pdf/2311.10413

$U_{\rm DM}$ induces pure phase rotation of its eigenstates

E.g. for $\alpha = 0$: $U_{\rm DM} \simeq \begin{pmatrix} 1 & i\delta \\ i\delta & 1 \end{pmatrix}$ with $\delta \equiv \eta\tau$

$\Rightarrow\ U_{\rm DM}|\pm\rangle = e^{\pm i\delta}|\pm\rangle$ with $|\pm\rangle \equiv \frac{1}{\sqrt{2}}\big(|g\rangle \pm |e\rangle\big)$

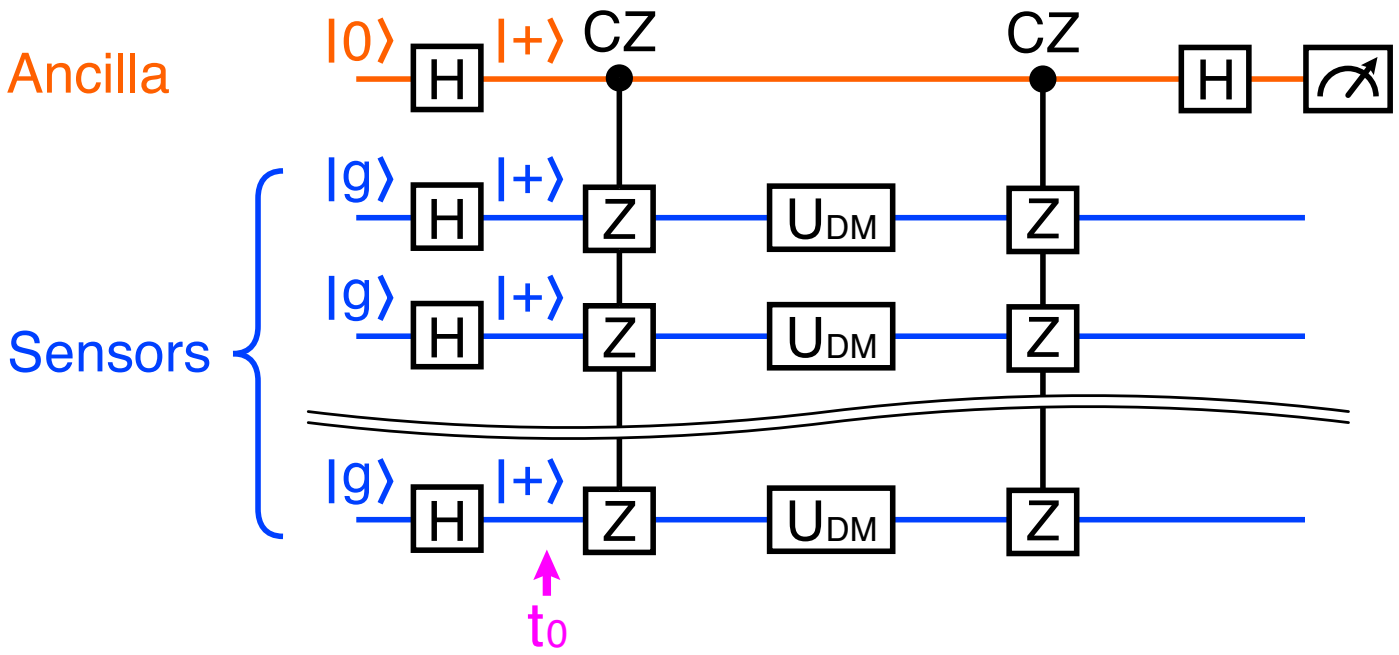
$\Rightarrow\ U_{\rm DM}^{\otimes N_{\rm q}}|+\rangle^{\otimes N_{\rm q}} = e^{iN_{\rm q}\delta}|+\rangle^{\otimes N_{\rm q}}$

We can design quantum operations to enhance the signal

$\Rightarrow$ Quantum enhanced parameter estimation

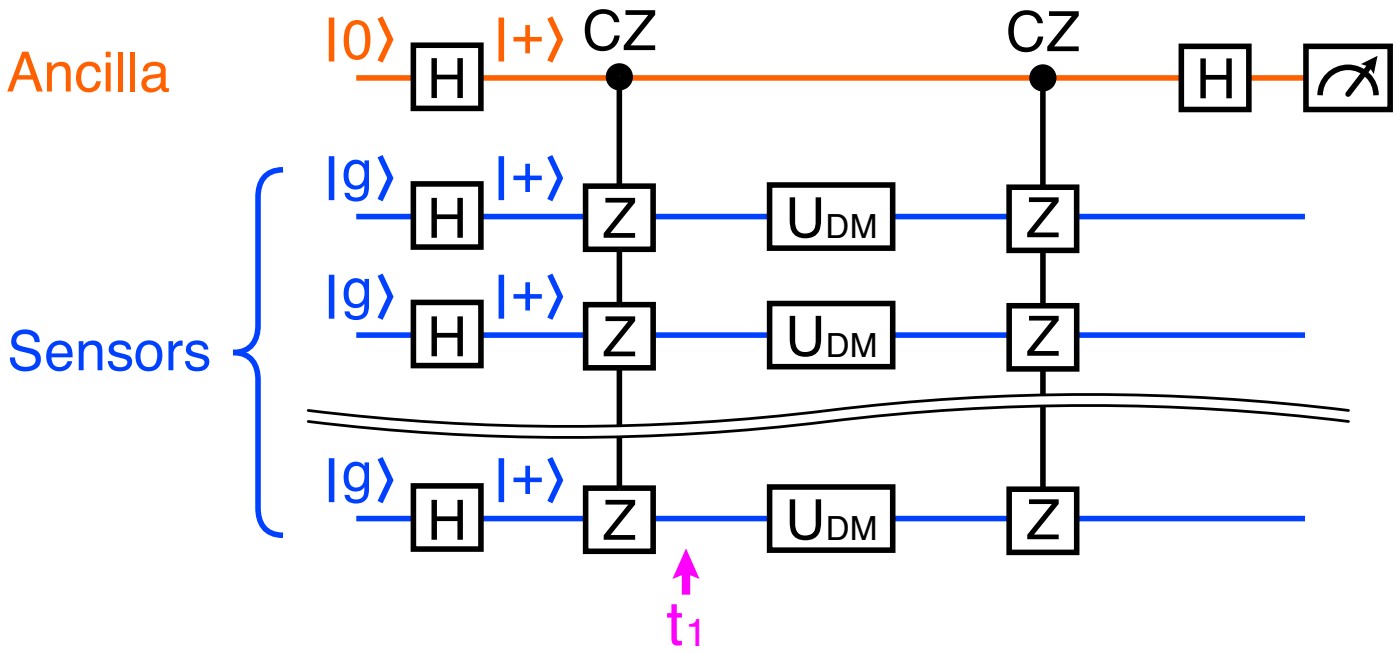
[Giovannetti, Lloyd, Maccone ('04)]

One measurement cycle for the signal enhancement



$$|\Psi(t_0)\rangle = |+\rangle \otimes |+\rangle^{\otimes N_{\rm q}} = \frac{1}{\sqrt{2}}|0\rangle \otimes |+\rangle^{\otimes N_{\rm q}} + \frac{1}{\sqrt{2}}|1\rangle \otimes |+\rangle^{\otimes N_{\rm q}}$$

One measurement cycle for the signal enhancement



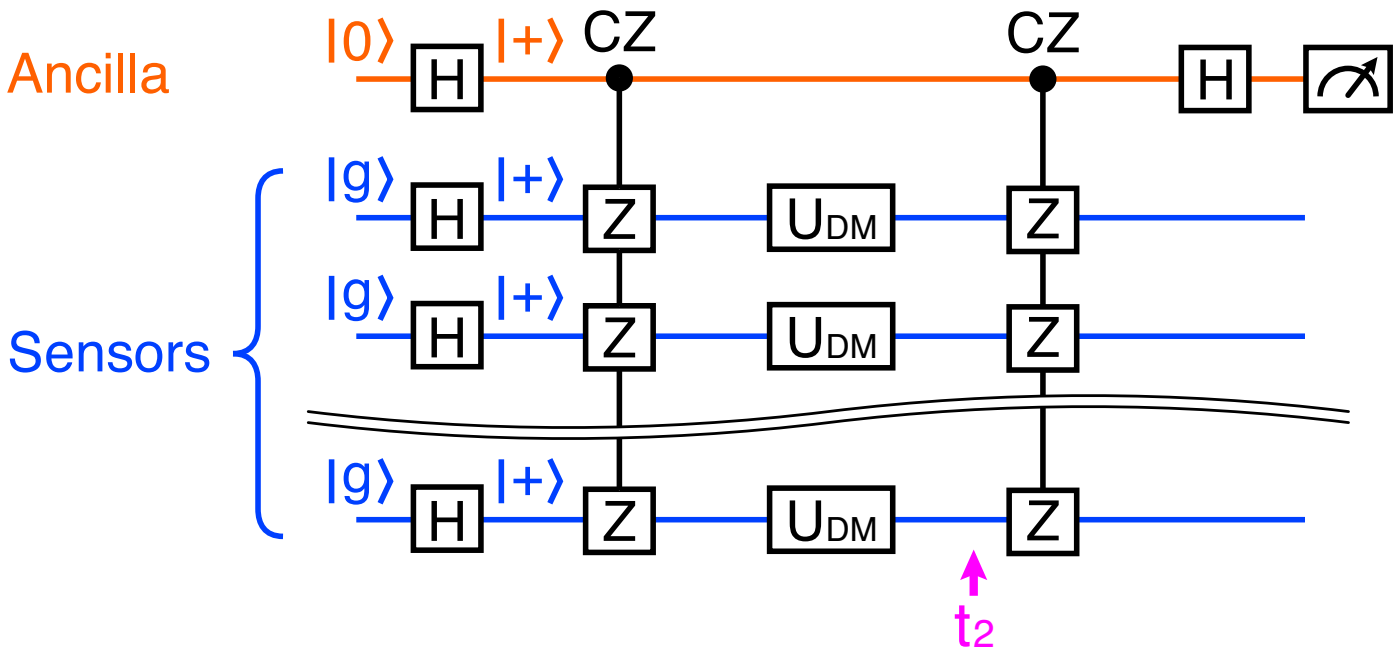
$$|\Psi(t_1)\rangle = \frac{1}{\sqrt{2}}|0\rangle \otimes |+\rangle^{\otimes N_{\rm q}} + \frac{1}{\sqrt{2}}|1\rangle \otimes |-\rangle^{\otimes N_{\rm q}}$$

- Controlled Z gate

$$CZ = |0\rangle\langle 0| \otimes \mathbf{1} + |1\rangle\langle 1| \otimes Z$$

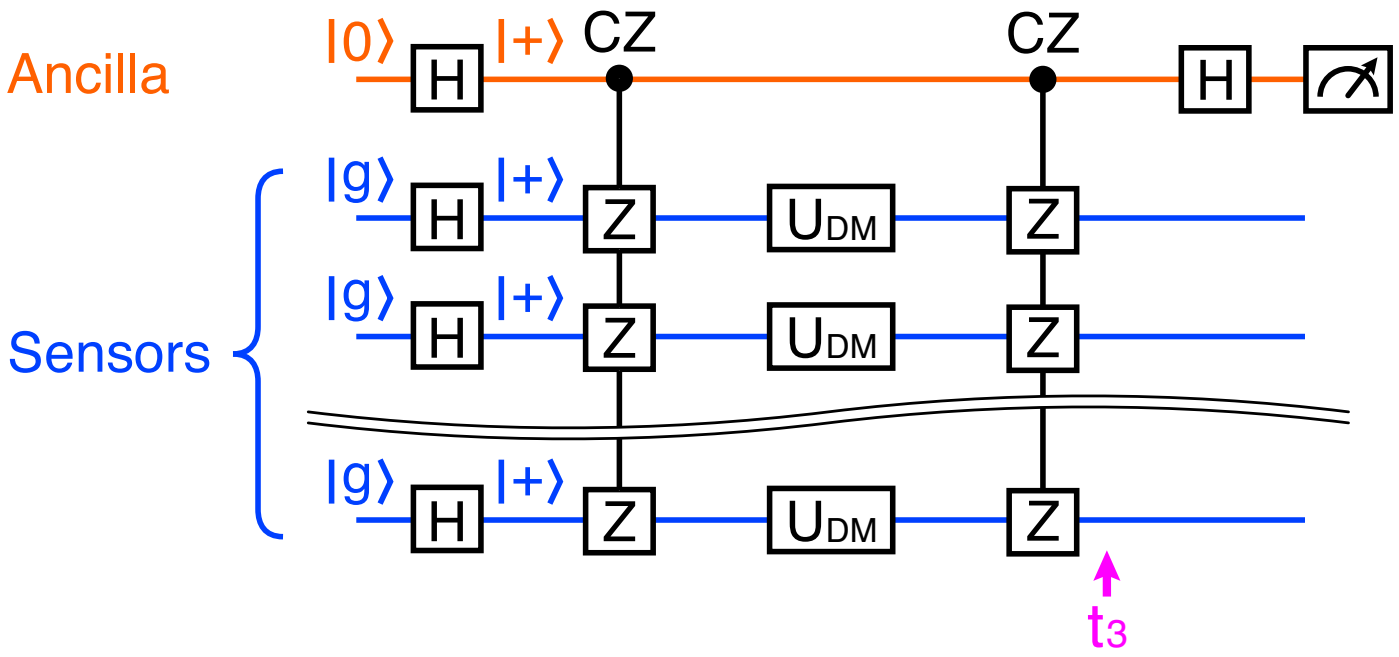
$$\Rightarrow \quad \frac{1}{\sqrt{2}}\big(|0\rangle + |1\rangle\big) \otimes |+\rangle \xrightarrow{CZ} \frac{1}{\sqrt{2}}|0\rangle \otimes |+\rangle + \frac{1}{\sqrt{2}}|1\rangle \otimes |-\rangle$$

One measurement cycle for the signal enhancement



$$|\Psi(t_2)\rangle = \frac{1}{\sqrt{2}} e^{iN_q\delta}|0\rangle \otimes |+\rangle^{\otimes N_q} + \frac{1}{\sqrt{2}} e^{-iN_q\delta}|1\rangle \otimes |-\rangle^{\otimes N_q}$$

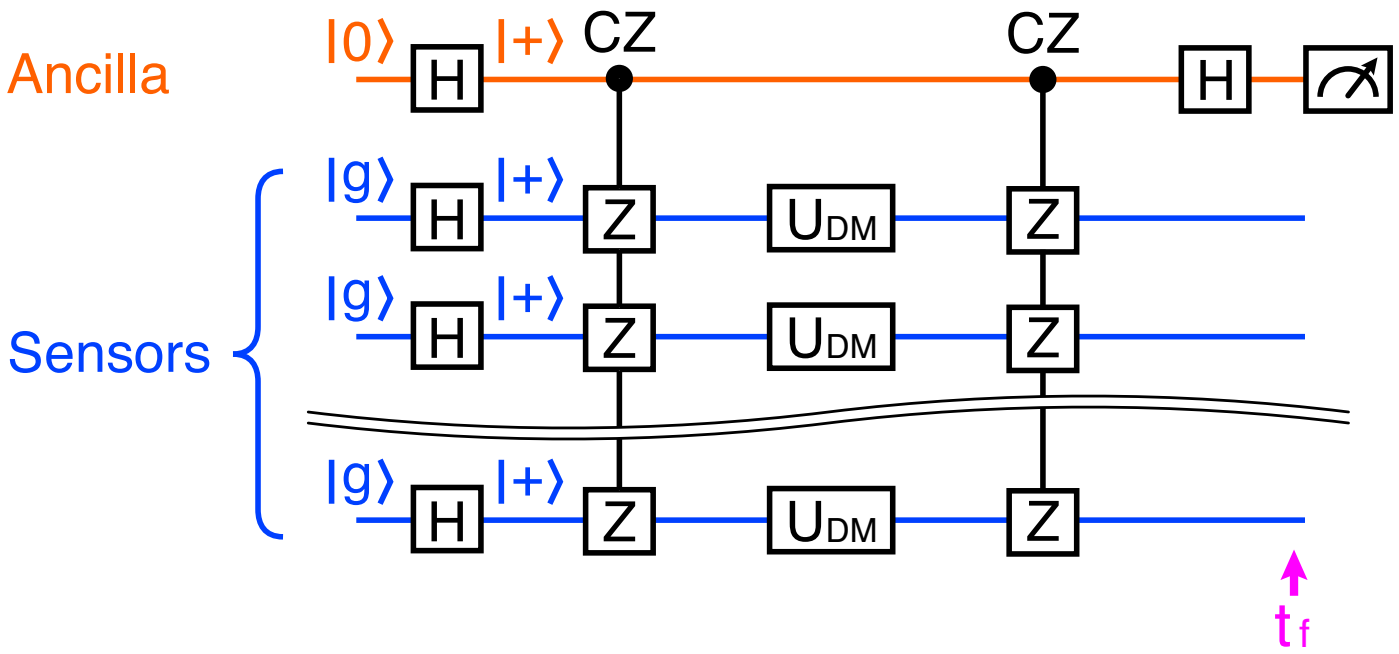
One measurement cycle for the signal enhancement



$$|\Psi(t_3)\rangle = \frac{1}{\sqrt{2}} e^{iN_{\rm q}\delta} |0\rangle \otimes |+\rangle^{\otimes N_{\rm q}} + \frac{1}{\sqrt{2}} e^{-iN_{\rm q}\delta} |1\rangle \otimes |+\rangle^{\otimes N_{\rm q}}$$

$$= \left(\cos\ N_{\rm q}\delta\, |+\rangle + i \sin\ N_{\rm q}\delta\, |-\rangle \right) \otimes |+\rangle^{\otimes N_{\rm q}}$$

One measurement cycle for the signal enhancement

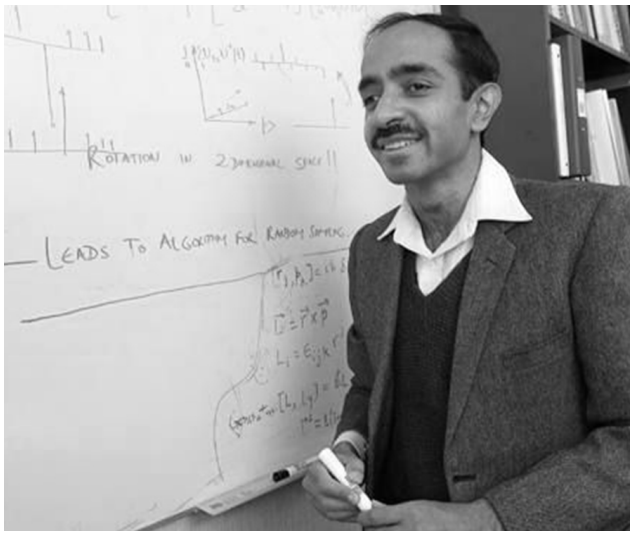


$$|\Psi(t_{\rm f})\rangle = \left(\cos\ N_{\rm q}\delta\,|0\rangle + i\sin\ N_{\rm q}\delta\,|1\rangle\right)\otimes|+\rangle^{\otimes N_{\rm q}}$$

$\Rightarrow$ Ancilla qubit can be excited: $P_{0\to1} \simeq \sin^2 N_{\rm q}\delta \simeq N_{\rm q}^2\delta^2$

Grover's search algorithm (Amplitude amplification)

Grover's search algorithm



- Grover's algorithm involves "amplitude amplification"
  - G. Brassard, P. Hoyer 1997, Lov Grover 1998
  - QFT is used for Shor's and Simon's algorithms
- Example: Find a name in a phone directory (ordered list)
  - Go to the midpoint of the list, see which half contains the name. Repeat the same → bisection method takes $\log_2 N$ operations until one of left.
- If we are given an unordered list, we will have to check all entries one a time. On average, this would take $N/2$ operations
  - For $N = 10^6,\ \log_2 N \approx 20$ and $N/2 \approx 5 \times 10^5$.
- Grover's algorithm (unstructured search): determines the special value with $p \approx 1$ (close to 1) by calling subroutine only $\frac{\pi}{4}\sqrt{N}$ times. → quadratic speed up compared with a classical computer.
  - (cf) exponential speed up is expected in Shor's algorithm.

Grover's algorithm: Black Box (Oracle)

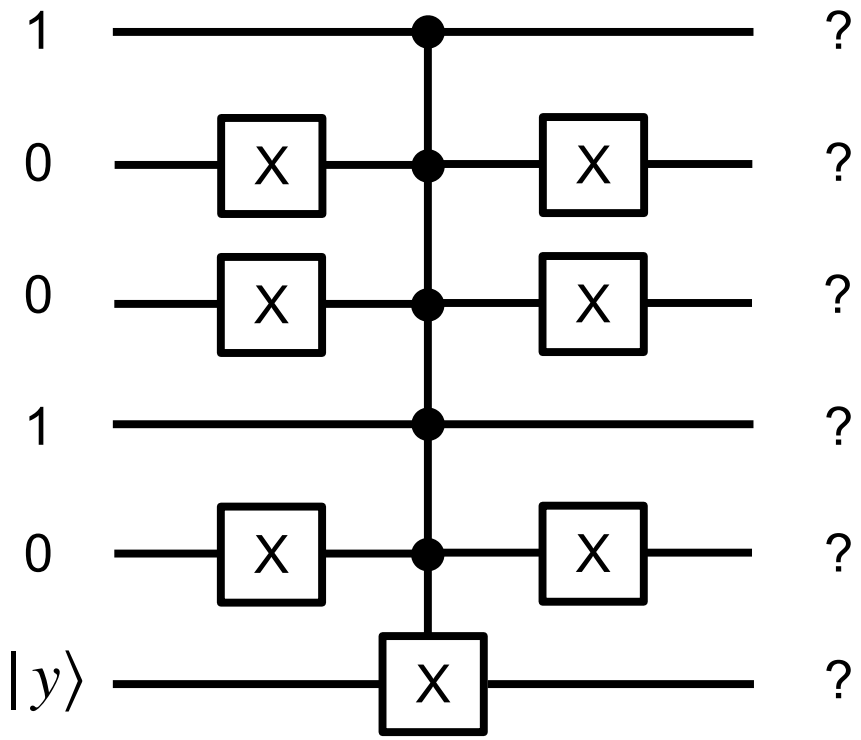
- Consider n-bit integers.
- "a" is a special number, and the goal is to find "a".
- Define a subroutine which output 1 if input value x is equal to a, and output 0 otherwise.

$$f(a) = 1, \quad f(x) = 0 \quad \text{for } x \neq a$$

$$U\,|x\rangle_n \otimes |y\rangle_1 = |x\rangle_n \otimes |y \oplus f(x)\rangle_1$$

n qubits — one qubit

Example: a=01001



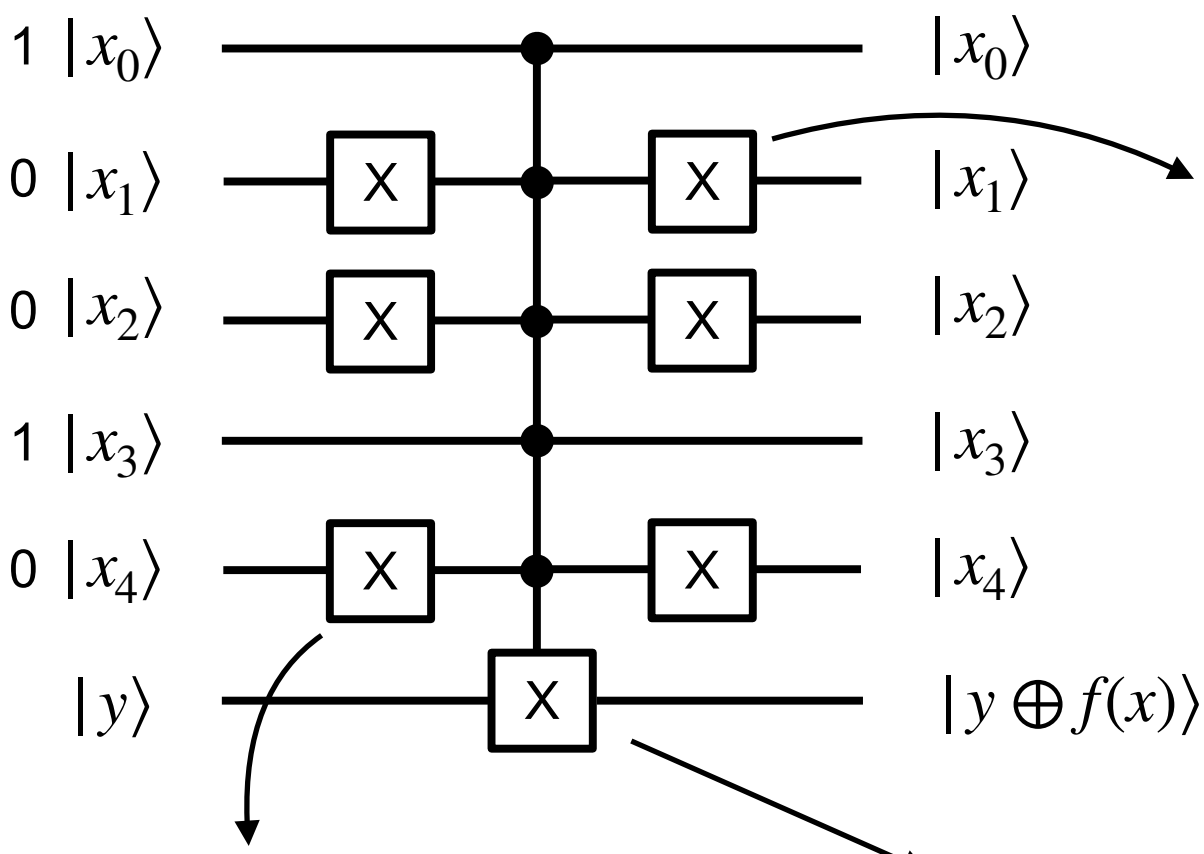
$$a = x_4 x_3 x_2 x_1 x_0 = 01001$$

Grover's algorithm: Black Box (Oracle)

$a = x_4 x_3 x_2 x_1 x_0 = 01001$

$f(a) = 1$

$f(x) = 0, \text{ if } x \neq a$



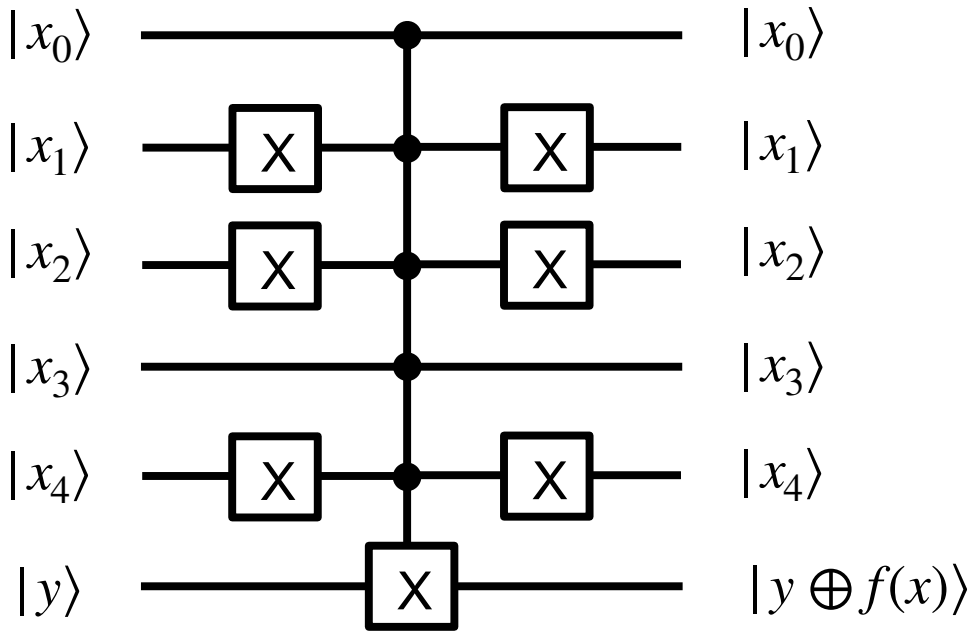
(3) X-gates on the right flip back to the original input.

(1) X-gates on the left flip qubits x_1, x_2 and $x_4 \rightarrow$ target qubit is flipped only if $x_4 x_3 x_2 x_1 x_0 = 01001$

(2) Five-fold-controlled NOT acts to flip the target qubit y, only if all control bits are 1.

Grover's algorithm: Black Box (Oracle)

$a = x_4 x_3 x_2 x_1 x_0 = 01001$



$$f(a) = 1$$

$$f(x) = 0, \text{ if } x \neq a$$

- Useful to initialize $|y\rangle = |1\rangle$ and apply H before U.
- The output qubit is

$$H|1\rangle = \frac{1}{\sqrt{2}}\Big(|0\rangle - |1\rangle\Big)$$

if $f(x) = 0$, $\quad |0 \oplus f(x)\rangle - |1 \oplus f(x)\rangle = |0\rangle - |1\rangle$

$f(x) = 1$, $\quad |0 \oplus f(x)\rangle - |1 \oplus f(x)\rangle = |1\rangle - |0\rangle = -\big(|0\rangle - |1\rangle\big)$

Target qubit changes the sign, depending on the function value.

$$U\Big(|x\rangle \otimes H|1\rangle\Big) = (-1)^{f(x)}\underbrace{|x\rangle \otimes H|1\rangle}$$

Output remains the same.

Grover's search algorithm

$$U\left(\,|x\rangle \otimes H|1\rangle\,\right) = (-1)^{f(x)}|x\rangle \otimes H|1\rangle$$

U and Q are linear operators.

$$\text{Define}: \quad Q|x\rangle = (-1)^{f(x)}|x\rangle = \begin{cases} |x\rangle, & \text{for } x \neq a \\ -|a\rangle, & \text{for } x = a \end{cases}$$

For a general state, $|\psi\rangle = \sum_x C_x|x\rangle$,

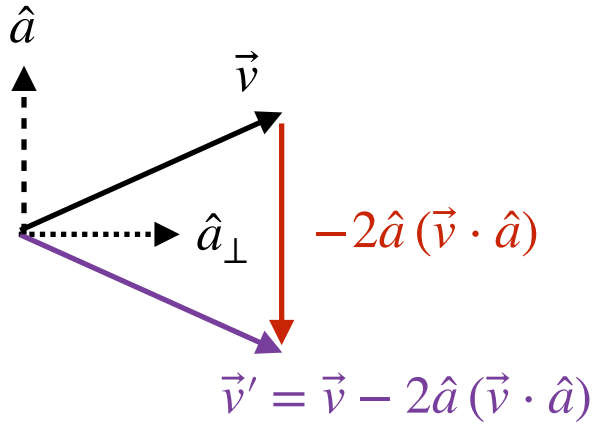
$$C_a \equiv \langle a|\psi\rangle$$

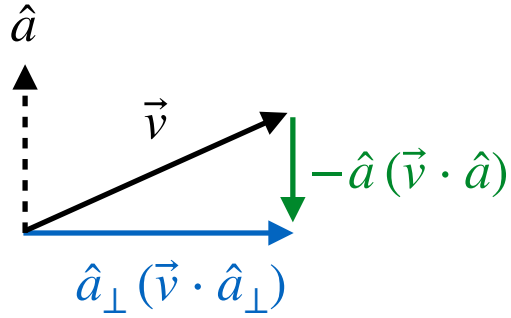
$$|\psi'\rangle = Q|\psi\rangle = \sum_{x\neq a} C_x|x\rangle - C_a|a\rangle = \sum_x C_x|x\rangle - 2C_a|a\rangle = |\psi\rangle - 2|a\rangle\langle a|\psi\rangle$$

$\langle a|\psi'\rangle = \langle a|\psi\rangle - 2\langle a|\psi\rangle = -\langle a|\psi\rangle$ → Suppose $|x\rangle$ satisfies $\langle x|a\rangle = 0$ for $x \neq a$

$\langle a_\perp|\psi'\rangle = \langle a_\perp|\psi\rangle$

Define such $|x\rangle$ as $|a_\perp\rangle$ with $\langle a|a_\perp\rangle = 0$





$\hat{a}\cdot\hat{a}_\perp = 0$

$\vec{v}\cdot\hat{a} = -\vec{v}'\cdot\hat{a}$

$\vec{v}\cdot\hat{a}_\perp = \vec{v}'\cdot\hat{a}_\perp$

reflection around the direction perpendicular to $\hat{a}$

Grover's search algorithm

- Consider uniform superposition of all possible basis states.

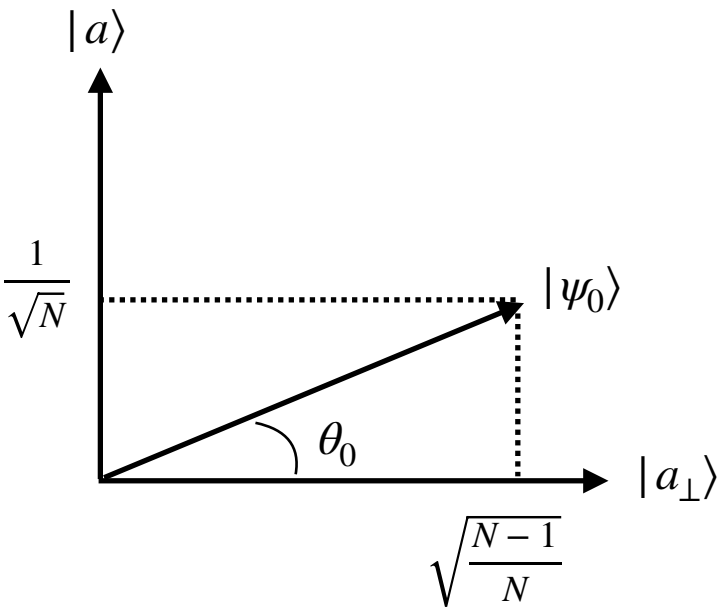
$$|\psi_0\rangle = H^{\otimes n}|0\rangle = \frac{1}{\sqrt{N}}\sum_{x=0}^{N-1}|x\rangle \qquad N = 2^n$$

$$\langle a|a_\perp\rangle = 0 \qquad \langle a|a\rangle = 1 \qquad \langle a_\perp|a_\perp\rangle = 1$$

$$|\psi_0\rangle = \frac{1}{\sqrt{N}}|a\rangle + \sqrt{\frac{N-1}{N}}|a_\perp\rangle = \sin\theta_0|a\rangle + \cos\theta_0|a_\perp\rangle$$

$$|a_\perp\rangle = \frac{1}{\sqrt{N-1}}\sum_{x\neq a,\, x=0}^{N-1}|x\rangle$$

$|a_\perp\rangle$ is the normalized uniform superposition of all basis states perpendicular to $|a\rangle$



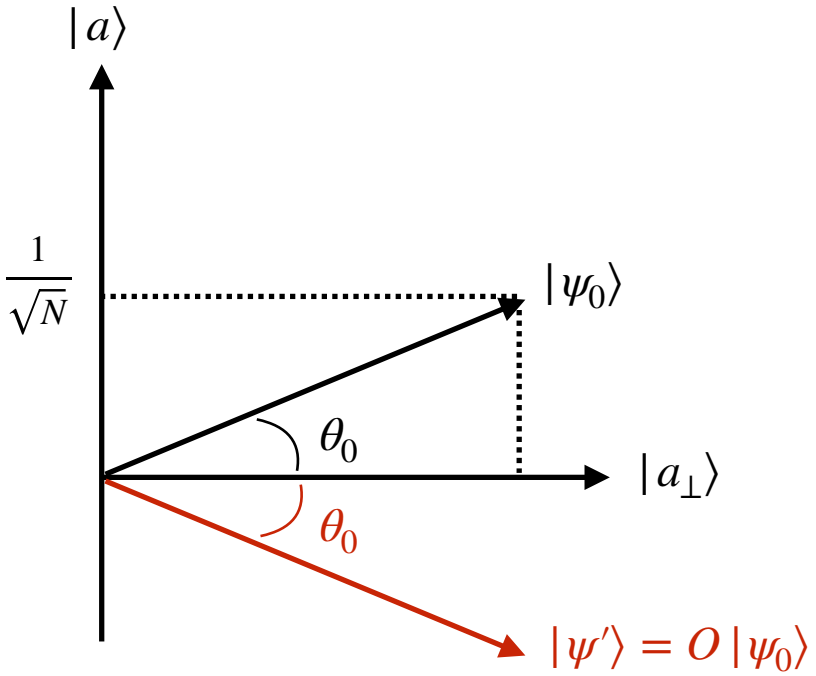
$$\langle a|\psi_0\rangle = \frac{1}{\sqrt{N}} \equiv \sin\theta_0 \qquad \langle a_\perp|\psi_0\rangle = \sqrt{\frac{N-1}{N}} \equiv \cos\theta_0$$

- Probability of $|\psi_0\rangle$ being in $|a\rangle$ = $|\langle a|\psi_0\rangle|^2 = \sin^2\theta_0 = \frac{1}{N}$

very small for a large N

- Grover's algorithm: iteratively rotate $|\psi_0\rangle$ (very close to $|a_\perp\rangle$ initially) to the direction close to $|a\rangle$ axis so that measurement returns a high probability → amplitude amplification

Grover's search algorithm



- Reflection about $|a_\perp\rangle$

$$|\psi'\rangle = O|\psi_0\rangle = |\psi\rangle - 2|a\rangle\langle a|\psi\rangle$$

- O reflects $|\psi_0\rangle$ about $|a_\perp\rangle$ axis
- $O|x\rangle = |x\rangle$ for $x \neq a$, $O|a_\perp\rangle = |a_\perp\rangle$

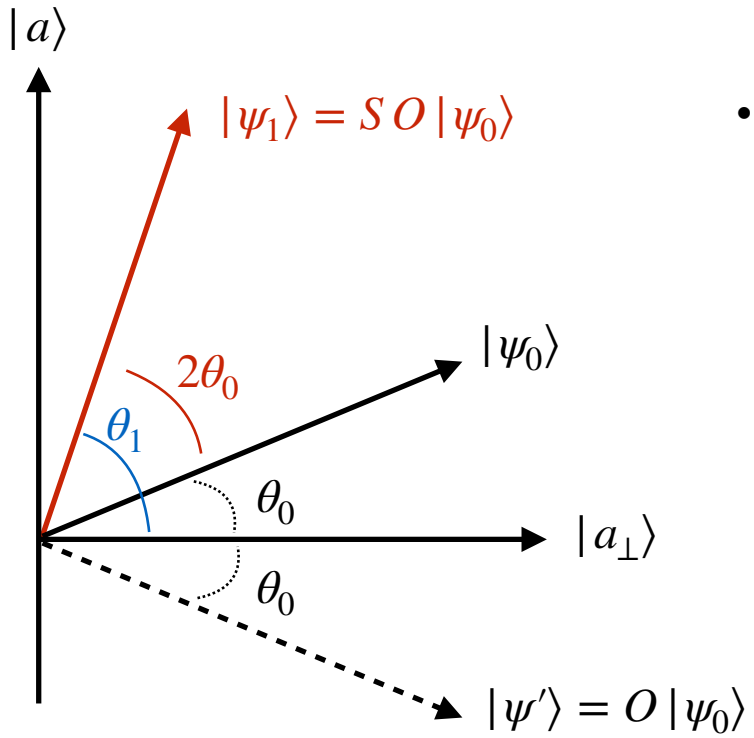
$O|a\rangle = -|a\rangle$ ⇨ flips $|a\rangle$ to $-|a\rangle$

$$|\psi_0\rangle = \sin\theta_0|a\rangle + \cos\theta_0|a_\perp\rangle$$

$$O|\psi_0\rangle = O\left(\sin\theta_0|a\rangle + \cos\theta_0|a_\perp\rangle\right)$$

$$= -\sin\theta_0|a\rangle + \cos\theta_0|a_\perp\rangle$$

Grover's search algorithm



- Reflection about $|\psi_0\rangle$ (initial state)

$$|\phi\rangle \longrightarrow |\phi'\rangle = S|\phi\rangle = 2|\psi_0\rangle\langle\psi_0|\phi\rangle - |\phi\rangle$$

$$\langle\psi_0|\phi'\rangle = 2\langle\psi_0|\psi_0\rangle\langle\psi_0|\phi\rangle - \langle\psi_0|\phi\rangle = \langle\psi_0|\phi\rangle$$

→ component along $|\psi_0\rangle$ does not change.

$$\langle\psi_{0\perp}|\phi'\rangle = 2\langle\psi_{0\perp}|\psi_0\rangle\langle\psi_0|\phi\rangle - \langle\psi_{0\perp}|\phi\rangle = -\langle\psi_{0\perp}|\phi\rangle$$

→ component perpendicular to $|\psi_0\rangle$ changes the sign.

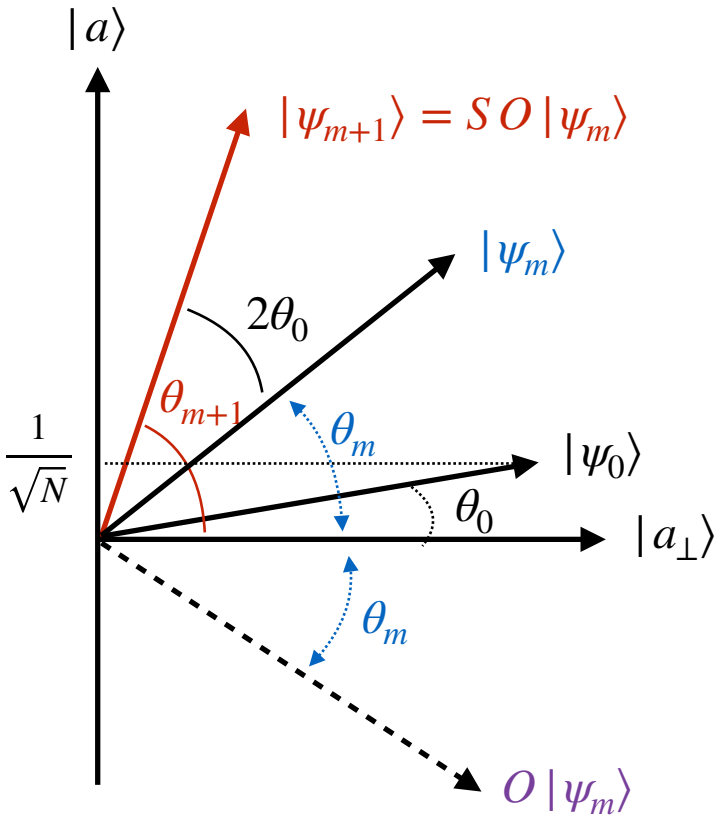
$G = SO$ O: reflection of $|\psi_0\rangle$ about $|a_\perp\rangle$

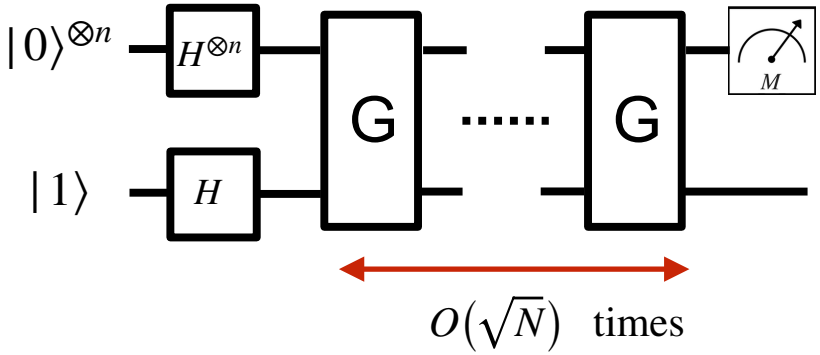
S: reflection of $|\psi'\rangle = O\,|\psi_0\rangle$ about initial state $|\psi_0\rangle$

G: Grover operator rotates the initial state $|\psi_0\rangle$ by $2\theta_0$ counterclockwise (toward the direction of $|a\rangle$ axis)

- Effect of 1st Grover iteration: rotate the initial state $|\psi_0\rangle$ by $2\theta_0$ counterclockwise.
- $|\psi_1\rangle$ making angle θ_1 to $|a_\perp\rangle$ axis, $\theta_1 = \theta_0 + 2\theta_0$

Grover's search algorithm





$|\psi_m\rangle = \cos\theta_m |a_\perp\rangle + \sin\theta_m |a\rangle$

$O|\psi_m\rangle$ rotates about $|a_\perp\rangle$ by angle $2\theta_m$

$SO|\psi_m\rangle$ rotates by angle $2(\theta_m + \theta_0)$ counterclockwise.

$$\theta_{m+1} = \theta_m + 2\theta_0 \qquad \theta_m = (2m+1)\theta_0$$

- $\langle a|\psi_m\rangle = \sin\theta_m = \sin[(2m+1)\theta_0]$
- Optimal number of Grover iteration: $\theta_m = \pi/2$

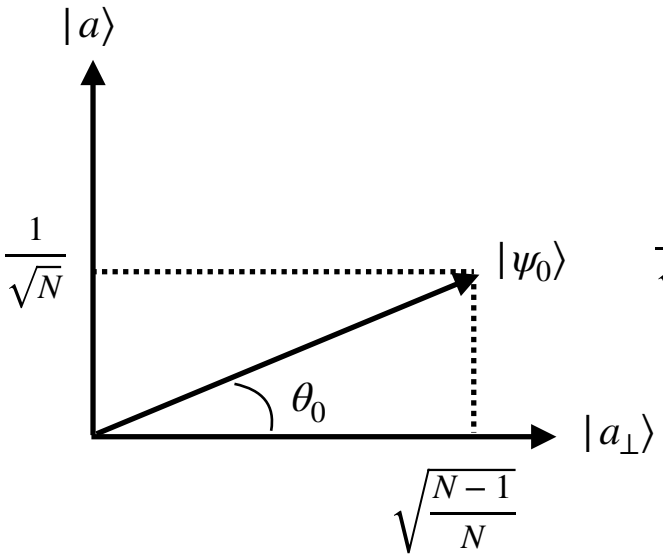
$$\frac{\pi}{2} = \theta_m = (2m+1)\theta_0 = (2m+1)\sin^{-1}\left(\frac{1}{\sqrt{N}}\right)$$

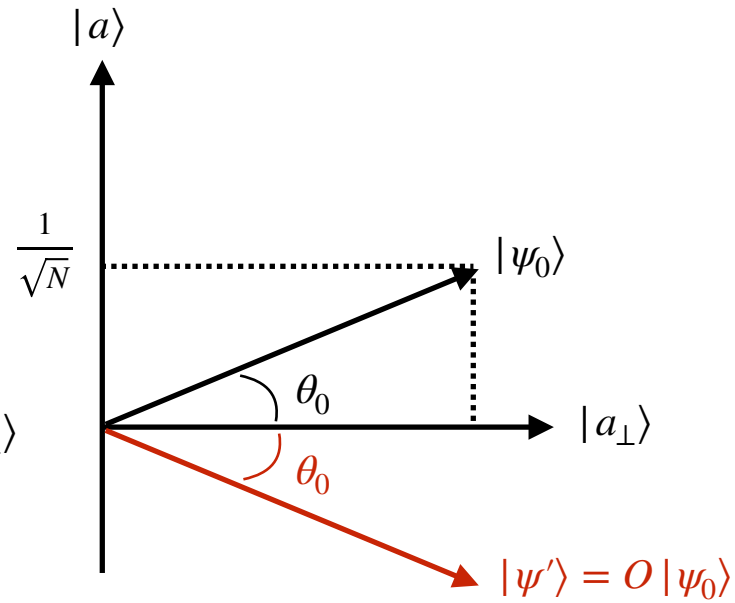
For a large N, $m = \frac{\pi}{4}\sqrt{N}$

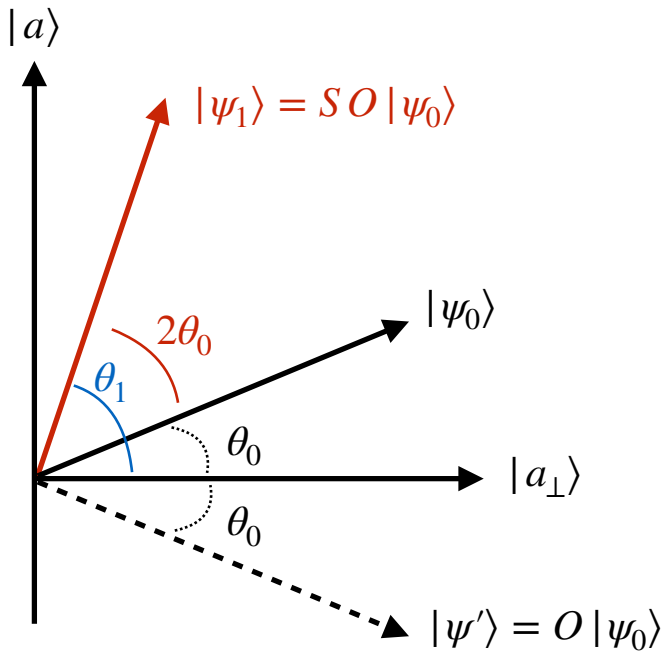
- When $\theta_m \approx \pi/2$, measurement gives a with high probability.
- For any value of θ_m such that $\frac{\pi}{4} < \theta_m \approx \frac{2m}{\sqrt{N}} < \frac{3\pi}{4}$

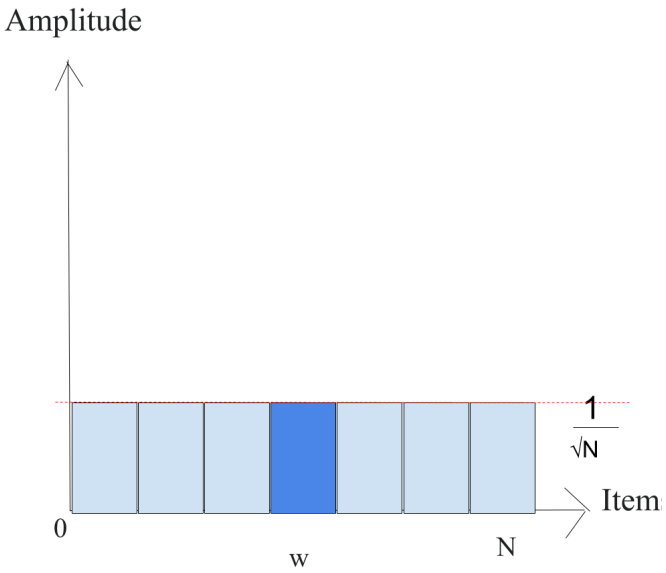
$\rightarrow \frac{\pi}{8}\sqrt{N} < m < \frac{3\pi}{8}\sqrt{N}$, Grover algorithm returns $|a\rangle$ with probability > 1/2.

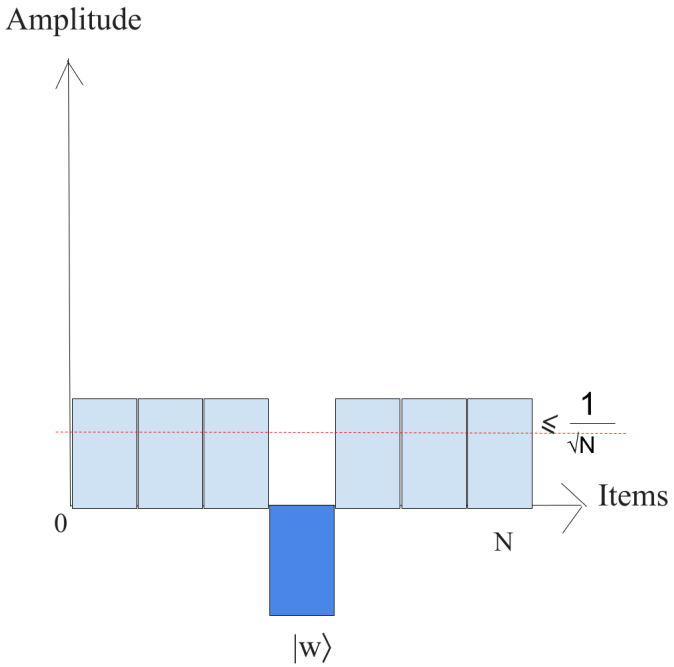
Grover's search algorithm

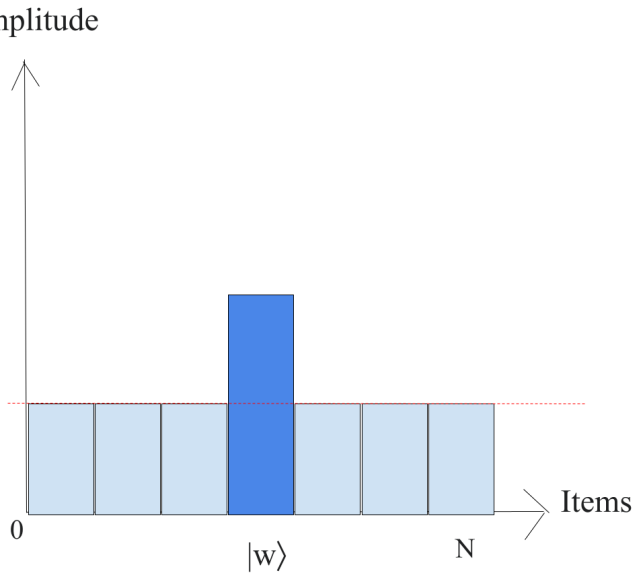




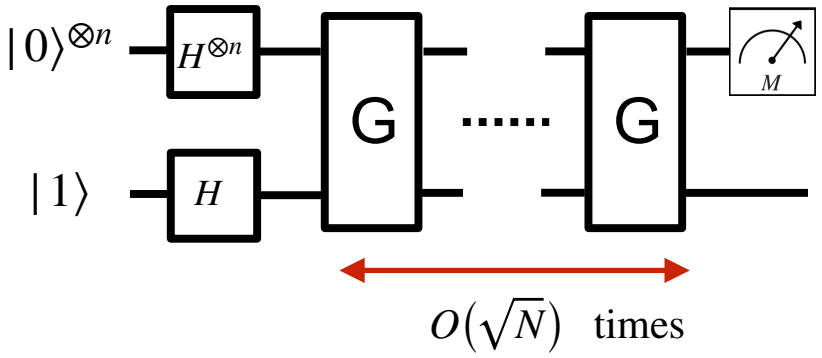








Grover's search algorithm



- Optimal number of Grover iteration: $\theta_m = \pi/2$

$$\frac{\pi}{2} = \theta_m = (2m+1)\,\theta_0 = (2m+1)\,\sin^{-1}\left(\frac{1}{\sqrt{N}}\right)$$

For a large N, $m = \frac{\pi}{4}\sqrt{N}$

- When $\theta_m \approx \pi/2$, measurement gives a with high probability.
- For any value of θ_m such that $\frac{\pi}{4} < \theta < \frac{3\pi}{4} \rightarrow \frac{\pi}{8}\sqrt{N} < m < \frac{3\pi}{8}\sqrt{N}$, Grover algorithm returns $|a\rangle$ with probability > 1/2.
- Probability decreases for $m > \frac{\pi}{4}\sqrt{N}$.
- Operational count of the Grover algorithm $\approx O(\sqrt{N}) \rightarrow$ quadratic speed up compared with $O(N)$ count on a classical computer.
- Quantum advantage: superposition and $N = 2^n$ values of $f(x)$ evaluated in parallel
  - Operation count of $O(1)$?
  - Measurement returns only one $(x, f(x))$ value
  - Requires additional operations $\rightarrow O(\sqrt{N})$

nature COMMUNICATIONS

https://arxiv.org/pdf/1703.10535.pdf

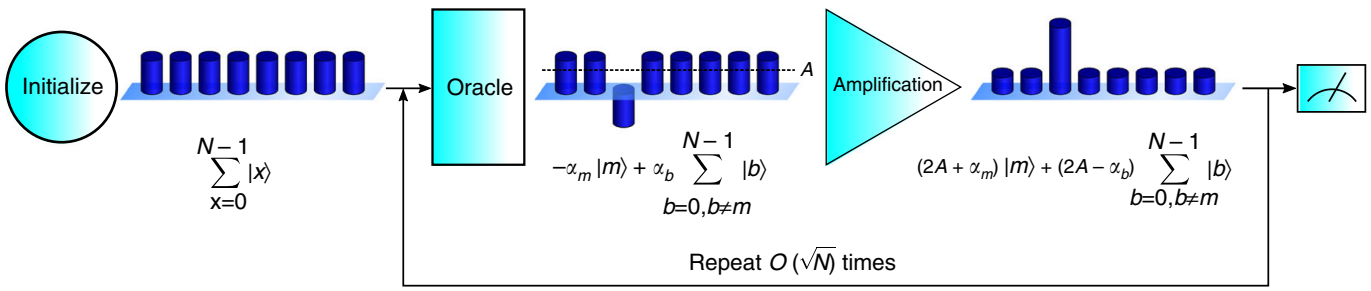
ARTICLE

DOI: 10.1038/s41467-017-01904-7 OPEN

Complete 3-Qubit Grover search on a programmable quantum computer

C. Figgatt [1], D. Maslov[1,2], K.A. Landsman[1], N.M. Linke[1], S. Debnath[1] & C. Monroe[1,3]

The Grover quantum search algorithm is a hallmark application of a quantum computer with a well-known speedup over classical searches of an unsorted database. Here, we report results for a complete three-qubit Grover search algorithm using the scalable quantum computing technology of trapped atomic ions, with better-than-classical performance. Two methods of state marking are used for the oracles: a phase-flip method employed by other experimental demonstrations, and a Boolean method requiring an ancilla qubit that is directly equivalent to the state marking scheme required to perform a classical search. We also report the deterministic implementation of a Toffoli-4 gate, which is used along with Toffoli-3 gates to construct the algorithms; these gates have process fidelities of 70.5% and 89.6%, respectively.



PHYSICAL REVIEW A **104**, 062438 (2021)

Editors' Suggestion

Quantum amplitude-amplification operators

Hyeokjea Kwon and Joonwoo Bae

School of Electrical Engineering, Korea Advanced Institute of Science and Technology, 291 Daehak-ro, Yuseong-gu, Daejeon 34141, Republic of Korea

(Received 21 July 2021; revised 17 November 2021; accepted 3 December 2021; published 27 December 2021)

In this work, we show the characterization of quantum iterations that would generally construct quantum amplitude-amplification algorithms with a quadratic speedup, namely, quantum amplitude-amplification operators (QAAOs). Exact quantum search algorithms that find a target with certainty and with a quadratic speedup can be composed of sequential applications of QAAOs: existing quantum amplitude-amplification algorithms thus turn out to be sequences of QAAOs. We show that an optimal and exact quantum amplitude-amplification algorithm corresponds to the Grover algorithm together with a single iteration of QAAO. We then realize three-qubit QAAOs with current quantum technologies via cloud-based quantum computing services, IBMQ and IonQ. Finally, our results show that the fixed-point quantum search algorithms known so far are not a sequence of QAAOs; for example, the amplitude of a target state may decrease during quantum iterations.

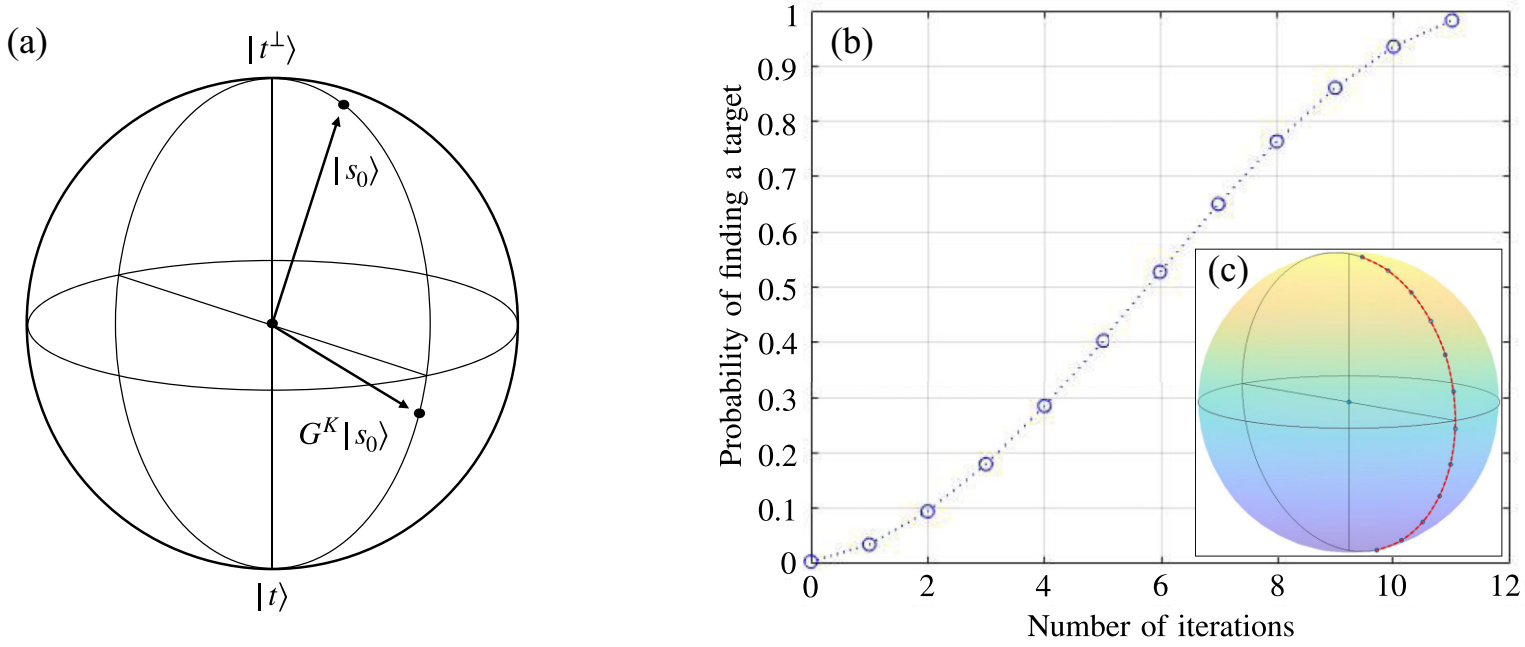


FIG. 1. (a) The Grover iteration corresponds to consecutive rotations in the space spanned by a target state $|t\rangle$ and its complement $|t^{\perp}\rangle$. (b) The probability of finding a target state is plotted in the case of eight qubits. The probability is monotonically increasing. (c) The path of an evolving state in the sphere is shown by Grover iterations from an initial to target states.

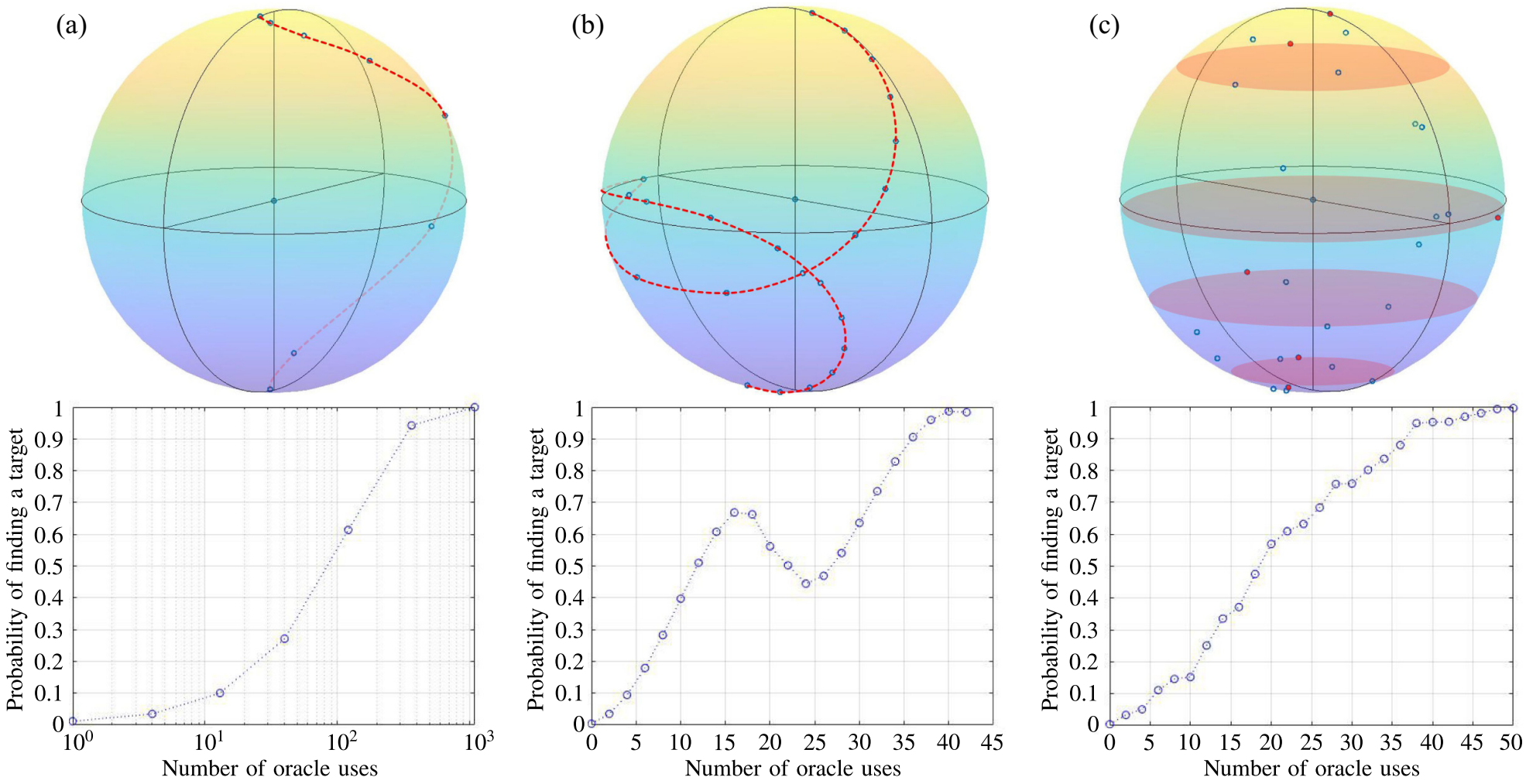
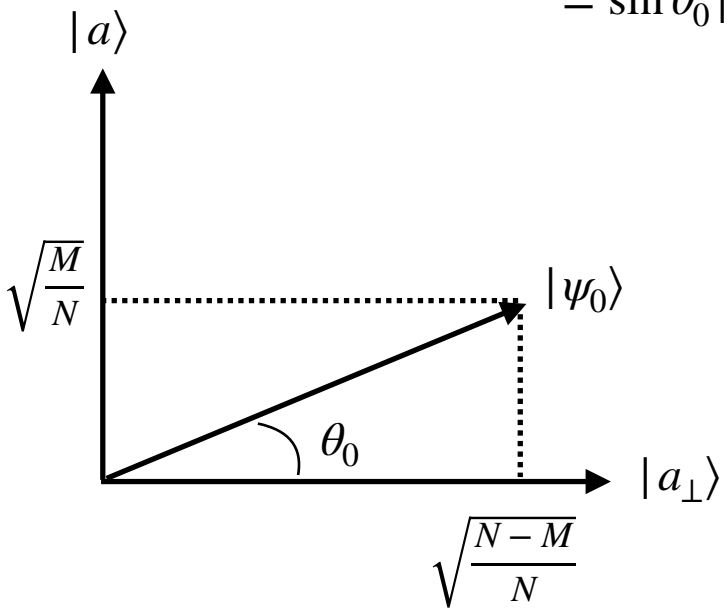


FIG. 4. Quantum amplitude amplification is performed in the case of eight qubits. The x axis shows the number of oracle uses, and the y axis shows the probability of finding a target state. (a) The $\pi/3$ algorithm is plotted [19]. The amplitude increases all the time until 10^3 oracle calls, without a quantum speedup. (b) A fixed-point quantum search with optimal query complexity is plotted [20]. The amplitude of the target state decreases in the meanwhile, and the oracle is called 45 times. (c) QAAOs are randomly generated and concatenated so that the amplitude keeps increasing until it reaches 1 after the oracle calls 50 times.

Extension to more than one special value

- What if three are M solutions, $a_i, \quad i = 1, 2, \cdots, M$
- Superposition of all special states: $|a\rangle = \frac{1}{\sqrt{M}} \sum_{x \in \{a_i\}} |x\rangle$ $\quad \langle a|a\rangle = 1$, $\quad \langle a|a_\perp\rangle = 0$
- Uniform superposition of all other states: $|a_\perp\rangle = \frac{1}{\sqrt{N-M}} \sum_{x \notin \{a_i\}} |x\rangle$ $\quad \langle a_\perp|a_\perp\rangle = 1$
- Initial state: $|\psi_0\rangle = \sqrt{\frac{M}{N}}|a\rangle + \sqrt{\frac{N-M}{N}}|a_\perp\rangle$ $\quad \langle\psi_0|\psi_0\rangle = 1$

$$= \sin\theta_0|a\rangle + \cos\theta_0|a_\perp\rangle \qquad \sin\theta_0 = \langle a|\psi_0\rangle = \sqrt{\frac{M}{N}}$$



- $M \ll N,\ \theta \approx \frac{\pi}{2} \rightarrow m = \frac{\pi}{4}\sqrt{\frac{N}{M}}$

Quantum Counting

- What if we had no prior knowledge of M?
- Grover operator G rotates vectors in $|a\rangle - |a_\perp\rangle$ plane by angle $2\theta_0$

$$\sin\theta_0 = \sqrt{\frac{M}{N}}$$

$$G = \begin{pmatrix} \cos 2\theta_0 & -\sin 2\theta_0 \\ \sin 2\theta_0 & \cos 2\theta_0 \end{pmatrix} \longrightarrow \text{eigenvalues } e^{\pm 2i\theta_0}$$

$\longrightarrow$ Phase estimation

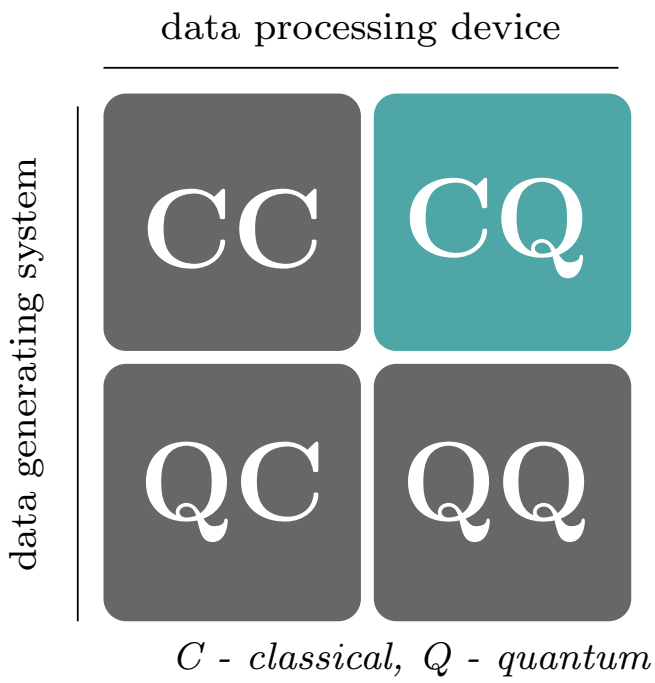
$\longrightarrow$ $\theta_0 \rightarrow M$

$\longrightarrow$ Quantum algorithm can tell us whether a special value exists at all, i.e., M=0.

Distance-based classifier

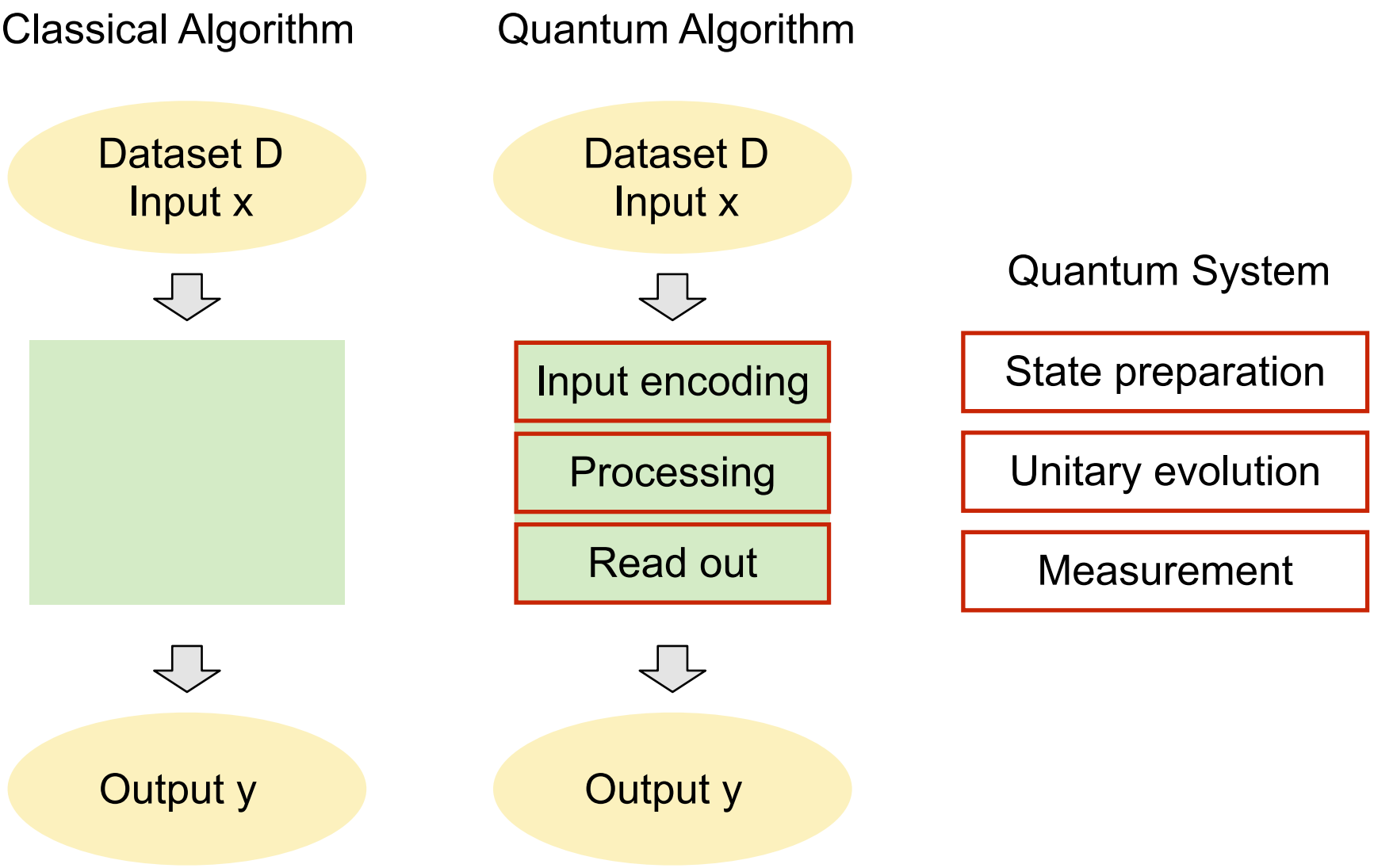
Quantum Machine Learning

- Artificial Intelligence: Statistical prediction
- Machine Learning: Learn from data
- Quantum Machine Learning: Learn from data with quantum algorithms
  - Subdiscipline of quantum computing and quantum information science



- CC: classical data being processed classically
- QC: how machine learning can help with quantum computing
- CQ: classical data fed into quantum computer for analysis (quantum machine learning)
- QQ: quantum data being processed by quantum computer (ex: Quantum simulation)

Quantum Algorithms and Data Embedding



Quantum Algorithms and Data Embedding

	Classical data	Requirement	Quantum state
Basis Encoding	$\vec{x} \in \{0,1\}^{\otimes n}$ $\vec{x} = (x_1, x_2, \cdots, x_n) \in \{0,1\}$		$\lvert\psi\rangle = \lvert x_1 x_2 \cdots x_n\rangle$ $= \lvert x_1\rangle \otimes \lvert x_2\rangle \otimes \cdots \otimes \lvert x_n\rangle$
Amplitude Encoding	$\vec{x} \in \mathbb{R}^{2^n}$ $x_i \in \mathbb{R}$	$\sum_{i=1}^{2^n} \lvert x_i\rvert^2 = 1$	$\lvert\psi_x\rangle = \sum_{i=1}^{2^n} x_i \lvert i\rangle$
	$A \in \mathbb{R}^{2^n \times 2^m}$ $i = 1,\cdots,2^n$ $A_{ij} \in \mathbb{R}$ $j = 1,\cdots,2^m$	$\sum_{i,j} \lvert A_{ij}\rvert^2 = 1$	$\lvert\psi_A\rangle = \sum_{i,j} A_{ij} \lvert i\rangle \otimes \lvert j\rangle$
	$A \in \mathbb{R}^{2^n \times 2^n}$	$\sum_i A_{ii} = 1$ $A^\dagger = A$ $A^*_{ij} = A_{ji}$	$\rho_A = \sum_{i,j} A_{ij} \lvert i\rangle\langle j\rvert$
Time-evolution Encoding	$x \in \mathbb{R}$	$x \in [0, 2\pi)$	$U(x) = e^{-ixH}$
Hamiltonian Encoding	$A \in \mathbb{R}^{2^n \times 2^n}$	$A^\dagger = A$	$H_A = A$
	$A \in \mathbb{R}^{2^n \times 2^n}$	$A^\dagger \neq A$ (in general)	$H_A = \begin{pmatrix} 0 & A \\ A^\dagger & 0 \end{pmatrix}$

Binary encoding into basis states

basis vector coefficient {0,1}

Represent numbers as binaries, each binary digit requires a qubit

$$x = \sum_{k=1}^{\tau-1} b_k \frac{1}{2^k}$$

binary fraction rep.

data vector

sign

$$x = (0.1, -0.6, 1.0) \longrightarrow \begin{matrix} 0.1 \to 0\,0001\ldots \\ -0.6 \to 1\,1001\ldots \\ 1.0 \to 0\,1111\ldots. \end{matrix} \longrightarrow |00001\ 11001\ 01111\rangle$$

quantum state

- Binary fraction = expression in power of 1/2

In decimal form: $0.j_\ell\, j_{\ell+1} \cdots j_m = \frac{j_\ell}{2} + \frac{j_{\ell+1}}{2^2} + \cdots + \frac{j_m}{2^{m-\ell+1}}$

$$j = j_1 2^7 + j_2 2^6 + j_3 2^5 + j_4 2^4 + j_5 2^3 + j_6 2^2 + j_7 2^1 + j_8 2^0$$

$$\frac{j}{2^3} = j_1 2^4 + j_2 2^3 + j_3 2^2 + j_4 2^1 + j_5 2^0 + j_6 2^{-1} + j_7 2^{-2} + j_8 2^{-3}$$

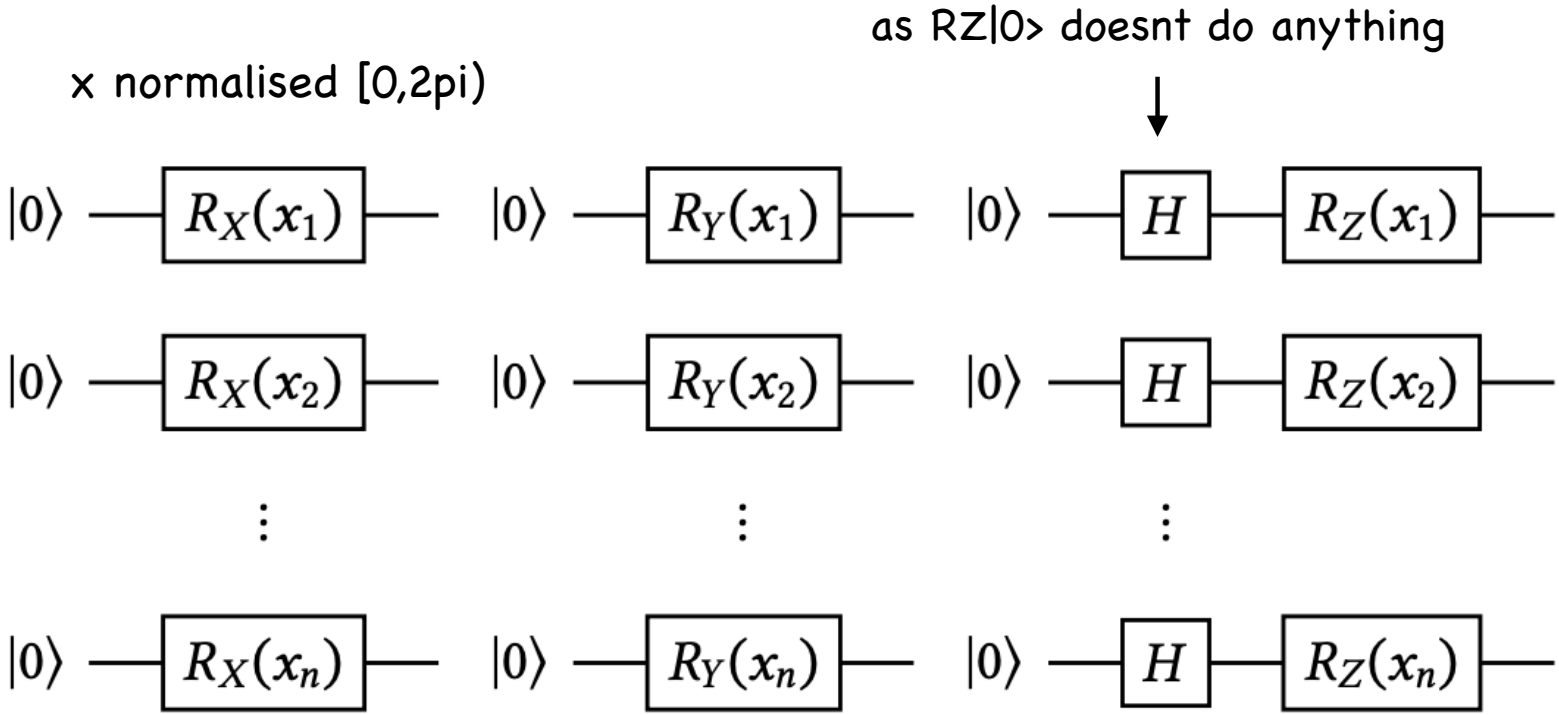
$j_1 j_2 j_3 j_4 j_5 \cdot j_6 j_7 j_8$

binary fraction: $0\,.j_6 j_7 j_8$

Angle/Rotation encoding

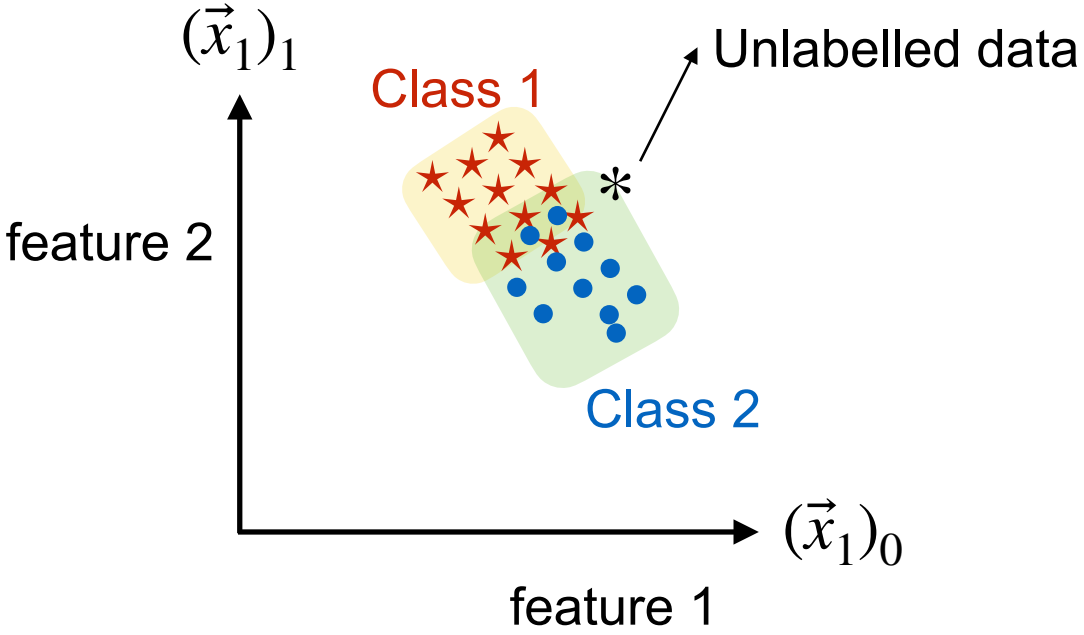
When used on an n-qubit circuit, this feature map of angle encoding can take up to n numerical inputs $x_1, \ldots, x_n$. The action of its circuit consists in the application of a rotation gate on each qubit j parametrised by the value x_j. In this feature map, we are using the x_j values as angles in the rotations, hence the name of the encoding.

Example



Distance-based classifier

- A distance-based classifier with a quantum interference circuit: arXiv:1703:10793 (supervised binary classification)



training data set

$$D = \left\{(\vec{x}_1, y_1),\ (\vec{x}_2, y_2),\ \cdots,\ (\vec{x}_M, y_M)\right\}$$

$$\vec{x}_m \in \mathbb{R}^N \quad y_m \in \{-1, +1\}$$

$$m = 1, 2, \cdots, M$$

M = the number of data

N = the number of features

$\vec{\tilde{x}}_m \in \mathbb{R}^N$: unlabelled data

→ Find the label $\tilde{y} \in \{-1, 1\}$

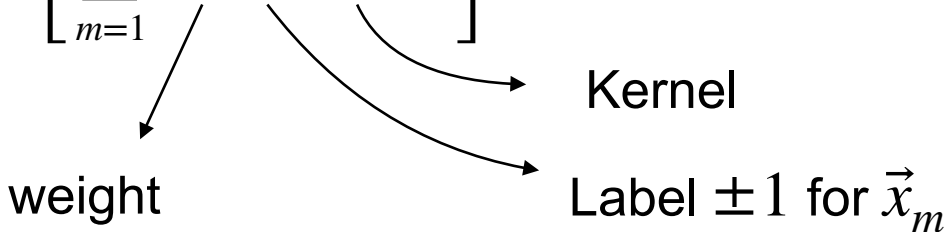
Classical Kernel Method

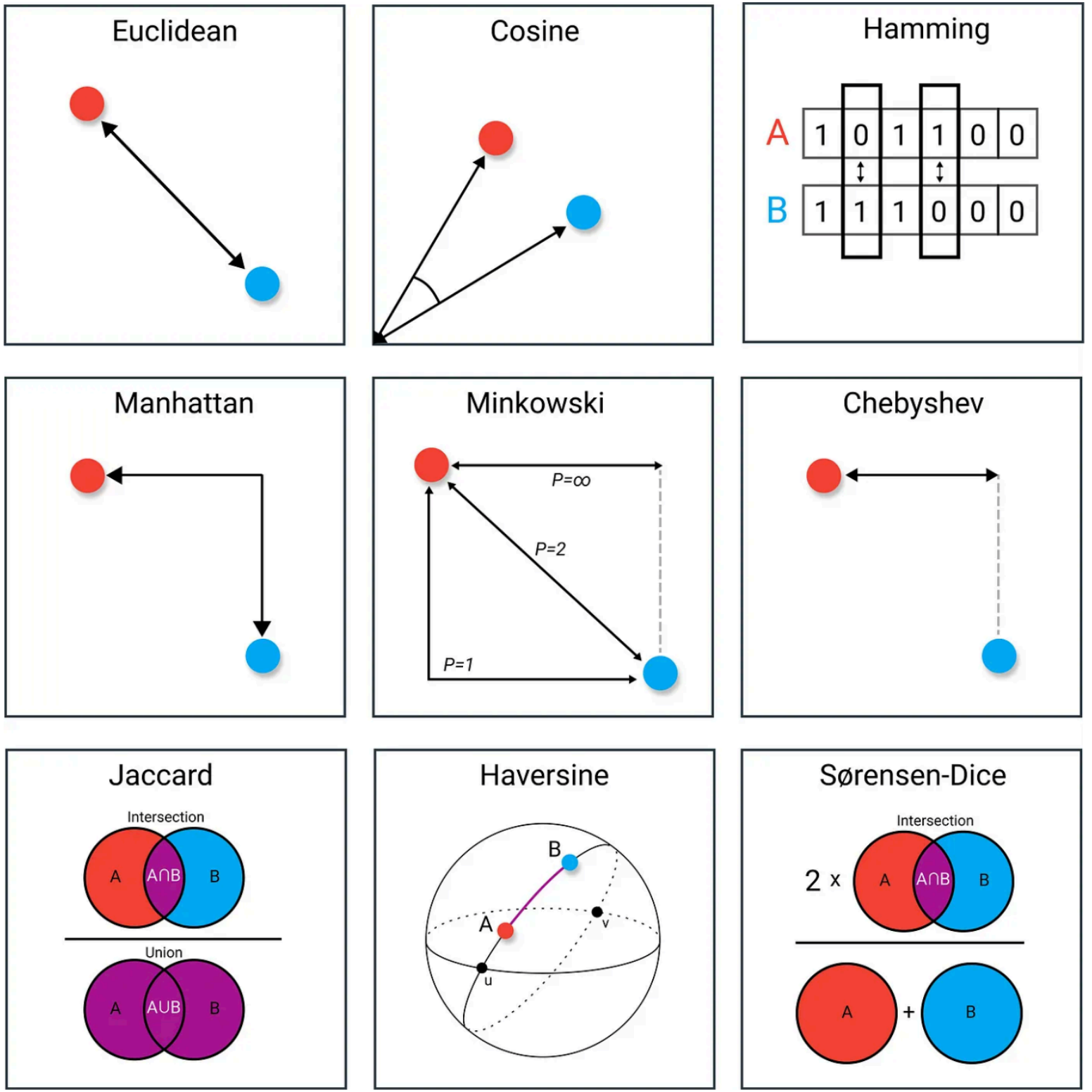
- Kernel methods: kNN (k-nearest neighborhood), KDE (kernel density estimation), SVM (support vector machine), Gaussian processes
  - Nearest neighborhood method: a new input data is given the same label as the data point closest to it → k-nearest neighborhood (kNN)
  - Closeness = distance measure
  - (ex) Euclidean distance $|\vec{\tilde{x}} - \vec{x}_m|^2$

$$\tilde{y} = \text{sign}\left[\sum_{m=1}^{M} y_m \left(1 - \frac{1}{4M}|\vec{\tilde{x}} - \vec{x}_m|^2\right)\right]$$

$$\tilde{y} = \text{sign}\left[\sum_{m=1}^{M} w_m\, y_m\, \kappa(\vec{\tilde{x}}, \vec{x}_m)\right]$$

- include all data but weigh influence of each data toward the decision by the weight $\kappa(\vec{\tilde{x}}, \vec{x}_m)$




Euclidean
Cosine
Hamming
A
1 0 1 1 0 0
B
1 1 1 0 0 0
Manhattan
Minkowski
P=∞
P=2
P=1
Chebyshev
Jaccard
Intersection
A
A∩B
B
Union
A
AUB
B
Haversine
B
A
v
u
Sørensen-Dice
Intersection
2 x
A
A∩B
B
A
+
B

Wasserstein distance (Kantorovich–Rubinstein metric)

- A distance function defined between probability distributions on a given metric space M (named after “Vaserštein" (Russian: Васерштейн))
- If P is an empirical measure with samples $X_1, \cdots, X_n$ and Q is an empirical measure with samples $Y_1, \cdots, Y_n$ the p-Wasserstein distance is a simple function of the order statistics:

$$W_p(P,Q) = \left(\frac{1}{n} \sum_{i=1}^{n} \|X_{(i)} - Y_{(i)}\|^p \right)^{1/p}.$$

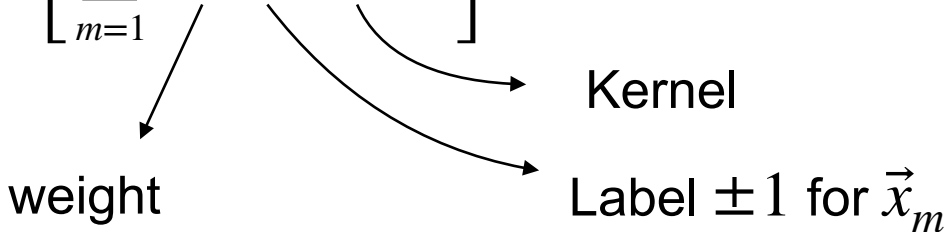
Classical Kernel Method

- Kernel methods: kNN (k-nearest neighborhood), KDE (kernel density estimation), SVM (support vector machine), Gaussian processes
  - Nearest neighborhood method: a new input data is given the same label as the data point closest to it → k-nearest neighborhood (kNN)
  - Closeness = distance measure
  - (ex) Euclidean distance $|\vec{\tilde{x}} - \vec{x}_m|^2$

$$\tilde{y} = \text{sign}\left[\sum_{m=1}^{M} y_m \left(1 - \frac{1}{4M}|\vec{\tilde{x}} - \vec{x}_m|^2\right)\right]$$

$$\tilde{y} = \text{sign}\left[\sum_{m=1}^{M} w_m\, y_m\, \kappa(\vec{\tilde{x}}, \vec{x}_m)\right]$$

- include all data but weigh influence of each data toward the decision by the weight $\kappa(\vec{\tilde{x}}, \vec{x}_m)$



Distance-based classifier

- Choose $w_m = 1$ for all equally important data

$$\kappa(\vec{\tilde{x}}, \vec{x}_m) = 1 - \frac{1}{4M} |\vec{\tilde{x}} - \vec{x}_m|^2$$

Close data (small distance) are weighted more importantly.

(1) Encode input data (features) into the amplitude of a quantum system (amplitude encoding). For classical vector $\vec{x} \in \mathbb{R}^N$, $(N = 2^n)$ Assume $x^T x = \vec{x} \cdot \vec{x} = 1$ (normalized to 1)

$N = 2^n$: number of features

$$|\psi_X\rangle = \sum_{i=0}^{N-1} x_i |i\rangle$$

i : index in the computational basis

Dimension of Hilbert space $\approx O(\log N)$

ancilla qubit is entangled with third register

(2) initial state: $$|D\rangle = \frac{1}{\sqrt{2M}} \sum_{m=1}^{M} |m\rangle \left(|0\rangle |\psi_{\tilde{x}}\rangle + |1\rangle |\psi_{x_m}\rangle \right) |y_m\rangle$$

data index
M = # of data

unlabelled data

labeled data

label of x_m class qubit

Distance-based classifier

(2) initial state: $|D\rangle = \frac{1}{\sqrt{2M}} \sum_{m=1}^{M} |m\rangle \left(|0\rangle |\psi_{\tilde{x}}\rangle + |1\rangle |\psi_{x_m}\rangle \right) |y_m\rangle$

ancilla qubit is entangled with third register

data index
M = # of data

unlabelled data

labeled data

label of x_m class qubit

$$|\psi_{x_m}\rangle = \sum_{i=0}^{N-1} x_m^i |i\rangle$$ encoding of m-th training data (labeled)

$$|\psi_{\tilde{x}}\rangle = \sum_{i=0}^{N-1} \tilde{x}^i |i\rangle$$ encoding of new data (unlabeled)

$$|y_m\rangle = \begin{cases} |0\rangle, & \text{if } y_m = -1 \\ |1\rangle, & \text{if } y_m = +1 \end{cases}$$

$|D\rangle$ contains all training data as well as M copies of new input.

Distance-based classifier

(3) Apply Hadamard gate on the ancilla (second) qubit.

$$|0\rangle \rightarrow \frac{1}{\sqrt{2}}\Big(|0\rangle + |1\rangle\Big) \qquad |1\rangle \rightarrow \frac{1}{\sqrt{2}}\Big(|0\rangle - |1\rangle\Big)$$

$$|D\rangle = \frac{1}{\sqrt{2M}} \sum_{m=1}^{M} |m\rangle \Big(|0\rangle\,|\psi_{\tilde{x}}\rangle + |1\rangle\,|\psi_{x_m}\rangle\Big)\,|y_m\rangle$$

$$\downarrow$$

$$|D'\rangle = \frac{1}{2\sqrt{M}} \sum_{m=1}^{M} |m\rangle \Big(|0\rangle\,|\psi_{\tilde{x}+x_m}\rangle + |1\rangle\,|\psi_{\tilde{x}-x_m}\rangle\Big)\,|y_m\rangle$$

$$|\psi_{\tilde{x}\pm x_m}\rangle = |\psi_{\tilde{x}}\rangle \pm |\psi_{x_m}\rangle = \sum_{i=0}^{M-1} \left(\tilde{x}^i \pm x_m^i\right)|i\rangle$$

(4) Conditional measurement selecting the branch with ancilla state $|0\rangle$.

Likely to succeed if the collective Euclidean distance b/w $\tilde{x}$ and training data set is small. For standard data, $p \geq 0.5$.

Probability is $p = \frac{1}{4M} \sum_{m} |\vec{\tilde{x}} + \vec{x}_m|^2$

$$|D''\rangle = \frac{1}{2\sqrt{Mp}} \sum_{m=1}^{M} \sum_{i=0}^{N-1} |m\rangle \left(\tilde{x}^i + x_m^i\right)|i\rangle\,|y_m\rangle$$

Distance-based classifier

(5) Probability of measuring the class qubit $|y_m\rangle = |0\rangle$

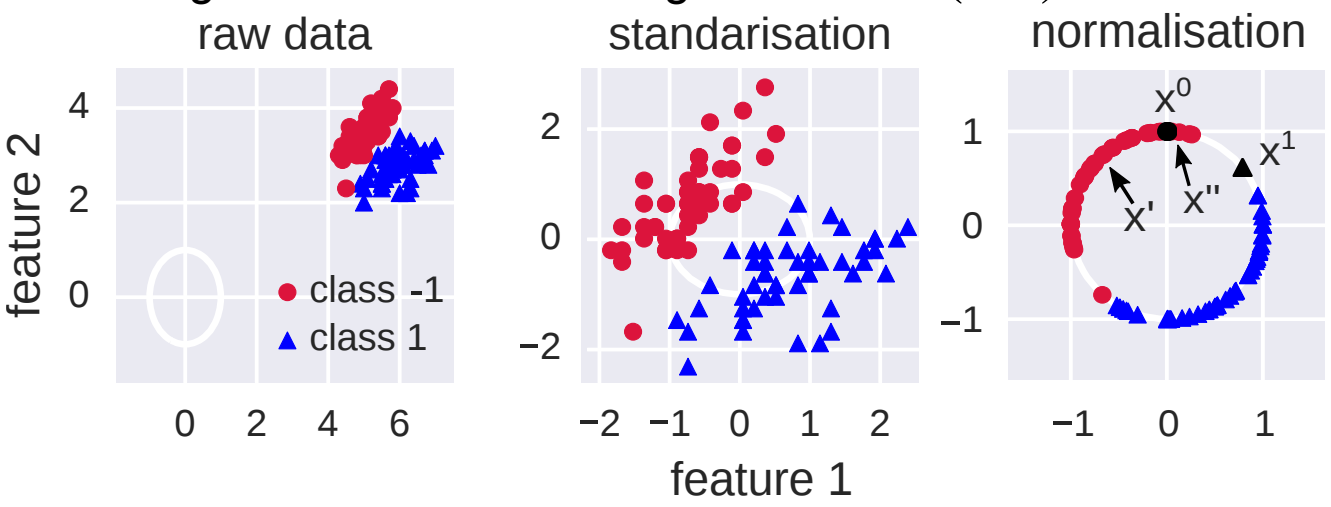
$$|D''\rangle = \frac{1}{2\sqrt{Mp}} \sum_{m=1}^{M} \sum_{i=0}^{N-1} |m\rangle \left(\tilde{x}^i + x_m^i\right) |i\rangle |y_m\rangle$$

$$P(\tilde{y}=0) = \frac{1}{4Mp} \sum_{y_m=0,\, m=1}^{M} |\vec{\tilde{x}} + \vec{x}_m|^2 = 1 - \frac{1}{4Mp} \sum_{y_m=0,\, m=1}^{M} |\vec{\tilde{x}} - \vec{x}_m|^2$$

Class 1

using normalization condition

→ choosing the class with the higher probability gives result of kernel method.
The # of measurement needed to estimate $P(\tilde{y}=0)$ to error ϵ with a reasonably high confidence interval grows with $O(\epsilon^{-1})$.



- arXiv:1703:10793 used Iris data

https://www.quantum-inspire.com/kbase/jupyter-classifier-part1/

The data set
Iris flower data set

consider only first two features of two samples

R. A. Fisher
(1936)

Iris setosa

Iris versicolor

Iris virginica

(Sepal length & width, petal length & width)

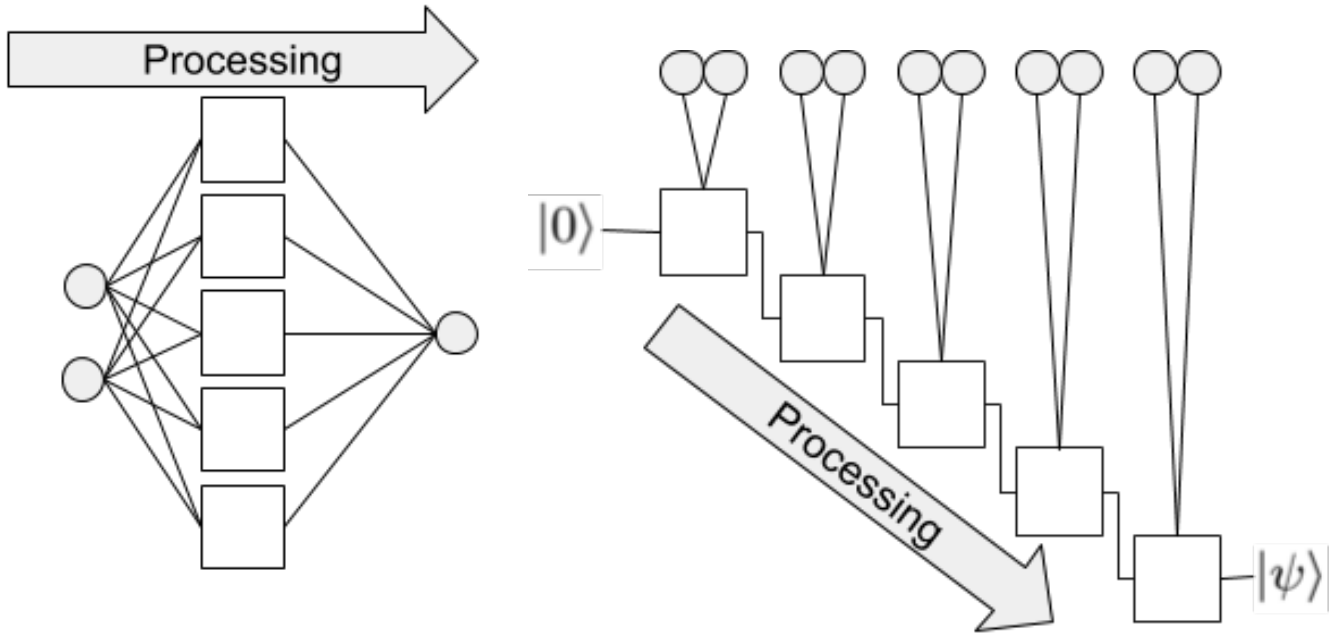
https://en.wikipedia.org/wiki/Iris_flower_data_set

Data re-uploading for a universal quantum classifier

1907.02085

Data re-uploading for a universal quantum classifier

1907.02085



(a) Neural network (b) Quantum classifier

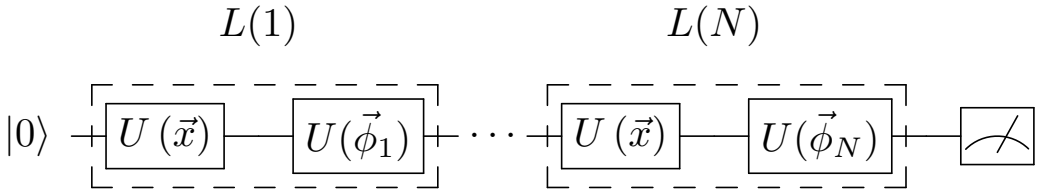
- Universal approximation theorem
- We can approximate a function $F(\vec{x})$ with $f(\vec{x}, \vec{\theta})$, where $\vec{x}$ is an input feature and $\vec{\theta}$ is a learnable parameter.
- The cost function (ex. MSE) to be minimized is $\sum_{i=1}^{n} |F(\vec{x}_i) - f(\vec{x}_i, \vec{\theta})|^2$

Single qubit classifier using data re-uploading

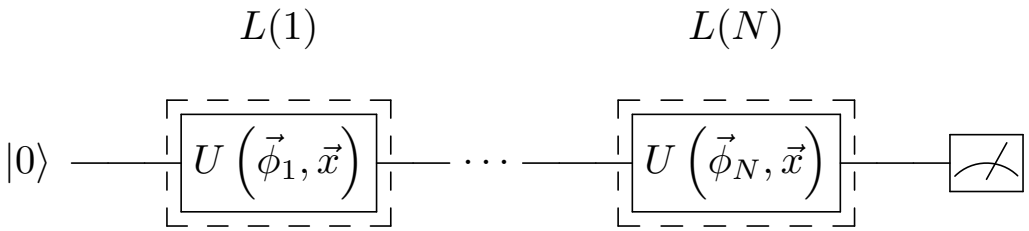
- Consider the three dimensional data, $\vec{x}$. (can be generalized.) 1907.02085
- Date can be re-uploaded using unitary transformation $U(\vec{x})$ rotating the qubit.
- The single-qubit classifier has the following structure: $|\psi\rangle = \mathcal{U}(\vec{\phi}, \vec{x})|0\rangle$

$$\mathcal{U}(\vec{\phi}, \vec{x}) = L(N)\dots L(1) \qquad L(i) \equiv U(\vec{\phi}_i)U(\vec{x}) \qquad \overrightarrow{\phi} = (\phi_1, \phi_2, \phi_3)$$

$$\mathcal{U}(\vec{\phi}, \vec{x}) \equiv U(\vec{\phi}_N)U(\vec{x})\dots U(\vec{\phi}_1)U(\vec{x}) \qquad U(\phi_1, \phi_2, \phi_3) \in SU(2)$$



(a) Original scheme



(b) Compressed scheme

$$L(i) = U\left(\vec{\theta}_i + \vec{w}_i \circ \vec{x}\right)$$

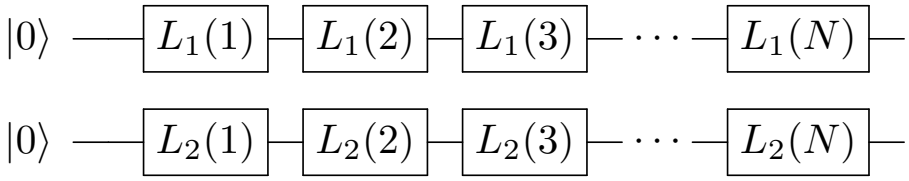
Hadamard product of $\overrightarrow{w}_i$ and $\vec{x}$:

$$\vec{w}_i \circ \vec{x} = \left(w_i^1 x^1, w_i^2 x^2, w_i^3 x^3\right)$$

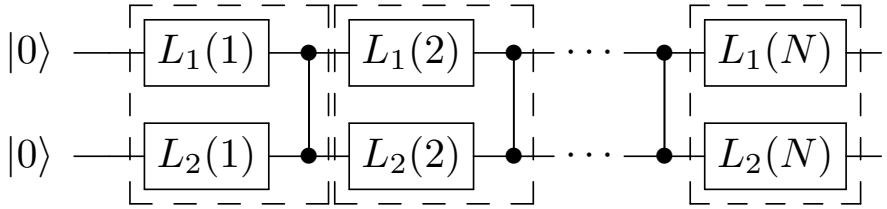
$$L(i) = U\left(\vec{\theta}_i^{(k)} + \vec{w}_i^{(k)} \circ \vec{x}^{(k)}\right) \cdots U\left(\vec{\theta}_i^{(1)} + \vec{w}_i^{(1)} \circ \vec{x}^{(1)}\right)$$

$$U(\vec{\phi}) = U(\phi_1, \phi_2, \phi_3) = e^{i\phi_2\sigma_z} e^{i\phi_1\sigma_y} e^{i\phi_3\sigma_z} \quad \text{or} \quad U(\vec{\phi}) = e^{i\vec{\omega}(\vec{\phi})\cdot\vec{\sigma}}$$

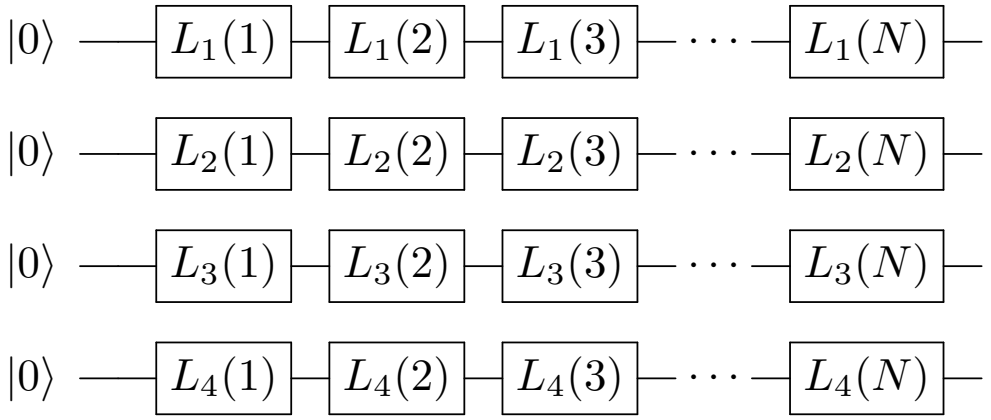
Multi-qubit using data re-uploading



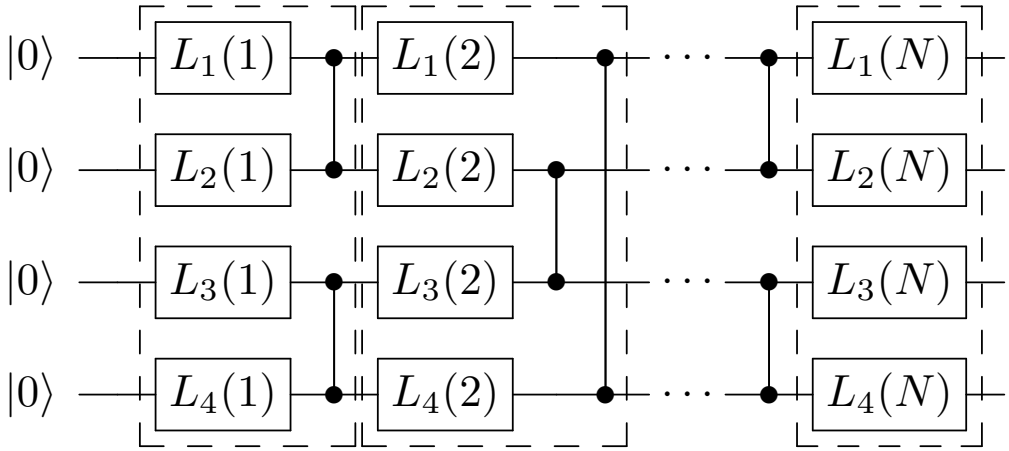
(a) Ansatz with no entanglement



(b) Ansatz with entanglement



(a) Ansatz with no entanglement



(b) Ansatz with entanglement

Single qubit classifier: measurements

- The quantum circuit characterized by a series of processing angle $\{\theta_i\}$ and weights $\{w_i\}$ delivers a final state $|\psi\rangle$. $|\psi\rangle = \mathcal{U}(\vec{\phi}, \vec{x})|0\rangle$
- The critical point in the quantum measurement is to find an optimal way to associate outputs from the observations to target classes.
- This is easily established for a dichotomic classification, where one of two classes A and B have to be assigned to the final measurement of the single qubit.
- In such a case it is possible to measure the output probabilities $P(0)$ for $|0\rangle$ and $P(1)$ for $|1\rangle$. A given pattern could be classified into the A class if $P(0) > P(1)$ and into B otherwise.
- We may refine this criterium by introducing a bias. That is, the pattern is classified as A if $P(0) > \lambda$, and as B otherwise. The λ is chosen to optimize the success of classification on a training set.
- The assignment of classes to the output reading of a single qubit becomes an involved issue when many classes are present.
    - For example, one possible strategy consists on comparing the probability $P(0)$ to four sectors with three thresholds: $0 \leq \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq 1$. Then, the value of $P(0)$ will fall into one of them, and classification is issued.

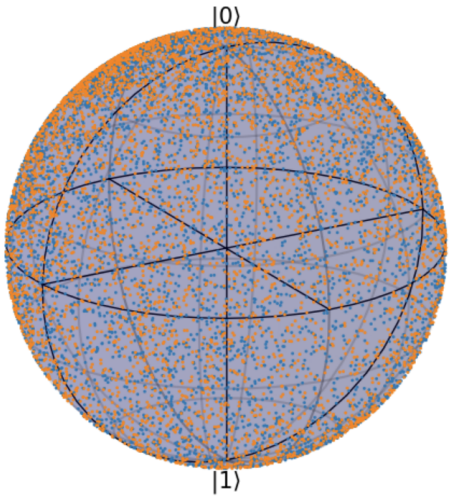
1907.02085

Single qubit classifier: cost function

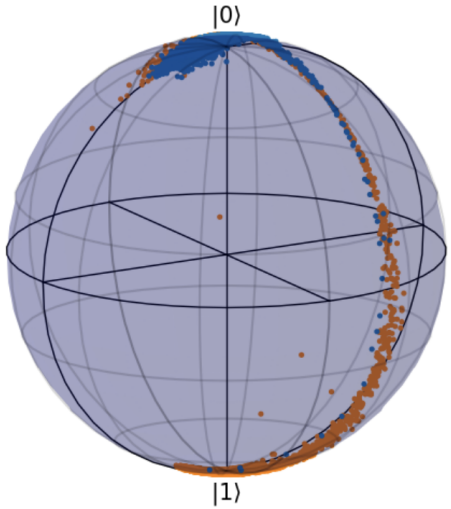
- A fidelity cost function: (fidelity is a measure of similarity of two things.)
  - We want to force the quantum state (data state) $|\psi(\vec{\theta}, \vec{w}, \vec{x})\rangle$ to be as near as possible to one particular state (label state) on the Bloch sphere.
  - The angular distance between the label state and the data state can be measured with the relative fidelity between the two states.
  - Goal is to maximize the average fidelity
  - $\chi_f^2(\vec{\theta}, \vec{w}) = \sum_{\mu=1}^{M} \left(1 - \left|\langle\tilde{\psi}_s|\psi(\vec{\theta}, \vec{w}, \vec{x}_\mu)\right|^2\right)$ where $|\tilde{\psi}_s\rangle$ is the correct label state of the data points. (M = total number of training data)

Easy choice: $|0\rangle$ and $|1\rangle$

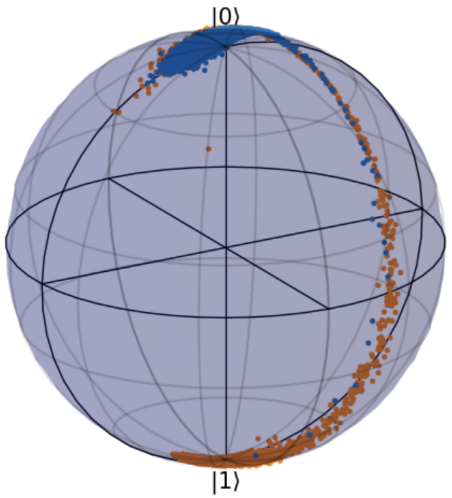
1907.02085



Before training

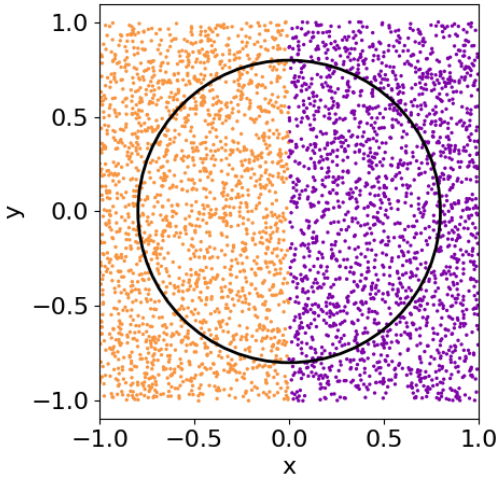


After training

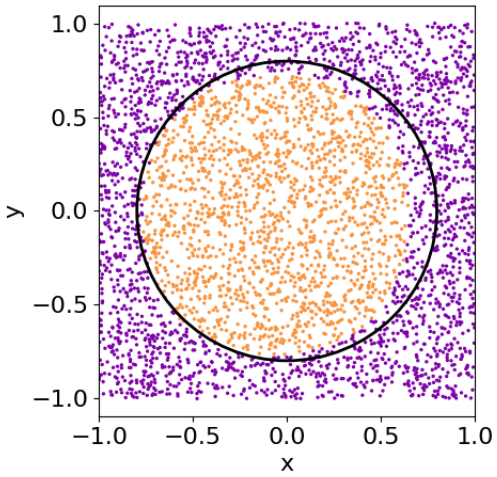


With testing data

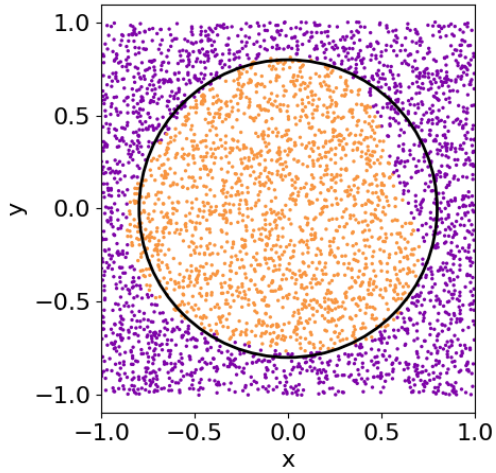
Example: binary classification



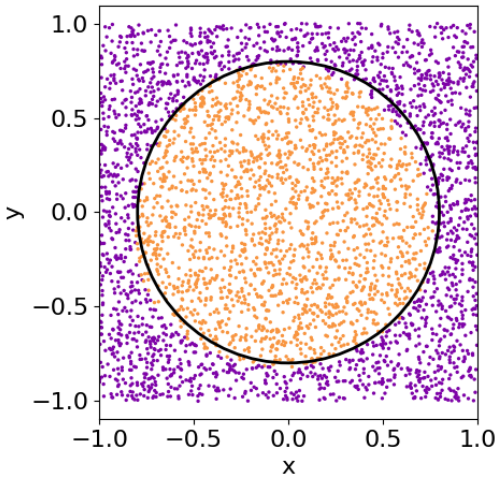
(a) 1 layer



(b) 2 layers



(c) 4 layers

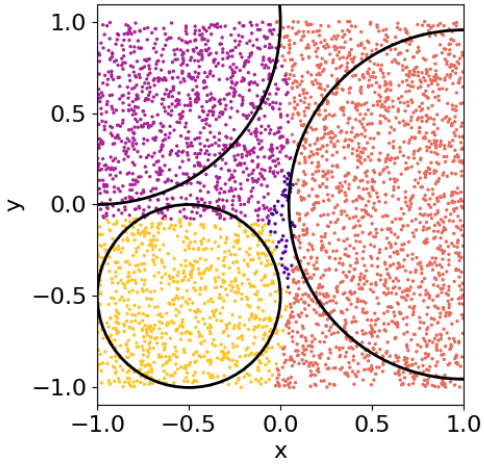


(d) 8 layers

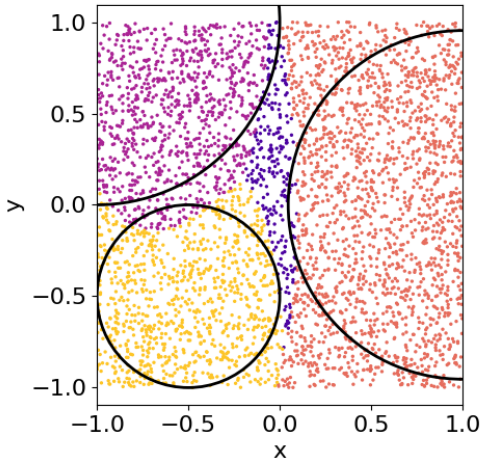
	χ_f^2		
Qubits	1	2	
Layers		No Ent.	Ent.
1	0.50	0.75	–
2	0.85	0.80	0.73
3	0.85	0.81	0.93
4	0.90	0.87	0.87
5	0.89	0.90	0.93
6	0.92	0.92	0.90
8	0.93	0.93	0.96
10	0.95	0.94	0.96

1907.02085

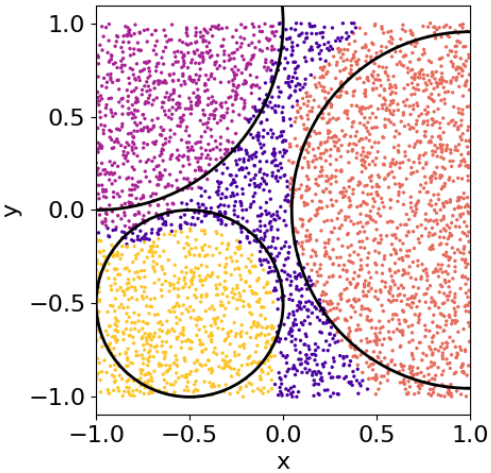
Example: 4 classes



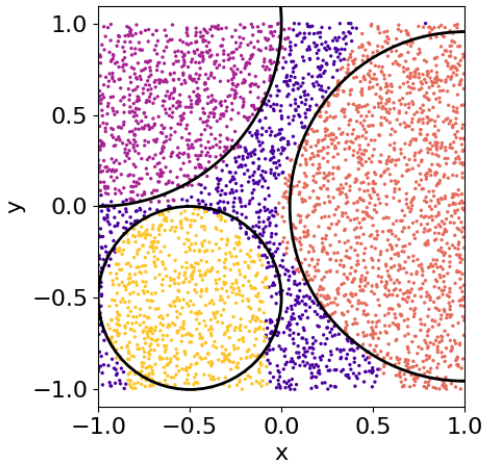
(a) 1 layer



(b) 3 layers



(c) 4 layers

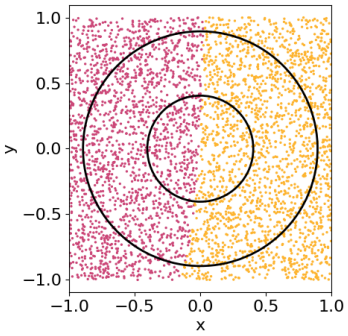


(d) 10 layers

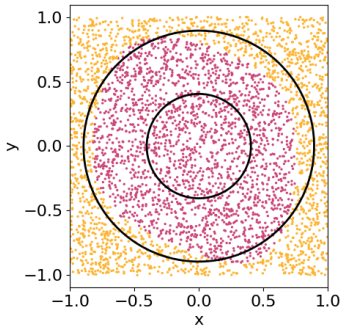
	χ_f^2		
Qubits	1	2	
Layers		No Ent.	Ent.
1	0.73	0.56	–
2	0.79	0.77	0.78
3	0.79	0.76	0.75
4	0.84	0.80	0.80
5	0.87	0.84	0.81
6	0.90	0.88	0.86
8	0.89	0.85	0.89
10	0.91	0.86	0.90

1907.02085

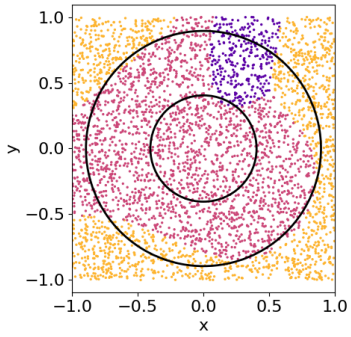
Example: 3 classes



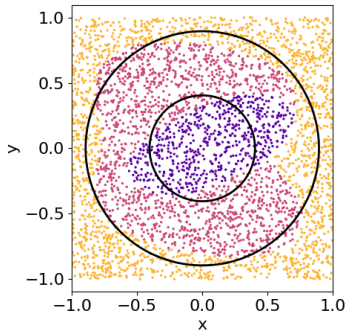
(a) 1 layer



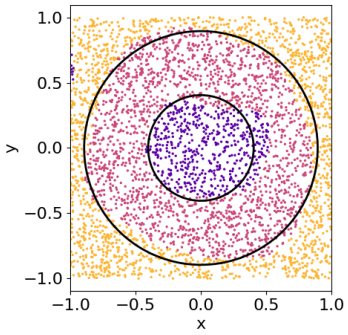
(b) 2 layers



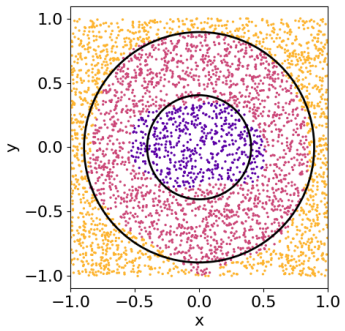
(c) 3 layers



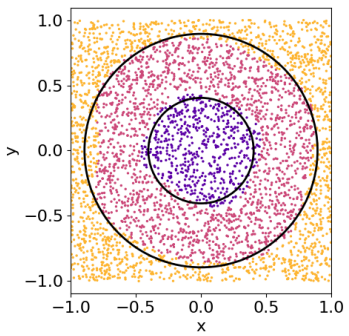
(d) 4 layers



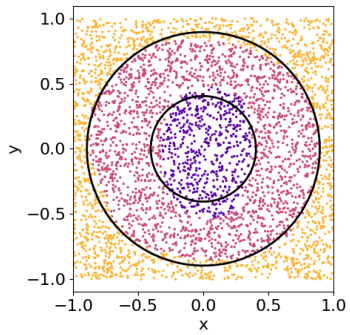
(e) 5 layers



(f) 6 layers



(g) 8 layers



(h) 10 layers

	χ^2_f		
Qubits	1	2	
Layers		No Ent.	Ent.
1	0.34	0.51	–
2	0.57	0.63	0.59
3	0.80	0.68	0.65
4	0.84	0.78	0.89
5	0.92	0.86	0.82
6	0.93	0.91	0.93
8	0.90	0.89	0.90
10	0.90	0.91	0.92

1907.02085

Single qubit classifier: example

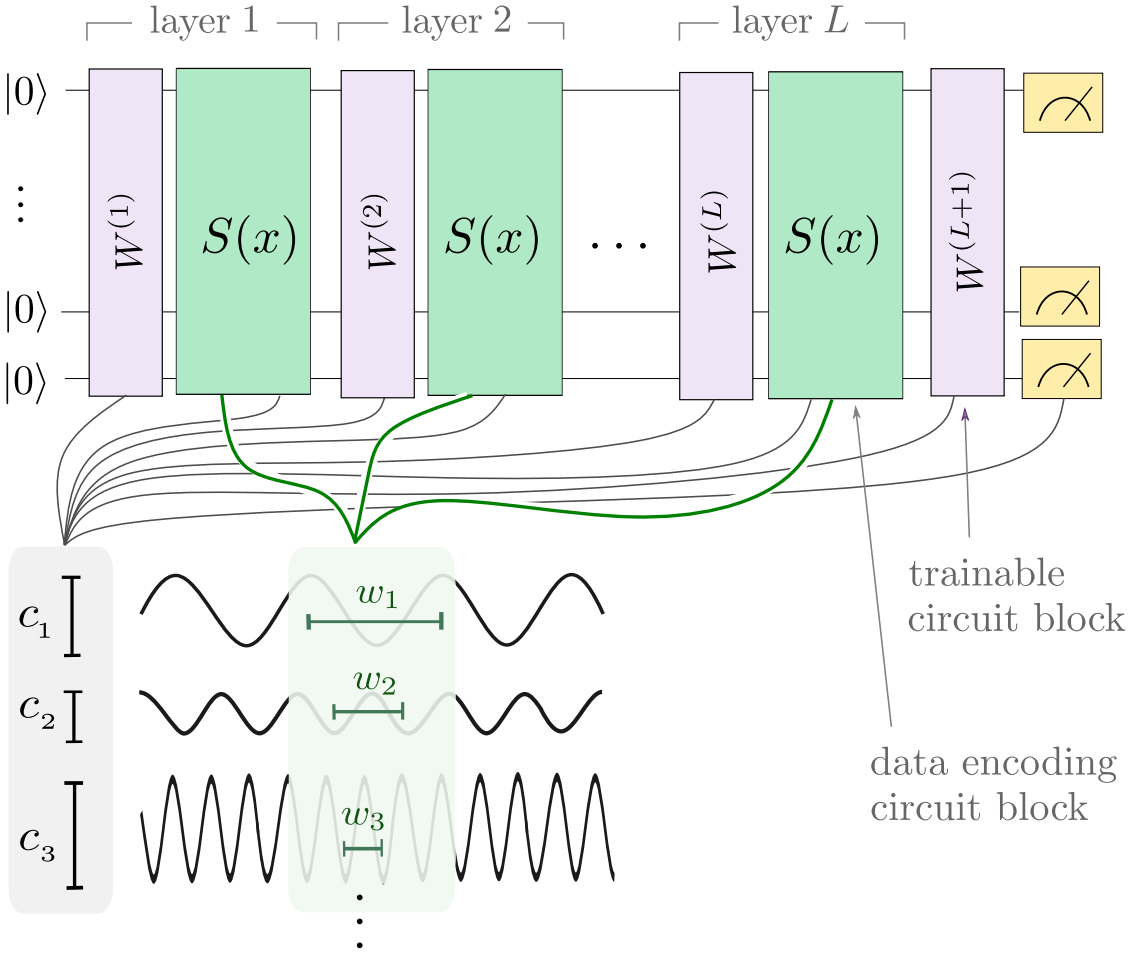
Problem	Classical classifiers NN	SVC	Quantum classifier χ^2_f	χ^2_{wf}
Circle	0.96	0.97	0.96	0.97
3 circles	0.88	0.66	0.91	0.91
Hypersphere	0.98	0.95	0.91	0.98
Annulus	0.96	0.77	0.93	0.97
Non-Convex	0.99	0.77	0.96	0.98
Binary annulus	0.94	0.79	0.95	0.97
Sphere	0.97	0.95	0.93	0.96
Squares	0.98	0.96	0.99	0.95
Wavy Lines	0.95	0.82	0.93	0.94

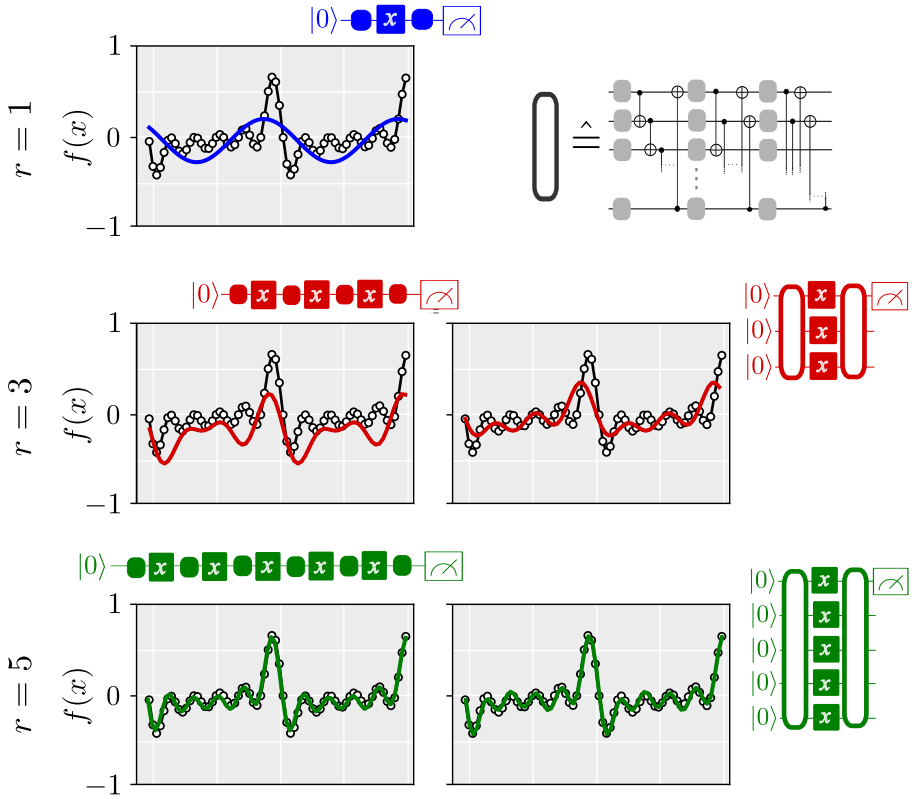
1907.02085

Comparison between single-qubit quantum classifier and two well-known classical classification techniques: a neural network (NN) with a single hidden layer composed of 100 neurons and a support vector classifier (SVC), both with the default parameters as defined in scikit-learn python package. This table shows the best success rate, being 1 the perfect classification, obtained after running ten times the NN and SVC algorithms and the best results obtained with single-qubit classifiers up to 10 layers.

The effect of data encoding on the expressive power of variational quantum machine learning models

- VQA as a Fourier series





$$f_{\boldsymbol{\theta}}(\boldsymbol{x}) = \sum_{\boldsymbol{n}\in\Omega} c_{\boldsymbol{n}}(\boldsymbol{\theta}) e^{i\boldsymbol{n}\boldsymbol{x}}$$

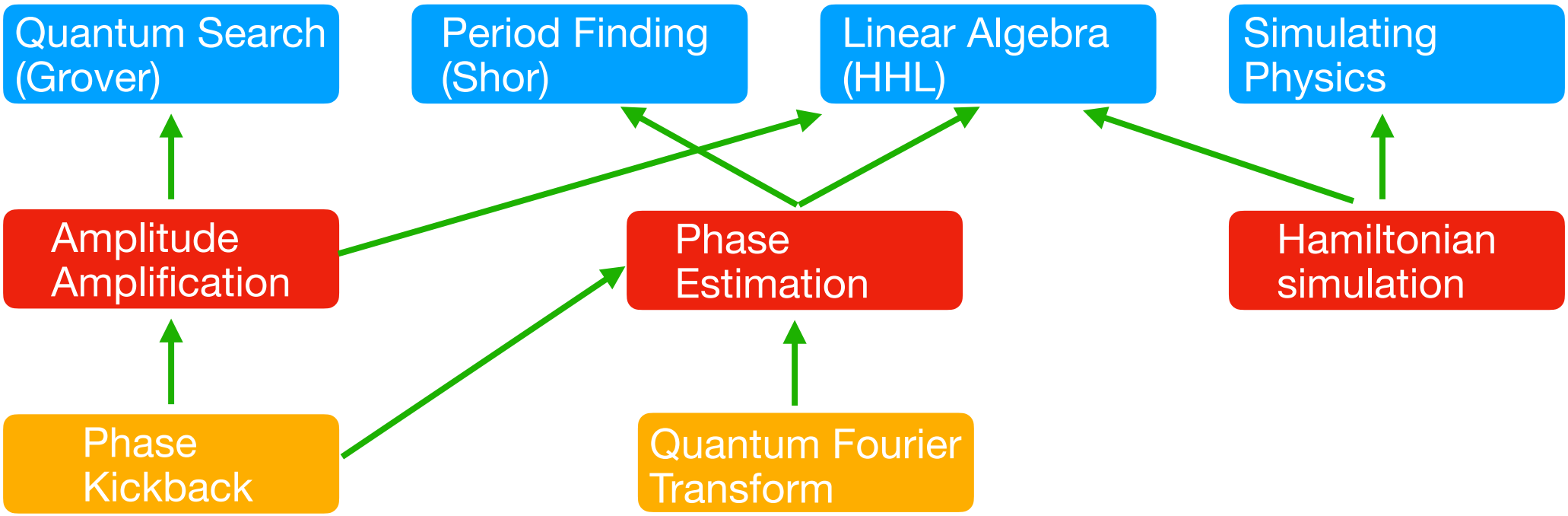
$$f_{\boldsymbol{\theta}}(x) = \langle 0| U^{\dagger}(x,\boldsymbol{\theta}) M U(x,\boldsymbol{\theta}) |0\rangle$$

$$U(x) = W^{(L+1)} S(x) W^{(L)} \dots W^{(2)} S(x) W^{(1)}$$

2008.08605

Quantum Algorithms

Algorithms, Tools, Tricks



Inspired by: https://cnls.lanl.gov/~baertschi/slides/Slides-QCA.pdf

The exponential Speed-Up

Description	Classical Algorithm	Quantum Algorithm	Reference
• **Grover's algorithm:** ✓ searches an unstructured database (or an unordered list) with N entries	$O(N)$	$O(\sqrt{N})$	L.K. Grover, "A fast quantum mechanical algorithm for database search." Proceedings of the twenty-eighth annual ACM symposium on Theory of computing. ACM (1996).
• **Shor's algorithm:** ✓ Integer factorization	$O\left(e^{1.9(\log N)^{\frac{1}{3}}(\log\log N)^{\frac{2}{3}}}\right)$	$O\left((\log N)^2(\log\log N)(\log\log\log N)\right)$	D. Beckman, et al. "Efficient networks for quantum factoring." Physical Review A 54.2 (1996).
• **Quantum Fourier Transform:** ✓ Discrete Fourier transform of size *N*	$O(N \log N)$	$O(\log N \ \log\log N)$	P. Shor, "Algorithms for quantum computation: discrete logarithms and factoring," Proc. 35th Annual Symp. on Foundations of Comp. Sci., pp.124-134(1994)
• **Eigenvalue solver:** ✓ To find eigenvalues and eigenvectors of a local Hamiltonian	$O(N^{2\sim 2.236})$	$O\left((logN)^4\right)$	Abrams and S. Lloyd, "Quantum Algorithm Providing Exponential Speed Increase for Finding Eigenvalues and Eigenvectorsg." Phys. Rev. Lett. 83, 24, p.5162 (1999).
• **Matrix inversion:** ✓ Finding inverse matrix. This can be applied to find a vector x satisfying Ax=b, where A and b are Hermitian N × N matrix and a unit vector, respectively	$O(Ns\kappa(\log 1/\epsilon))$	$\widetilde{O}\left(\frac{\log N\, s^2\kappa^2}{\epsilon}\right)$	A.W. Harrow, Avinatan Hassidim, and S. Lloyd. "Quantum algorithm for linear systems of equations." Physical review letters 103.15 (2009): 150502.
• **Distance (inner product) evaluation:** ✓ Calculating inner product between a given N dimensional vector and each N dimensional vector of M samples	$O(MN)$	$O(\log NM\ )$	S. Lloyd, Masoud Mohseni, and Patrick Rebentrost. "Quantum algorithms for supervised and unsupervised machine learning." arXiv preprint arXiv:1307.0411 (2013).

The exponential Speed-Up

Learning Problem	Classical Algorithm	Quantum Algorithm	Reference
• **k-means problem:** ✓ Assigning M vectors to k clusters in a way that minimizes the average distance to the centroid of the cluster	$O(M^2N)$	$O(Mlog(MN))$	Lloyd, Seth, Masoud Mohseni, and Patrick Rebentrost. "Quantum algorithms for supervised and unsupervised machine learning." arXiv preprint arXiv:1307.0411 (2013).
• **Principle component analysis (PCA) problem:** ✓ To convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated	$O(d^2 \log R + d R^2)$	$O(R \log d)$	Lloyd, Seth, Masoud Mohseni, and Patrick Rebentrost. "Quantum principal component analysis." Nature Physics 10.9 (2014): 631-633.
• **Support vector machine (SVM) problem:** ✓ To classify data clusters with support vector learning	$O(NM)$	$O(log(NM))$	P. Rebentrost, M. Mohseni, S. Lloyd, "Quantum Support Vector Machine for Big Data Classification," PRL 113, p.130503, 2014.
• **Quantum Neural Network (QNN) problem:** ✓ Qubit (or Node) requirement for neural network machine learning	$O(ND)$	$O(\log(N))$	S. Gupta and R.K.P. Zia, "Quantum Neural Network," Journal of Computer and System Sciences 63, 355–383 (2001).
• **Classification Problem:** ✓ Instant measure of hamming distance among training vector data and query vector	$O(M^3)$	$O(1)$?	M. Schuld, M. Fingerhuth, and F. Petruccione, "Implementing a distance-based classifier with a quantum interference circuit," EPL, v119,n6, 60002,2017
• **Learning parity with noise (LPN) problem:** ✓ For given some samples $(x, f(x))$, estimating the function f computing the parity of bits at some fixed locations	**N queries in a noiseless channel**	**O(logN) queries in a noisy (depolarizing) channel**	A.W. Cross, S. Graeme, and J.A. Smolin. "Quantum learning robust against noise." Physical Review A 92.1 (2015): 012327.

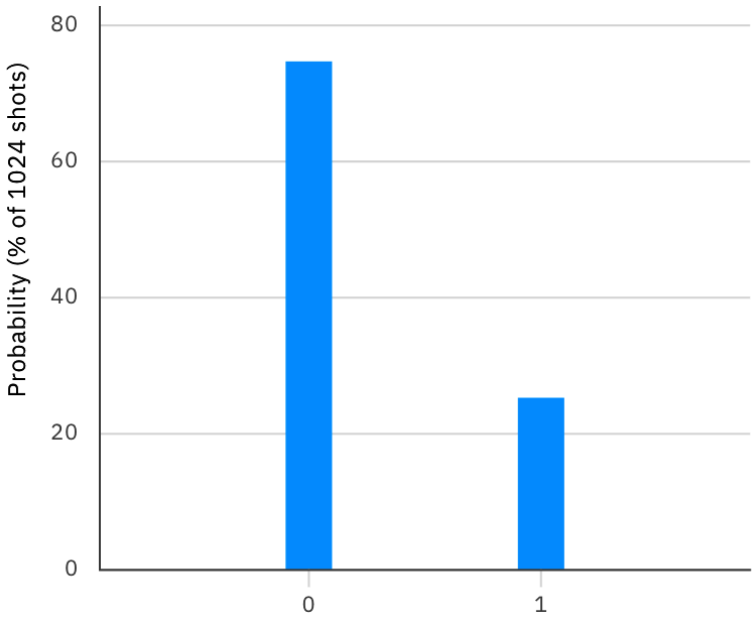
Inner product

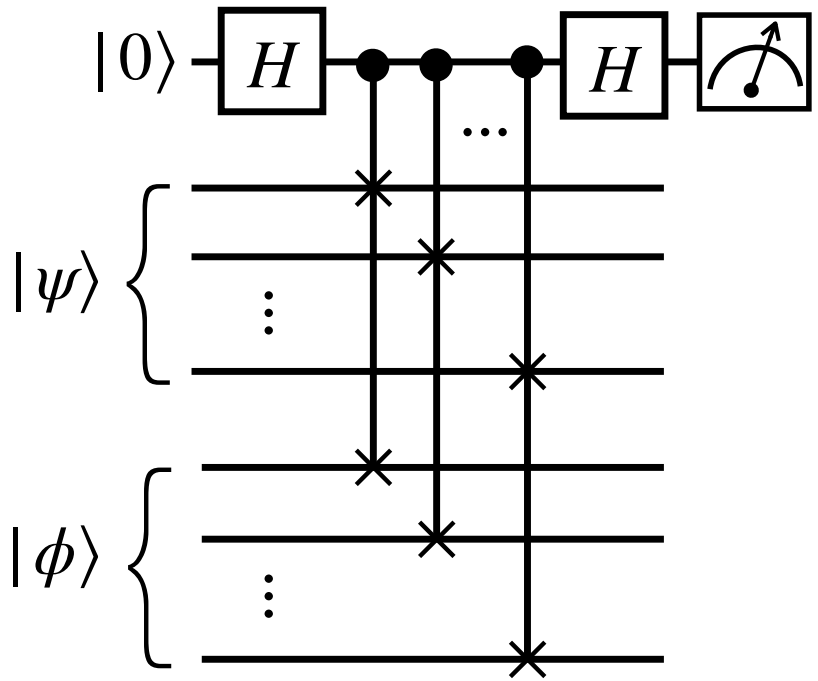
- Let $|\psi\rangle, |\phi\rangle \in \mathbb{C}^{2^n}$ be two $N = 2^n$ dimensional vectors. How to compute the magnitude of the inner product $|\langle\phi|\psi\rangle|^2$?
- Classical
  - $N = 2^n$ multiplications and additions
  - Decompose multiplications and additions as NAND gates
- Quantum
  - Run the following circuits with $2n + 1$ qubits and $n + 2$ gates
  - $\text{Prob}(0) - \text{Prob}(1) = |\langle\phi|\psi\rangle|^2$

$$|\psi\rangle = (\psi_1, \cdots, \psi_N)$$

$$|\phi\rangle = (\phi_1, \cdots, \phi_N)$$

$$\langle\phi|\psi\rangle = \sum_{i=1}^{N} \phi_i^* \psi_i$$

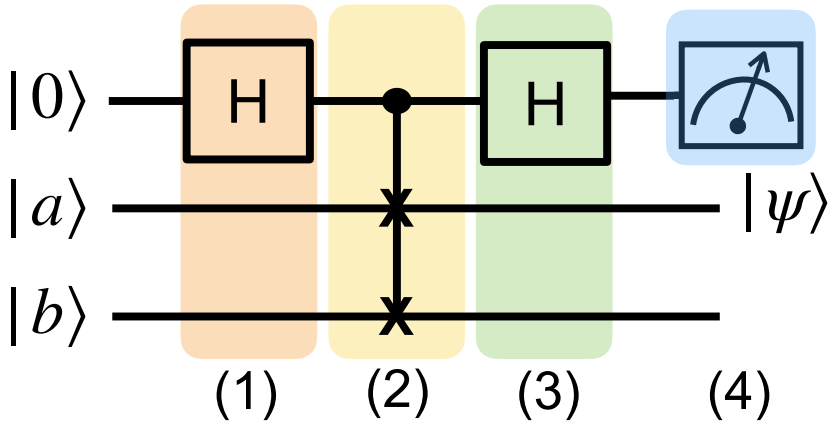




Proof: Inner product

$$|a\rangle \otimes |b\rangle = |a\rangle|b\rangle$$

$$H|x\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + (-1)^x|1\rangle\right)$$



(1) $$H\left(|0\rangle|a\rangle|b\rangle\right) = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|a\rangle|b\rangle$$

(2) $$SWAP\,H\left(|0\rangle|a\rangle|b\rangle\right) = \frac{1}{\sqrt{2}}SWAP(|0\rangle + |1\rangle)|a\rangle|b\rangle = \frac{1}{\sqrt{2}}(|0\rangle|a\rangle|b\rangle + |1\rangle|b\rangle|a\rangle)$$

(3) $$|\psi\rangle = HSWAP\,H\left(|0\rangle|a\rangle|b\rangle\right) = \frac{1}{\sqrt{2}}H(|0\rangle|a\rangle|b\rangle + |1\rangle|b\rangle|a\rangle)$$

$$= \frac{1}{2}|0\rangle\left(|a\rangle|b\rangle + |b\rangle|a\rangle\right) + \frac{1}{2}|1\rangle\left(|a\rangle|b\rangle - |b\rangle|a\rangle\right)$$

sum of the “unswapped” and the “swapped”

difference

(4) $$P(0) = |(\langle 0| \otimes I)|\psi\rangle|^2 = \frac{1}{2} - \frac{1}{2}|\langle a|b\rangle|^2$$

$$P(1) = |(\langle 0| \otimes I)|\psi\rangle|^2 = \frac{1}{2} + \frac{1}{2}|\langle a|b\rangle|^2$$

$$\Rightarrow \quad P(1) - P(0) = |\langle a|b\rangle|^2$$

Measuring Overlap of Quantum States

- Prepare two qubits registers: $|a\rangle \otimes |b\rangle = |a\rangle|b\rangle$ and want to compute $|\langle a|b\rangle|^2$

I. Swap test

(1) add ancilla qubit: $|0\rangle|a\rangle|b\rangle$

(2) apply Hadamard gate to ancilla: $\frac{1}{\sqrt{2}}\Big(|0\rangle + |1\rangle\Big)|a\rangle|b\rangle$

(3) apply swap operator on two registers $|a\rangle$ and $|b\rangle$ conditioned on the ancilla being in state 1: $\frac{1}{\sqrt{2}}\Big(|0\rangle|a\rangle|b\rangle + |1\rangle|b\rangle|a\rangle\Big)$

(4) apply Hadamard gate to ancilla:

$$|\psi\rangle = \frac{1}{2}\Big(|0\rangle + |1\rangle\Big)|a\rangle|b\rangle + \frac{1}{2}\Big(|0\rangle - |1\rangle\Big)|b\rangle|a\rangle$$

$$|\psi\rangle = \frac{1}{2}|0\rangle \otimes \Big(|a\rangle|b\rangle + |b\rangle|a\rangle\Big) + \frac{1}{2}|1\rangle \otimes \Big(|a\rangle|b\rangle - |b\rangle|a\rangle\Big)$$

sum of the "unswapped" and the "swapped" — difference

Measuring Overlap of Quantum States

I. Swap test

(5) $P_0 = P(\text{last qubit} = 0) = \frac{1}{2}\Big(\langle a|\langle b| + \langle b|\langle a|\Big)\frac{1}{2}\Big(|a\rangle|b\rangle + |b\rangle|a\rangle\Big)$

$$= \frac{1}{2} + \frac{1}{2}|\langle a|b\rangle|^2$$

if $\langle a|b\rangle = 0$, $P_0 = \frac{1}{2}$

if $\langle a|b\rangle = 1$, $P_0 = 1$

$\longrightarrow$ $|\langle a|b\rangle|^2 = 2P_0 - 1$

$\vec{a} = (a_1, a_2, \cdots, a_N) \in \mathbb{R}^N \longrightarrow |a\rangle$ $\longrightarrow |\langle a|b\rangle|^2 = |a^T b|^2$

$\vec{b} = (b_1, b_2, \cdots, b_N) \in \mathbb{R}^N \longrightarrow |b\rangle$ sign ambiguity?

Assume $|\vec{a}| = 1 = |\vec{b}|$

$$-1 \le \vec{a}\cdot\vec{b} \le 1$$

$$|a\rangle = \Big(\frac{1}{\sqrt{2}}a_1, \cdots \frac{1}{\sqrt{2}}a_N, \frac{1}{\sqrt{2}}\Big)$$

$$|b\rangle = \Big(\frac{1}{\sqrt{2}}b_1, \cdots \frac{1}{\sqrt{2}}b_N, \frac{1}{\sqrt{2}}\Big)$$

$$P_0 = \frac{1}{2} + \frac{1}{2}|\langle a|b\rangle|^2 = \frac{1}{2} + \frac{1}{2}\Big|\frac{1}{2}a_1b_1 + \cdots + \frac{1}{2}a_Nb_N + \frac{1}{2}\Big|^2$$

$$= \frac{1}{2} + \frac{1}{2}\Big|\frac{1}{2}a^Tb + \frac{1}{2}\Big|^2 = \frac{1}{2} + \frac{1}{2}\Big(\frac{1}{2}a^Tb + \frac{1}{2}\Big)^2$$

$\therefore$ $a^Tb = 2\sqrt{2P_0 - 1} - 1$ $\quad P_0 \in [\frac{1}{2}, 1]$ $\quad a^Tb \in [-1, 1]$

Measuring Overlap of Quantum States

II. Hadamard test

(1) Prepare initial state: $|\psi\rangle = \frac{1}{\sqrt{2}}\Big(|0\rangle|a\rangle + |1\rangle|b\rangle \Big)$

$\rightarrow$ similar to $\frac{1}{\sqrt{2}}(\alpha, \beta)$ for the representation of $|a\rangle$ and $|b\rangle$ rather than $\alpha \otimes \beta$

(2) apply Hadamard: $|\psi\rangle = \frac{1}{2}|0\rangle\Big(|a\rangle + |b\rangle \Big) + \frac{1}{2}\Big(|a\rangle - |b\rangle \Big)$

III. Inversion test $\quad |a\rangle = A|0\rangle \qquad |b\rangle = B|0\rangle$

Run the circuit $B^\dagger A|0\rangle$

Probability of observing the final quantum state in the initial state $|0\rangle$

$= \Big| \langle 0|B^\dagger A|0\rangle \Big|^2 = |\langle b|a\rangle|^2$

For many operators, $U = U^\dagger$ or $U^\dagger(\theta) = U(-\theta)$

The First Wave of Quantum Machine Learning?

PRL **103,** 150502 (2009) PHYSICAL REVIEW LETTERS week ending 9 OCTOBER 2009

Quantum Algorithm for Linear Systems of Equations

Aram W. Harrow,[1] Avinatan Hassidim,[2] and Seth Lloyd[3]

[1]*Department of Mathematics, University of Bristol, Bristol, BS8 1TW, United Kingdom*
[2]*Research Laboratory for Electronics, MIT, Cambridge, Massachusetts 02139, USA*
[3]*Research Laboratory for Electronics and Department of Mechanical Engineering, MIT, Cambridge, Massachusetts 02139, USA*

(Received 5 July 2009; published 7 October 2009)

Solving linear systems of equations is a common problem that arises both on its own and as a subroutine in more complex problems: given a matrix A and a vector $\vec{b}$, find a vector $\vec{x}$ such that $A\vec{x} = \vec{b}$. We consider the case where one does not need to know the solution $\vec{x}$ itself, but rather an approximation of the expectation value of some operator associated with $\vec{x}$, e.g., $\vec{x}^\dagger M\vec{x}$ for some matrix M. In this case, when A is sparse, $N \times N$ and has condition number κ, the fastest known classical algorithms can find $\vec{x}$ and estimate $\vec{x}^\dagger M\vec{x}$ in time scaling roughly as $N\sqrt{\kappa}$. Here, we exhibit a quantum algorithm for estimating $\vec{x}^\dagger M\vec{x}$ whose runtime is a polynomial of $\log(N)$ and κ. Indeed, for small values of κ [i.e., poly $\log(N)$], we prove (using some common complexity-theoretic assumptions) that any classical algorithm for this problem generically requires exponentially more time than our quantum algorithm.

$$Ax = b$$

Complexity of inversion of a regular matrix=$O(N^3)$

Complexity of inversion of a sparse matrix=$O(N)$

Harrow-Hassidim-Lloyd Algorithm

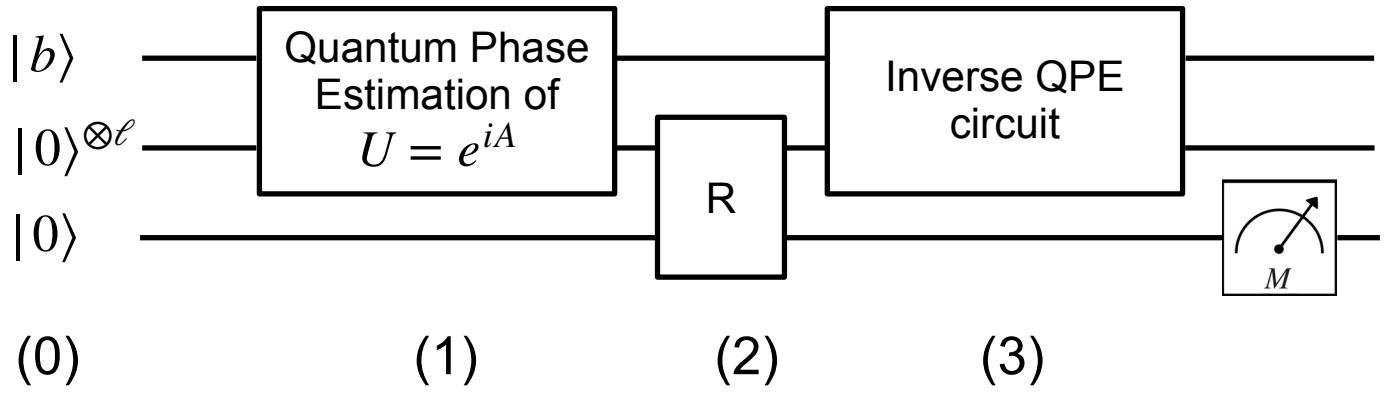
- An algorithm that solves exponentially large linear system in polynomial time. Given a matrix $A \in \mathbb{R}^{N\times M}$ and a vector $b \in \mathbb{R}^N$, find $x \in \mathbb{R}^M$ such that $Ax = b$.
- Let us assume $N = M$ and A is invertible $\rightarrow x = A^{-1}b \rightarrow$ problem is reduced to inverting A.

$$\begin{aligned} x &\longleftrightarrow |x\rangle \\ b &\longleftrightarrow |b\rangle \end{aligned} \qquad x = A^{-1}b \quad \Rightarrow \quad |x\rangle = \frac{A^{-1}|b\rangle}{\left|A^{-1}|b\rangle\right|}$$

- Additional assumptions
  - A is hermitian (and squared and invertible), $A^\dagger = A$.
  - For eigenvalues $0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N \leq 1$, $\min_j |\lambda_j| \geq \alpha > 0$ and $\max_j |\lambda_j| \leq 1$.
  - Algorithm has access to many copies of $|b\rangle$.
  - Can implement unitary transformation e^{iAt} for any value of t such that $0 \leq t \leq \text{polylog}\, N$.

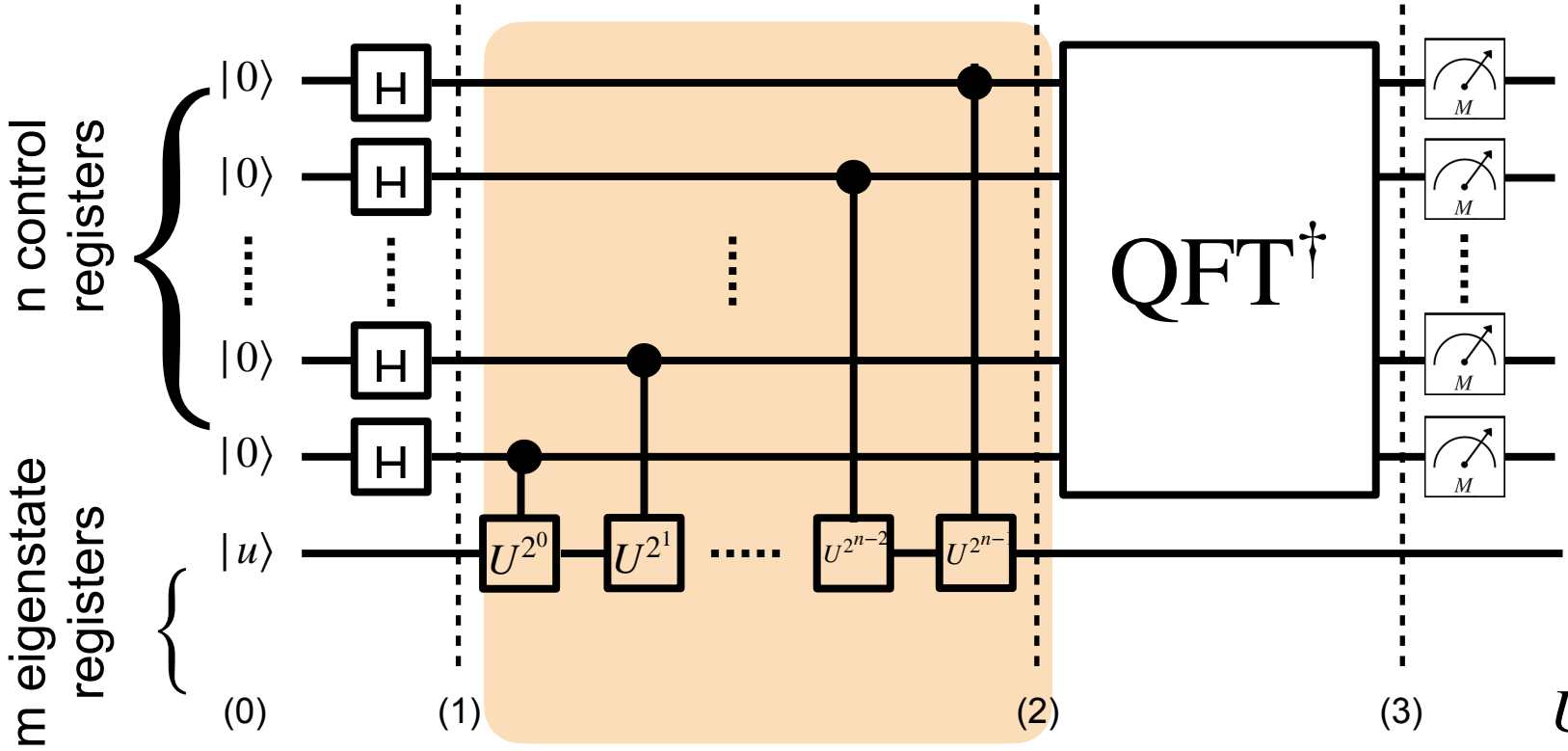
some polynomial in log(N) = $(\log N)^k = \log^k N$

Harrow-Hassidim-Lloyd Algorithm



- $|u_j\rangle = |j\rangle$: eigenvectors of A with eigenvalues $0 < \alpha \leq \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_N \leq 1$
- $\{|u_j\rangle\}$ forms orthonormal basis, $\langle i|j\rangle = \delta_{ij}$
- $A = \sum_j \lambda_j |j\rangle\langle j|$, $U = e^{iA} = \sum_j e^{i\lambda_j} |j\rangle\langle j|$ and $A^{-1} = \sum_j \frac{1}{\lambda_j} |j\rangle\langle j|$
- $|b\rangle = \sum_{j=1}^{n} b_j |j\rangle$ with $\sum_j |b_j|^2 = 1$ (Algorithm has access to many copies of $|b\rangle$.)

Quantum Circuit for QPE



$$U|u\rangle = e^{i\phi}|u\rangle$$

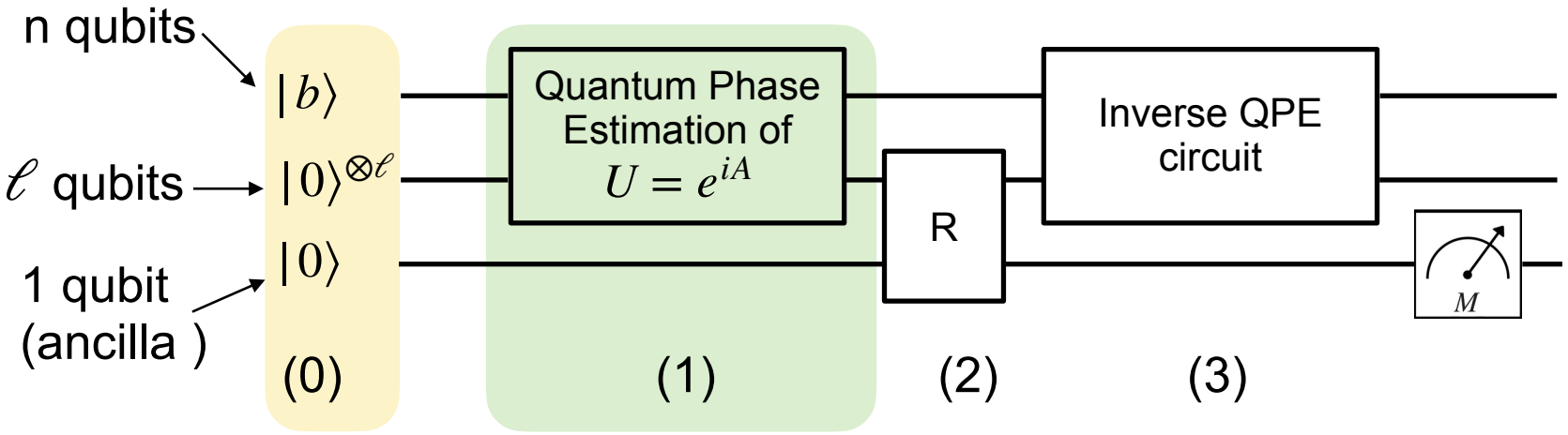
$$|\psi_2\rangle = \prod_{j=0}^{n-1} \mathrm{CU}^{2^j} \frac{1}{\sqrt{2}^n}\Big(|0\rangle + |1\rangle\Big)^{\otimes n} \otimes |u\rangle$$

$$U^{2j}|u\rangle = \left(e^{i\phi}\right)^{2^j}|u\rangle = e^{i\phi\, 2^j}|u\rangle$$

$$\frac{1}{\sqrt{2}}\Big(|0\rangle + |1\rangle\Big) \otimes |u\rangle \xrightarrow{\mathrm{CU}^{2^j}} \frac{1}{\sqrt{2}}\Big(|0\rangle \otimes |u\rangle + U^{2^j}|1\rangle \otimes |u\rangle\Big)$$

$$= \frac{1}{\sqrt{2}}\Big(|0\rangle + e^{i\phi\, 2^j}|1\rangle\Big) \otimes |u\rangle$$

Harrow-Hassidim-Lloyd Algorithm



$|\psi_0\rangle = |b\rangle \otimes |0\rangle^{\otimes \ell} \otimes |0\rangle$

$QPE = H + \text{conditional } U^{2^k} + \text{QFT}^\dagger$

$U = e^{2\pi i A}$

Target qubits

$$|b\rangle \otimes |0\rangle^{\otimes \ell} \xrightarrow{H} |b\rangle \otimes \frac{1}{\sqrt{2^\ell}}\Big(|0\rangle + |1\rangle\Big)^{\otimes \ell} \xrightarrow{\text{CU}} \prod_{k=0}^{\ell-1} \text{CU}^{2^k}\Big(\sum_{j=1}^{n} b_j |j\rangle\Big) \otimes \frac{1}{\sqrt{2^\ell}}\Big(|0\rangle + |1\rangle\Big)^{\otimes \ell}$$

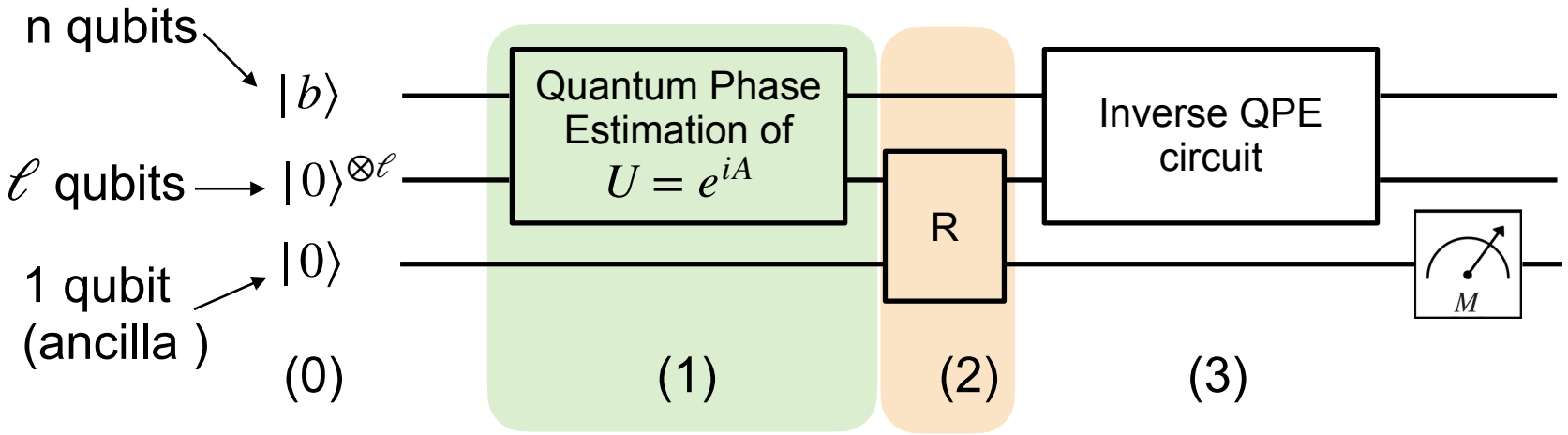
Control qubits

$$= \sum_{j=1}^{n} b_j \prod_{k=0}^{\ell-1} \text{CU}^{2^k} |j\rangle \otimes \frac{1}{\sqrt{2^\ell}}\Big(|0\rangle + |1\rangle\Big)^{\otimes \ell}$$

$$= \sum_{j=1}^{n} b_j |j\rangle \otimes \frac{1}{\sqrt{2^\ell}} \Big(|0\rangle + |1\rangle e^{2\pi i \lambda_j 2^0}\Big) \otimes \Big(|0\rangle + |1\rangle e^{2\pi i \lambda_j 2^1}\Big) \otimes \cdots \otimes \Big(|0\rangle + |1\rangle e^{2\pi i \lambda_j 2^{\ell-1}}\Big)$$

$$= \sum_{j=1}^{n} b_j |j\rangle \otimes \frac{1}{\sqrt{2^\ell}} \sum_{y=0}^{2^\ell-1} e^{2\pi i \lambda_j y} |y\rangle \xrightarrow{\text{QFT}^{-1}} \cdots\cdots$$

Harrow-Hassidim-Lloyd Algorithm



$$QPE = H + \text{conditional } U^{2^k} + \text{QFT}^\dagger$$

$$\lambda_j = x/2^\ell$$

$$\cdots\cdots = \sum_{j=1}^{n} b_j |j\rangle \otimes \frac{1}{\sqrt{2^\ell}} \sum_{y=0}^{2^\ell-1} e^{2\pi i \lambda_j y} |y\rangle \xrightarrow{\text{QFT}^{-1}} \sum_{j=1}^{n} b_j \sum_{y=0}^{2^\ell-1} \frac{1}{\sqrt{2^\ell}} \frac{1}{\sqrt{2^\ell}} \sum_{x=0}^{2^\ell-1} e^{2\pi i y(\lambda_j - x/2^\ell)} |j\rangle \otimes |x\rangle$$

$$= \sum_{j=1}^{n} b_j |j\rangle \otimes |\lambda_j\rangle$$

ℓ-bit string = ℓ qubits approximation of λ_j

$$\frac{1}{2^\ell} \sum_{y=0}^{2^\ell-1} e^{2\pi i y(a-x)/2^\ell} = \delta_{ax}$$

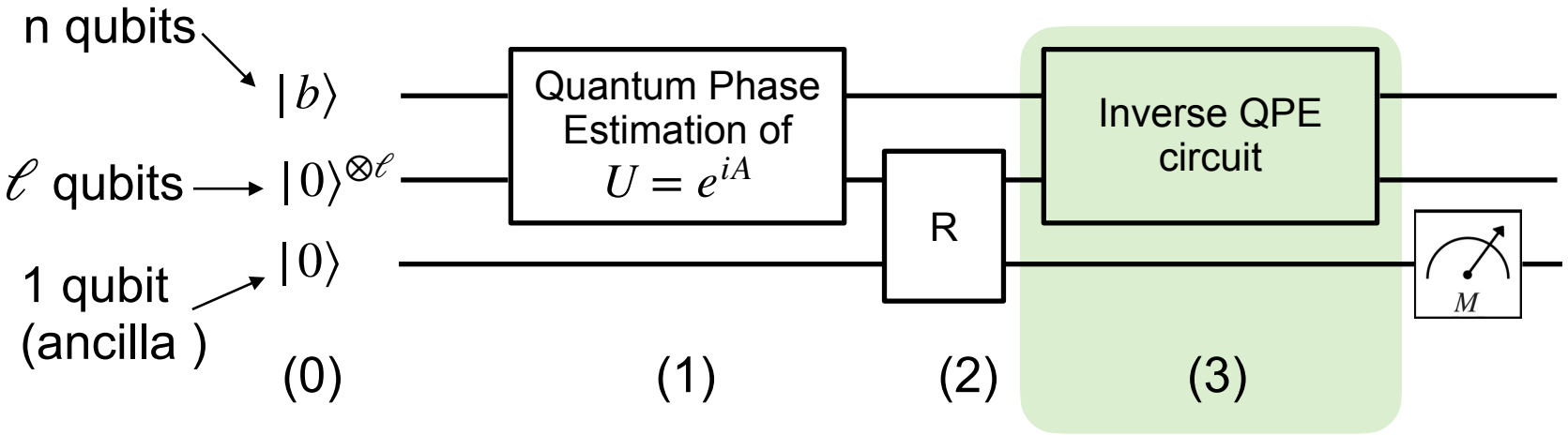
$$|\psi_1\rangle = \text{QPE}\,|\psi_0\rangle = \sum_{j=1}^{n} b_j |j\rangle \otimes |\lambda_j\rangle \otimes |0\rangle$$

$$\text{QFT}^{-1}|y\rangle = \frac{1}{\sqrt{2^\ell}} \sum_{x=0}^{2^\ell-1} e^{-2\pi i x y/2^\ell} |x\rangle$$

$$\xrightarrow{\text{R}} \quad |\psi_2\rangle = \sum_{j=1}^{n} b_j |j\rangle \otimes |\lambda_j\rangle \otimes \left[\sqrt{1 - \frac{\alpha^2}{\lambda_j^2}}\,|0\rangle + \frac{\alpha}{\lambda_j}|1\rangle\right]$$

$$0 < \alpha \le \min|\lambda_j| \le \max|\lambda_j| \le 1$$

Harrow-Hassidim-Lloyd Algorithm



$$\xrightarrow{\text{R}} \quad |\psi_2\rangle = \sum_{j=1}^{n} b_j |j\rangle \otimes |\lambda_j\rangle \otimes \left[\sqrt{1-\frac{\alpha^2}{\lambda_j^2}}\,|0\rangle + \frac{\alpha}{\lambda_j}|1\rangle\right] \qquad 0 < \alpha \le \min|\lambda_j| \le \max|\lambda_j| \le 1$$

$$\xrightarrow{\text{QPE}^{-1}} \quad |\psi_3\rangle = \sum_{j=1}^{n} b_j |j\rangle \otimes |0\rangle^{\otimes \ell} \otimes \left[\sqrt{1-\frac{\alpha^2}{\lambda_j^2}}\,|0\rangle + \frac{\alpha}{\lambda_j}|1\rangle\right]$$

"uncomputing"

$$|\psi_3\rangle = \left(\sum_{j=1}^{n} b_j \sqrt{1-\frac{\alpha^2}{\lambda_j^2}}\,|j\rangle \otimes |0\rangle^{\otimes \ell}\right) \otimes |0\rangle + \left(\sum_{j=1}^{n} b_j \frac{\alpha}{\lambda_j}|j\rangle \otimes |0\rangle^{\otimes \ell}\right) \otimes |1\rangle$$

$$P(|1\rangle) = \left\| \sum_j b_j \frac{\alpha}{\lambda_j}|j\rangle \otimes |0\rangle^{\otimes \ell} \right\|^2 = \sum_j \left| b_j \frac{\alpha}{\lambda_j} \right|^2 = \sum_j \left| b_j \right|^2 \left| \frac{\alpha}{\lambda_j} \right|^2 \ge \sum_j \left| b_j \right|^2 |\alpha|^2 = |\alpha|^2 > 0$$

$$0 < \alpha \le \min|\lambda_j| \le \max|\lambda_j| \le 1$$

Harrow-Hassidim-Lloyd Algorithm

$$P(|1\rangle) = \left\| \sum_j b_j \frac{\alpha}{\lambda_j} |j\rangle \otimes |0\rangle^{\otimes \ell} \right\|^2 = \sum_j \left| b_j \frac{\alpha}{\lambda_j} \right|^2 = \sum_j \left| b_j \right|^2 \left| \frac{\alpha}{\lambda_j} \right|^2 \geq \sum_j \left| b_j \right|^2 |\alpha|^2 = |\alpha|^2 > 0$$

$$0 < \alpha \leq \min |\lambda_j| \leq \max |\lambda_j| \leq 1$$

- If we measure $|0\rangle$, repeat the measurement.
- How many times do we repeat? $\rightarrow O\left(\frac{1}{\alpha^2}\right)$ $\rightarrow$ amplitude application

  $\rightarrow O\left(\frac{1}{\alpha}\right)$ calls of the algorithm for having a probability of success $p \sim 1$.

- Post-measurement state conditioned on $|1\rangle = \frac{1}{\sqrt{p}} \sum_j b_j \frac{\alpha}{\lambda_j} |j\rangle \times |0\rangle^{\otimes \ell} \otimes |1\rangle$

$$= \frac{\alpha}{\sqrt{p}} \left(\sum_j \frac{b_j}{\lambda_j} |j\rangle \right) \times |0\rangle^{\otimes \ell} \otimes |1\rangle$$

$$= \frac{\alpha}{\sqrt{p}} \left(\underbrace{A^{-1}|b\rangle}_{|x\rangle} \right) \times |0\rangle^{\otimes \ell} \otimes |1\rangle$$

$$|\psi_3\rangle = \left(\sum_{j=1}^{n} b_j \sqrt{1 - \frac{\alpha^2}{\lambda_j^2}} |j\rangle \otimes |0\rangle^{\otimes \ell} \right) \otimes |0\rangle$$

$$+ \left(\sum_{j=1}^{n} b_j \frac{\alpha}{\lambda_j} |j\rangle \otimes |0\rangle^{\otimes \ell} \right) \otimes |1\rangle$$

HHL Algorithm: revisit assumptions

- What if $A \in \mathbb{R}^{N\times M}$? $\quad Ax = b \quad A \in \mathbb{R}^{N\times M} \quad x \in \mathbb{R}^{M} \quad b \in \mathbb{R}^{N}$

$$\tilde{A} = \begin{pmatrix} 0 & A \\ A^{\dagger} & 0 \end{pmatrix} \in \mathbb{R}^{(N+M)\times(N+M)}$$

any solution $\tilde{x}$ to $\tilde{A}\tilde{x} = \tilde{b}$ will be such that $\tilde{x} = (x, 0, \cdots 0)^{T}$ satisfies $Ax = b$.

- Bounds on eigenvalues
  - Why can we assume $\lambda_{\max}(A) \leq 1$? If not, work with $\tilde{A} = \dfrac{A}{\lambda_{\max}(A)}$ and $\tilde{b} = \dfrac{b}{\lambda_{\max}(A)}$
  - What about $\lambda_{\min}(A) \geq \alpha > 0$? This condition is known as "A is a well-conditioned matrix" $\rightarrow$ usually needed in classical algorithms to avoid issues with non-convertible matrix because $\lambda_{\min}$ is zero or very close to zero,
- $|\psi(t)\rangle = e^{-iAt}|\psi(0)\rangle$ and $i\dfrac{d}{dt}|\psi(t)\rangle = A|\psi(t)\rangle$ need to simulate dynamics of Hamiltonian.
  - Can we efficiently implement e^{-iAt} for a range of t.
  - If A is sparse, i.e., each row of A contains at most $s \ll N$ entries, each entry can be captured in $O(s)$ time.

HHL Algorithm: 4 qubit example

$A = \begin{pmatrix} 1 & -1/3 \\ -1/3 & 1 \end{pmatrix}$ $\quad |b\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $\quad n_b = 1$ to store $|b\rangle$ and $|x\rangle$ $\qquad A|x\rangle = |b\rangle$

$\lambda_1 = \frac{2}{3}$ $\quad |u_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $\qquad n_\ell = 2$ to store binary approximation of eigenvalues

$\lambda_2 = \frac{4}{3}$ $\quad |u_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ $\qquad n = 1$ auxiliary (ancilla) qubit for conditional rotation

- QPE will output an n_ℓ bit (2 bit) binary approximation to $\frac{\lambda_j t}{2\pi}$.
- If we set $t = 2\pi\frac{3}{8}$, QPE will give 2-bit binary approximation to $\frac{\lambda_1 t}{2\pi} = \frac{1}{4}$ and $\frac{\lambda_2 t}{2\pi} = \frac{1}{2}$, which are $|01\rangle$ and $|10\rangle$, respectively.
  - $01 = 0 \cdot 2^{-1} + 1 \cdot 2^{-2} = 1/4$
  - $10 = 1 \cdot 2^{-1} + 0 \cdot 2^{-2} = 1/2$

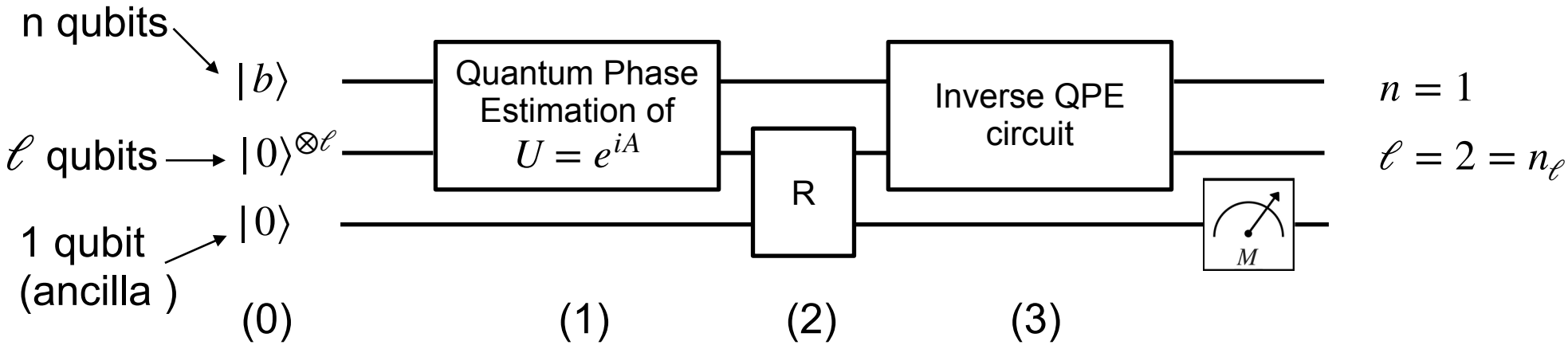
$|01\rangle = |\lambda_1\rangle$

$|11\rangle = |\lambda_2\rangle$

$$|b\rangle = \sum_{j=1}^{2} b_j |u_j\rangle = \frac{1}{\sqrt{2}} \sum_{j=1}^{2} |u_j\rangle = \frac{1}{\sqrt{2}}\Big(|u_1\rangle + |u_2\rangle\Big) = |b\rangle = \sum_{j=1}^{n} b_j |j\rangle$$

$$b_1 = \frac{1}{\sqrt{2}} = b_2$$

HHL Algorithm: 4 qubit example



(0) $|\psi_0\rangle = |b\rangle \otimes |0\rangle^{\otimes\ell} \otimes |0\rangle$ $\quad |b\rangle = \frac{1}{\sqrt{2}}\Big(|u_1\rangle + |u_2\rangle\Big)$ $\quad \lambda_1 = \frac{2}{3}$ $\quad |u_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}$

(1) $|\psi_1\rangle = \mathrm{QPE}\,|\psi_0\rangle = \sum_{j=1}^{\ell} b_j |j\rangle \otimes |\lambda_j\rangle \otimes |0\rangle$ $\quad \lambda_2 = \frac{4}{3}$ $\quad |u_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$

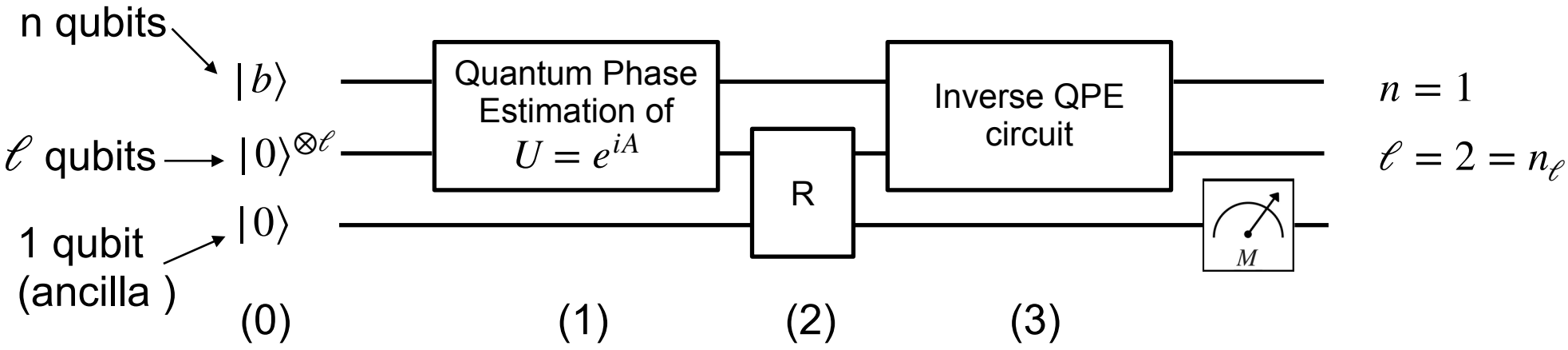
$= \frac{1}{\sqrt{2}}\Big(|u_1\rangle \otimes |01\rangle + |u_2\rangle \otimes |10\rangle\Big) \otimes |0\rangle$ $\quad b_1 = \frac{1}{\sqrt{2}} = b_2$

(2) conditional rotation with $\alpha = \frac{1}{8} \le \frac{\lambda_1 t}{2\pi} = \frac{1}{4} \le \frac{\lambda_2 t}{2\pi} = \frac{1}{2} \le 1$

$$|\psi_2\rangle = \frac{1}{\sqrt{2}}|u_1\rangle|01\rangle\left[\sqrt{1-\left(\frac{1/8}{1/4}\right)^2}|0\rangle + \left(\frac{1/8}{1/4}\right)|1\rangle+\right] + \frac{1}{\sqrt{2}}|u_2\rangle|10\rangle\left[\sqrt{1-\left(\frac{1/8}{1/2}\right)^2}|0\rangle + \left(\frac{1/8}{1/2}\right)|1\rangle+\right]$$

$$|\psi_2\rangle = \frac{1}{\sqrt{2}}|u_1\rangle|01\rangle\left[\sqrt{1-\frac{1}{4}}|0\rangle + \frac{1}{2}|1\rangle+\right] + \frac{1}{\sqrt{2}}|u_2\rangle|10\rangle\left[\sqrt{1-\frac{1}{16}}|0\rangle + \frac{1}{4}|1\rangle+\right]$$

HHL Algorithm: 4 qubit example



(2) $$|\psi_2\rangle = \frac{1}{\sqrt{2}}|u_1\rangle|01\rangle\left[\sqrt{1-\frac{1}{4}}|0\rangle+\frac{1}{2}|1\rangle\right]+\frac{1}{\sqrt{2}}|u_2\rangle|10\rangle\left[\sqrt{1-\frac{1}{16}}|0\rangle+\frac{1}{4}|1\rangle\right]$$

(3) $$|\psi_3\rangle = \mathrm{QPE}^\dagger|\psi_2\rangle = \frac{1}{\sqrt{2}}|u_1\rangle|00\rangle\left[\sqrt{1-\frac{1}{4}}|0\rangle+\frac{1}{2}|1\rangle\right]+\frac{1}{\sqrt{2}}|u_2\rangle|00\rangle\left[\sqrt{1-\frac{1}{16}}|0\rangle+\frac{1}{4}|1\rangle\right]$$

(4) on outcome 1 when measuring the ancilla qubit, the state is

$$\frac{\frac{1}{\sqrt{2}}|u_1\rangle|00\rangle\frac{1}{2}|1\rangle+\frac{1}{\sqrt{2}}|u_2\rangle|00\rangle\frac{1}{4}|1\rangle}{\sqrt{5/32}} = \frac{\left(\frac{1}{2\sqrt{2}}|u_1\rangle+\frac{1}{4\sqrt{2}}|u_2\rangle\right)|00\rangle|1\rangle}{\sqrt{5/32}} = \frac{|x\rangle}{\||x\rangle\|}$$

(5) how to compute the norm of $|x\rangle$ → probability of measuring 1 in the ancilla qubit

$$P(|1\rangle) = \left(\frac{1}{2\sqrt{2}}\right)^2+\left(\frac{1}{4\sqrt{2}}\right)^2 = \frac{5}{32} = \left\||x\rangle\right\|^2$$

Quantum Error Correction

- quant-ph/9705052, Stabilizer codes and quantum error correction, Caltech PhD thesis by D. Gottesman

- https://github.com/qiskit-community/qiskit-community-tutorials/blob/master/awards/teach_me_quantum_2018/intro2qc/10.Quantum error correction.ipynb

Simple Classical (Bitflip) Error Correction

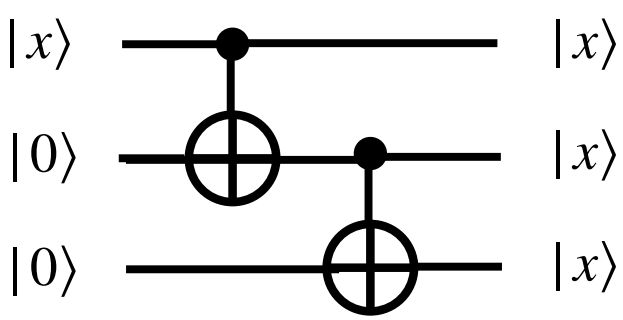
- Classically error correction is not necessary
  - Hardware for one bit is huge on an atomic scale
  - State 0 and 1 are so different that the probability of an unwanted flip is tiny.
- Error correction is needed for transmitting signal over long distance where it attenuates and can be corrupted by noise.
- Suppose we send one bit through a channel.
- Use redundancy: $|0\rangle \longrightarrow |000\rangle$, $|1\rangle \longrightarrow |111\rangle$ called codewords
- Apply majority rule: $\{000,001,010,100\} \rightarrow 0$, $\{111,110,101,011\} \rightarrow 1$
- Flip probability is p: $p^3 + 3(1-p)p^2 = 3p^2 - 2p^3 \leq p, \text{ if } p < 1/2$

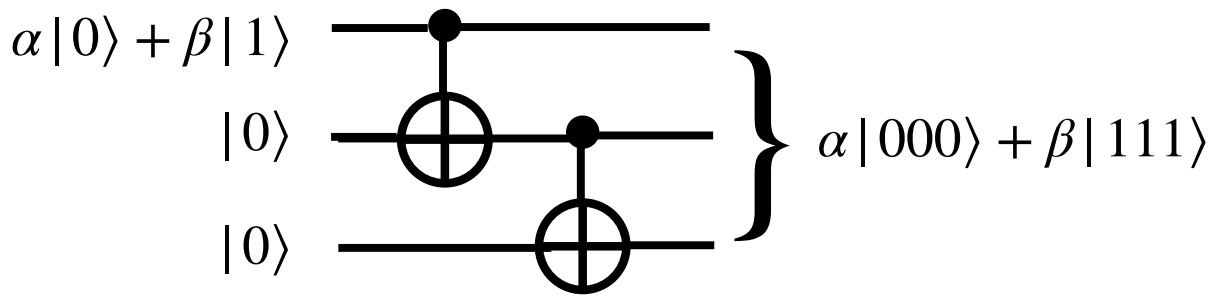
Quantum Error Correction

- QEC is essential and QC requires error correction
  - Physical system for a single qubit is small (often on an atomic scale) so any small external interference can disrupt the quantum system
- Measurement destroys quantum information
  - Checking for error is problematic.
  - Monitoring means measuring which would alter quantum states
- More general types of error can occur
  - (ex) phase error: $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \longrightarrow \frac{1}{\sqrt{2}}(|0\rangle + e^{i\phi}|1\rangle)$
- Errors are continuous
  - Unlike all or nothing bit flip errors for classical bits, errors on qubits can grow continuously out of the uncorrupted state.

Bit Flip Error Correction

- If the error rate is low, we hope to correct them by tailing the number of qubits as the classical case.





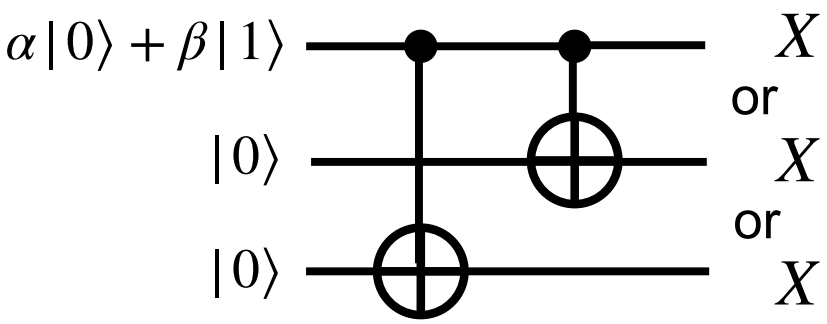
$\alpha|0\rangle+\beta|1\rangle \longrightarrow \alpha|000\rangle+\beta|111\rangle$ is not a clone of the input state

$$\left(\alpha|0\rangle+\beta|1\rangle\right)^{\otimes 3} = \alpha^3|000\rangle+\alpha^2\beta(|001\rangle+|010\rangle+|100\rangle)$$

$$+\alpha\beta^2(|110\rangle+|101\rangle+|011\rangle)+\beta^3|111\rangle$$

Bit Flip Error Correction

- Assume that no more than one qubit is flipped (reasonable approximation if the error rate is small)



$$|\psi\rangle = \alpha|000\rangle + \beta|111\rangle$$

$$|\psi_1\rangle = \alpha|100\rangle + \beta|011\rangle = X_1|\psi\rangle \quad \text{qubit 1 flipped}$$

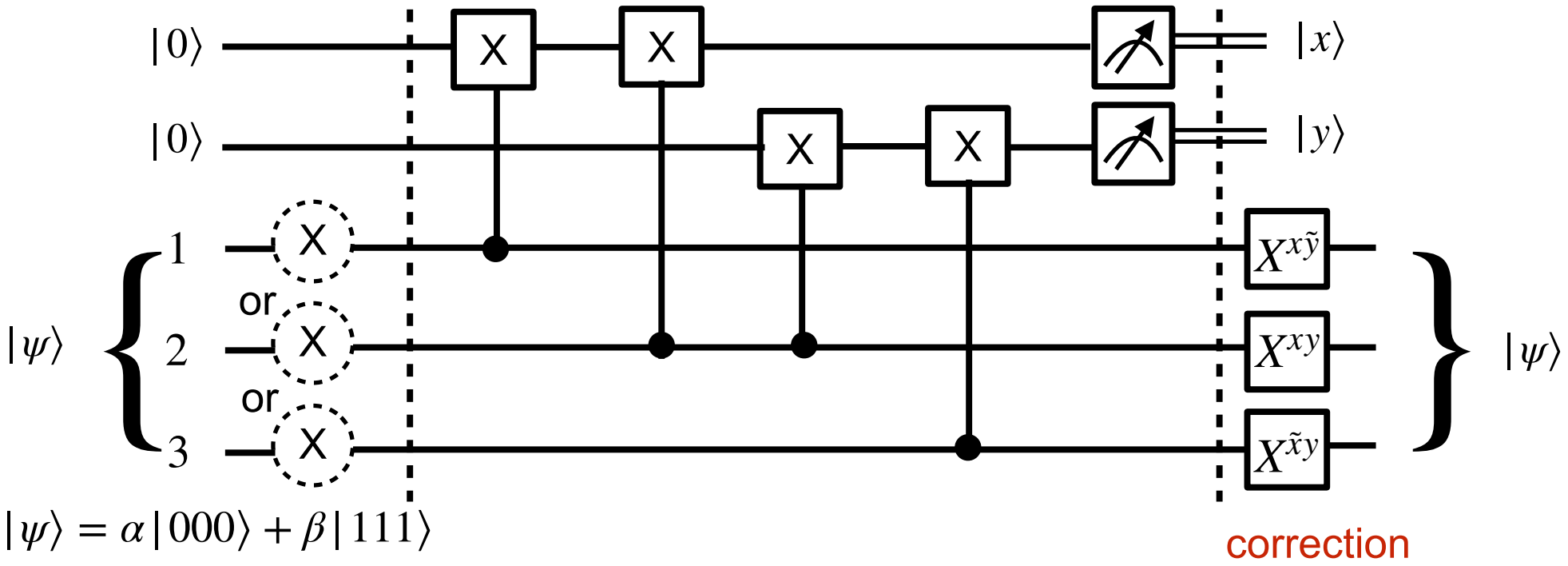
$$|\psi_2\rangle = \alpha|010\rangle + \beta|101\rangle = X_2|\psi\rangle \quad \text{qubit 2 flipped}$$

$$|\psi_3\rangle = \alpha|001\rangle + \beta|110\rangle = X_3|\psi\rangle \quad \text{qubit 3 flipped}$$

$\longrightarrow$ four states are called “syndromes”

- Classically to determine if one of the bits is flipped, we just have to look at them. However quantum mechanically, if we measure $|\psi\rangle$, we get $|000\rangle$ with probability $|\alpha|^2$ and $|111\rangle$ with $|\beta|^2$ which destroys the coherent superposition.
- Need to couple the codeword qubits to ancilla qubits and measure those, which does not destroy the coherent superposition.

Bit Flip Error Correction



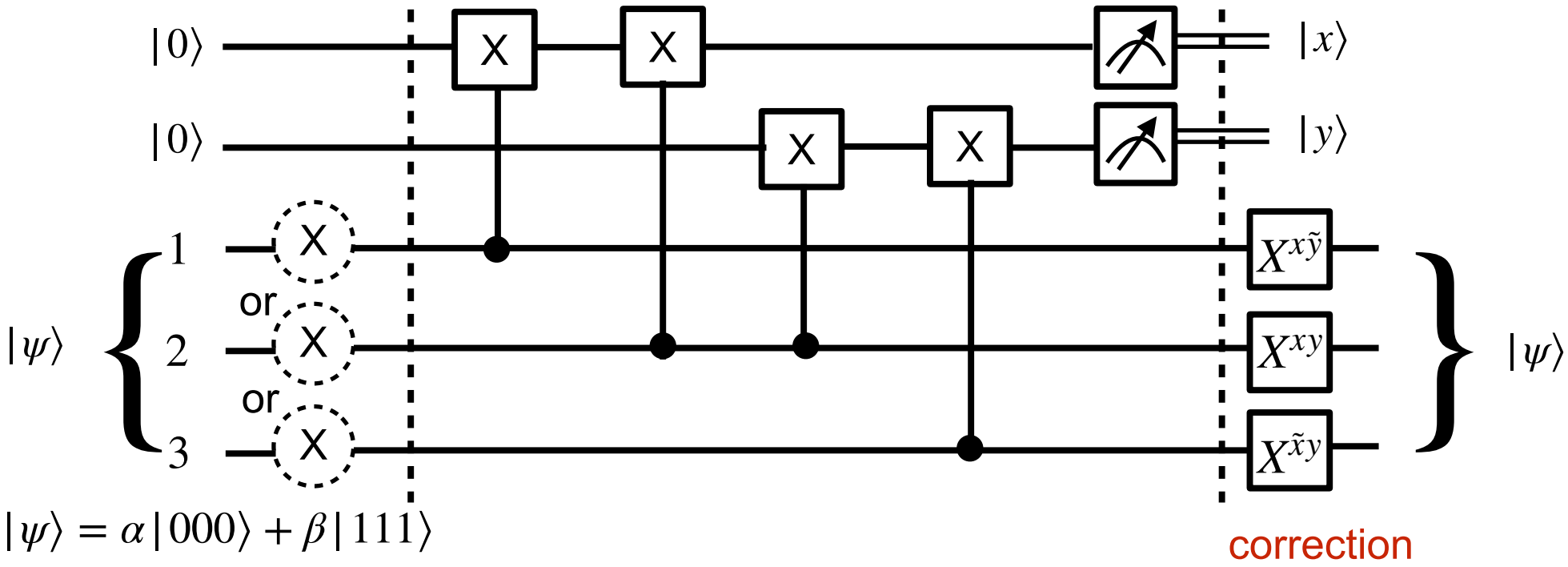
$|\psi\rangle$: codeword $|000\rangle \rightarrow$ no ancilla flipped $\rightarrow$ $x = 0 = y$
codeword $|111\rangle \rightarrow$ both ancillas flipped $\rightarrow$ $x = 0 = y$

$|\psi_1\rangle$: codeword $|100\rangle \rightarrow$ x flipped, y not flipped $\rightarrow$ $x = 1, y = 0$
codeword $|011\rangle \rightarrow$ x flipped, y flipped twice $\rightarrow$ $x = 1, y = 0$

$|\psi_2\rangle$: codeword $|010\rangle \rightarrow$ x and y flipped once $\rightarrow$ $x = 1 = y$
codeword $|101\rangle \rightarrow$ x and y flipped once $\rightarrow$ $x = 1 = y$

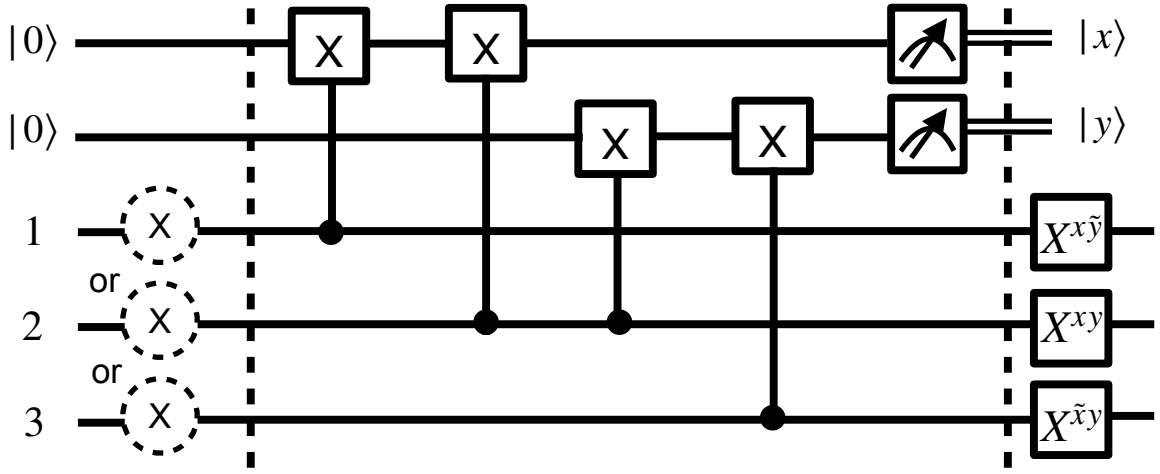
$|\psi_3\rangle$: codeword $|001\rangle \rightarrow$ x not flipped, y flipped $\rightarrow$ $x = 0, y = 1$
codeword $|110\rangle \rightarrow$ x flipped twice, y flipped $\rightarrow$ $x = 0, y = 1$

Bit Flip Error Correction



Syndromes	Bit flipped	x	y
$\|\psi\rangle = \alpha\|000\rangle + \beta\|111\rangle$	None	0	0
$\|\psi_1\rangle = \alpha\|100\rangle + \beta\|011\rangle$	1	1	0
$\|\psi_2\rangle = \alpha\|010\rangle + \beta\|101\rangle$	2	1	1
$\|\psi_3\rangle = \alpha\|001\rangle + \beta\|110\rangle$	3	0	1

Bit Flip Error Correction



correction

Syndromes	Bit flipped	x	y
$\|\psi\rangle = \alpha\|000\rangle + \beta\|111\rangle$	None	0	0
$\|\psi_1\rangle = \alpha\|100\rangle + \beta\|011\rangle$	1	1	0
$\|\psi_2\rangle = \alpha\|010\rangle + \beta\|101\rangle$	2	1	1
$\|\psi_3\rangle = \alpha\|001\rangle + \beta\|110\rangle$	3	0	1

$$|\psi\rangle = \alpha|000\rangle + \beta|111\rangle$$

$X^{x\tilde{y}}$ gate on qubit 1, only if x=1 and y=0 $\rightarrow$ correcting $|\psi_1\rangle$

$$x = 1,\ \tilde{y} = 1$$

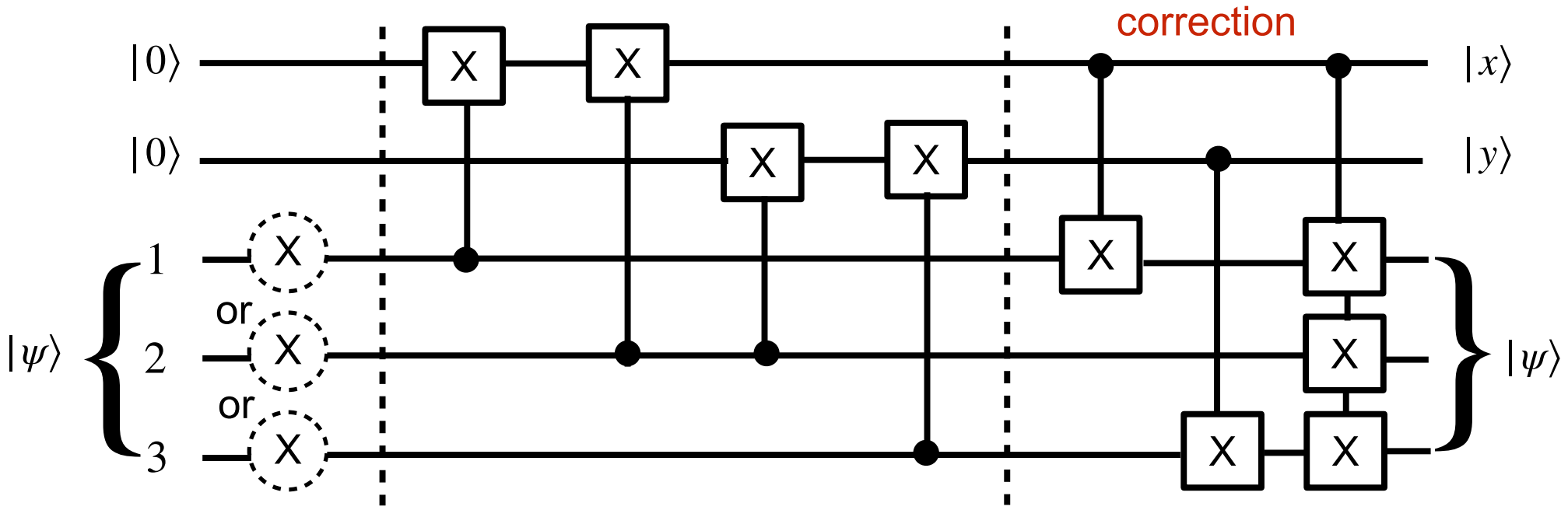
X^{xy} gate on qubit 2, only if x=1 and y=1 $\rightarrow$ correcting $|\psi_2\rangle$

$$x = 1,\ y = 1$$

$X^{\tilde{x}y}$ gate on qubit 3, only if x=0 and y=0 $\rightarrow$ correcting $|\psi_3\rangle$

$$\tilde{x} = 1,\ \tilde{y} = 1$$

Bit Flip Error Correction



$$|\psi\rangle = \alpha|000\rangle + \beta|111\rangle$$

$X^{x\tilde{y}}$ gate on qubit 1, only if x=1 and y=0 $\rightarrow$ correcting $|\psi_1\rangle$

X^{xy} gate on qubit 2, only if x=1 and y=1 $\rightarrow$ correcting $|\psi_2\rangle$

$X^{\tilde{x}y}$ gate on qubit 3, only if x=0 and y=0 $\rightarrow$ correcting $|\psi_3\rangle$

- What if errors in quantum circuits can arise continuously from zero? (Assume the error rate is small)

$$|\psi\rangle \longrightarrow \left[1 + \left(\epsilon_1 X_1 + \epsilon_2 X_2 + \epsilon_3 X_3\right)\right]|\psi\rangle \qquad \epsilon_i \in \mathbb{C},\ |\epsilon_i| \ll 1$$

Stabilizer Formalism

- Useful method for error correction of arbitrary error.
- Consider two Hermitian operators, Z_1Z_2 and Z_2Z_3

$$Z_i^2 = I_{2\times 2} \qquad Z_1Z_2 = Z_2Z_1 \qquad (Z_1Z_2)^2 = I_{2\times 2} \qquad (Z_2Z_3)^2 = I_{2\times 2}$$

$$\longrightarrow \quad A^2 = I_{2\times 2} \quad \longrightarrow \quad \text{eigenvalues} = \pm 1 \qquad Ax = \lambda x \qquad A^2x = \lambda^2 x = x \qquad \lambda^2 = 1$$

$$\longrightarrow \quad [\, Z_1Z_2, Z_2Z_3 \,] = 0 \qquad Z_1Z_3 \text{ and } Z_2Z_3 \text{ have the same eigenvectors.}$$

Syndromes	Z_1Z_2	Z_2Z_3	x	y
$\lvert\psi\rangle = \alpha\lvert 000\rangle + \beta\lvert 111\rangle$	1	1	0	0
$\lvert\psi_1\rangle = \alpha\lvert 100\rangle + \beta\lvert 011\rangle = X_1\lvert\psi\rangle$	-1	1	1	0
$\lvert\psi_2\rangle = \alpha\lvert 010\rangle + \beta\lvert 101\rangle = X_2\lvert\psi\rangle$	-1	-1	1	1
$\lvert\psi_3\rangle = \alpha\lvert 001\rangle + \beta\lvert 110\rangle = X_3\lvert\psi\rangle$	1	-1	0	1

$$Z_1Z_2 = (-1)^x$$

$$Z_2Z_3 = (-1)^y$$

- Syndromes are eigenvectors of Z_1Z_2 and Z_2Z_3.
- Stabilizers are operators whose eigenvalues distinguish the different syndromes.

Properties of Stabilizers and Syndromes

- Syndromes are eigenvectors of $Z_1 Z_2$ and $Z_2 Z_3$.
- Stabilizers are operators whose eigenvalues distinguish the different syndromes.
- Eigenvalues of a stabilizer in a syndrome is +1 or -1.
- Eigenvalues of all stabilizers are +1 in the uncorrupted syndrome $|\psi\rangle$.
- Operators for the stabilizers are built out of the single qubit operators Z_i and X_i.
- Syndromes with a single qubit error are obtained by acting on the uncorrupted syndrome with X_i, Y_i and Z_i operators.
- For a general stabilizer A_α and a syndrome state $|\psi_\beta\rangle = B_\beta |\psi\rangle$, A_α either commutes or anti-commutes with B_β.
  - B_β involves a single Pauli's operator (X, Y or Z).
  - A_α involves a product of Pauli's operators (X's, and Z's b/c $Y = iXZ$).

Properties of Stabilizers and Syndromes

- If $[A_\alpha, B_\beta] = 0,\ A_\alpha |\psi_\beta\rangle = +1 |\psi_\beta\rangle$ and eigenvalue of the stabilizer A_α in state $|\psi_\beta\rangle$ is +1.

  $$A_\alpha |\psi_\beta\rangle = A_\alpha B_\beta |\psi\rangle = B_\beta A_\alpha |\psi\rangle = B_\beta |\psi\rangle = |\psi_\beta\rangle$$

- If $\{A_\alpha, B_\beta\} = 0,\ A_\alpha |\psi_\beta\rangle = -1 |\psi_\beta\rangle$

  $$A_\alpha |\psi_\beta\rangle = A_\alpha B_\beta |\psi\rangle = -B_\beta A_\alpha |\psi\rangle = -B_\beta |\psi\rangle = -|\psi_\beta\rangle$$

- Syndromes must be eigenvectors of all stabilizers → stabilizers must commute each other
- How to determine efficiently if a stabilizer commutes or anti-commutes with the operator which generates a corrupted syndrome out of the uncorrupted syndrome?
- For the case of 3-qubit bit-flip code, stabilizers are $Z_1 Z_2$ and $Z_2 Z_3$.
- Operators which generate the corrupted syndromes from the uncorrupted syndrome: X_1, X_2 and X_3.

Properties of Stabilizers and Syndromes

- How to determine efficiently if a stabilizer commutes or anti-commutes with the operator which generates a corrupted syndrome out of the uncorrupted syndrome?
- For the case of 3-qubit bit-flip code, stabilizers are Z_1Z_2 and Z_2Z_3.
- Operators which generate the corrupted syndromes from the uncorrupted syndrome: X_1, X_2 and X_3.
  - X_1 commutes with $Z_2Z_3 \longleftrightarrow [X_1, Z_2Z_3] = 0$. $\because$ no sites in common $\rightarrow \; Z_2Z_3 |\psi_1\rangle = +1 |\psi_1\rangle$
  - X_2 has one common site with Z_2Z_3.
    $\rightarrow \; X_2Z_2Z_3 = -Z_2X_2Z_3 = -Z_2Z_3X_2$
    $\rightarrow \; \{X_2, Z_2Z_3\} = 0 \rightarrow \; Z_2Z_3 |\psi_2\rangle = -|\psi_2\rangle$

Stabilizer Formalism

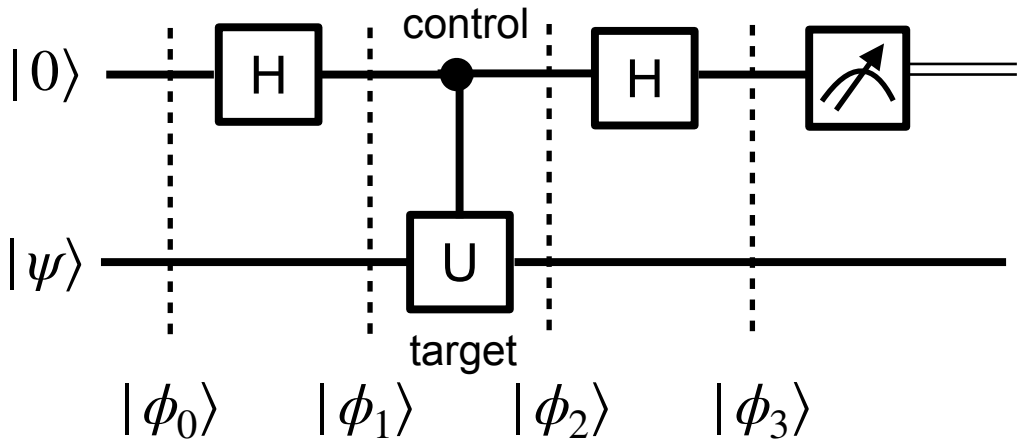
- In the stabilizer formalism, we need to construct a set of Hermitian operators (stabilizers) which satisfy the following properties
  - They square to 1 (so eigenvalues are ± 1).
  - They mutually commute (so they have the same eigenvectors).
  - The syndromes are eigenstates.
  - The uncorrupted syndrome has eigenvalue +1 for all stabilizers.
  - The set of ± 1 eigenvalues of the stabilizers uniquely specifies the syndrome.
  - Whether the eigenvalue is +1 or -1 is easily determined from the commutation properties of the stabilizer with respect to the operator which generate the corruption in the syndrome.

Stabilizer Formalism: Circuits

- Circuit which will measure the eigenvalues of stabilizers and hence determine which syndromes have occurred.

$$U = U^{\dagger}$$

$$U|\psi_{\pm}\rangle = \pm|\psi_{\pm}\rangle$$



$$|\psi\rangle \equiv \alpha_+|\psi_+\rangle + \alpha_-|\psi_-\rangle$$

$$|\phi_0\rangle = |0\rangle \otimes |\psi\rangle = \alpha_+|0\psi_+\rangle + \alpha_-|0\psi_-\rangle$$

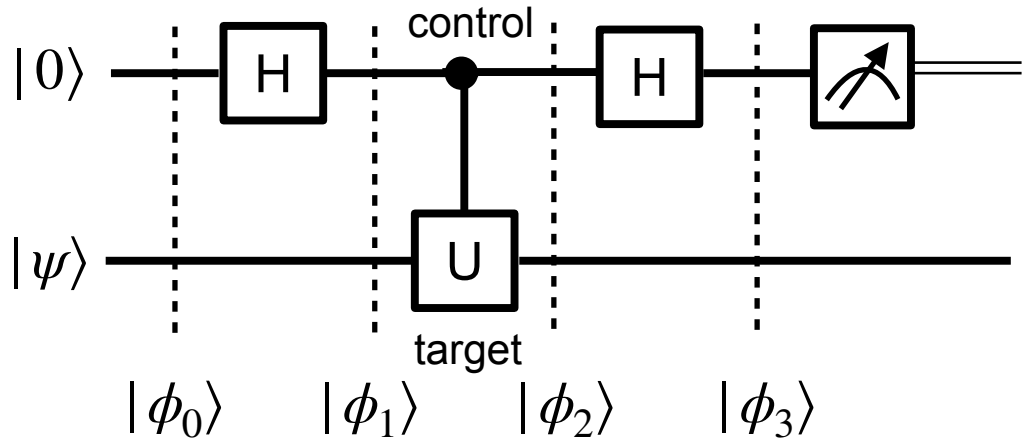
$$|\phi_1\rangle = \frac{1}{\sqrt{2}}\big(|0\rangle + |1\rangle\big) \otimes |\psi\rangle = \frac{\alpha_+}{\sqrt{2}}\Big[|0\psi_+\rangle + |1\psi_+\rangle\Big] + \frac{\alpha_-}{\sqrt{2}}\Big[|0\psi_-\rangle + |1\psi_-\rangle\Big]$$

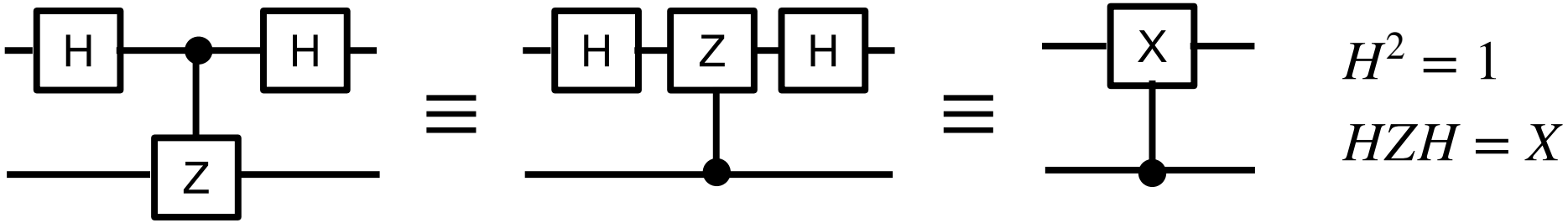
$$|\phi_2\rangle = \frac{\alpha_+}{\sqrt{2}}\Big(|0\psi_+\rangle + |1\psi_+\rangle\Big) + \frac{\alpha_-}{\sqrt{2}}\Big(|0\psi_-\rangle - |1\psi_-\rangle\Big)$$

$$|\phi_3\rangle = \alpha_+|0\psi_+\rangle + \alpha_-|1\psi_-\rangle$$

Stabilizer Formalism: Circuits

- If a measurement of the upper qubit gives $|0\rangle$ (with probability $|\alpha_+|^2$), the lower qubit will be in state $|\psi_+\rangle$.
- If a measurement of the upper qubit gives $|1\rangle$ (with probability $|\alpha_-|^2$), the lower qubit will be in state $|\psi_-\rangle$.
- ∴ control bit tells us which eigenstates of U the target qubit is in.

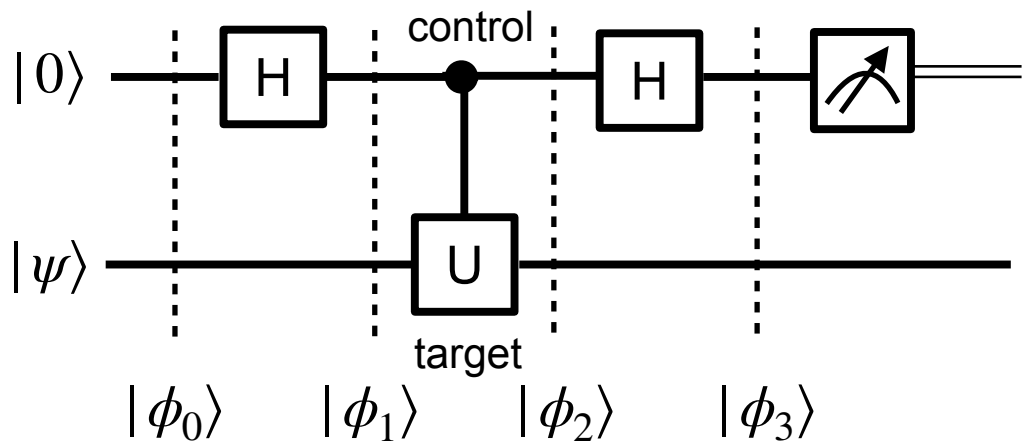




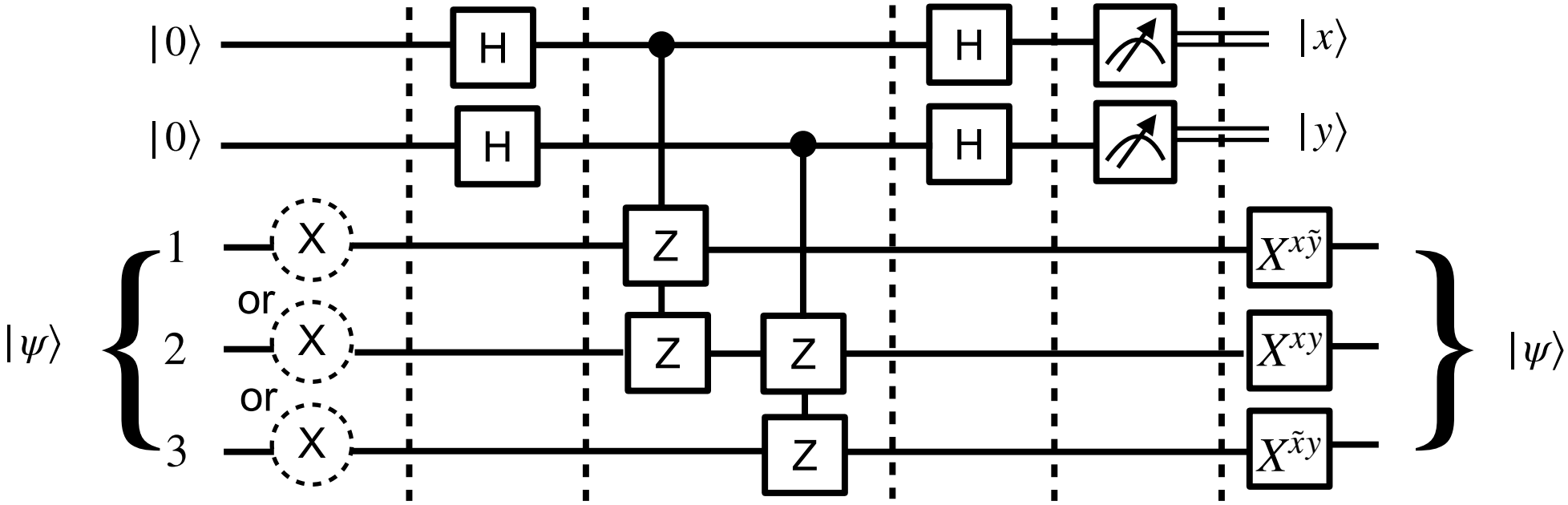
$H^2 = 1$

$HZH = X$

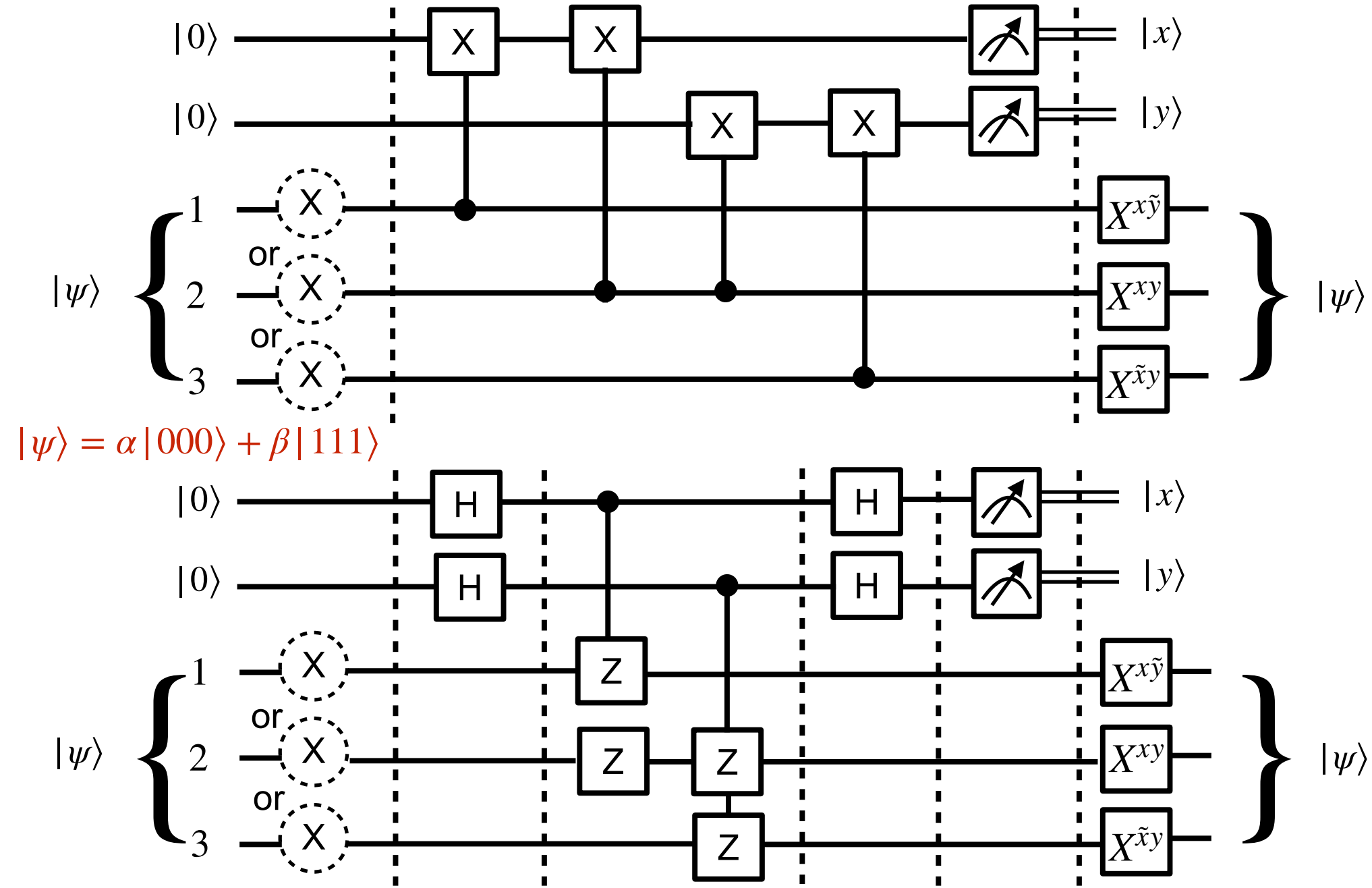
Bitflip code for 3 qubits



$$|\psi\rangle = \alpha|000\rangle + \beta|111\rangle$$



Bitflip code for 3 qubits



Phase Flip

- With some probability p, the relative phase of $|0\rangle$ and $|1\rangle$ is flipped.

Phase Flip

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

$$\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \longrightarrow Z\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \begin{pmatrix}\alpha\\ -\beta\end{pmatrix}$$ in Z-basis (computational basis)

Bit Flip $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \longrightarrow \alpha|1\rangle + \beta|0\rangle$

$$\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \longrightarrow X\begin{pmatrix}\alpha\\ \beta\end{pmatrix} = \begin{pmatrix}\beta\\ \alpha\end{pmatrix}$$

$$X|0\rangle = |1\rangle$$
$$X|1\rangle = |0\rangle$$

- Phase flip error model can be turned into the bit-flip error model by transforming to the ± basis (X basis).

$$|+\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \qquad |-\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle - |1\rangle\right)$$

Transformation is Hadamard:

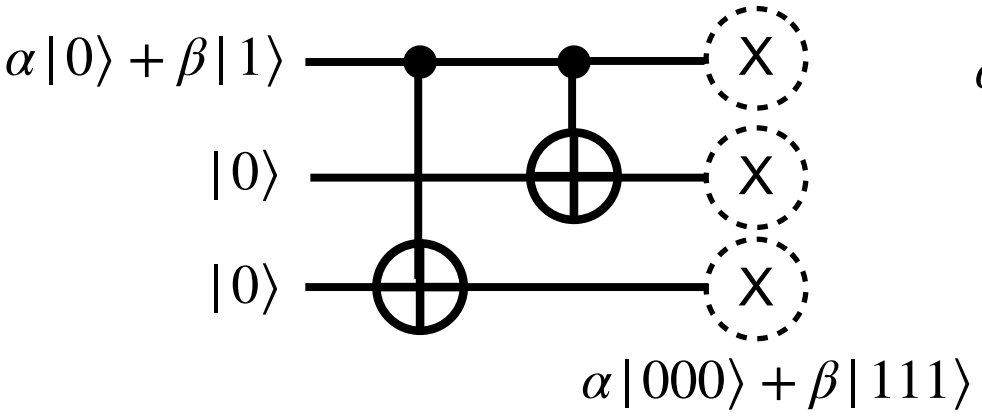
$$H|0\rangle = |+\rangle \qquad H|+\rangle = |0\rangle$$
$$H|1\rangle = |-\rangle \qquad H|-\rangle = |1\rangle$$

Phase Flip

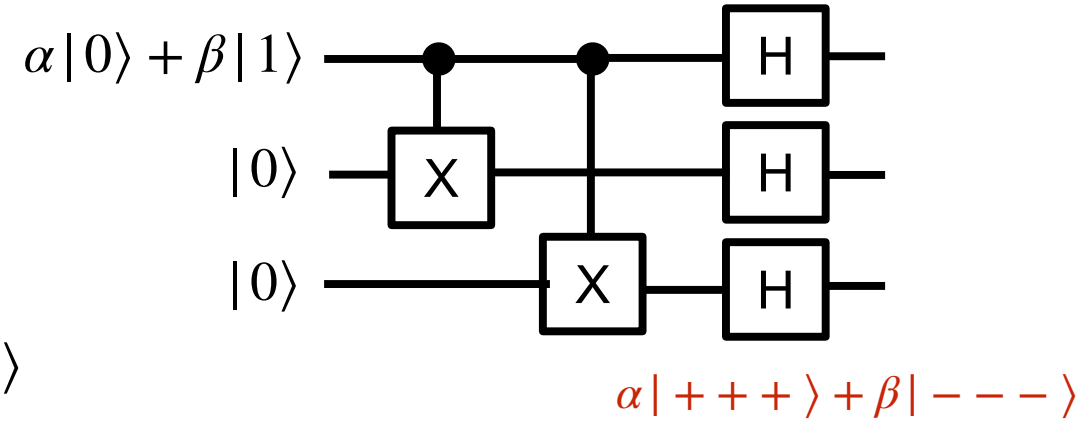
- In the X-basis, roles of X and Z are interchanged.

Bit-flip	$X\|0\rangle = \|1\rangle$ $X\|1\rangle = \|0\rangle$	$Z\|+\rangle = \|-\rangle$ $Z\|-\rangle = \|+\rangle$	Phase-flip
Phase-flip	$Z\|0\rangle = \|0\rangle$ $Z\|1\rangle = -\|1\rangle$	$X\|+\rangle = \|+\rangle$ $X\|-\rangle = -\|-\rangle$	Bit-flip
	In computational basis (Z-basis)	In X-basis	

- Stabilizers to detect phase errors involve X-operations as opposed to those used to detect bit-flip errors which involve Z-operators.



Circuit to encode 3-qubit bit-flip code acting on a linear combination of $|0\rangle$ and $|1\rangle$



Encoding circuit for the 3-qubit phase flip

Spin Correlation Measurements and Bell's Inequality

Correlations in Spin-Singlet States

- Consider a two-electron system in a spin-singlet state, that is, with a total spin of zero.

$$|\text{spin singlet}\rangle = \left(\frac{1}{\sqrt{2}}\right)(|\hat{\mathbf{z}}+;\hat{\mathbf{z}}-\rangle - |\hat{\mathbf{z}}-;\hat{\mathbf{z}}+\rangle),$$

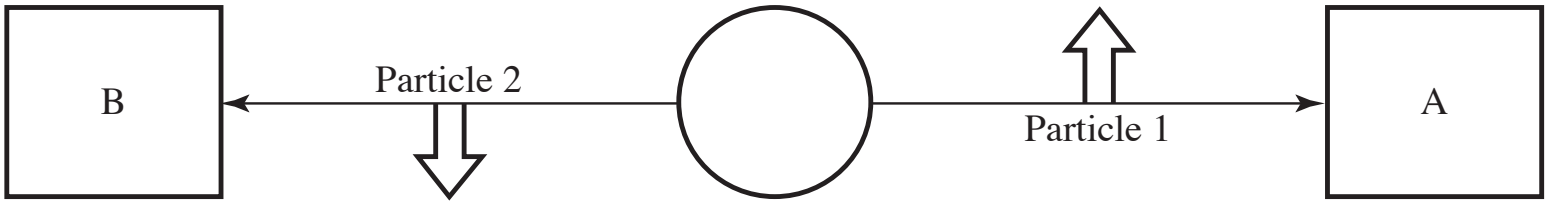
- If one of the components is shown to be in the spin-up state, the other is necessarily in the spin-down state, and vice versa.
- This correlation can persist even if the two particles are well separated and have ceased to interact provided that as they fly apart, there is no change in their spin states.
- This is certainly the case for a J = 0 system disintegrating spontaneously into two spin 1/2 particles with no relative orbital angular momentum, because angular momentum conservation must hold in the disintegration / decay process. A good example is the decay of eta:

$$\eta \rightarrow \mu^+ + \mu^-$$

- Proton-proton scattering at a very low energy can lead to $^1S_0$
  - No orbital angular momentum, spin-singlet state.

Correlations in Spin-Singlet States

- Consider a system of two spin 1/2 particles moving in opposite direction.



- Observer A specializes in measuring Sz of particle 1, while observer B specializes in measuring Sz of particle 2.
- If A finds S_z to be positive for particle 1,

  what does A predict about B's measurement?

what does B measure for Sz?

Correlations in Spin-Singlet States

- Pick one ball out of a bag of a black ball and a white ball.
- When we pick one ball, there is 50%-50% chance if getting black or white. If the 1st ball is black, then we predict with certainty that the second ball must be white.

Correlations in Spin-Singlet States

- Quantum case is more complicated, because observers may choose to measure S_x in place of S_z.
- Sx and Sz eigenstates ae related via:

$$|\hat{x}\pm\rangle = \left(\frac{1}{\sqrt{2}}\right)(|\hat{\mathbf{z}}+\rangle \pm |\hat{\mathbf{z}}-\rangle),$$

$$|\hat{\mathbf{z}}\pm\rangle = \left(\frac{1}{\sqrt{2}}\right)(|\hat{\mathbf{x}}+\rangle \pm |\hat{\mathbf{x}}-\rangle).$$

$$|\text{spin singlet}\rangle = \left(\frac{1}{\sqrt{2}}\right)(|\hat{\mathbf{z}}+;\hat{\mathbf{z}}-\rangle - |\hat{\mathbf{z}}-;\hat{\mathbf{z}}+\rangle),$$

$$|\text{spin singlet}\rangle = \left(\frac{1}{\sqrt{2}}\right)(|\hat{\mathbf{x}}-;\hat{\mathbf{x}}+\rangle - |\hat{\mathbf{x}}+;\hat{\mathbf{x}}-\rangle).$$

Correlations in Spin-Singlet States

$$|\text{spin singlet}\rangle = \left(\frac{1}{\sqrt{2}}\right)(|\hat{\mathbf{z}}+;\hat{\mathbf{z}}-\rangle - |\hat{\mathbf{z}}-;\hat{\mathbf{z}}+\rangle), \qquad |\text{spin singlet}\rangle = \left(\frac{1}{\sqrt{2}}\right)(|\hat{\mathbf{x}}-;\hat{\mathbf{x}}+\rangle - |\hat{\mathbf{x}}+;\hat{\mathbf{x}}-\rangle).$$

- If A determines S_z of particle 1 to be positive, what does B measure for Sx?

- If A chooses to measure S_x and determines S_x of particle 1 to be positive, then B will measure S_x of particle 2 to be
- If A makes no measurement, what can B say about Sx measurement?

Spin Correlation Measurements

Spin component measured by A	A's result	Spin component measured by B	B's result
z	$+$	z	$-$
z	$-$	x	$+$
x	$-$	z	$-$
x	$-$	z	$+$
z	$+$	x	$-$
x	$+$	x	$-$
z	$+$	x	$+$
x	$-$	x	$+$
z	$-$	z	$+$
z	$-$	x	$-$
x	$+$	z	$+$
x	$+$	z	$-$

- The outcome of B's measurement appears to depend on what kind of measurement A decides to perform: an S_x measurement, an S_z measurement, or no measurement.
- A and B can be miles apart with no possibility of communications or mutual interactions.

Einstein's Locality Principle

- Einstein's locality principle: "But on one supposition we should, in my opinion, absolutely hold fast: The real factual situation of the system S_2 is independent of what is done with the system S_1, which is spatially separated from the former." - A. Einstein, B. Podolsky, and N. Rosen, 1935 - EPR paradox
- Some have argued that the difficulties encountered here are inherent in the probabilistic interpretations of quantum mechanics and that the dynamic behavior at the microscopic level appears probabilistic only because some yet unknown parameters, so-called hidden variables, have not been specified.
- Still 1964, people thought that such theorists would give no prediction other than usual QM prediction that would be verified experimentally. The whole debate belonged to the realm of metaphysics….
- Bell pointed out that alternative theories based on Einstein's locality principle actually predict a testable inequality relation among the observables of spin-correlation experiments that disagrees with the predictions of quantum mechanics.

Bell's Inequality with a simple spin 1/2 model by Wigner (alternative theories)

- It is impossible to determine S_x and S_z simultaneously.
- When we have a large number of spin 1/2 particles, we assign a certain fraction of them to have the following property.

  If S_z is measured, we obtain a plus sign with certainty.

  If S_x is measured, we obtain a minus sign with certainty.

- A particle satisfying this property is said to belong to type (**z**ˆ+,**x**ˆ−).
- we can not simultaneously measure S_z and S_x to be + and −, respectively.
- When we measure S_z, we do not measure S_x, and vice versa.
- We are assigning definite values of spin components in more than one direction with the understanding that only one or the other of the components can actually be measured. Even though this approach is fundamentally different from that of quantum mechanics, the quantum-mechanical predictions for S_z and S_x measurements performed on the spin-up (S_z+) state are reproduced provided there are as many particles belonging to type (**z**ˆ+,**x**ˆ+) as to type (**z**ˆ+ , **x**ˆ−) .

Spin-Singlet States (alternative theories)

- There must be a perfect matching between particle 1 and particle 2 to ensure zero total angular momentum: if particle 1 is of type (**z**ˆ+,**x**ˆ−), then particle 2 must belong to type (**z**ˆ−,**x**ˆ+), and so forth, with equal populations, 25% each.

particle 1 particle 2

$$(\hat{\mathbf{z}}+,\hat{\mathbf{x}}-) \leftrightarrow (\hat{\mathbf{z}}-,\hat{\mathbf{x}}+),$$

$$(\hat{\mathbf{z}}+,\hat{\mathbf{x}}+) \leftrightarrow (\hat{\mathbf{z}}-,\hat{\mathbf{x}}-),$$

$$(\hat{\mathbf{z}}-,\hat{\mathbf{x}}+) \leftrightarrow (\hat{\mathbf{z}}+,\hat{\mathbf{x}}-),$$

$$(\hat{\mathbf{z}}-,\hat{\mathbf{x}}-) \leftrightarrow (\hat{\mathbf{z}}+,\hat{\mathbf{x}}+),$$

- If observer A decides to measure S_z of particle 1, then he or she necessarily obtains a plus sign regardless of whether B decides to measure S_z or S_x.
- It is in this sense that Einstein's locality principle is incorporated in this model: A's result is predetermined independently of B's choice as to what to measure.

Spin-Singlet States (alternative theories)

- Consider three unit vectors, a, b and c (in general, not mutually orthogonal).
- One particle belongs to some definite type, say (**a**ˆ−,**b**ˆ+,**c**ˆ+), which means that if **S**·**a** is measured, we obtain a minus sign with certainty; if **S** · **b** is measured, we obtain a plus sign with certainty; if **S** · **c** is measured, we obtain a plus sign with certainty.
- There must be a perfect matching in the sense that the other particle necessarily belongs to type (**a**ˆ+,**b**ˆ−,**c**ˆ−) to ensure zero total angular momentum.
- These eight possibilities are mutually exclusive and disjoint.

Population	Particle 1	Particle 2
N_1	$(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+,\hat{\mathbf{c}}+)$	$(\hat{\mathbf{a}}-,\hat{\mathbf{b}}-,\hat{\mathbf{c}}-)$
N_2	$(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+,\hat{\mathbf{c}}-)$	$(\hat{\mathbf{a}}-,\hat{\mathbf{b}}-,\hat{\mathbf{c}}+)$
N_3	$(\hat{\mathbf{a}}+,\hat{\mathbf{b}}-,\hat{\mathbf{c}}+)$	$(\hat{\mathbf{a}}-,\hat{\mathbf{b}}+,\hat{\mathbf{c}}-)$
N_4	$(\hat{\mathbf{a}}+,\hat{\mathbf{b}}-,\hat{\mathbf{c}}-)$	$(\hat{\mathbf{a}}-,\hat{\mathbf{b}}+,\hat{\mathbf{c}}+)$
N_5	$(\hat{\mathbf{a}}-,\hat{\mathbf{b}}+,\hat{\mathbf{c}}+)$	$(\hat{\mathbf{a}}+,\hat{\mathbf{b}}-,\hat{\mathbf{c}}-)$
N_6	$(\hat{\mathbf{a}}-,\hat{\mathbf{b}}+,\hat{\mathbf{c}}-)$	$(\hat{\mathbf{a}}+,\hat{\mathbf{b}}-,\hat{\mathbf{c}}+)$
N_7	$(\hat{\mathbf{a}}-,\hat{\mathbf{b}}-,\hat{\mathbf{c}}+)$	$(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+,\hat{\mathbf{c}}-)$
N_8	$(\hat{\mathbf{a}}-,\hat{\mathbf{b}}-,\hat{\mathbf{c}}-)$	$(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+,\hat{\mathbf{c}}+)$

- Suppose that observer A finds $\mathbf{S}_1 \cdot \mathbf{a}$ to be plus and observer B finds $\mathbf{S}_2 \cdot \mathbf{b}$ to be plus also. It is clear from Table in the left that the pair belong to either type 3 or type 4, so the number of particle pairs for which this situation is realized is N3+N4

Bell's Inequality from Einstein's locality principle (alternative theories)

- Let P(**a**ˆ+; **b**ˆ+) be the probability that, in a random selection, observer A measures $\mathbf{S}_1 \cdot \mathbf{a}$ to be + and observer B measures $\mathbf{S}_2 \cdot \mathbf{b}$ to be +, and so on.

$$N_3 + N_4 \le (N_2 + N_4) + (N_3 + N_7). \quad \Rightarrow \quad P(\hat{\mathbf{a}}+;\hat{\mathbf{b}}+) \le P(\hat{\mathbf{a}}+;\hat{\mathbf{c}}+) + P(\hat{\mathbf{c}}+;\hat{\mathbf{b}}+).$$

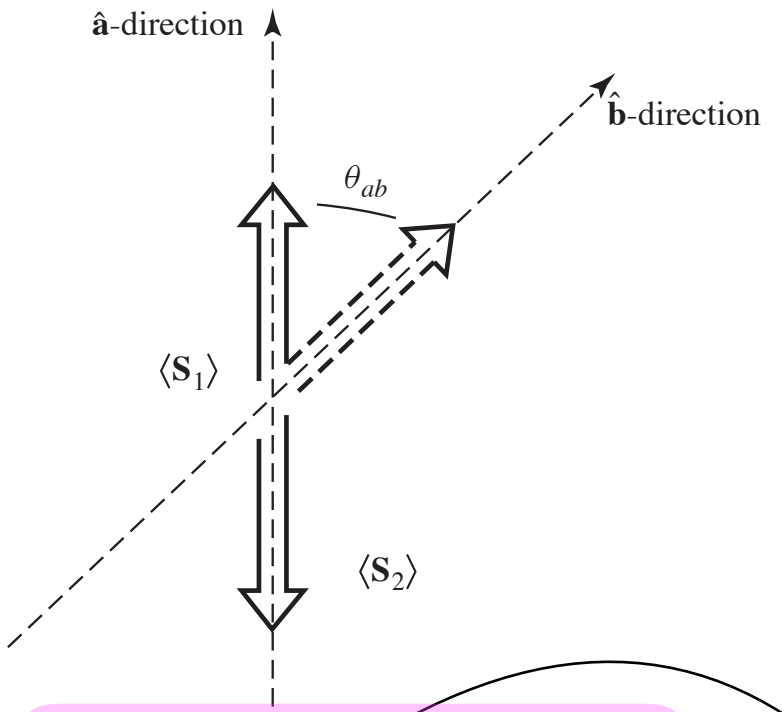
$$P(\hat{\mathbf{a}}+;\hat{\mathbf{c}}+) = \frac{(N_2+N_4)}{\sum_i^8 N_i} \qquad P(\hat{\mathbf{c}}+;\hat{\mathbf{b}}+) = \frac{(N_3+N_7)}{\sum_i^8 N_i} \qquad P(\hat{\mathbf{a}}+;\hat{\mathbf{b}}+) = \frac{(N_3+N_4)}{\sum_i^8 N_i}.$$

Population	Particle 1	Particle 2
N_1	$(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+,\hat{\mathbf{c}}+)$	$(\hat{\mathbf{a}}-,\hat{\mathbf{b}}-,\hat{\mathbf{c}}-)$
N_2	$(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+,\hat{\mathbf{c}}-)$	$(\hat{\mathbf{a}}-,\hat{\mathbf{b}}-,\hat{\mathbf{c}}+)$
N_3	$(\hat{\mathbf{a}}+,\hat{\mathbf{b}}-,\hat{\mathbf{c}}+)$	$(\hat{\mathbf{a}}-,\hat{\mathbf{b}}+,\hat{\mathbf{c}}-)$
N_4	$(\hat{\mathbf{a}}+,\hat{\mathbf{b}}-,\hat{\mathbf{c}}-)$	$(\hat{\mathbf{a}}-,\hat{\mathbf{b}}+,\hat{\mathbf{c}}+)$
N_5	$(\hat{\mathbf{a}}-,\hat{\mathbf{b}}+,\hat{\mathbf{c}}+)$	$(\hat{\mathbf{a}}+,\hat{\mathbf{b}}-,\hat{\mathbf{c}}-)$
N_6	$(\hat{\mathbf{a}}-,\hat{\mathbf{b}}+,\hat{\mathbf{c}}-)$	$(\hat{\mathbf{a}}+,\hat{\mathbf{b}}-,\hat{\mathbf{c}}+)$
N_7	$(\hat{\mathbf{a}}-,\hat{\mathbf{b}}-,\hat{\mathbf{c}}+)$	$(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+,\hat{\mathbf{c}}-)$
N_8	$(\hat{\mathbf{a}}-,\hat{\mathbf{b}}-,\hat{\mathbf{c}}-)$	$(\hat{\mathbf{a}}+,\hat{\mathbf{b}}+,\hat{\mathbf{c}}+)$

$$N_i \ge 0$$

Quantum Mechanics and Bell's Inequality

- In quantum mechanics we do not talk about a certain fraction of particle pairs, belonging to a particular type. Instead, we characterize all spin-singlet systems by the same ket.



$$P(\hat{\mathbf{a}}+;\hat{\mathbf{b}}+) = \left(\frac{1}{2}\right)\sin^2\left(\frac{\theta_{ab}}{2}\right)$$

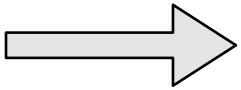
- P(**a**ˆ+;**b**ˆ+):
  - A finds $\mathbf{S}_1 \cdot \mathbf{a}$ to be positive; because of the 100% (opposite sign) correlation we discussed earlier, B's measurement of $\mathbf{S}_2 \cdot \mathbf{a}$ will yield a minus sign with certainty.
  - To calculate P(**a**ˆ+;**b**ˆ+) we must consider a new quantization axis **b** that makes an angle θ_{ab} with **a,** the probability that the $\mathbf{S}_2 \cdot \mathbf{b}$ measurement yields + when particle 2 is known to be in an eigenket of $\mathbf{S}_2 \cdot \mathbf{a}$ with negative eigenvalue is given by

$$\left| \langle \vec{S}_2 \cdot \vec{b} + \mid \vec{S}_2 \cdot \vec{a} - \rangle \right|^2 = \cos^2\left(\frac{\pi - \theta_{ab}}{2}\right) = \sin^2\left(\frac{\theta_{ab}}{2}\right)$$

where 1/2 is from the probability of initially obtaining $\mathbf{S}_1 \cdot \mathbf{a}$ with +.

Quantum Mechanics and Bell's Inequality

$$P(\hat{\mathbf{a}}+;\hat{\mathbf{b}}+) \le P(\hat{\mathbf{a}}+;\hat{\mathbf{c}}+) + P(\hat{\mathbf{c}}+;\hat{\mathbf{b}}+).$$



$$\sin^2\left(\frac{\theta_{ab}}{2}\right) \le \sin^2\left(\frac{\theta_{ac}}{2}\right) + \sin^2\left(\frac{\theta_{cb}}{2}\right)$$

- For simplicity let us choose **a**, **b**, and **c** to lie in a plane, and let **c** bisect the two directions defined by **a** and **b**: $\theta_{ab} = 2\theta, \quad \theta_{ac} = \theta_{cb} = \theta$
- Inequality is violated for $0 < \theta < \frac{\pi}{2}.$
- Note that we can not use the spin-correlation measurement to transmit any useful information between two macroscopically separated points.
- Superluminal communications are impossible.

Bell's Inequality Revisited

- N(A, not B) = number of objects which have parameters A but not B.

$$N(A, \text{not } B, C) + N(\text{not } A, B, \text{not } C) \geq 0$$

Add: $N(A, \text{not } B, \text{not } C) + N(A, B, \text{not } C)$

$$N(A, \text{not } B, C) + N(A, \text{not } B, \text{not } C) + N(\text{not } A, B, \text{not } C) + N(A, B, \text{not } C) \geq N(A, \text{not } B, \text{not } C) + N(A, B, \text{not } C)$$

$$N(A, \text{not } B) + N(B, \text{not } C) \geq N(A, \text{not } C)$$

- We assumed that parameters exist whether they are measured or not. i.e., either B or not B is true for every member.
- A: spin-up along a, B: spin-up along b, C: spin-up along c
- Consider spin1/2 system:
- Not B for particle 1 = spin-down along b for particle 1 = spin-up along b for particle 2

Bell's Inequality Revisited

- A: spin-up along a, B: spin-up along b, C: spin-up along c
- Consider spin1/2 system:
- Not B for particle 1 = spin-down along b for particle 1 = spin-up along b for particle 2

N(A, not B) + N(B, not C) ≥ N(A, not C)

$$N(\hat{a}+,\hat{b}+)+N(\hat{b}+,\hat{c}+)\geq N(\hat{a}+,\hat{c}+)$$

$$P(\hat{a}+,\hat{b}+)+P(\hat{b}+,\hat{c}+)\geq P(\hat{a}+,\hat{c}+)$$

- Assumption
  - Logic is valid.
  - Electrons have spin in a given direction, even if we do not measure it → there is a reality separate from its observation →Hidden variables exist.
  - No information can travel faster than the speed of light → Locality → Hidden variables are local.

Bell's Inequality Revisited

- Assumptions
  1. Logic is valid.
  2. Electrons have spin in a given direction, even if we do not measure it → there is a reality separate from its observation →Hidden variables exist.
  3. No information can travel faster than the speed of light → Locality → Hidden variables are local.
- Are we sure about the assumptions?
  1. It has been suspected since long before Bell that Quantum Mechanics is in conflict with classical logic. Deductive logic has proved that logic is incomplete (1931, Kurt Gödel), Self-reference: "This statement is false."
  2. What if there is no reality separate from its observation? "The path of the electron comes into existence3 only when we observe it" - Heisenberg for the double slit. Ontology of QM?
  3. Non-locality? Quantum teleportation, quantum entanglement.

Bell's Inequality and experimental tests

- QM prediction is not compatible with Bell's inequality.
- Experimental tests showing violation of Bell inequalities have been performed for pairs of two-outcome measurements using photons, ions, superconducting systems and nitrogen vacancy centers, and in pairs of three-outcome measurements using photons.
- Proposals have also been made to test Bell inequalities in e+e- collisions and positronium decays. Recently it has been proposed to make such tests in entangled $t + \bar{t}$ decays, in the Higgs decay to WW, and in systems of B0-B0bar mesons at the LHC.

Testing Bell inequalities at the LHC with top-quark pairs

- The quantum state of a two spin-1/2 pair, as the one formed by a top-quark pair system, can be expressed by the density matrix:

$$\rho \;=\; \frac{1}{4}\Big[\mathbb{1}\otimes\mathbb{1} + \sum_i A_i(\sigma_i\otimes\mathbb{1}) + \sum_j B_j(\mathbb{1}\otimes\sigma_j) + \sum_{ij} C_{ij}(\sigma_i\otimes\sigma_j)\Big]$$

Polarization matrices

$$A_i = \mathrm{Tr}[\rho\,(\sigma_i\otimes\mathbf{1})]$$
$$B_j = \mathrm{Tr}[\rho\,(\mathbf{1}\otimes\sigma_j)]$$

Correlation matrix

$$C_{ij} \;=\; \mathrm{Tr}[\rho\,(\sigma_i\otimes\sigma_j)]$$

In the CM reference frame of the top-quark pair system as produced at a pp collider, the two spin-1/2 quarks fly apart in opposite directions. One can then extract the probability $\mathcal{P}(\uparrow_{\hat{n}};-)$ of finding the spin of one quark in the state $\uparrow_{\hat{n}}$, with the projection of the spin along the axis determined by the unit vector $\hat{n}$ pointing in the up direction. Similarly, one can consider double probabilities, like $\mathcal{P}(\uparrow_{\hat{n}};\downarrow_{\hat{m}})$, of finding the projection of the spin of the quark along the unit vector $\hat{n}$ pointing in the up state, while the companion antiquark has the projection of its spin along the direction of a different unit vector $\hat{m}$ pointing in the down state.

In classical physics, these probabilities involve averages over suitable distributions of variables and obey the following (generalized) Bell inequality [34]:

$$\mathcal{P}(\uparrow_{\hat{n}_1};\uparrow_{\hat{n}_2}) - \mathcal{P}(\uparrow_{\hat{n}_1};\uparrow_{\hat{n}_4}) + \mathcal{P}(\uparrow_{\hat{n}_3};\uparrow_{\hat{n}_2}) + \mathcal{P}(\uparrow_{\hat{n}_3};\uparrow_{\hat{n}_4}) \leq \mathcal{P}(\uparrow_{\hat{n}_3};-) + \mathcal{P}(-;\uparrow_{\hat{n}_2})\;, \qquad (2)$$

Generalized Bell's inequality or Clauser-Horne-Shimony-Holt (CHSH) inequality for spin-1/2

Testing Bell inequalities at the LHC with top-quark pairs

Generalized Bell's inequality or Clauser-Horne-Shimony-Holt (CHSH) inequality for spin-1/2

$$\mathcal{P}(\uparrow_{\hat{n}_1};\uparrow_{\hat{n}_2}) - \mathcal{P}(\uparrow_{\hat{n}_1};\uparrow_{\hat{n}_4}) + \mathcal{P}(\uparrow_{\hat{n}_3};\uparrow_{\hat{n}_2}) + \mathcal{P}(\uparrow_{\hat{n}_3};\uparrow_{\hat{n}_4}) \leq \mathcal{P}(\uparrow_{\hat{n}_3};-) + \mathcal{P}(-;\uparrow_{\hat{n}_2}) \ , \qquad (2)$$

where $\hat{n}_1$, $\hat{n}_2$, $\hat{n}_3$ and $\hat{n}_4$ are four different three-dimensional unit vectors determining four spatial directions along which the spins of the quark and antiquark can be measured. In quantum mechanics the same probabilities are computed as expectation of suitable spin-observable operators (Eq. (1)), so that the previous inequality reduces to the following constraint

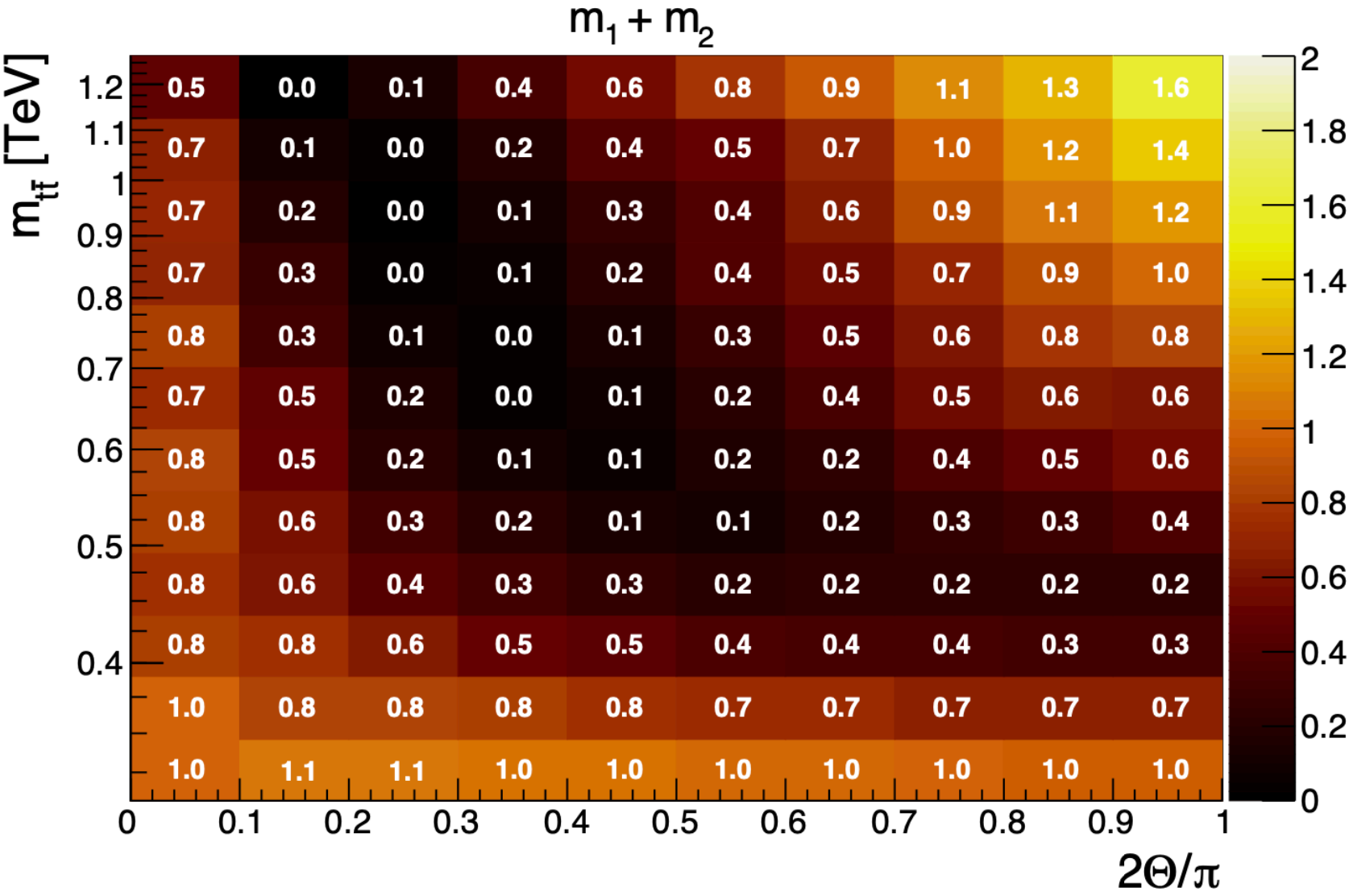
$$\left|\hat{n}_1 \cdot C \cdot \left(\hat{n}_2 - \hat{n}_4\right) + \hat{n}_3 \cdot C \cdot \left(\hat{n}_2 + \hat{n}_4\right)\right| \leq 2 \ , \qquad (3)$$

involving only the spin correlation matrix C_{ij} and not the polarization coefficients A_i and B_j.

consider the matrix C and its transpose C^T and form the symmetric, positive, 3×3 matrix $M = C^T C$ whose three eigenvalues m_1, m_2, m_3 can be ordered by decreasing magnitude: $m_1 \geq m_2 \geq m_3$. The two-spin state density matrix ρ in (1) violates the inequality (3), or equivalently (2), if and only if the sum of the two greatest eigenvalues of M is strictly larger than 1, that is

$$m_1 + m_2 > 1 \, . \qquad (4)$$

Testing Bell inequalities at the LHC with top-quark pairs



Barren plateaus in QNN training landscapes

- 1803.11173

Quantum Metric Learning