Quantum Information Processing

Quick notes

Lorenzo Laneve

Contents

Utility

0.1 Overlap with superposition vs overlap with subspace

To write: this bound is already present in the analysis of Grover's algorithm

0.2 Metric between unitaries for approximation algorithms

Let $U, \bar{U}$ be unitaries. We say that $\bar{U}$ ε-approximates U if:

$$D(U, \bar{U}) := ||U - \bar{U}||_2 = \max_{\langle\psi|\psi\rangle=1} ||(U - \bar{U})\,|\psi\rangle\,||_2 \leq \varepsilon$$

Theorem 0.1. $D(U_1U_2, V_1V_2) \leq D(U_1, V_1) + D(U_2, V_2)$.

Proof.

$$\begin{aligned}
D(U_1U_2, V_1V_2) &= \max_{\langle\psi|\psi\rangle=1} ||(U_1U_2 - V_1V_2)\,|\psi\rangle\,||_2 \\
&= \max_{\langle\psi|\psi\rangle=1} ||(U_1U_2 - V_1U_2 + V_1U_2 - V_1V_2)\,|\psi\rangle\,||_2 \\
&= \max_{\langle\psi|\psi\rangle=1} ||(U_1 - V_1)U_2\,|\psi\rangle + V_1(U_2 - V_2)\,|\psi\rangle\,||_2 \\
&\leq \max_{\langle\psi|\psi\rangle=1} ||(U_1 - V_1)U_2\,|\psi\rangle\,||_2 + \max_{\langle\psi|\psi\rangle=1} ||V_1(U_2 - V_2)\,|\psi\rangle\,||_2 \\
&= \max_{\langle\psi|\psi\rangle=1} ||(U_1 - V_1)\,|\psi\rangle\,||_2 + \max_{\langle\psi|\psi\rangle=1} ||(U_2 - V_2)\,|\psi\rangle\,||_2 \\
&= D(U_1, V_1) + D(U_2, V_2)
\end{aligned}$$

□

This claim tells us that local changes are the only ones contributing to the approximation error, in an additive fashion.

Theorem 0.2. $D(\bigwedge G, \mathbb{1}_4) = D(G, \mathbb{1}_2)$.

Proof.

$$\begin{aligned}
D(\bigwedge G, \mathbb{1}_4) &= \max_{\langle\psi|\psi\rangle=1} ||(\bigwedge G - \mathbb{1}_4)\,|\psi\rangle\,||_2 \\
&= \max_{\langle\psi|\psi\rangle=1} ||(|0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes G - |0\rangle\langle 0| \otimes \mathbb{1} - |1\rangle\langle 1| \otimes \mathbb{1})\,|\psi\rangle\,||_2 \\
&= \max_{\langle\psi|\psi\rangle=1} ||(|1\rangle\langle 1| \otimes (G - \mathbb{1})\,|\psi\rangle\,||_2 \\
&= \max_{\langle\psi|\psi\rangle=1} ||(G - \mathbb{1})\,|\psi\rangle\,||_2 = D(G, \mathbb{1}_2)
\end{aligned}$$

The last inequality holds because we can always take $|\psi\rangle = |1\rangle \otimes |\phi\rangle$. Any entangled state would result in a convex combination of norms of the form $||(G - \mathbb{1})\,|\phi_i\rangle\,||_2$. □

1 Introduction

1.1 No-cloning theorem

Theorem 1.1. *No unitary U is such that*

$$U|0\rangle|\psi\rangle = |\psi\rangle|\psi\rangle$$

for every $|\psi\rangle \in \mathcal{H}$.

Proof. Suppose for a contradiction to have U such that:

$$U|\psi\rangle|0\rangle|0\rangle = |\psi\rangle|\psi\rangle|\gamma_\psi\rangle$$

Notice that the existence of a unitary as in the claim implies the one we are assuming here, for some choice of $|\gamma_\psi\rangle$. Take $|\psi\rangle, |\phi\rangle$ with $\langle\psi|\phi\rangle \in (0,1)$. Since we have:

$$\begin{cases} U|\psi\rangle|0\rangle|0\rangle = |\psi\rangle|\psi\rangle|\gamma_\psi\rangle \\ U|\phi\rangle|0\rangle|0\rangle = |\phi\rangle|\phi\rangle|\gamma_\phi\rangle \end{cases}$$

This implies that:

$$\begin{aligned} \langle\psi|\phi\rangle &= \langle\psi|\langle 0|\langle 0|U^\dagger U|\phi\rangle|0\rangle|0\rangle \\ &= \langle\psi|\phi\rangle\langle\psi|\phi\rangle\langle\gamma_\psi|\gamma_\phi\rangle \end{aligned}$$

which means $\langle\psi|\phi\rangle\langle\gamma_\psi|\gamma_\phi\rangle = 1$, which is a contradiction. □

Theorem 1.2. *A cloning machine allows signalling, i.e. a channel with transmission faster than the speed of light can be achieved.*

Proof. Let $\mathcal{C}$ be a machine such that:

$$\mathcal{C} : |\psi\rangle \otimes |0\rangle \otimes \cdots \otimes |0\rangle \mapsto |\psi\rangle \otimes \cdots \otimes |\psi\rangle$$

Consider two spatially separated agents Alice and Bob, each holding a qubit of the entangled pair $|\Psi^{11}\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}$.

Now suppose Alice uses the measurement basis $\{|\psi\rangle, |\psi^\perp\rangle\}$, where

$$\begin{cases} |\psi\rangle = \alpha|0\rangle + \beta|1\rangle \\ |\psi^\perp\rangle = \beta^*|0\rangle - \alpha^*|1\rangle \end{cases}$$

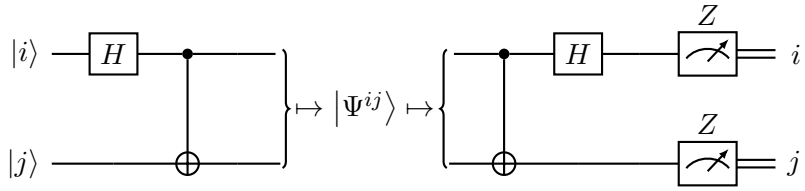
The possible outcomes for Bob's qubit are:

$$\propto \begin{cases} (|\psi\rangle\langle\psi| \otimes \mathbb{1})|\Psi^{00}\rangle = -|\psi\rangle \otimes |\psi^\perp\rangle \\ (|\psi^\perp\rangle\langle\psi^\perp| \otimes \mathbb{1})|\Psi^{00}\rangle = |\psi^\perp\rangle \otimes |\psi\rangle \end{cases}$$

Suppose Alice wants to signal a bit x to Bob: if $x = 0$ then she measures her qubit in the computational basis, otherwise she will do so in the Hadamard basis. Bob will now have a state in $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$ and, in order to learn x, he will just have to discriminate between the case $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle, |-\rangle\}$. This can be achieved by cloning the state n times using $\mathcal{C}$ and doing a state tomography, e.g. measuring in the computational basis: if the measurement outcome never changes, then we will have $x = 0$ with probability $1 - \frac{1}{2^n}$.

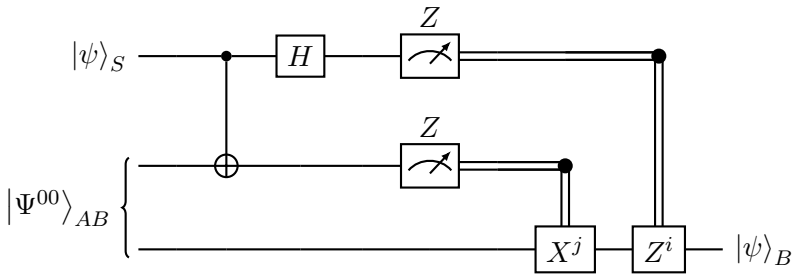
Notice that the cloning here is necessary: if we just repeated the procedure with n copies of $|\Psi^{11}\rangle$, Bob would receive a series of qubits at random between $|\psi\rangle$ and $|\psi^\perp\rangle$, which would give random outcomes regardless of the basis. □

1.2 Bell basis



- $\left|\Psi^{00}\right\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$
- $\left|\Psi^{01}\right\rangle = \frac{|01\rangle+|10\rangle}{\sqrt{2}}$
- $\left|\Psi^{10}\right\rangle = \frac{|00\rangle-|11\rangle}{\sqrt{2}}$
- $\left|\Psi^{11}\right\rangle = \frac{|01\rangle-|10\rangle}{\sqrt{2}}$

1.3 Quantum teleportation



Claim 1.3. *The above circuit correctly teleports the state* $|\psi\rangle$*.*

Proof. Let $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. The initial state is:

$$\begin{aligned}
|\phi\rangle_{SAB} &= |\psi\rangle_S \otimes \frac{|00\rangle + |11\rangle}{\sqrt{2}} \\
&= \frac{1}{\sqrt{2}}(\alpha|000\rangle + \alpha|011\rangle + \beta|100\rangle + \beta|111\rangle) \\
&= \frac{1}{2\sqrt{2}}(\alpha|000\rangle + \alpha|011\rangle + \beta|100\rangle + \beta|111\rangle + \alpha|000\rangle + \alpha|011\rangle + \beta|100\rangle + \beta|111\rangle) \\
&= \frac{1}{2\sqrt{2}}(\alpha|000\rangle + \alpha|011\rangle + \beta|100\rangle + \beta|111\rangle + \alpha|000\rangle + \alpha|011\rangle + \beta|100\rangle + \beta|111\rangle) \\
&+ \frac{1}{2\sqrt{2}}(\alpha|110\rangle + \alpha|101\rangle + \beta|010\rangle + \beta|001\rangle - \alpha|110\rangle - \alpha|101\rangle - \beta|010\rangle - \beta|001\rangle) \\
&= \frac{1}{2}\frac{|00\rangle + |11\rangle}{\sqrt{2}} \otimes (\alpha|0\rangle + \beta|1\rangle) + \frac{1}{2}\frac{|01\rangle + |10\rangle}{\sqrt{2}} \otimes (\alpha|1\rangle + \beta|0\rangle) \\
&+ \frac{1}{2}\frac{|00\rangle - |11\rangle}{\sqrt{2}} \otimes (\alpha|0\rangle - \beta|1\rangle) + \frac{1}{2}\frac{|01\rangle - |10\rangle}{\sqrt{2}} \otimes (\alpha|1\rangle - \beta|0\rangle) \\
&= \frac{1}{2}|\Psi^{00}\rangle \otimes |\psi\rangle + \frac{1}{2}|\Psi^{01}\rangle \otimes X|\psi\rangle + \frac{1}{2}|\Psi^{10}\rangle \otimes Z|\psi\rangle + \frac{1}{2}|\Psi^{11}\rangle \otimes XZ|\psi\rangle
\end{aligned}$$

□

Claim 1.4. *Without* i *and* j*, no information can be retrieved after the Bell measurement.*

Proof. The Bell measurement corresponds to the following quantum channel:

$$\mathcal{B}(\rho_{SAB}) = \sum_{i,j} \left(|\Psi^{ij}\rangle\langle\Psi^{ij}| \otimes \mathbb{1}_B\right) \rho_{SAB} \left(|\Psi^{ij}\rangle\langle\Psi^{ij}| \otimes \mathbb{1}_B\right)$$

Applied to the initial state $|\phi\rangle\langle\phi|$ we obtain:

$$\begin{aligned}
\mathcal{B}(|\phi\rangle\langle\phi|) = \frac{1}{4}|\Psi^{00}\rangle\langle\Psi^{00}| \otimes |\psi\rangle\langle\psi| + \frac{1}{4}|\Psi^{10}\rangle\langle\Psi^{10}| \otimes X|\psi\rangle\langle\psi|X \\
+ \frac{1}{4}|\Psi^{01}\rangle\langle\Psi^{01}| \otimes Z|\psi\rangle\langle\psi|Z + \frac{1}{4}|\Psi^{11}\rangle\langle\Psi^{11}| \otimes XZ|\psi\rangle\langle\psi|ZX
\end{aligned}$$

Since we only consider Bob's qubit, we can trace out Alice's system:

$$\mathrm{tr}_{SA}(\mathcal{B}(|\phi\rangle\langle\phi|)) = \frac{1}{4}|\psi\rangle\langle\psi| + \frac{1}{4}X|\psi\rangle\langle\psi|X + \frac{1}{4}Z|\psi\rangle\langle\psi|Z + \frac{1}{4}XZ|\psi\rangle\langle\psi|ZX = \frac{\mathbb{1}}{2}$$

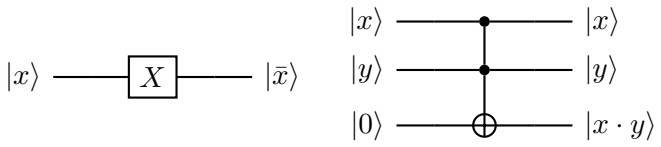
Hence, we have a fully mixed state, giving no information. □

2 General Reversible Computation

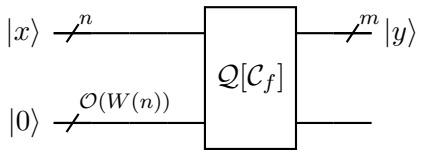
Theorem 2.1. *Quantum computing can carry out arbitrary classical computation.*

Proof. Let $f : \{0,1\}^n \to \{0,1\}^m$ be a boolean function, and let $\mathcal{C}_f$ be a (bounded fan-in) classical circuit implementing f with $d(n)$ depth and $W(n)$ total work. Each gate can be implemented by $\mathcal{O}(1)$ AND and NOT gates, since they are complete for propositional logic. Thus, assume without loss of generality these are the only gates in the circuit.

We are going to replace these classical gates with quantum gates acting similarly on the computational basis.



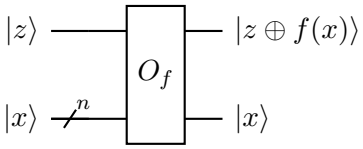
This gives us a circuit $\mathcal{Q}[\mathcal{C}_f]$ with $d(n)$ depth and $W(n)$ total work.



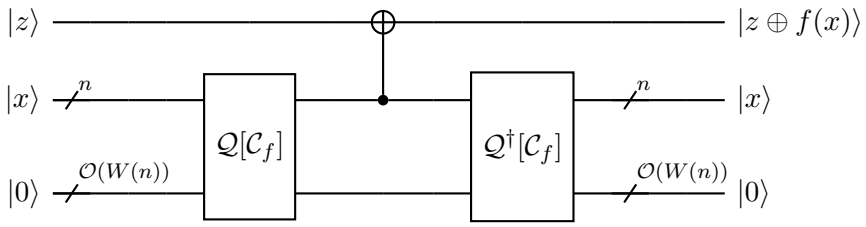
□

2.1 Bit Oracle

Consider a function $f : \{0,1\}^n \to \{0,1\}$. Considering $\mathcal{Q}[\mathcal{C}_f]$ as above we construct the following circuit:



$\equiv$

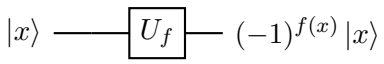


One can see that:

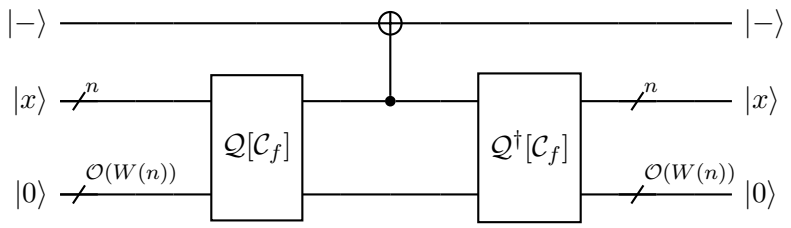
$$O_f |x\rangle |z\rangle = |x\rangle |z \oplus f(x)\rangle$$

2.2 Phase Oracle

We fix the $|z\rangle$ in the bit oracle above with a $|-\rangle$:



$\equiv$



Claim 2.2. $U_f |x\rangle = (-1)^{f(x)} |x\rangle$.

Proof.

$$\begin{aligned}(\mathbb{1} \otimes \mathcal{Q}^\dagger[\mathcal{C}_f])(X\Lambda \otimes \mathbb{1})(\mathbb{1} \otimes \mathcal{Q}[\mathcal{C}_f]) |-\rangle |x\rangle |0\rangle &= (\mathbb{1} \otimes \mathcal{Q}^\dagger[\mathcal{C}_f])(X\Lambda \otimes \mathbb{1}) |-\rangle |f(x)\rangle |\psi_x\rangle \\ &= (\mathbb{1} \otimes \mathcal{Q}^\dagger[\mathcal{C}_f])(-1)^{f(x)} |-\rangle |f(x)\rangle |\psi_x\rangle \\ &= (-1)^{f(x)} |-\rangle |x\rangle |0\rangle\end{aligned}$$

□

2.3 Deutsch-Josza Algorithm

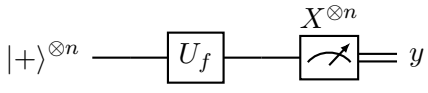
Problem 2.3 (Constant or balanced)**.** *Given a function $f : \{0,1\}^n \to \{0,1\}$ as oracle, which is either constant or balanced (exactly half of the assignments give 1), discriminate between the two cases.*

Claim 2.4. *Any classical algorithm requires $\Omega(2^n)$ queries to the oracle to achieve a non-trivial success probability.*

Proof. Consider an arbitrary randomized algorithm $\mathcal{A}$ using q queries. Take a uniformly random function F, and consider $\mathcal{A}$ as a random variable determined by the randomness used by the algorithm. By using Yao's principle:

$$\begin{aligned}\min_f \mathbb{P}(\mathcal{A}(f) \text{ succeeds}) &\leq \sum_f \mathbb{P}(\mathcal{A}(F) \text{ succeeds} \mid F = f)\, \mathbb{P}(F = f) \\ &= \mathbb{P}(\mathcal{A}(F) \text{ succeeds}) \\ &= \sum_a \mathbb{P}(\mathcal{A}(F) \text{ succeeds} \mid \mathcal{A} = a)\, \mathbb{P}(\mathcal{A} = a) \\ &\leq \max_a \mathbb{P}(a(F) \text{ succeeds})\end{aligned}$$

Notice that, if $q \leq 2^{n-1}$, F still has $\frac{1}{2}$ probability of being constant or balanced, thus the above is at most $\frac{1}{2}$. This means that $q > 2^{n-1} = \Omega(2^n)$ in order to have a non-trivial success probability. If $\mathcal{A}$ has probability p of querying more than 2^{n-1} times, then the success probability is bounded by $\frac{1}{2}(1-p) + p = \frac{1+p}{2}$, by the law of total probability. For p close to 0, the success probability is close to $\frac{1}{2}$. □



The algorithm returns 'balanced' if and only if $y = 0$. One can see that exactly one query is made to the oracle.

Claim 2.5. *The Deutsch-Josza algorithm always succeeds.*

Proof. We have $y = 0$ if and only if $|+\rangle^{\otimes n}$ is measured.

$$\begin{aligned}
\mathbb{P}(y=0) &= \left|\langle+|^{\otimes n} U_f |+\rangle^{\otimes n}\right|^2 \\
&= \left|\frac{1}{2^n}\sum_{x'}\langle x'| U_f \sum_x |x\rangle\right|^2 \\
&= \left|\frac{1}{2^n}\sum_{x',x}(-1)^{f(x)}\langle x'|x\rangle\right|^2 \\
&= \left|\frac{1}{2^n}\sum_{x}(-1)^{f(x)}\right|^2
\end{aligned}$$

If f is balanced, the sum (and thus the probability) is 0. If f is constant, the sum is 2^n and the overall probability becomes 1. □

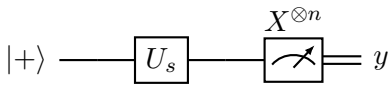
2.4 Parity of the Hamming weight

Problem 2.6. *Given a string s of N bits, decide whether its Hamming weight is even or odd. In other words, compute $\bigoplus_i s_i$.*

Construct a phase oracle U_s such that:

$$U_s : |i\rangle \mapsto (-1)^{s_i} |i\rangle$$

where s_i denotes the i-th bit of s. If we have a string of two bits, the input of U_s is a single qubit, and we can compute $s_0 \oplus s_1$ in the following way:



This yields, right before the measurement:

$$|\phi\rangle = \frac{(-1)^{s_0}|0\rangle + (-1)^{s_1}|1\rangle}{\sqrt{2}} = \begin{cases} (-1)^{s_0}|+\rangle & s_0 = s_1 \\ (-1)^{s_0}|-\rangle & s_0 \neq s_1 \end{cases}$$

and measuring with the Hadamard basis yields $y = s_0 \oplus s_1$. If $N > 2$, we input in the above circuit the state:

$$|x\rangle \otimes |+\rangle = \frac{|2x\rangle + |2x+1\rangle}{\sqrt{2}}$$

For $x = 0, \ldots, \left\lfloor \frac{N}{2} \right\rfloor - 1$. Each step yields $y_x = s_{2x} \oplus s_{2x+1}$. If N is odd, another query is needed to retrieve s_{N-1}. This gives the result in $\left\lfloor \frac{N}{2} \right\rfloor$ queries to the memory.

2.5 The Z-Y decomposition

Lemma 2.7. *Any 2×2 unitary matrix U can be written as:*

$$U = \begin{bmatrix} e^{i(\alpha-\beta-\delta)}\cos\gamma & -e^{i(\alpha-\beta+\delta)}\sin\gamma \\ e^{i(\alpha+\beta-\delta)}\sin\gamma & e^{i(\alpha+\beta+\delta)}\cos\gamma \end{bmatrix}$$

for some $\alpha, \beta, \gamma, \delta \in \mathbb{R}$.

Proof. A unitary matrix must have a pair of orthogonal unit vectors as its columns. Consider two arbitrary such vectors:

$$|\psi\rangle = ae^{i\theta_a}|0\rangle + be^{i\theta_b}|1\rangle$$
$$|\phi\rangle = e^{ic}(-be^{-i\theta_b}|0\rangle + ae^{-i\theta_a}|1\rangle)$$

We can set $a = \cos\gamma, b = \sin\gamma$ in order to match the amplitudes, obtaining:

$$|\psi\rangle = e^{i\theta_a}\cos\gamma|0\rangle + e^{i\theta_b}\sin\gamma|1\rangle$$
$$|\phi\rangle = e^{ic}(-e^{-i\theta_b}\sin\gamma|0\rangle + e^{-i\theta_a}\cos\gamma|1\rangle)$$

By matching the phases we solve a linear system of equations, whose solution is:

$$\begin{cases} \alpha = \frac{c}{2} \\ \beta = \frac{\theta_b - \theta_a}{2} \\ \delta = \frac{c - \theta_b - \theta_a}{2} \end{cases}$$

Thus, we found an explicit representation for an arbitrary unitary operator. □

One can see that:

$$\begin{bmatrix} e^{i(\alpha-\beta-\delta)}\cos\gamma & -e^{i(\alpha-\beta+\delta)}\sin\gamma \\ e^{i(\alpha+\beta-\delta)}\sin\gamma & e^{i(\alpha+\beta+\delta)}\cos\gamma \end{bmatrix} = e^{i\alpha}R_z(2\beta)R_y(2\gamma)R_z(2\delta)$$

implying that every single qubit unitary can be expressed as three rotations on the Bloch sphere.

2.6 Measurement-based quantum computation

Suppose to have two qubits, where the first controls the second through a CZ gate. Notice that:

$$\begin{aligned} \bigwedge Z &= |0\rangle\langle 0| \otimes \mathbb{1} + |1\rangle\langle 1| \otimes Z \\ &= |0\rangle\langle 0| \otimes (|0\rangle\langle 0| + |1\rangle\langle 1|) + |1\rangle\langle 1| \otimes (|0\rangle\langle 0| - |1\rangle\langle 1|) \\ &= |00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle|10\rangle - |11\rangle|11\rangle \\ Z\bigwedge &= \mathbb{1} \otimes |0\rangle\langle 0| + Z \otimes |1\rangle\langle 1| \\ &= (|0\rangle\langle 0| + |1\rangle\langle 1|) \otimes |0\rangle\langle 0| + (|0\rangle\langle 0| - |1\rangle\langle 1|) \otimes |1\rangle\langle 1| \\ &= |00\rangle\langle 00| + |10\rangle\langle 10| + |01\rangle\langle 01| - |11\rangle\langle 11| \end{aligned}$$

Thus, $\bigwedge Z = Z\bigwedge \equiv CZ$, i.e. the role of the qubits does not matter. Now, let us take three qubits:

$$\begin{aligned} CZ_{12} \cdot CZ_{23} &= (|0\rangle\langle 0| \otimes \mathbb{1} \otimes \mathbb{1} + |1\rangle\langle 1| \otimes Z \otimes \mathbb{1}) \cdot (\mathbb{1} \otimes |0\rangle\langle 0| \otimes \mathbb{1} + \mathbb{1} \otimes |1\rangle\langle 1| \otimes Z) \\ &= |00\rangle\langle 00| \otimes \mathbb{1} + |10\rangle\langle 10| \otimes \mathbb{1} + |01\rangle\langle 01| \otimes Z - |11\rangle\langle 11| \otimes Z \\ CZ_{23} \cdot CZ_{12} &= (\mathbb{1} \otimes |0\rangle\langle 0| \otimes \mathbb{1} + \mathbb{1} \otimes |1\rangle\langle 1| \otimes Z) \cdot (|0\rangle\langle 0| \otimes \mathbb{1} \otimes \mathbb{1} + |1\rangle\langle 1| \otimes Z \otimes \mathbb{1}) \\ &= |00\rangle\langle 00| \otimes \mathbb{1} + |01\rangle\langle 01| \otimes Z + |10\rangle\langle 10| \otimes \mathbb{1} - |11\rangle\langle 11| \otimes Z \end{aligned}$$

Thus the order of pairs of qubits we choose to apply the CZ gates does not matter either. Therefore, we can define an undirected graph $G = (V, E)$ where each vertex $v \in V$ is a qubit, and $(u, v) \in E$ if and only if we apply a CZ on the qubits defined by u, v. The state $|G\rangle$ is the one obtained from $|+\rangle^{\otimes|V|}$ by applying the CZ gates whenever an edge is present. If we have two qubits connected with an edge:

$$CZ|+\rangle|+\rangle = \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle - |11\rangle)$$

Now we measure the first qubit in the basis $M(\theta) = \{|0\rangle \pm e^{2i\theta} |1\rangle\}$. This is equivalent to applying the unitary $X^s HR_z(\theta)$ to the second qubit, where s is the outcome of the measurement. Indeed, we can see that:

$$(\langle 0| + (-1)^s e^{-2i\theta} \langle 1|)(|0\rangle |+\rangle + |1\rangle |-\rangle) = |+\rangle + (-1)^s e^{-2i\theta} |-\rangle$$

On the other hand, we have:

$$X^s HR_z(\theta) |+\rangle = X^s(|+\rangle + e^{-2i\theta} |-\rangle) = |+\rangle + (-1)^s e^{-2i\theta} |-\rangle$$

Thus, entangling with a different qubit and measuring with $M(\theta)$ is the same as directly applying $X^s HR_z(\theta)$.

Using a path with four qubits, measuring the first three qubits with $M(\theta_1), M(\theta_2), M(\theta_3)$ gives the same result as the following unitary on the fourth qubit, by the same reasoning as above:

$$U = X^{s_3} HR_z(\theta_3) X^{s_2} HR_z(\theta_2) X^{s_1} HR_z(\theta_1)$$

By using $HR_z H = R_x, HR_x H = R_z, XR_z(\theta) = R_z(-\theta) = X, ZR_x(\theta) = R_x(-\theta)Z$ we obtain:

$$U = X^{s_3} Z^{s_2} X^{s_1} HR_z((-1)^{s_2}\theta_3) R_x((-1)^{s_1}\theta_2) R_z(\theta_1)$$

If we assume $s_1 = s_2 = s_3 = 1$ (which happens with probability 1/8, thus we can repeat the process $\mathcal{O}(1)$ times in expectation until this does not happen), we have:

$$U = XZXHR_z(\theta_3) R_x(\theta_2) R_z(\theta_1)$$

The last three terms form a Z-X decomposition, which we know can represent any single-qubit unitary. With a similar argument we can show that a rectangular grid can achieve an arbitrary n-qubit gate, thus this measurement-based model is complete for quantum computation.

3 Unstructured Search

Problem 3.1 (Unstructured search, single marked element)**.** *Given a function* $f : [N] \to 0, 1$ *as oracle such that:*

$$f(x) = \begin{cases} 1 & x = x_0 \\ 0 & x \neq x_0 \end{cases}$$

Find x_0*.*

Claim 3.2. *Any classical algorithm achieving* $\omega(\frac{1}{N})$ *success probability requires* $q = \Omega(N)$ *queries.*

Proof. Consider an arbitrary randomized algorithm $\mathcal{A}$ making q queries. An input is fully determined by the marked element x_0, thus let X be an element chosen uniformly at random. By Yao's principle:

$$p := \min_{x_0} \mathbb{P}\left(\mathcal{A}(x_0) \text{ succeeds}\right) \leq \max_{a} \mathbb{P}\left(a(X) \text{ succeeds}\right) \leq \frac{q}{N} + \frac{1}{N-q}$$

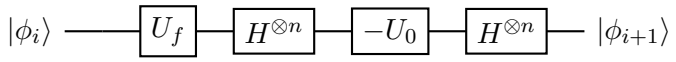
The upper bound on the right is given by a union bound between the probability of finding the marked element among the queried elements and the probability of guessing it among the non-queried ones. If we choose $q = o(N)$, we have that:

$$p \leq o(1) + \frac{1}{N - o(N)} = \mathcal{O}\left(\frac{1}{N}\right)$$

□

3.1 Grover's Algorithm: analysis for single marked element

Assume without loss of generality $N = 2^n$, or extend f to the next power of two so that it returns 0 for $x > N$ (increasing N by at most a factor of 2). We apply the following circuit r times, with $|\phi_0\rangle = |+\rangle^{\otimes n}$:



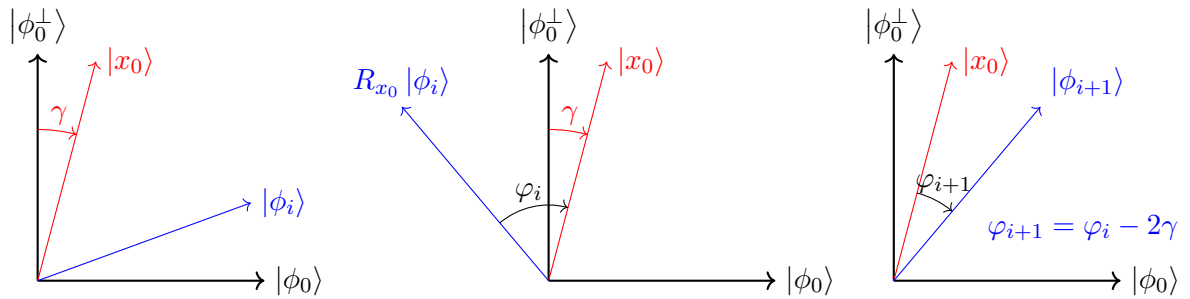
and measure in the computational basis.

Claim 3.3. *For* $r = \Theta(\sqrt{N})$*, Grover's algorithm succeeds with probability* $1 - \mathcal{O}\left(\frac{1}{N}\right)$*.*

Proof. Notice that:

$$U_f = \mathbb{1} - 2\left|x_0\right\rangle\left\langle x_0\right| =: I_{x_0}$$
$$H^{\otimes n} U_0 H^{\otimes n} = \mathbb{1} - 2\left|+\right\rangle^{\otimes n}\left\langle +\right|^{\otimes n} = \mathbb{1} - 2\left|\phi_0\right\rangle\left\langle \phi_0\right| =: I_+$$

where I_ψ geometrically represents the inversion operator with respect to $|\psi\rangle$. Adding a minus we obtain $-I_\psi =: R_\psi$, the reflection operator around $|\psi\rangle$. Now consider $\mathcal{H}^* = \text{span}\{|\phi_0\rangle, |x_0\rangle\}$. Observe that if $|\phi_i\rangle \in \mathcal{H}^*$ then also $I_{x_0}|\phi_i\rangle, I_+|\phi_i\rangle \in \mathcal{H}^*$, thus we can restrict our analysis to this two-dimensional subspace. Let $\left|\phi_0^\perp\right\rangle$ be the state orthogonal to $|\phi_0\rangle$ in $\mathcal{H}^*$.



Therefore, we have that the angle between $|\phi_i\rangle$ and $|x_0\rangle$ decreases each step by a factor 2γ, where

$$\sin\gamma = \langle x_0|\phi_0\rangle = \frac{1}{\sqrt{N}} \Longrightarrow \gamma = \arcsin\frac{1}{\sqrt{N}}$$

Notice that, since $\varphi_0 = \frac{\pi}{2} - \gamma$, we have:

$$\varphi_r = \frac{\pi}{2} - (2r+1)\gamma$$

Thus, if we take

$$r = \left\lfloor \frac{\pi}{4\arcsin\frac{1}{\sqrt{N}}} - \frac{1}{2} \right\rfloor = \Theta(\sqrt{N}) \Longrightarrow 0 \leq \varphi_r \leq 2\gamma$$

And the probability of failure becomes:

$$\mathbb{P}\left(y \neq x_0\right) = \sin^2\varphi_r \leq \sin^2 2\gamma \sim \frac{4}{N}$$

□

3.2 Grover's Algorithm: analysis for multiple known marked elements

Problem 3.4 (Unstructured search, multiple known marked element)**.** *Given a function* $f : [N] \to \{0,1\}$ *as oracle such that:*

$$f(x) = \begin{cases} 1 & x \in S \\ 0 & x \notin S \end{cases}$$

for $S \subseteq [N]$*, with* $|S| = M$*. Find a value in* S*.*

Claim 3.5. *For* $r = \Theta\left(\sqrt{\frac{N}{M}}\right)$*, Grover's algorithm succeeds with probability* $1 - \mathcal{O}\left(\frac{M}{N}\right)$*.*

Proof. Let $|S\rangle = \frac{1}{\sqrt{M}}\sum_{x\in S}|x\rangle$, and consider $\mathcal{H}^* = \text{span}\{|S\rangle, |\phi_0\rangle\}$. Notice that:

$$\begin{aligned} I_S|\phi_0\rangle &= (\mathbb{1} - 2|S\rangle\langle S|)|\phi_0\rangle \\ &= |\phi_0\rangle - 2\left(\frac{1}{M}\sum_{x,y\in S}|x\rangle\langle y|\right)\left(\frac{1}{\sqrt{N}}\sum_x|x\rangle\right) \\ &= |\phi_0\rangle - \frac{2}{\sqrt{N}}\sum_{x\in S}|x\rangle \\ &= (\mathbb{1} - 2\Pi_S)|\phi_0\rangle = U_f|\phi_0\rangle \end{aligned}$$

$$I_S|S\rangle = -|S\rangle = (\mathbb{1} - 2\Pi_S)|S\rangle = U_f|S\rangle$$

Thus, U_f and I_S are equivalent in the subspace $\mathcal{H}^*$, and we can apply the analysis for the single marked element case. If φ_i is the angle between $|\phi_i\rangle$ and $|S\rangle$, then:

$$\begin{cases} \sin\gamma = \langle\phi_0|S\rangle = \sqrt{\frac{M}{N}} \\ \varphi_0 = \frac{\pi}{2} - \gamma \\ \varphi_{i+1} = \varphi_i - 2\gamma \end{cases}$$

Hence, $\varphi_r = \frac{\pi}{2} - (2r+1)\arcsin\sqrt{\frac{M}{N}}$ and

$$r = \left\lfloor \frac{\pi}{4\arcsin\sqrt{\frac{M}{N}}} - \frac{1}{2} \right\rfloor = \Theta\left(\sqrt{\frac{N}{M}}\right) \Longrightarrow 0 \leq \varphi_r \leq 2\gamma$$

This gives

$$|\langle \phi_r | S \rangle|^2 = 1 - \sin^2 \varphi_r \geq 1 - \sin^2 2\gamma \sim 1 - 4\frac{M}{N}$$

and the probability of success is, by the Cauchy-Schwarz inequality:

$$\begin{aligned}\mathbb{P}(y \in S) &= \langle \phi_r | \, \Pi_S \, | \phi_r \rangle = \sum_{x \in S} |\langle x | \phi_r \rangle|^2 \\ &\geq \frac{1}{M} \left| \sum_{x \in S} \langle x | \phi_r \rangle \right|^2 = |\langle S | \phi_r \rangle|^2 = 1 - \mathcal{O}\left(\frac{M}{N}\right)\end{aligned}$$

□

Note that, if $M = \Omega(\sqrt{N})$, the bound on the success probability may loosen too much. However, if M becomes too large (say $M = \Theta(N)$), the trivial algorithm guessing an element at random has $\Omega(1)$ success probability.

3.3 Grover's Algorithm: analysis for multiple unknown marked elements

From the previous section, we have an algorithm which succeeds with probability $\Omega(1)$ for any possible M, provided we know it.

Exponential search. Run this algorithm $\mathcal{O}(\log N)$ times, with $M = 2^m$ up to $m = \lfloor \log_2 N \rfloor$. The total number of queries is:

$$\stackrel{\Theta}{=} \sum_{m=0}^{\lfloor \log_2 N \rfloor} \sqrt{\frac{N}{2^m}} \leq \sqrt{N} \sum_{m=0}^{\infty} \frac{1}{\sqrt{2^m}} = \mathcal{O}(\sqrt{N})$$

Claim 3.6. *Grover's Algorithm with exponential search succeeds with probability* $\Omega(1)$.

Proof. Among all the runs, there must be at least one where:

$$\frac{M}{2} \leq 2^m \leq 2M \Longrightarrow \frac{r(M)}{\sqrt{2}} \leq r(2^m) \leq r(M)\sqrt{2}$$

The probability of success of this run is lower bounded by:

$$\begin{aligned}\left|\left\langle S \middle| \phi_{r(2^m)} \right\rangle\right|^2 &= \sin^2((2r(2^m)+1)\gamma) \\ &= \sin^2\left(\frac{2r(2^m)+1}{2r(M)+1}(2r(M)+1)\gamma\right) \\ &= \sin^2\left(\frac{2r(2^m)+1}{2r(M)+1}\left(\frac{\pi}{2} + \mathcal{O}\left(\sqrt{\frac{M}{N}}\right)\right)\right) \\ &=: \sin^2 \eta\end{aligned}$$

Let us now bound η:

$$\frac{2r(2^m)+1}{2r(M)+1} \sim \frac{r(2^m)}{r(M)} \in \left[\frac{1}{\sqrt{2}}, \sqrt{2}\right]$$

This gives us the following two asymptotic bounds:

$$\frac{\pi}{2\sqrt{2}} + \mathcal{O}\left(\sqrt{\frac{M}{N}}\right) \leq \eta \leq \frac{\pi\sqrt{2}}{2} + \mathcal{O}\left(\sqrt{\frac{M}{N}}\right)$$

Assuming that $M = o(N)$ (otherwise, we would use the trivial algorithm guessing at random to achieve the constant success probability), we have that:

$$\left|\left\langle S \middle| \phi_{r(2^m)} \right\rangle\right|^2 \geq 0.63 - o(1)$$

This concludes the proof. □

3.4 Optimality of Grover's search

Claim 3.7. *Any quantum algorithm for unstructured search over single marked element achieving $\Omega(1)$ success probability requires $\Omega(\sqrt{N})$ queries.*

Proof. Suppose to have an arbitrary quantum algorithm using q queries. This can be represented as a unitary, starting from an initial state $|\psi_0\rangle$

$$\left|\psi_q^{x_0}\right\rangle = G_q U_{x_0} G_{q-1} U_{x_0} \cdots G_1 U_{x_0} |\psi_0\rangle$$

where U_{x_0} is the phase oracle of the function marking x_0. Consider the same algorithm without the oracle calls, i.e.

$$|\psi_q\rangle = G_q G_{q-1} \cdots G_1 |\psi_0\rangle$$

We take the following quantity:

$$D_q = \sum_{x_0 \in [N]} \left\| \left|\psi_q^{x_0}\right\rangle - |\psi_q\rangle \right\|^2$$

Lemma 3.8. $D_q = \mathcal{O}(q^2)$.

Proof. We prove this claim by induction: $q = 0$ is trivial. Assume $D_{q-1} \leq c(q-1)^2$

$$\begin{aligned}
D_q &= \sum_{x_0 \in [N]} \left\| G_q U_{x_0} \left|\psi_{q-1}^{x_0}\right\rangle - G_q |\psi_{q-1}\rangle \right\|^2 \\
&= \sum_{x_0 \in [N]} \left\| U_{x_0} \left|\psi_{q-1}^{x_0}\right\rangle - |\psi_{q-1}\rangle \right\|^2 \\
&= \sum_{x_0 \in [N]} \left\| U_{x_0} \left|\psi_{q-1}^{x_0}\right\rangle - U_{x_0} |\psi_{q-1}\rangle + U_{x_0} |\psi_{q-1}\rangle - |\psi_{q-1}\rangle \right\|^2 \\
&= \sum_{x_0 \in [N]} \left\| U_{x_0} \left(\left|\psi_{q-1}^{x_0}\right\rangle - |\psi_{q-1}\rangle \right) + (U_{x_0} - \mathbb{1}) |\psi_{q-1}\rangle \right\|^2 \\
&\leq \sum_{x_0 \in [N]} \left\| U_{x_0} \left(\left|\psi_{q-1}^{x_0}\right\rangle - |\psi_{q-1}\rangle \right) \right\|^2 + \left\| (U_{x_0} - \mathbb{1}) |\psi_{q-1}\rangle \right\|^2 \\
&+ 2 \left\| U_{x_0} \left(\left|\psi_{q-1}^{x_0}\right\rangle - |\psi_{q-1}\rangle \right) \right\| \left\| (U_{x_0} - \mathbb{1}) |\psi_{q-1}\rangle \right\| \\
&= \sum_{x_0 \in [N]} \left\| \left|\psi_{q-1}^{x_0}\right\rangle - |\psi_{q-1}\rangle \right\|^2 + 4 \left|\langle x_0 | \psi_{q-1} \rangle\right|^2 + 4 \left|\langle x_0 | \psi_{q-1} \rangle\right| \left\| \left|\psi_{q-1}^{x_0}\right\rangle - |\psi_{q-1}\rangle \right\|
\end{aligned}$$

Where we used Cauchy-Schwarz inequality and the fact that $U_{x_0} - \mathbb{1} = -2\left|x_0\right\rangle\left\langle x_0\right|$. By breaking the sum we obtain:

$$\begin{aligned}
D_q &\leq D_{q-1} + 4\sum_{x_0\in[N]} \left|\langle x_0|\psi_{q-1}\rangle\right|^2 + 4\sum_{x_0\in[N]} \left|\langle x_0|\psi_{q-1}\rangle\right| \left\|\left|\psi_{q-1}^{x_0}\right\rangle - \left|\psi_{q-1}\right\rangle\right\| \\
&\leq D_{q-1} + 4\sum_{x_0\in[N]} \left|\langle x_0|\psi_{q-1}\rangle\right|^2 + 4\sqrt{\sum_{x_0\in[N]} \left|\langle x_0|\psi_{q-1}\rangle\right|^2}\sqrt{\sum_{x_0\in[N]} \left\|\left|\psi_{q-1}^{x_0}\right\rangle - \left|\psi_{q-1}\right\rangle\right\|^2} \\
&= D_{q-1} + 4 + 4\sqrt{D_{q-1}}
\end{aligned}$$

If we take $c = 4$, we have $D_q \leq 4(q-1)^2 + 4 + 8(q-1) = 4((q-1)^2 + 2(q-1) + 1) = 4q^2$. $\square$

We show the claim by proving that $D_q = \Omega(N)$. Assume that the algorithm has $\Omega(1)$ success probability, i.e. for any choice of x_0:

$$\left|\left\langle x_0\middle|\psi_q^{x_0}\right\rangle\right|^2 \geq p$$

For some constant p. We can bound D_q as follows:

$$\begin{aligned}
D_q &= \sum_{x_0} \left\|\left|\psi_q^{x_0}\right\rangle - \left|x_0\right\rangle + \left|x_0\right\rangle - \left|\psi_q\right\rangle\right\|^2 \\
&= \sum_{x_0} \left\|\left|\psi_q^{x_0}\right\rangle - \left|x_0\right\rangle + \left|x_0\right\rangle - \left|\psi_q\right\rangle\right\|^2 \\
&\geq \sum_{x_0} \left\|\left|\psi_q^{x_0}\right\rangle - \left|x_0\right\rangle\right\|^2 + \left\|\left|x_0\right\rangle - \left|\psi_q\right\rangle\right\|^2 - 2\left\|\left|\psi_q^{x_0}\right\rangle - \left|x_0\right\rangle\right\| \left\|\left|x_0\right\rangle - \left|\psi_q\right\rangle\right\| \\
&= \sum_{x_0} \left\|\left|\psi_q^{x_0}\right\rangle - \left|x_0\right\rangle\right\|^2 + \sum_{x_0} \left\|\left|x_0\right\rangle - \left|\psi_q\right\rangle\right\|^2 - 2\sum_{x_0} \left\|\left|\psi_q^{x_0}\right\rangle - \left|x_0\right\rangle\right\| \left\|\left|x_0\right\rangle - \left|\psi_q\right\rangle\right\| \\
&\geq \sum_{x_0} \left\|\left|\psi_q^{x_0}\right\rangle - \left|x_0\right\rangle\right\|^2 + \sum_{x_0} \left\|\left|x_0\right\rangle - \left|\psi_q\right\rangle\right\|^2 - 2\sqrt{\sum_{x_0} \left\|\left|\psi_q^{x_0}\right\rangle - \left|x_0\right\rangle\right\|^2}\sqrt{\sum_{x_0} \left\|\left|x_0\right\rangle - \left|\psi_q\right\rangle\right\|^2} \\
&=: E_q + F_q - 2\sqrt{E_q F_q} = \left(\sqrt{F_q} - \sqrt{E_q}\right)^2
\end{aligned}$$

We now bound these two sums:

$$E_q = \sum_{x_0} \left\|\left|\psi_q^{x_0}\right\rangle - \left|x_0\right\rangle\right\|^2 = \sum_{x_0} 2 - 2\left\langle\psi_q^{x_0}\middle|x_0\right\rangle \leq (2 - 2\sqrt{p})N = 2N(1 - \sqrt{p})$$

Here we assumed without loss of generality that the inner products are real (otherwise we can add a global phase to each $\left|x_0\right\rangle$ in the whole argument so that this is always true).

$$\begin{aligned}
F_q &= \sum_{x_0} \left\|\left|\psi_q\right\rangle - \left|x_0\right\rangle\right\|^2 \\
&= \sum_{x_0} \left\|\left|\psi_q\right\rangle\right\|^2 - \left\|\left|x_0\right\rangle\right\|^2 - \langle x_0|\psi_q\rangle - \langle\psi_q|x_0\rangle \\
&= 2N - \left(\sum_{x_0}\langle x_0|\right)|\psi_q\rangle - \langle\psi_q|\left(\sum_{x_0}|x_0\rangle\right) \\
&\geq 2N - 2\left\|\sum_{x_0}|x_0\rangle\right\| \left\|\left|\psi_q\right\rangle\right\| = 2N - 2\sqrt{N} = 2N\left(1 - \frac{1}{\sqrt{N}}\right)
\end{aligned}$$

This gives the final bound:

$$\begin{aligned} D_q &\geq 2N\left(\sqrt{1-\frac{1}{\sqrt{N}}}-\sqrt{1-\sqrt{p}}\right)^2 \\ &\geq 2N\left(\sqrt{1-\frac{1}{\sqrt{N}}}-1+\frac{\sqrt{p}}{2}\right)^2 && \text{using } \sqrt{1-x}\leq 1-\frac{x}{2} \\ &\sim 2N\left(\frac{\sqrt{p}}{2}\right)^2 = \Omega(pN) = \Omega(N) \end{aligned}$$

□

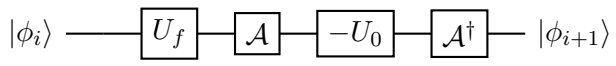
3.5 Amplitude amplification

Suppose that we have a heuristic $\mathcal{A}$ such that:

$$\mathcal{A}\left|0\right\rangle^{\otimes n} = \sum_x \alpha_x \left|x\right\rangle =: \left|\alpha\right\rangle$$

Let $p = \sum_{x\in S}|\alpha_x|^2$. Notice that a classical probability amplification uses $\mathcal{O}\left(\frac{1}{p}\right)$ independent runs of $\mathcal{A}$ and yields an element of S with probability $\Omega(1)$.

We generalize Grover's algorithm as follows:



setting $\left|\phi_0\right\rangle = \left|\alpha\right\rangle$ and then measuring $\left|\phi_r\right\rangle$ with the computational basis.

Claim 3.9. *For $r = \mathcal{O}\left(\frac{1}{\sqrt{p}}\right)$, the amplitude amplification algorithm succeeds with probability $1-\mathcal{O}(p)$.*

Proof. We follow the same argument as in Grover's algorithm for multiple marked elements. Let $\left|S\right\rangle = \frac{1}{\sqrt{p}}\sum_{x\in S}\alpha_x\left|x\right\rangle$.

$$\mathcal{A}U_0\mathcal{A}^\dagger = \mathbb{1} - 2\left|\alpha\right\rangle\left\langle\alpha\right| = I_\alpha$$

considering the subspace $\mathcal{H}^* = \text{span}\{\left|\alpha\right\rangle, \left|S\right\rangle\}$[1], and denoting φ_i the angle between $\left|S\right\rangle$ and $\left|\phi_i\right\rangle$:

$$\begin{cases} \sin\gamma = \langle\phi_0|S\rangle = \sqrt{p} \\ \varphi_0 = \frac{\pi}{2} - \gamma \\ \varphi_{i+1} = \varphi_i - 2\gamma \end{cases}$$

giving $\varphi_r = \frac{\pi}{2} - (2r+1)\gamma$ and

$$r = \left\lfloor \frac{\pi}{4\arcsin\sqrt{p}} - \frac{1}{2}\right\rfloor = \Theta\left(\frac{1}{\sqrt{p}}\right) \Longrightarrow 0 \leq \varphi_r \leq 2\gamma$$

This gives a success probability of at least:

$$|\langle\phi_r|S\rangle|^2 = 1-\sin^2\varphi_r \geq 1-\sin^2 2\gamma \sim 1-4p$$

□

[1]These two vectors may be linearly dependent, but this is certainly not true if $\mathcal{A}$ fails at least once.

3.6 Oblivious Amplitude Amplification

The algorithm above for amplitude amplification requires us to know how to construct the state $|\alpha\rangle$ in order to carry out an iteration of the algorithm. Sometimes, the success of the algorithm is not given directly by the state, but rather it is written on some ancilla register. For example, whatever state we have in our main register is a good state for us, as long as the ancilla register is in the all-zero state. Here we show that it is possible to amplify the amplitude of such component.

Theorem 3.10 (Oblivious Amplitude Amplification)**.** *Suppose that a unitary U acts as:*

$$U|0\rangle^{\otimes\mu}|\psi\rangle = \sin(\theta)|0\rangle^{\otimes\mu}|\phi\rangle + \cos(\theta)\left|\Phi^{\perp}\right\rangle$$

for some $\theta \in [0, \pi/2]$, where $(\langle 0|^{\otimes\mu} \otimes \mathbb{1})\left|\Phi^{\perp}\right\rangle = 0$. Let $R = 2(|0\rangle^{\otimes\mu}\langle 0|^{\otimes\mu} \otimes \mathbb{1}) - \mathbb{1}$ be a reflection about the target subspace. Then we have, for any $\ell \in \mathbb{Z}$:

$$(-URU^{\dagger}R)^{\ell}|\psi\rangle = \sin((2\ell+1)\theta)|0\rangle^{\otimes\mu}|\phi\rangle + \cos((2\ell+1)\theta)\left|\Phi^{\perp}\right\rangle$$

Proof. The state $\left|\Psi^{\perp}\right\rangle$ defined by the equation:

$$U\left|\Psi^{\perp}\right\rangle = \cos(\theta)|0\rangle^{\otimes\mu}|\phi\rangle - \sin(\theta)\left|\Phi^{\perp}\right\rangle$$

is orthogonal to $|0\rangle^{\otimes\mu}|\psi\rangle$ (check!). Moreover, it is possible to show that $(\langle 0|^{\otimes\mu}\otimes\mathbb{1})\left|\Psi^{\perp}\right\rangle = 0$ (see original paper). Therefore, U acts as a two-dimensional rotation like in the original amplitude amplification settings, except for the fact that the input and output subspaces are different. Defining $|\Psi\rangle = |0\rangle^{\otimes\mu}|\psi\rangle, |\Phi\rangle = |0\rangle^{\otimes\mu}|\phi\rangle$:

$$\begin{aligned} U|\Psi\rangle &= \sin(\theta)|\Phi\rangle + \cos(\theta)\left|\Phi^{\perp}\right\rangle \\ U\left|\Psi^{\perp}\right\rangle &= \cos(\theta)|\Phi\rangle - \sin(\theta)\left|\Phi^{\perp}\right\rangle \end{aligned}$$

and, similarly

$$\begin{aligned} U^{\dagger}|\Phi\rangle &= \sin(\theta)|\Psi\rangle + \cos(\theta)\left|\Psi^{\perp}\right\rangle \\ U^{\dagger}\left|\Phi^{\perp}\right\rangle &= \cos(\theta)|\Psi\rangle - \sin(\theta)\left|\Psi^{\perp}\right\rangle \end{aligned}$$

Therefore, the constructed unitary acts as:

$$\begin{aligned} -URU^{\dagger}R|\Phi\rangle &= \sin(2\theta)|\Phi\rangle + \cos(2\theta)\left|\Phi^{\perp}\right\rangle \\ -URU^{\dagger}R\left|\Phi^{\perp}\right\rangle &= \cos(2\theta)|\Phi\rangle - \sin(2\theta)\left|\Phi^{\perp}\right\rangle \end{aligned}$$

and the standard amplitude amplification argument concludes the proof. □

4 The QFT Paradigm

4.1 Quantum Fourier Transform

We want to construct the following operation, for some natural number N:

$$Q_N : |x\rangle \mapsto \frac{1}{\sqrt{N}} \sum_{y=0}^{N-1} \omega_N^{xy} |y\rangle$$

where $\omega_N = e^{2\pi i/N}$ is the N-th root of unity.

Claim 4.1. *Q_N is unitary.*

Proof.

$$\begin{aligned} Q_N^\dagger Q_N |x\rangle &= \frac{1}{\sqrt{N}} \sum_{y=0}^{N-1} \omega_N^{xy} Q_N^\dagger |y\rangle \\ &= \frac{1}{\sqrt{N}} \sum_{y=0}^{N-1} \omega_N^{xy} \left(\frac{1}{\sqrt{N}} \sum_{x'=0}^{N-1} (\omega_N^{yx'})^* \left|x'\right\rangle \right) \\ &= \frac{1}{N} \sum_{x'=0}^{N-1} \sum_{y=0}^{N-1} \omega_N^{y(x-x')} \left|x'\right\rangle \\ &=: \frac{1}{N} \sum_{x'=0}^{N-1} c_{x,x'} \left|x'\right\rangle \end{aligned}$$

Notice that:

$$c_{x,x'} = \begin{cases} \sum_{y=0}^{N} 1 = N & x = x' \\ \frac{e^{2\pi i(x-x')}-1}{e^{2\pi i(x-x')/N}-1} = 0 & x \neq x' \end{cases}$$

This means that $\frac{1}{N} c_{x,x'} = \delta_{x,x'}$ is the Kronecker delta, and the above expression becomes $|x\rangle$, giving $Q_N^\dagger Q_N |x\rangle = \mathbb{1} |x\rangle$ for any element of the computational basis. A similar argument holds for $Q_N Q_N^\dagger$. □

4.2 The QFT algorithm

We want to implement Q_N, for $N = 2^n$. We decompose the elements:

$$|y\rangle = |y_{n-1}\rangle \otimes \cdots \otimes |y_0\rangle$$

Hence, $y = \sum_k y_k 2^k$

$$\begin{aligned} Q_N |x\rangle &= \frac{1}{\sqrt{N}} \sum_{y=0}^{N-1} \omega_N^{xy} |y\rangle = \frac{1}{\sqrt{N}} \sum_{y=0}^{N-1} \omega_N^{x \sum_k y_k 2^k} |y\rangle \\ &= \frac{1}{\sqrt{N}} \sum_{y=0}^{N-1} \bigotimes_{k=0}^{n-1} \omega_N^{x y_{n-k} 2^{n-k}} |y_{n-k}\rangle = \frac{1}{\sqrt{N}} \sum_{y=0}^{N-1} \bigotimes_{k=0}^{n-1} \omega_{2^k}^{x y_{n-k}} |y_{n-k}\rangle \\ &= \frac{1}{\sqrt{N}} \bigotimes_{k=0}^{n-1} \sum_{y_{n-k} \in \{0,1\}} \omega_{2^k}^{x y_{n-k}} |y_{n-k}\rangle = \bigotimes_{k=0}^{n-1} \frac{1}{\sqrt{2}} \left(|0\rangle + \omega_{2^k}^x |1\rangle \right) \end{aligned}$$

We can rewrite $\omega_{2^k}^x$:

$$\omega_{2^k}^x = \omega_1^{x/2^k} = \omega_1^{\sum_{j=0}^{k-1} x_j 2^{j-k}} =: \omega_1^{(.x_{k-1}\cdots x_0)}$$

More explicitly, $(.a_{k-1}\cdots a_0) = \frac{a_{k-1}}{2^1} + \cdots + \frac{a_0}{2^k}$. We define a gate R_d as

$$R_d = \begin{bmatrix} 1 & 0 \\ 0 & e^{\pi i/2^d} \end{bmatrix}$$

On the other hand, one can see that the Hadamard gate applies the following transformation:

$$H|b\rangle \propto |0\rangle + e^{\pi i b}|1\rangle = |0\rangle + e^{2\pi i(.b)}|1\rangle$$

Thus, the idea of the algorithm is to start from the most significant qubit $|x_{n-1}\rangle$:

- Apply H, so to add a phase of $\pi i x_{n-1}$;
- Apply a R_k controlled by $|x_{n-1-k}\rangle$. If this qubit is in the state $|0\rangle$, nothing changes, otherwise a phase of $2\pi i/2^{k+1}$ is added. In general, we have

$$\left(\bigwedge R_k\right)|x_{n-1-k}\rangle\left(|0\rangle + e^{\pi i\phi}|1\rangle\right) \propto |x_{n-1-k}\rangle\left(|0\rangle + e^{\pi i(\phi + \frac{x_{n-1-k}}{2^k})}|1\rangle\right)$$

- This gives the last qubit $|y_0\rangle$ (remember to invert the order with SWAP operators at the end!).

$$|y_0\rangle \propto |0\rangle + e^{\pi i\left(\sum_{k=0}^{n-1}\frac{x_{n-1-k}}{2^k}\right)}|1\rangle = |0\rangle + e^{2\pi i(.x_{n-1}\cdots x_0)}|1\rangle$$

  We can exclude $|x_{n-1}\rangle$ and apply this algorithm recursively to $|x_{n-2}\rangle\cdots|x_0\rangle$.

This algorithm yields the quantum Fourier transform in $\mathcal{O}(\log^2 N)$ time and $\mathcal{O}(\log^2 N)$ total work. *Dovrei disegnare il circuito qui ma un c'ho vojaaaa*

4.3 $\mathcal{O}(1/n^c)$-approximation in $\mathcal{O}(n\log n)$

In order to drop the time complexity from $\mathcal{O}(n^2)$ to $\mathcal{O}(n\log n)$ we can simply omit all the controlled R_d gates, for $d \geq k$, $k = \Theta(\log n)$. Each qubit will only pass $\mathcal{O}(k)$ gates, having a total $\mathcal{O}(nk) = \mathcal{O}(n\log n)$ time. We want to argue that this yields a state which has $1-\mathcal{O}(1/n^c)$ overlap with the real Fourier-transformed state. If Q_N^k is the approximatin circuit, the approximation error is certainly bounded by:

$$D(Q_N, Q_N^k) \leq n\sum_{i=0}^{\infty} D(\bigwedge R_{k+i}, \mathbb{1}) = n\sum_{i=0}^{\infty} D(R_{k+i}, \mathbb{1})$$

Now notice that:

$$\begin{aligned} D(R_d, \mathbb{1}) &= \max_{|\alpha|^2+|\beta|^2=1} ||(R_d - \mathbb{1})(\alpha|0\rangle + \beta|1\rangle)||_2 \\ &= \max_{|\alpha|^2+|\beta|^2=1} ||(e^{i\pi/2^d} - 1)\beta|1\rangle||_2 \\ &= |e^{i\pi/2^d} - 1| = |e^{i\pi/2^{d+1}}(e^{i\pi/2^{d+1}} - e^{-i\pi/2^{d+1}})| \\ &= 2\sin\frac{\pi}{2^{d+1}} \sim \frac{\pi}{2^d} = \mathcal{O}(2^{-d}) \end{aligned}$$

This implies that:

$$D(Q_N, Q_N^k) = n\sum_{i=0}^{\infty}\mathcal{O}\left(\frac{1}{2^{k+i}}\right) = \mathcal{O}\left(\frac{n}{2^k}\right)$$

Taking $k = (c+1)\log n$, the error is bounded by $\mathcal{O}(1/n^c)$.

4.4 The Period Finding Algorithm

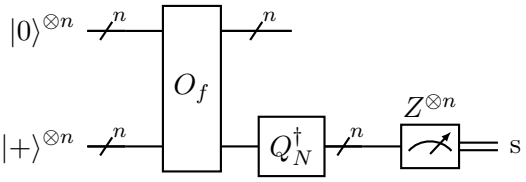
Problem 4.2 (Period Finding)**.** *Given $f : \mathbb{Z}_N \to \mathbb{Z}_N$ as oracle, find the period r such that:*

(i) $f(x) = f(x + r)$ for any $x \in \mathbb{Z}_N$;

(ii) $f(x) \neq f(x + s)$ for any $0 < s < r$, i.e. $f|_{\mathbb{Z}_r}$ is injective.

Condition (ii) in particular ensures the minimality (and thus the uniqueness) of the period.

We will assume $r|N$ for simplicity.



Run the above circuit two times: this gives two values s, s'. Using the Euclidean algorithm, simplify the fractions $\frac{s}{N} \mapsto \frac{a}{b}, \frac{s'}{N} \mapsto \frac{a'}{b'}$ so that $a \perp b, a' \perp b'$. Then return the least common multiple of b and b'. The whole algorithm runs in $\mathcal{O}(T_f(n) + n^2)$ time, where $T_f(n)$ is the time required to compute O_f.

Lemma 4.3. *The period finding circuit yields $s = k\frac{N}{r}$ for some k with probability 1.*

Proof. Let $|\phi\rangle$ be the final state right before measuring.

$$\begin{aligned}
|\phi\rangle &= (\mathbb{1} \otimes Q_N^\dagger) O_f |0\rangle^{\otimes n} |+\rangle^{\otimes n} \\
&= (\mathbb{1} \otimes Q_N^\dagger) \frac{1}{\sqrt{N}} \sum_x |f(x)\rangle |x\rangle \\
&= (\mathbb{1} \otimes Q_N^\dagger) \frac{1}{\sqrt{N}} \sum_z \sum_{x:f(x)=z} |z\rangle |x\rangle \\
&= \frac{1}{\sqrt{N}} \sum_z |z\rangle \sum_{x:f(x)=z} Q_N^\dagger |x\rangle \\
&= \frac{1}{\sqrt{N}} \sum_z |z\rangle \sum_{t=0}^{N/r-1} Q_N^\dagger |x_z + rt\rangle \\
&= \frac{1}{N} \sum_z |z\rangle \sum_{t=0}^{N/r-1} \sum_{y=0}^{N-1} \omega_N^{-(x_z+rt)y} |y\rangle \\
&= \frac{1}{N} \sum_z |z\rangle \sum_{y=0}^{N-1} \omega_N^{-x_z y} \sum_{t=0}^{N/r-1} \omega_N^{-rty} |y\rangle \\
&=: \frac{1}{N} \sum_z |z\rangle \sum_{y=0}^{N-1} \omega_N^{-x_z y} c_y |y\rangle
\end{aligned}$$

Now notice that:

$$c_y = \sum_{t=0}^{N/r-1} e^{-2\pi i r t y/N} = \begin{cases} \sum_t 1 = \frac{N}{r} & N \mid ry \\ \frac{e^{-2\pi i t y}-1}{e^{-2\pi i r t y/N}-1} = 0 & N \nmid ry \end{cases}$$

This gives us:

$$|\phi\rangle = \frac{1}{r} \sum_z |z\rangle \sum_{y:N|ry} \omega_N^{-x_z y} |y\rangle$$

Notice that the values of z in this sum are exactly r, since over a period, all the elements of a function must be distinct. Measuring using $\mathbb{1} \otimes Z^{\otimes n}$ yields the following probability:

$$\mathbb{P}(s = y) = \begin{cases} \frac{1}{r} & N \mid ry \\ 0 & N \nmid ry \end{cases}$$

This concludes the proof. □

Also notice that the probability distribution is uniform among the values of the form $k\frac{N}{r}$.

Lemma 4.4. *Two independent runs of the period finding algorithm yield $s = k\frac{N}{r}, s' = k'\frac{N}{r}$ with $k \perp k'$ with probability $\Omega(1)$.*

Proof. Since $s, s' \in [N]$, we must have $k, k' \sim \mathcal{U}[r]$.

$$\begin{aligned} \mathbb{P}\left(k \not\perp k'\right) &\leq \sum_{p \in \mathbb{P}} \mathbb{P}\left(p|k \wedge p|k'\right) \\ &= \sum_{p \in \mathbb{P}} \mathbb{P}\left(p|k\right)^2 \\ &\leq \sum_{p \in \mathbb{P}} \left(\frac{r}{p} \cdot \frac{1}{r}\right)^2 && \text{Union bound, } p \text{ divides} \leq \frac{r}{p} \text{ numbers in } [r] \\ &\leq \sum_{p=2}^{\infty} \frac{1}{p^2} = \frac{\pi^2}{6} - 1 \leq 0.7 \end{aligned}$$

□

Lemma 4.5. *$r \geq lcm(b, b')$. If k, k' are coprime, then the claim holds with equality.*

Proof. Since $\frac{k}{r} = \frac{s}{N} = \frac{a}{b}$ and thus $bk = ar$, we must have that $b|r$ (since $a \perp b$). The same holds for b', implying r is a common multiple of b, b'.

Now let us assume $k \perp k'$: suppose for a contradiction that $r > \text{lcm}(b, b')$: this means there is a factor q of r that does not divide neither b nor b'. Since we have that

$$\begin{cases} ar = bk \\ ar' = b'k' \end{cases}$$

q divides both bk and $b'k'$, i.e. divides both k, k' contradicting their co-primality. □

Claim 4.6. *The period finding algorithm retrieves r with probability $\Omega(1)$.*

Notice that this is a Las Vegas algorithm: we can prove the correctness in $\mathcal{O}(T_f(n) + n)$ by checking $f(0) \overset{!}{=} f(r)$, thus without changing the asymptotic complexity (the output of the algorithm cannot be higher than the real period, as proven above).

4.5 Shor's factoring algorithm

Problem 4.7 (Factorization)**.** *Given a composite integer N, output one of its non-trivial factors.*

The above algorithm runs in $\mathcal{O}(\log^2 N \log\log N)$

- Line 2 can be done in $\mathcal{O}(\log^2 N)$;
- Line 3 runs in expected $\mathcal{O}\left(\log^2 N \frac{N}{\varphi(N)}\right) = \mathcal{O}\left(\log^2 N \log\log N\right)$;

Algorithm 1 SHOR'S FACTORING ALGORITHM

Input: A composite number N
Output: $x \in \mathbb{N}$ such that $x \mid N$ and $1 < x < N$

1: **if** N is even **then return** 2
2: **if** $\sqrt[i]{N} \in \mathbb{N}$ for some $i \in [\lfloor \log_2 N \rfloor]$ **then return** $\sqrt[i]{N}$
3: Choose $a \sim \mathcal{U}(\mathbb{Z}_N^*)$
4: $r \leftarrow$ PERIOD-FINDING$(x \mapsto a^x \bmod N)$
5: **if** r is odd or $a^{r/2} \equiv_N -1$ **then** declare failure
6: **return** $gcd(a^{r/2} + 1, N)$

- The period finding algorithm runs in expected $\mathcal{O}(\log^2 N)$, since the function $x \mapsto a^x \bmod N$ can be computed in $\mathcal{O}(\log^2 N)$;
- The greatest common divisor is computed in $\mathcal{O}(\log^2 N)$ using the Euclidean algorithm.

Now we analyze the success probability. Notice that r is the multiplicative order of a modulo N. We prove the following:

Lemma 4.8. *Let a be a randomly chosen totative of an odd composite number $N = p_1^{\alpha_1} \cdots p_m^{\alpha_m}$, and let r be its multiplicative order in $\mathbb{Z}_N^*$. Then with probability $1 - \frac{2}{2^m}$ we have that:*

- *r is even;*
- *$a^{r/2} + 1$ and $a^{r/2} - 1$ are not multiples of N.*

Proof. Suppose $m = 1$ for now. This means that $\varphi(N) = p^{\alpha-1}(p-1)$ is necessarily even, and we can rewrite it as $\varphi(N) = c2^d$ for some odd c. Since N is a power of a prime, $\mathbb{Z}_N^*$ is a cyclic group. Thus taking a generator g we have:

$$a^r \equiv_N g^{k_a r} \equiv_N 1$$

implying that $\varphi(N) = c2^d$ divides $k_a r$. Since g is a generator for $\mathbb{Z}_N^*$, we have half of the elements yielding an odd k, and this implies that r is even.

For $m > 1$ we use the Chinese Remainder Theorem: taking a random totative is equivalent to taking random $x_i \sim \mathcal{U}(\mathbb{Z}_{p_i^{\alpha_i}})$, and the CRT will unambiguously determine x. If $x_i \equiv_{p_i^{\alpha_i}} x$ and r_i is its multiplicative order in $\mathbb{Z}_{p_i^{\alpha_i}}$ we have that

$$x^{r/2} \equiv_N -1 \implies x^{r/2} \equiv_{p_i^{\alpha_i}} -1 \implies x_i^{r/2} \equiv_{p_i^{\alpha_i}} -1$$

Since for each x_i this happens with probability $\leq \frac{1}{2}$, $x^{r/2} + 1$ would be a multiple of N with probability $\leq \frac{1}{2^m}$. The same applies with the parity of each r_i. *Non c'hovvoja di formalizzarlo meglio* □

Claim 4.9. *Shor's factoring algorithm succeeds with probability $\Omega(1)$.*

Proof. If we end up with r even and $a^{r/2} + 1$ which is not a multiple of N, then:

$$kN = a^r - 1 = (a^{r/2} - 1)(a^{r/2} + 1)$$

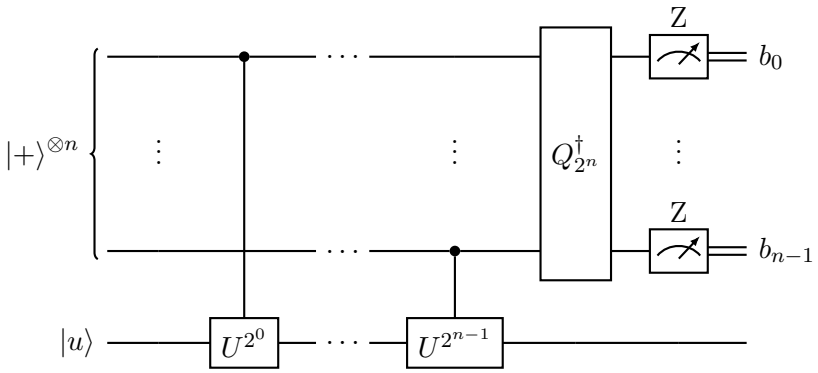
Since N does not divide $a^{r/2} + 1$, we have $a^{r/2} + 1 = k'N'$, where $k' \mid k, N' \mid N$ with N' a non-trivial factor of N, and the greatest common divisor should return it. □

4.6 Phase estimation

Problem 4.10 (Phase estimation)**.** *Given a unitary U and one of its eigenstates $|u\rangle$*

$$U|u\rangle = e^{i2\pi\varphi}|u\rangle$$

Estimate φ.



Claim 4.11. *Suppose $\varphi = \frac{\ell}{2^n}$. The phase estimation algorithm will return ℓ with probability* 1*.*

Proof. Notice that:

$$\left(\bigwedge U^k\right)|0\rangle|u\rangle = |0\rangle|u\rangle$$
$$\left(\bigwedge U^k\right)|1\rangle|u\rangle = e^{2\pi k\varphi}|1\rangle|u\rangle$$

which means, in general:

$$\left(\bigwedge U^k\right)|b\rangle|u\rangle = e^{2\pi b\varphi}|b\rangle|u\rangle$$

This implies that, if we have the value $|x\rangle$ in the ancilla qubits, right before applying the inverse Fourier transform:

$$\begin{aligned}|x\rangle|u\rangle &\mapsto \left(\bigotimes_{j=0}^{n-1} e^{2\pi x_j 2^j\varphi}|x_j\rangle\right)\otimes|u\rangle \\ &= e^{2\pi(\sum_j x_j 2^j)\varphi}|x\rangle|u\rangle = e^{2\pi x\varphi}|x\rangle|u\rangle\end{aligned}$$

Since we have $|+\rangle^{\otimes n}$ as input, the state before applying the inverse Fourier transform is:

$$\left(\frac{1}{\sqrt{2^n}}\sum_{x=0}^{2^n-1} e^{2\pi x\varphi}|x\rangle\right)\otimes|u\rangle = \left(\frac{1}{\sqrt{2^n}}\sum_{x=0}^{2^n-1} e^{2\pi x\ell/2^n}|x\rangle\right)\otimes|u\rangle$$

Applying the inverse Fourier transform gives

$$\begin{aligned}Q^\dagger_{2^n}\left(\frac{1}{\sqrt{2^n}}\sum_{x=0}^{2^n-1} e^{2\pi x\ell/2^n}|x\rangle\right)\otimes|u\rangle &= \left(\frac{1}{\sqrt{2^n}}\sum_{x=0}^{2^n-1} e^{2\pi x\ell/2^n}Q^\dagger_{2^n}|x\rangle\right)\otimes|u\rangle \\ &= \left(\frac{1}{2^n}\sum_{j=0}^{2^n-1}\sum_{x=0}^{2^n-1} e^{2\pi x(\ell-j)/2^n}|j\rangle\right)\otimes|u\rangle \\ &= |\ell\rangle\otimes|u\rangle\end{aligned}$$

□

Claim 4.12. *Suppose $\varphi = \frac{\ell+\delta}{2^n}$, for some $\delta \in [-\frac{1}{2}, \frac{1}{2}]$. The phase estimation algorithm will return ℓ' such that:*

$$|\ell' - \ell| > d$$

with probability at most $\frac{1}{2(d-1)}$.

Proof. With a derivation similar to the above, the coefficients of the final state are:

$$\alpha_j = \frac{1}{2^n} \sum_{k=0}^{2^n-1} e^{2\pi i k(2^n\varphi - j)/2^n} = \frac{1}{2^n} \sum_{k=0}^{2^n-1} e^{2\pi i k(\ell+\delta-j)/2^n} = \frac{1}{2^n} \frac{e^{2\pi i(\ell+\delta-j)} - 1}{e^{2\pi i(\ell+\delta-j)/2^n} - 1}$$

Thus the probability of returning j is:

$$|\alpha_j|^2 = \frac{1}{2^{2n}} \frac{|e^{2\pi i(\ell+\delta-j)} - 1|^2}{|e^{2\pi i(\ell+\delta-j)/2^n} - 1|^2}$$

We now rewrite $j = \ell - j'$, so that j' represents the distance of j from the best estimation ℓ. Moreover, we take advantage of modularity to let j' go from -2^{n-1} to $+2^{n-1}$.

$$\begin{aligned}
\mathbb{P}(|j-\ell| > d) &= \sum_{p<|j'|\leq 2^{n-1}} |\alpha_j|^2 \\
&= \frac{1}{2^{2n}} \sum_{p<|j'|\leq 2^{n-1}} \frac{|e^{2\pi i(j'+\delta)} - 1|^2}{|e^{2\pi i(j'+\delta)/2^n} - 1|^2} \\
&\leq \frac{1}{2^{2n}} \sum_{p<|j'|\leq 2^{n-1}} \frac{4}{|e^{2\pi i(j'+\delta)/2^n} - 1|^2} && \text{since } |e^{i\theta} - 1| \leq 2 \\
&\leq \sum_{p<|j'|\leq 2^{n-1}} \frac{1}{4|j'+\delta|^2} && \text{since } |e^{i\theta} - 1| \geq \frac{2|\theta|}{\pi} \\
&\leq 2 \sum_{t=d+1}^{2^{n-1}} \frac{1}{4|t+\delta|^2} \leq \sum_{t=d+1}^{2^{n-1}} \frac{1}{2(t-\frac{1}{2})^2} && \text{since } \delta \geq \frac{1}{2} \\
&\leq \int_d^\infty \frac{1}{2(t-\frac{1}{2})^2}\, dt = \frac{1}{2(d-\frac{1}{2})}
\end{aligned}$$

□

4.7 Shor's factoring algorithm from phase estimation

In Shor's factoring algorithm, we use period finding to solve the order finding problem, and then we use the classical, well-known reduction from factorization to order finding.

Here we compute the order of a totative a of a number M in a different way. Define a unitary U such that:

$$U|x\rangle = \begin{cases} |xa \bmod M\rangle & x < M \\ |x\rangle & x \geq M \end{cases}$$

One can see that this is a unitary, as $U^\dagger$ would simply multiply by a^{-1}, which exists and is unique since a is a totative. We want to find its eigenstates: by the definition above, we have that $|x\rangle$ are eigenstates with eigenvalue 1 in the subspace spanned by $x \geq M$, but we are only interested in the other subspace.

Claim 4.13. *The following are eigenstates of U, for every $s \in [r]$:*

$$|u_s\rangle = \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} e^{-\frac{2\pi i s k}{r}} \left| a^k \bmod M \right\rangle$$

with eigenvalue $e^{2\pi i s/r}$.

Proof. Let $|u_s\rangle = \sum \alpha_x |x\rangle$ be an eigenstate of U. Therefore:

$$\begin{aligned}
0 &= (U - e^{2\pi i s/r}\mathbb{1}) \sum_x \alpha_x |x\rangle \\
&= \sum_x \alpha_x \left(U|x\rangle - e^{2\pi i s/r}|x\rangle \right) \\
&= \sum_x \alpha_x \left(|ax \bmod M\rangle - e^{2\pi i s/r}|x\rangle \right) \\
&= \sum_x \alpha_x |ax \bmod M\rangle - \sum_x \alpha_x e^{2\pi i s/r}|x\rangle \\
&= \sum_x \alpha_{a^{-1}x \bmod M}|x\rangle - \sum_x \alpha_x e^{2\pi i s/r}|x\rangle \\
&= \sum_x \left(\alpha_{a^{-1}x \bmod M} - \alpha_x e^{2\pi i s/r} \right)|x\rangle
\end{aligned}$$

By linear independence of the basis states, we conclude that this holds if and only if:

$$\begin{aligned}
\alpha_x &= \alpha_{a^{-1}x \bmod M} e^{-2\pi i s/r} \\
\alpha_{ax \bmod M} &= \alpha_x e^{-2\pi i s/r}
\end{aligned}$$

which means:

$$\alpha_{a^k \bmod M} = \alpha_1 e^{-2\pi i s k/r}$$

while the choice of α_1 would only give a global phase. Therefore, after normalization, and choosing $\alpha_1 = 1$, we get our eigenstate:

$$|u_s\rangle = \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} e^{-2\pi i s k/r} \left| a^k \bmod M \right\rangle$$

□

Claim 4.14. $|1\rangle = \frac{1}{\sqrt{r}} \sum_s |u_s\rangle$

Proof.

$$\begin{aligned}
\frac{1}{\sqrt{r}} \sum_s |u_s\rangle &= \frac{1}{r} \sum_s \sum_{k=0}^{r-1} e^{-\frac{2\pi i s k}{r}} \left| a^k \bmod M \right\rangle \\
&= \frac{1}{r} \sum_{k=0}^{r-1} \sum_s e^{-\frac{2\pi i s k}{r}} \left| a^k \bmod M \right\rangle \\
&= \frac{1}{r} \sum_{k=0}^{r-1} r \cdot \delta_{k,0} \left| a^k \bmod M \right\rangle = |1\rangle
\end{aligned}$$

□

Therefore, if we input $|1\rangle$ to the phase estimation algorithm (taking n such that $2^n \geq M$) we obtain a final state of:

$$\frac{1}{\sqrt{r}} \sum_s |u_s\rangle \mapsto \frac{1}{\sqrt{r}} \sum_s \left| \frac{s}{r} \cdot 2^n \right\rangle$$

and measuring gives exactly the same result as the period finding algorithm. The above assumes that r divides 2^n, but in case it does not, the phase estimation algorithm will return something that is as close as possible to one of these values.

4.8 Grover's search algorithm from phase estimation

We can use the phase estimation algorithm to find the number M of marked elements in an unstructured search setting. By doing a Monte Carlo sampling, i.e. try q random choices of x and count the number l of the ones who give $f(x) = 1$, we get an estimate $\bar{M} = \frac{\ell N}{q}$ such that:

$$\mathbb{P}\left(|\bar{M} - M| > t\right) = \mathbb{P}\left(\left|\bar{\ell} - \frac{Mq}{N}\right| > \frac{qt}{N}\right) \leq e^{-qt^2/3MN}$$

as we apply an additive Chernoff bound, since $\ell \sim Binom(q, \frac{M}{N})$. if we take $t = \omega(\sqrt{\frac{MN}{q}})$, we obtain:

$$\mathbb{P}\left(|\bar{M} - M| = \omega\left(\sqrt{\frac{MN}{q}}\right)\right) = e^{-\omega(1)} = o(1)$$

This achieves deviation σ after $q = \mathcal{O}\left((\frac{M^2N^2}{\sigma^2})\right)$ queries. Here we want to use a different way to estimate M. First, consider the Grover iteration $G = -U_+U_f$. Since this is a rotation in the two-dimensional subspace $\mathcal{H} = \text{span}\{|S\rangle, |+\rangle\}$, G has two eigenstates with eigenvalues $e^{\pm 2i\gamma}$ spanning $\mathcal{H}$, and the phase estimation algorithm can be used to estimate γ. At this point:

$$\sin\gamma = \sqrt{\frac{M}{N}} \Longrightarrow M = N\sin^2\gamma$$

The phase estimation algorithm will find an r-digit γ with constant probability using $\mathcal{O}(2^r)$ queries to G (and thus also to the phase oracle U_f). If $\gamma \to 0$, then we have $\gamma \simeq M/N$ therefore, we obtain an estimation of $\frac{M}{N}$ with deviation $\mathcal{O}(1/q)$ using q queries. If we want a deviation of σ on the estimation for M, then:

$$\frac{\sigma}{N} = \mathcal{O}\left(\frac{1}{q}\right) \Longrightarrow q = \Theta\left(\frac{N}{\sigma}\right)$$

i.e. we obtain an estimation with deviation σ quadratically faster.

5 Quantum Error Correction

In quantum information we consider two types of error on a qubit:

- Bit flip errors: $\mathcal{N}_X(\rho) = pX\rho X^\dagger + (1-p)\rho$;
- Phase flip errors: $\mathcal{N}_Z(\rho) = pZ\rho Z^\dagger + (1-p)\rho$;

We will see that an arbitrary single-qubit error can be rewritten as a superposition of these two errors. We would like to devise a quantum operation $\mathcal{C}$ such that:

$$\mathcal{C}(\mathcal{N}(\rho)) = \rho$$

We encode $|0\rangle, |1\rangle$ with two code words. The correction is done by a syndrome measurement, which in the quantum case is an observable which commutes with the logical X, Z gates (i.e. shares an eigenbasis with them, and preserves the quantum coherences between the logical states). The outcome of the measurement gives classical information about which unitary must be applied in order to correct the syndrome.

$$\mathcal{C}(\rho) = \sum_s U_s \Pi_s \rho \Pi_s^\dagger U_s^\dagger$$

where $\mathcal{M}(\rho) = \sum_s \Pi_s \Pi_s^\dagger$ is the quantum operation of the syndrome measurement. If s is measured, the quantum operation $\mathcal{U}_s(\rho) = U_s \rho U_s^\dagger$ is applied.

Theorem 5.1. *An error correcting code which corrects single-qubit bit flips (X), single-qubit phase flips (Z) and combined flips (ZX) can correct an arbitrary single-qubit error.*

Proof. Notice that $ZX = -iY$, therefore $ZX|\psi\rangle \propto Y|\psi\rangle$. An arbitrary single-qubit error is a quantum channel $\mathcal{N}(\rho)$ with Kraus decomposition:

$$\mathcal{N}(\rho) = \sum_k E_k \rho E_k^\dagger$$

The space of single-qubit matrices is spanned by the basis $\{\mathbb{1}, X, Y, Z\}$, thus we rewrite E_k as:

$$E_k = x_k X + y_k Y + z_k Z + w_k \mathbb{1}$$

by doing the multiplications, and applying the cyclic properties of the Pauli matrices, we obtain:

$$\mathcal{N}(\rho) = x^* X\rho X^\dagger + y^* Y\rho Y^\dagger + z^* Z\rho Z^\dagger + w^*\rho$$

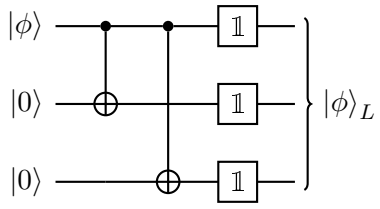
By linearity and trace preservation, the error correction gives:

$$\mathcal{C}(\mathcal{N}(\rho)) = x^*\mathcal{C}(X\rho X^\dagger) + y^*\mathcal{C}(Y\rho Y^\dagger) + z^*\mathcal{C}(Z\rho Z^\dagger) + w^*\mathcal{C}(\rho) = (x^* + y^* + z^* + w^*)\rho = \rho$$

□

5.1 3-qubit codes

Bit flip code



$$|0\rangle \mapsto |000\rangle\,, |1\rangle \mapsto |111\rangle$$

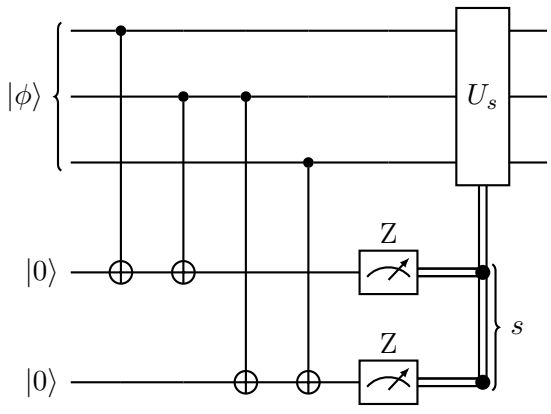
$$X_L = X^{\otimes 3}, Z_L = Z^{\otimes 3}$$

Syndrome measurements:

- Z_1Z_2: checks whether $x_1 \neq x_2$;
- Z_2Z_3: checks whether $x_2 \neq x_3$;

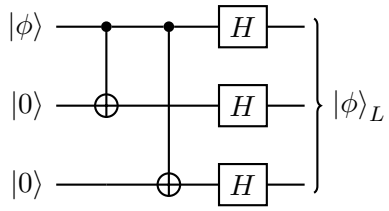
Syndrome outcomes:

- `00`: $x_1 = x_2 = x_3$, no error;
- `01`: $x_1 = x_2 \neq x_3$, apply $U_{01} = X_3$;
- `10`: $x_1 \neq x_2 = x_3$, apply $U_{10} = X_1$;
- `11`: $x_1 \neq x_2 \neq x_3$, apply $U_{11} = X_2$.



Corrects any single-qubit bit flip.

Phase flip code



$$|0\rangle \mapsto |+++\rangle\,, |1\rangle \mapsto |---\rangle$$

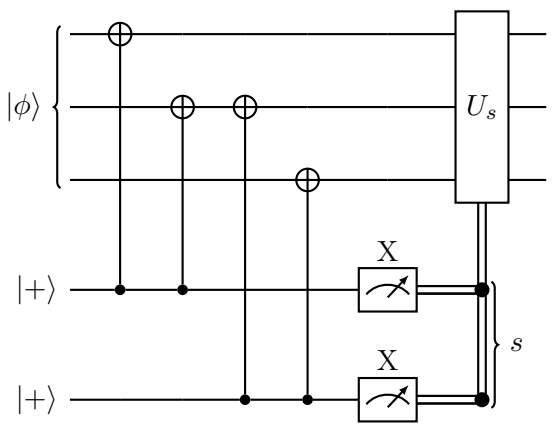
$$X_L = Z^{\otimes 3}, Z_L = X^{\otimes 3}$$

Syndrome measurements:

- X_1X_2: checks whether $x_1 \neq x_2$;
- X_2X_3: checks whether $x_2 \neq x_3$;

Syndrome outcomes:

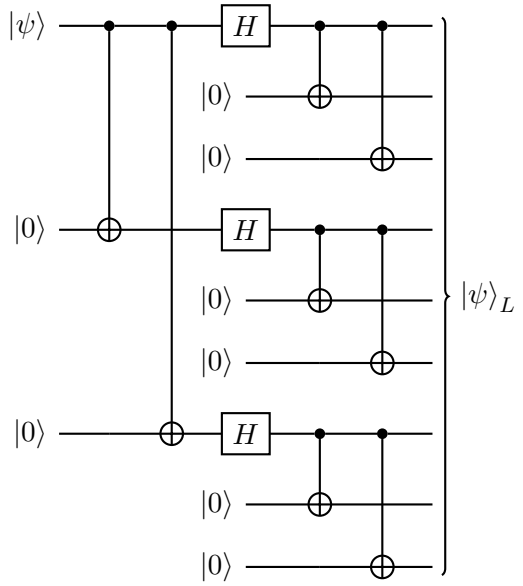
- `00`: $x_1 = x_2 = x_3$, no error;
- `01`: $x_1 = x_2 \neq x_3$, apply $U_{01} = Z_3$;
- `10`: $x_1 \neq x_2 = x_3$, apply $U_{10} = Z_1$;
- `11`: $x_1 \neq x_2 \neq x_3$, apply $U_{11} = Z_2$.



Corrects any single-qubit phase flip.

5.2 Shor code

Apply a 3-qubit phase flip code, and then duplicate each of the three qubit three times with a bit flip code.



$$|0\rangle \overset{\wedge X}{\mapsto} |000\rangle \overset{H}{\mapsto} |+++\rangle \overset{\wedge X}{\mapsto} \left(\frac{|000\rangle + |111\rangle}{\sqrt{2}}\right)^{\otimes 3}$$
$$|1\rangle \overset{\wedge X}{\mapsto} |111\rangle \overset{H}{\mapsto} |---\rangle \overset{\wedge X}{\mapsto} \left(\frac{|000\rangle - |111\rangle}{\sqrt{2}}\right)^{\otimes 3}$$

$$X_L = Z_1 Z_4 Z_7$$
$$Z_L = X_1 X_2 X_3$$

Shor's code corrects both phase flips and bit flips, thus it correct any single-qubit error by Theorem 5.1

Syndrome measurements Measure with Z_1Z_2, Z_2Z_3 to check and correct the equality of the bits of the first block (same with the other three blocks). After the correction of the bits, each block is in the subspace $\mathcal{H}^* = \text{span}\{|000\rangle, |111\rangle\}$. Now we measure with $X_1 \cdots X_6, X_4 \cdots X_9$. Notice that:

$$X^{\otimes 3}\left(\frac{|000\rangle + |111\rangle}{\sqrt{2}}\right) = \frac{X^{\otimes 3}|000\rangle + X^{\otimes 3}|111\rangle}{\sqrt{2}} = \frac{|000\rangle + |111\rangle}{\sqrt{2}}$$
$$X^{\otimes 3}\left(\frac{|000\rangle - |111\rangle}{\sqrt{2}}\right) = \frac{X^{\otimes 3}|000\rangle - X^{\otimes 3}|111\rangle}{\sqrt{2}} = -\frac{|000\rangle - |111\rangle}{\sqrt{2}}$$

This means that $X^{\otimes 3}$ measures the phase of a block, and $X^{\otimes 6}$ checks whether two blocks have the same phase. With an analogous reasoning, we can check and correct the equality of the phases of the blocks with $X_1 \cdots X_6, X_4 \cdots X_9$. Notice, however, that these two bits of information do not tell us which qubit had the phase flipped, but it is sufficient for us to apply Z to an arbitrary qubit of the block in order to correct its phase.

5.3 The stabilizer formalism

Definition 5.2 (Stabilizer)**.** *A quantum state $|\phi\rangle$ is said to be stabilized by an operator K if $K|\phi\rangle = |\phi\rangle$.*

The single-qubit Pauli group is the set

$$\mathcal{P} = \{\pm\mathbb{1}, \pm i\mathbb{1}, \pm X, \pm iX, \pm Y, \pm iY, \pm Z, \pm iZ\}$$

equipped with the product between operators $\cdot$. The n-qubit Pauli group is the set:

$$\mathcal{P}_n = \{A_1 \otimes \cdots \otimes A_n \mid A_i \in \mathcal{P}\}$$

Notice that any two elements of $\mathcal{P}_n$ commute or anti-commute.

Definition 5.3. *A stabilizer group is an abelian subgroup of $\mathcal{P}_n$ not containing $-\mathbb{1}^{\otimes n}$.*

We can specify a stabilizer group $\mathcal{S}$ in terms of a set of its generators:

$$\mathcal{S} = \langle S_1, \ldots, S_k \rangle = \left\{ \prod A_i \mid A_i \in \{S_1, \ldots, S_k\} \right\}$$

Now notice that, if two operators S_1, S_2 stabilize a state $|\phi\rangle$ then

$$S_1 S_2 |\phi\rangle = S_1 |\phi\rangle = |\phi\rangle$$

implying that also their product stabilize $|\phi\rangle$. More generally, a stabilizer group $\mathcal{S}$ induces a stabilizer subspace:

$$\mathcal{H}_S = \{|\phi\rangle \mid S|\phi\rangle = |\phi\rangle \ \ \forall S \in \mathcal{S}\}$$

Theorem 5.4. $\dim \mathcal{H}_S = 2^{n-k}$ *where k is the size of a minimal set of generators for $\mathcal{S}$.*

Proof. For any $S \in \mathcal{P}_n$ we have $S^2 = \pm \mathbb{1}^{\otimes n}$ but, since $-\mathbb{1}^{\otimes} \notin \mathcal{S}$, we get $S^2 = \mathbb{1}$. This implies that any eigenvalue of S must satisfy $\lambda^2 = 1$, i.e. $\lambda = \pm 1$. Moreover, since any $S \in \mathcal{S}$ is an element of $\mathcal{P}_n$, we can rewrite $S = i^k \bigotimes_j P_j$:

$$\operatorname{tr} S = \operatorname{tr}(i^k \bigotimes_j P_j) = i^k \prod_j \operatorname{tr}(P_j) = 0$$

since Pauli operators all have zero trace. This means that the $+1$- and -1- eigenspaces divide the total space in half. Since the stabilizer group is abelian, all operators pairwise commute, and thus they share an eigenbasis.

Suppose to have a stabilizer code with generators $S_1, \ldots, S_k$, and denote with $\mathcal{H}_i$ the subspace stabilized by $S_1, \ldots, S_i$. We want to show that S_{i+1} divides the subspace $\mathcal{H}_i$ in half. If $\operatorname{tr}(S_i \Pi_{\mathcal{H}_i}) = 0$, where $\Pi_{\mathcal{H}_i}$ is the projector onto $\mathcal{H}_i$, and $|\phi_1\rangle, \ldots, |\phi_{2^i}\rangle$ is a basis for $\mathcal{H}_i$:

$$\operatorname{tr}(S_{i+1}\Pi_{\mathcal{H}_i}) = \operatorname{tr}\left(\sum_j S_{i+1} |\phi_j\rangle \langle\phi_j| \right) = \sum_j \operatorname{tr}(S_{i+1} |\phi_j\rangle \langle\phi_j|) = \sum_j \operatorname{tr} S_{i+1} = 0$$

in particular, this means that the non-zero eigenvalues of $S_{i+1}\Pi_{\mathcal{H}_i}$ (i.e. the ones for the eigenstates spanning $\mathcal{H}_i$) sum up to 0. This is sufficient to conclude that the eigenspaces of S_{i+1} divide $\mathcal{H}_i$ in half. By induction, we infer that $\mathcal{H}_k$ must have dimension $2^n/2^k$, as claimed. □

Stabilizer codes In order to define an error correcting code from a stabilizer group $\mathcal{S}$ we do as follows:

- The logical subspace of our code will be the stabilizer subspace induced by $\mathcal{S}$;
- If an error E anti-commutes with a generator S then:

  $$S(E|\phi\rangle) = -E(S|\phi\rangle) = -E|\phi\rangle$$

  i.e. the faulty state is a -1-eigenstate of S, and any single-qubit error can be seen as a superposition of errors which anti-commute with one of the generators. Therefore, the generators of $\mathcal{S}$ can be taken as the syndrome measurements: if $+1$ is measured on all of them, then the state has no error;

Different (minimal) sets of generators give different syndrome measurements, which are all equivalent in terms of the information they give about the errors. If we want to encode a qubit, then we will need $n-1$ independent generators

- Take two unitaries U, V such that $[U, S] = [V, S] = 0$ for all $S \in \mathcal{S}$;

- Set $X_L := U, Z_L = V$, and take $|0\rangle_L, |1\rangle_L$ as the eigenstates of Z_L in the stabilizer subspace of $\mathcal{S}$.

Some examples include:

- *3-qubit bit flip code*: $\mathcal{S} = \langle ZZ\mathbb{1}, \mathbb{1}ZZ\rangle$;
- *3-qubit phase flip code*: $\mathcal{S} = \langle XX\mathbb{1}, \mathbb{1}XX\rangle$;
- *5-qubit code*: $\mathcal{S} = \langle XZZX\mathbb{1}, \mathbb{1}XZZX, X\mathbb{1}XZZ, ZX\mathbb{1}XZ, ZZX\mathbb{1}X\rangle$.

5.4 Constructing quantum error correcting codes from linear codes

Let C be a $[n, k]$ linear code with generator matrix $G \in \mathbb{Z}_2^{n\times k}$ and parity matrix $H \in \mathbb{Z}_2^{(n-k)\times n}$. Take the dual code $C^\perp$ with generator $G^\perp = H^T$ and parity matrix $H^\perp = G^T$.

We define the following state:

$$\left|x \oplus C^\perp\right\rangle = \frac{1}{\sqrt{|C^\perp|}} \sum_{y\in C^\perp} |x \oplus y\rangle$$

One can notice that $\left\langle x \oplus C^\perp \middle| x' \oplus C^\perp\right\rangle = 0$ if $x \oplus x' \notin C^\perp$: this because $C^\perp$ is closed under addition (any combination of the columns of $G^\perp$ is a valid code), thus adding something that is not in $C^\perp$ must give something that is not in $C^\perp$. Since $\left|x \oplus C^\perp\right\rangle$ is a superposition of elements in $C^\perp$, and $\left|x' \oplus C^\perp\right\rangle$ is a superposition of elements not in $C^\perp$, the terms must be pairwise orthogonal. Therefore we can choose two such states to be $|0\rangle_L, |1\rangle_L$. X_L will be implemented by applying an X gate whenever $x - x'$ gives 1, and Z_L will has to flip the sign of all the members not in $C^\perp$.

A more straightforward way to see such construction is by using the stabilizer formalism: suppose to have a stabilizer group $\mathcal{S}$. The following states can be chosen:

$$|0\rangle_L = \sum_{S\in\mathcal{S}} S\,|0\rangle^{\otimes n}, |1\rangle_L = X_L\,|0\rangle_L$$

where X_L, Z_L can be chosen as specified in the last section. Notice that they are both in the stabilizer subspace of $\mathcal{S}$:

$$S'\,|0\rangle_L = \sum_{S\in\mathcal{S}} S'S\,|0\rangle^{\otimes n} = \sum_{S\in\mathcal{S}} S\,|0\rangle^{\otimes n} = |0\rangle_L$$
$$S'\,|1\rangle_L = S'X_L\,|0\rangle_L = X_L S'\,|0\rangle_L = X_L\,|0\rangle_L = |1\rangle_L$$

where we used the fact that $S' \cdot \mathcal{S} = \mathcal{S}$, by the properties of a group, and the fact that X_L has to commute with any element of $\mathcal{S}$.

5.5 Knill-Laflamme condition

Theorem 5.5. *Let $\mathbb{E} = \{E_a\}_a$ be a set of errors. There exists a code correcting all the errors in $\mathbb{E}$ if and only if:*

$$\langle i|_L\, E_a^\dagger E_b\, |j\rangle_L = c_{ab}\delta_{ij} \quad \text{for any } a, b, i, j$$

where $c_{ab} = \langle i|\, E_a^\dagger E_b\, |i\rangle$ form an Hermitian operator.

Sketch of proof. Suppose two different errors map orthogonal vectors to non-orthogonal ones, e.g. $\langle 0|_L\, E_a^\dagger E_b\, |1\rangle_L \neq 0$. There is an overlap, i.e. a probability of measuring something that could result both from a $|0\rangle$ followed by the error E_a and from a $|1\rangle$ followed by the error E_b.

Thus, if E_a, E_b cannot be corrected by the same unitary, no error correcting code will be able to correct all the errors in $\mathbb{E}$.

On the other hand, if any two errors map different codewords to orthogonal subspaces, a set of syndrome observables which check in which of these subspaces the faulty state lies is sufficient to correct any error. □

Corollary 5.6. *Let $\mathcal{S} = \langle S_1, \ldots, S_k \rangle$ define a stabilizer code, and let $\mathbb{E} = \{E_a\}_a$. If for any a, b either:*

- $E_b^\dagger E_a \in \mathcal{S}$ *or*
- $\exists S \in \mathcal{S}$ *that anti-commutes with $E_b^\dagger E_a$.*

then the Knill-Laflamme condition is satisfied, and the stabilizer code corrects any error in $\mathbb{E}$.

Proof. If $E_b^\dagger E_a \in \mathcal{S}$, then notice that the codewords are stabilized by this operator, hence

$$\langle i|_L E_b^\dagger E_a |j\rangle_L = \langle i|j\rangle = \delta_{ij}$$

If $E_b^\dagger E_a$ anti-commutes with $S \in \mathcal{S}$, then we have

$$\langle i|_L E_b^\dagger E_a |j\rangle_L = \langle i|_L E_b^\dagger E_a S |j\rangle_L = -\langle i|_L S E_b^\dagger E_a |j\rangle_L = -\langle i|_L E_b^\dagger E_a |j\rangle_L$$

implying $\langle i|_L E_b^\dagger E_a |j\rangle_L = 0$. In both cases, the Knill-Laflamme conditions are fulfilled. □

5.6 Clifford groups

We see a different application of the stabilizer formalism. A stabilizer state $|\phi\rangle$ is uniquely determined by the n operators in the Pauli group $\mathcal{P}_n$ that stabilize it (this because $\dim \mathcal{H} = 2^{n-n} = 1$!). This means that such state can be encoded in $n(2n+1) = \mathcal{O}(n^2)$ bits of information (each stabilizing operator is a tensor product of n operators in $\{X, Y, Z, \mathbb{1}\}$, plus a bit representing the sign, $\pm i$ cannot appear as they would give anti-Hermitian operators and not stabilize any state).

Now suppose we apply a unitary U to a state $|\phi\rangle$ stabilized by $P \in \mathcal{P}_n$ then:

$$UPU^\dagger(U|\phi\rangle) = UP|\phi\rangle = U|\phi\rangle$$

implying $U|\phi\rangle$ is stabilized by $P' = UPU^\dagger$. The Clifford group is the normalizer of $\mathcal{P}_n$:

$$\mathcal{C}_n = \{U \in \mathcal{U}(2^n) \mid U\mathcal{P}_n U^\dagger = \mathcal{P}_n\}$$

We also call unitaries from the Clifford group *Clifford gates*.

Theorem 5.7 (Gottesmann-Knill)**.** *Quantum computation starting from a stabilized state involving only Clifford gates and Z-measurements can be efficiently simulated by a classical computer.*

Proof. Suppose $\{H, S, CNOT\}$ can generate any Clifford gate in $\mathcal{O}(n^2)$ time. *[Proof missing here]*

Applying one of these three gates to a stabilizer state takes $\mathcal{O}(1)$ time (they act independently on a constant number of Pauli operators in the representation of the state), thus the action of a Clifford gate on a stabilizer state can be reproduced in $\mathcal{O}(n^2)$ time on a classical computer.

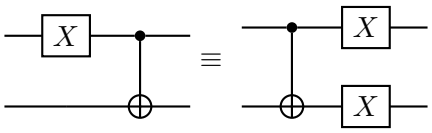
The constraint on the measurement basis is needed, as otherwise an arbitrary unitary can be integrated in the measurement basis. □

5.7 Fault-tolerant quantum information processing

Suppose that we want to apply an ideal gate $\mathcal{U}(\rho) = U\rho U^\dagger$ but we only have an imperfect gate $\mathcal{E}(\rho)$. Since we can rewrite:

$$\mathcal{E} = \mathcal{E} \circ \mathcal{U}^{-1} \circ \mathcal{U} = \bar{\mathcal{E}} \circ \mathcal{U}$$

We can always formalize an imperfect gate as the ideal gate followed by some noise. The idea is that error correcting codes can correct the noises introduced by the gate. But it is possible that a single-qubit error (say) may be propagated to multiple qubits with a gate. For example:



In order to prevent this, we would like to have stabilizer code which allows gates to be implemented in a *transversal* way. For example, in the Steane code we have:

$$X_L = X^{\otimes 7}, Z_L = Z^{\otimes 7}$$

In this way, single-qubit errors will not propagate to different qubits.

Assuming that a single-qubit error happens with probability p (and no other error can happen), the probability of an error becomes:

$$p \mapsto cp^2$$

For example, the Steane code gives $c \simeq 10^4$. A way to get lower and lower probabilities is to concatenate codes: we use the logical qubits of a code as physical codes to another code. For example, concatenating two Steane codes gives:

$$p \mapsto cp^2 \mapsto c(cp^2)^2 = c^3p^4$$

In general k concatenations of the Steane code gives $p \mapsto \frac{1}{c}(cp)^{2^k}$ error probability using 7^k physical qubits, and this tends to 0 as long as $cp < 1$.

The so-called *threshold theorems* give bounds on the accuracy of physical qubits in order to be able to achieve fault-tolerant QIP.

5.8 State and process tomography

Suppose we would like to experimentally determine a single-qubit state ρ, assuming we can create a sufficient number of copies for it (remember we cannot clone a state!). Notice that $\langle A, B\rangle = tr(AB)$ is a valid inner product for the space of 2×2 Hermitian matrices, and that $\{\mathbb{1}, X, Y, Z\}$ is an orthonormal basis for this space. Therefore, ρ can be written as:

$$\rho = w\mathbb{1} + xX + yY + zZ$$

where $w = \mathrm{tr}(\rho) = 1, x = \mathrm{tr}(X\rho), y = \mathrm{tr}(Y\rho), z = \mathrm{tr}(Z\rho)$. Also notice that, given an observable $A = \sum_i \lambda_i |i\rangle \langle i|$:

$$\mathrm{tr}(A\rho) = \sum_i \lambda_i \,\mathrm{tr}(|i\rangle \langle i| \rho) = \sum_i \lambda_i \langle i| \rho |i\rangle = \sum_i \lambda_i \mathbb{P}(i)_\rho$$

Thus, these inner products are the expected labels of measurements of ρ in, respectively, X, Y and Z bases. Therefore, we can estimate these quantities with a Monte Carlo sampling. Using

each of the three Pauli matrices, we measure a sufficient number of copies of ρ and take the mean value of the received label.

Now suppose we have an unknown quantum channel $\mathcal{E}$ operating from a system A to a system B: we show here that we can fully determine it by doing a state tomography on a particular state. Take the following:

$$|\phi\rangle = \frac{1}{\dim A} \sum_j |j\rangle_A \otimes |j\rangle_{A'}$$

where A' is a copy of A. We now apply $\mathcal{E}$ to the system A:

$$\rho_{BA'} = [\mathcal{E} \otimes \mathbb{1}_{A'}](|\phi\rangle \langle\phi|)$$

Now notice that:

$$\begin{aligned}
\mathrm{tr}_{A'}[\rho_{BA'}(\mathbb{1}_B \otimes |k\rangle \langle\ell|_{A'})] &= \mathrm{tr}_{A'} \left[[\mathcal{E} \otimes \mathbb{1}_{A'}](|\phi\rangle \langle\phi|)(\mathbb{1}_B \otimes |k\rangle \langle\ell|_{A'})\right] \\
&= \mathrm{tr}_{A'} \left[[\mathcal{E} \otimes \mathbb{1}_{A'}] \left(\frac{1}{(\dim A)^2} \sum_{i,j} |ii\rangle \langle jj| \right) (\mathbb{1}_B \otimes |k\rangle \langle\ell|_{A'}) \right] \\
&= \frac{1}{(\dim A)^2} \sum_{i,j} \mathrm{tr}_{A'} \left[[\mathcal{E} \otimes \mathbb{1}_{A'}] \left(|i\rangle \langle j|_A \otimes |i\rangle \langle j|_{A'} \right) (\mathbb{1}_B \otimes |k\rangle \langle\ell|_{A'}) \right] \\
&= \frac{1}{(\dim A)^2} \sum_{i,j} \mathrm{tr}_{A'} \left[\mathcal{E}(|i\rangle \langle j|)_B \otimes |i\rangle \langle j|_{A'} \left(\mathbb{1}_B \otimes |k\rangle \langle\ell|_{A'} \right) \right] \\
&= \frac{1}{(\dim A)^2} \sum_{i,j} \mathcal{E}(|i\rangle \langle j|)_B \, \mathrm{tr}(|i\rangle \langle j|k\rangle \langle\ell|_{A'}) \\
&= \frac{1}{(\dim A)^2} \sum_{i,j} \mathcal{E}(|i\rangle \langle j|)_B \langle j|k\rangle \langle i|\ell\rangle = \frac{1}{(\dim A)^2} \mathcal{E}(|\ell\rangle \langle k|)_B
\end{aligned}$$

Now let us consider an arbitrary input state $\sigma_A = \sum_{ij} a_{ij} |i\rangle \langle j|$. If we apply our quantum operation:

$$\begin{aligned}
\mathcal{E}(\sigma_A) &= \sum_{i,j} a_{ij} \mathcal{E}(|i\rangle \langle j|) \\
&= (\dim^2 A) \sum_{i,j} a_{ij} \, \mathrm{tr}_{A'}[\rho_{BA'}(\mathbb{1}_B \otimes |j\rangle \langle i|_{A'})] \\
&= (\dim^2 A) \, \mathrm{tr}_{A'}[\rho_{BA'}(\mathbb{1}_B \otimes \sum_{i,j} a^*_{ji} |j\rangle \langle i|_{A'})] \\
&= (\dim^2 A) \, \mathrm{tr}_{A'}[\rho_{BA'}(\mathbb{1}_B \otimes \sigma^*_A)]
\end{aligned}$$

This shows that, in order to carry out a process tomography on $\mathcal{E}$, it is sufficient to do a state tomography on $\rho_{BA'}$.

6 Hamiltonian Simulation

Problem 6.1. *Given a time-independent Hamiltonian H, an initial state $|\psi(0)\rangle$, and a time t, find an ε-approximation for $|\psi(t)\rangle$.*

By the Schrödinger equation:

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = H|\psi(t)\rangle \Longrightarrow |\psi(t)\rangle = e^{-iHt/\hbar}|\psi(0)\rangle$$

From now on, we will absorb the Planck constant in the Hamiltonian $H \leftarrow H/\hbar$ for simplicity. The problem can be translated to the construction of an algorithm implementing a unitary $\bar{U}$ such that:

$$D(\bar{U}, e^{-iHt}) \leq \varepsilon$$

This problem is QMA-complete for the general case, therefore we restrict ourselves to a special class of Hamiltonians.

Definition 6.2 (k-locality)**.** *An Hamiltonian H is said to be k-local if it can be expressed as:*

$$H = \sum_{i=1}^{m} H_i$$

where H_i acts non-trivially on k qubits.

Sub-Hamiltonians acting on the same subset of qubits can be added together, therefore we can assume the number m of sub-Hamiltonians is at most:

$$m \leq \binom{n}{k} \leq n^k$$

which is polynomial in n for $\mathcal{O}(1)$-local Hamiltonians. Moreover, we have the following result:

Theorem 6.3 (Solovay-Kitaev)**.** *Any unitary acting non-trivially on $\mathcal{O}(1)$ qubits can be ε-approximated in $\mathcal{O}(\log^4 \varepsilon)$ time.*

6.1 Simulation for $\mathcal{O}(1)$-local, non-interacting Hamiltonians

We start with a simple example: assume that all the sub-Hamiltonians pairwise commute. This implies that the sub-systems are non-interacting, and

$$U = e^{-iHt} = e^{-it\sum_j H_j} = \prod_j e^{-iH_j t} =: \prod_j U_j$$

If we replace each U_j with an approximation $\bar{U}_j$ made with the Solovay-Kitaev construction, we obtain by Theorem 0.1:

$$D\left(\prod_j U_j, \prod_j \bar{U}_j\right) \leq \sum_j D(U_j, \bar{U}_j) \leq m\varepsilon$$

And m Solovay-Kitaev constructions take $\mathcal{O}(m\log^4 \varepsilon)$ time. In order to obtain a ε-approximation we need to replace $\varepsilon \leftarrow \varepsilon/m$. This gives an ε-approximation in $\mathcal{O}(m\log^4(m/\varepsilon))$ time.

6.2 Trotterisation

If two operators A, B do not pairwise commute, then we have $e^{A+B} \neq e^A e^B$ in general. However, we can see the following:

Theorem 6.4 (Lie-Trotter)**.** $e^{x(A+B)} = e^{xA}e^{xB} - \frac{1}{2}x^2[A,B] + \mathcal{O}(x^3)$.

Proof.

$$\begin{aligned}
e^{x(A+B)} - e^{xA}e^{xB} &= (\mathbb{1} + x(A+B) + \frac{1}{2}x^2(A+B)^2 + \mathcal{O}(x^3)) + \\
&\quad - (\mathbb{1} + xA + \frac{1}{2}x^2A^2 + \mathcal{O}(x^3)) \cdot (\mathbb{1} + xB + \frac{1}{2}x^2B^2 + \mathcal{O}(x^3)) \\
&= (\mathbb{1} + x(A+B) + \frac{1}{2}x^2(A+B)^2 + \mathcal{O}(x^3)) + \\
&\quad - (\mathbb{1} + x(A+B) + \frac{1}{2}x^2A^2 + \frac{1}{2}x^2B^2 + x^2AB + \mathcal{O}(x^3)) \\
&= \frac{1}{2}x^2(A+B)^2 - \frac{1}{2}x^2A^2 - \frac{1}{2}x^2B^2 - x^2AB + \mathcal{O}(x^3) \\
&= \frac{1}{2}x^2(AB+BA) - x^2AB + \mathcal{O}(x^3) \\
&= -\frac{1}{2}x^2[A,B] + \mathcal{O}(x^3)
\end{aligned}$$

□

Corollary 6.5. *For a $\mathcal{O}(1)$-local Hamiltonian $H = \sum_j H_j$ with $||H_j|| \leq h$ we have:*

$$e^{-iHt} = \left(e^{-iH_1t/K} \cdots e^{-iH_mt/K}\right)^K + \mathcal{O}\left(\frac{m^2t^2h^2}{K}\right)$$

Proof. We inductively apply the Lie-Trotter decomposition:

$$\begin{aligned}
e^{-iHt/K} &= e^{-iH_1t/K}e^{-i\sum_{j=2}^m H_jt/K} + \mathcal{O}\left(\frac{t^2}{K^2}\left\|\left[H_1, \sum_{i=2}^m H_j\right]\right\|\right) \\
&= e^{-iH_1t/K}e^{-i\sum_{j=2}^m H_jt/K} + \mathcal{O}\left(\frac{t^2(m-1)h^2}{K^2}\right) \\
&= e^{-iH_1t/K} \cdots e^{-iH_mt/K} + \sum_{j=1}^m \mathcal{O}\left(\frac{t^2(m-j)h^2}{K^2}\right) \\
&= e^{-iH_1t/K} \cdots e^{-iH_mt/K} + \mathcal{O}\left(\frac{m^2t^2h^2}{K^2}\right)
\end{aligned}$$

Applying this unitary K times amplifies the error to a factor K, by Theorem 0.1. □

Notice that, if we have $m = \Theta(n^k)$ (i.e. the maximum possible), for a fixed subset of k qubits, the number of subsets of k qubits having a non-empty intersection with this subset is:

$$\sum_{\ell=1}^k \binom{k}{\ell}\binom{n-k}{k-\ell} = \mathcal{O}(n^{k-1}) = \mathcal{O}(m^{1-1/k})$$

This implies that each sub-Hamiltonian has at most $\mathcal{O}(m^{1-1/k})$ interacting sub-Hamiltonians, i.e. the total number of pairs of interacting sub-Hamiltonians is $\mathcal{O}(m \cdot m^{1-1/k})$. If we now use the Solovay-Kitaev construction to approximate each element of the above, we have:

$$D\left(\left(\bar{U}_1 \cdots \bar{U}_m\right)^K, \left(e^{-iH_1t/K} \cdots e^{-iH_mt/K}\right)^K\right) \leq mK\varepsilon_L + \mathcal{O}\left(\frac{m^2t^2h^2}{K}\right)$$

where ε_L is the approximation error chosen for each Solovay-Kitaev construction. If we want a total error of ε we can impose:

$$\frac{\varepsilon}{2} \stackrel{!}{=} mK\varepsilon_L \stackrel{!}{=} \mathcal{O}\left(\frac{m^2t^2h^2}{K}\right) \Longrightarrow \begin{cases} K \stackrel{\Theta}{=} \frac{m^2t^2h^2}{\varepsilon} \\ \varepsilon_L \stackrel{\Theta}{=} \frac{\varepsilon}{mK} \stackrel{\Theta}{=} \frac{\varepsilon^2}{m^3t^2h^2} \end{cases}$$

This leads to a $\mathcal{O}\left(\frac{m^3t^2h^2}{\varepsilon}\log^4(\frac{mth}{\varepsilon})\right)$ total time.

6.3 Higher-order approximants

One could achieve a better error by using second-order approximants:

$$e^{xA/2}e^{xB}e^{xA/2} = e^{x(A+B)} + \mathcal{O}(x^3)$$

This trotterisation gives:

$$\prod_{j=1}^{m} e^{xH_j/2} \prod_{j=1}^{m} e^{xH_{m-j}/2} = e^{iHt} + \mathcal{O}\left(\frac{m^3t^3h^3}{K^2}\right)$$

6.4 Application: estimating the expectation of an observable

Suppose we have a polynomial-time observable O, with $||O|| \leq 1$. We want to estimate

$$\langle O \rangle = \langle \phi | \, e^{iHt} O e^{-iHt} \, | \phi \rangle$$

This can be done with a Monte Carlo sampling: we compute $e^{-iHt} \, |\psi\rangle$ a sufficient number of times, and measure with O, taking the sample mean of the results.

6.5 Application: estimating the energy eigenvalues

Assuming we can simulate e^{-iHt} in polynomial-time, we can find the eigenvalues and eigenstates of H. For this, we use phase estimation: if we already have an eigenstate $|\mathcal{E}_k\rangle$ such that

$$H \, |\mathcal{E}_k\rangle = E_k \, |\mathcal{E}_k\rangle$$

using this state with the phase estimation for the unitary $U = e^{-iHt}$ will give an estimation of the φ_k such that:

$$e^{-2\pi i \varphi_k} \simeq e^{-iE_k t}$$

and we can compute the energy eigenvalue $E_k = \frac{2\pi\varphi_k}{t}$. The problem is that, however, we do not know the energy eigenstate $|\mathcal{E}_k\rangle$. The solution is to prepare a state:

$$|\phi\rangle = \sum_k \alpha_k \, |\mathcal{E}_k\rangle$$

and, by linearity, the phase estimation algorithm will return an estimation of E_k with probability $|\alpha_k|^2$ (while the state will collapse to $|\mathcal{E}_k\rangle$!). The only problem is that it is hard to find a state $|\psi\rangle$ with $| \langle \mathcal{E}_k | \phi \rangle |^2 = \frac{1}{\text{poly}(n)}$.

Adiabatic state preparation Suppose we have an Hamiltonian H for which it is hard to find the ground state $|\mathcal{E}_0\rangle$, and an Hamiltonian H^* for which it is easy. The idea is to compute the ground state $|\mathcal{E}_0^*\rangle$ of H^* and then adiabatically evolve using the Hamiltonian

$$H(t) = \frac{t}{T}H + \left(1 - \frac{t}{T}\right) H^*$$

The adiabatic theorem states that the system will always be in the ground state of $H(t)$ at any point in time. It remains to see how to simulate the evolution of such time-dependent Hamiltonian. The idea is to use a sufficiently large Trotter number K and update the Hamiltonian at each time slice.

6.6 Application: particle on the line

Suppose to have a particle on the x axis. For simplicity, assume it is in the segment $[0, 1]$.

$$|\psi\rangle = \int_0^1 |x\rangle \langle x|\psi\rangle \, dx$$

We can discretize this interval: using n qubits, we represent $x_j = j/2^n$ with the state $|j\rangle$. Therefore, the position operator becomes:

$$\hat{X} = \int_0^1 x \, |x\rangle \langle x| \, dx \mapsto \tilde{X} = \sum_{j=0}^{2^n-1} \frac{j}{2^n} |j\rangle \langle j|$$

The momentum basis can be retrieved from the position basis through a Fourier transform, and we use the Quantum Fourier transform in the discrete case:

$$\hat{P} = Q\hat{X}Q^\dagger \mapsto \tilde{P} = Q_n \tilde{X} Q_n^\dagger$$

Suppose to have a general Hamiltonian for a particle of mass μ subject to a potential $V(x)$, i.e.

$$H = \frac{\hat{P}^2}{2\mu} + V(\hat{X}) = Q\frac{\hat{X}^2}{2\mu}Q^\dagger + V(\hat{X}) \mapsto \tilde{H} = \frac{\tilde{P}^2}{2\mu} + V(\tilde{X}) = Q_n \frac{\tilde{X}^2}{2\mu} Q_n^\dagger + V(\tilde{X})$$

Since we have the sum of two terms, we trotterize:

$$\begin{aligned} e^{-i\tilde{H}t/K} &= e^{-i\left(Q_n \frac{\tilde{X}^2}{2\mu} Q_n^\dagger + V(\tilde{X})\right)t/K} \\ &= e^{-iQ_n \frac{\tilde{X}^2}{2\mu} Q_n^\dagger t/K} e^{-iV(\tilde{X})t/K} + \mathcal{O}\left(\frac{t^2}{K^2}\right) \\ &= Q_n e^{-i\tilde{X}^2 t/2\mu K} Q_n^\dagger \cdot e^{-iV(\tilde{X})t/K} + \mathcal{O}\left(\frac{t^2}{K^2}\right) \end{aligned}$$

Thus the unitary $e^{-i\hat{H}t}$ can be trotterised with $\mathcal{O}(t^2/K)$ error. We only need to implement an approximation for $e^{-i\hat{X}^2 t/2\mu K}$ and $e^{-iV(\hat{X})t/K}$.