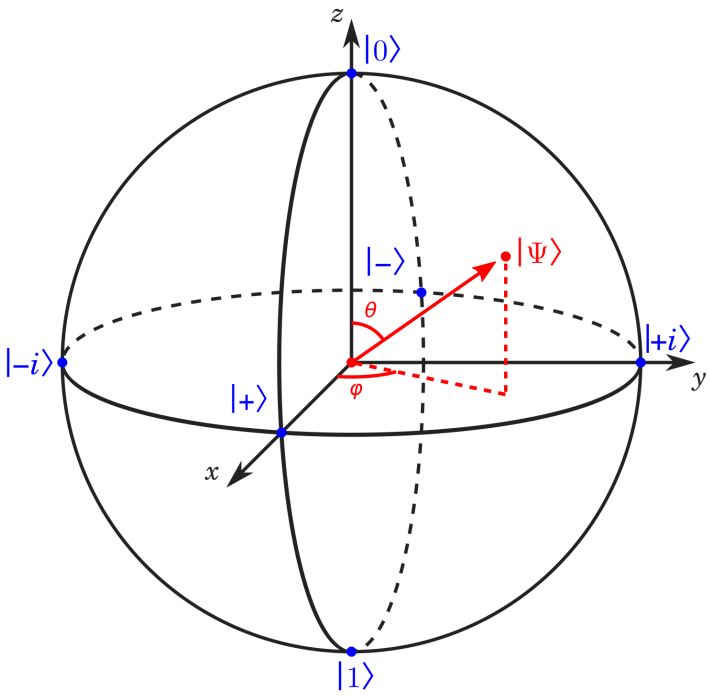
Quantum Information and Computation

Ryan Hsiang

June 11, 2025



Preface

This note is based on the lecture slides of the Quantum Information and Computation course taught by Prof. Hao-Chung Cheng at NTU in the Spring of 2025, as well as lecture notes on Quantum Information by Prof. John Preskill of Caltech. Prof. Hao-Chung Cheng approaches this topic from a more "CS" perspective, whereas Prof. Preskill provides more physical context and is more detailed in his extremely long lecture notes. This document serves three main purposes:

1. As a less condensed version of the midterm cheatsheet.
2. To help me learn the topic better.
3. For quick and easy reference on relevant equations, definitions, and theorems in the future.

Thus, this note is not written to be easy to understand, and beginners of the topic may find parts of it perplexing. However, If you are an NTU student taking QIC, this might help you prepare for your midterm.

1 Mathematical Background

1.1 Linear Algebra

A *vector space* over a field F is a non-empty set V together must define addition and multiplication. In addition, we can define the *inner product*, where it must satisfy for all $\mathbf{u}, \mathbf{v} \in V$ and $c \in F$

- Linearity in the second argument: $\langle \mathbf{u}, c\mathbf{v} + \mathbf{w} \rangle = c\langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{w} \rangle$.
- Conjugate symmetry: $\langle \mathbf{u}, \mathbf{v} \rangle^* = \langle \mathbf{v}, \mathbf{u} \rangle$.
- Positive definiteness: $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$, with equality if and only if $\mathbf{u} = \mathbf{0}$.

Note that linearity in the first argument follows from rewriting axioms 1 and 2. A complete inner product is called a *Hilbert Space*. The *dimension* of a Hilbert space is defined as the cardinality of some orthonormal basis. Given such a basis $\{\mathbf{e}_i\}_i$, we can represent any vector as

$$\mathbf{u} = \sum_i \langle \mathbf{e}_i, \mathbf{u} \rangle \mathbf{e}_i.$$

We prove this by contradiction, if there were some vector that can not be written as a linear combination of the basis, we add this vector to form a basis of size $n+1$, which contradicts the definition of the dimension. After which, the coefficients of each term of the expansion can be easily derived.

Given two vectors $u, v \in V$, we can find the define $\mathbf{z}$ as

$$\mathbf{z} = \mathbf{u} - \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{v}\|^2} \mathbf{v}.$$

where $\mathbf{z}$ and $\mathbf{u}$ is orthogonal,

$$\langle \mathbf{z}, \mathbf{v} \rangle = \mathbf{0}.$$

We can thus show that

$$\|\mathbf{u}\|^2 = \frac{|\langle \mathbf{u}, \mathbf{v} \rangle|^2}{\|\mathbf{v}\|^2} + \|\mathbf{z}\|^2,$$

or

$$\|\mathbf{u}\|^2 \|\mathbf{v}\|^2 \geq |\langle \mathbf{u}, \mathbf{v} \rangle|^2.$$

This is the *Cauchy-Schwarz inequality.*

A linear map $\mathbf{A} : \mathcal{H} \to \mathcal{H}$ that maps a vector space to itself and preserves linear combinations,

$$\mathbf{A}(a\mathbf{u} + b\mathbf{v}) = a\mathbf{A}\mathbf{u} + b\mathbf{A}\mathbf{v}$$

is called an *operator*. In particular, an operator is bounded if

$$\|\mathbf{A}\mathbf{u}\| \leq t\|\mathbf{u}\|, \quad \forall u \in \mathcal{H}.$$

The set $\mathcal{B}(\mathcal{H})$ of all bounded operators in $\mathcal{H}$ forms a vector space. It is a normed space with a norm defined by

$$\|\mathbf{A}\|_\infty = \sup_{\mathbf{u}:\,\|\mathbf{u}\|=1} \|\mathbf{A}\mathbf{u}\|.$$

In other words, the number t mentioned above. Given an operator $\mathbf{A}$, the *adjoint* of $\mathbf{A}$ is defined as the conjugate transpose of $\mathbf{A}$,

$$[\mathbf{A}^\dagger]_{ij} = [\mathbf{A}]^*_{ji}.$$

An operator is *self-adjoint* or *Hermitian* if $\mathbf{A}^\dagger = \mathbf{A}$. The eigenvalues of a Hermitian operator are real,

$$\langle \mathbf{v}, \mathbf{A}\mathbf{v}\rangle = \lambda\langle \mathbf{v}, \mathbf{v}\rangle = \langle \mathbf{A}\mathbf{v}, \mathbf{v}\rangle^* = \lambda^*\langle \mathbf{v}, \mathbf{v}\rangle.$$

In addition, we can apply this technique to distinct vectors $\mathbf{u}, \mathbf{v}$, such that,

$$\langle \mathbf{u}, \mathbf{A}\mathbf{v}\rangle = \lambda\langle \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{A}\mathbf{v}, \mathbf{u}\rangle^* = \mu^*\langle \mathbf{u}, \mathbf{v}\rangle.$$

which suggests, when $\mu \neq \lambda$,

$$\langle \mathbf{u}, \mathbf{v}\rangle = 0.$$

The eigenvectors of a Hermitian operator are mutually orthogonal. As a result, if there are N distinct eigenvalues in an N-dimensional Hilbert space, then the eigenvectors form a complete basis.

An operator is *Unitary* if $\mathbf{A}^\dagger\mathbf{A} = \mathbf{I}$, note that a unitary operator preserves the norm of a vector.

Consider a normal operator $\mathbf{A}$, with a set of eigenvectors $\{\mathbf{q}_i\}_i$ and eigenvalues $\{\lambda_i\}_i$ forming matrices $\mathbf{Q}$ and $\mathbf{\Lambda}$, we can write,

$$\mathbf{A}\mathbf{Q} = \mathbf{Q}\mathbf{\Lambda},$$

or

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^{-1},$$

$\mathbf{Q}$ consists of the orthonormal eignenvectors of $\mathbf{A}$ and is thus unitary, so

$$\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^*.$$

We can also show that, equivalently,

$$\mathbf{A} = \sum_i \lambda_i \mathbf{q}_i \mathbf{q}_i^*.$$

From this, we can easily see that,

$$\mathbf{A}\mathbf{q}_i = \lambda_i \mathbf{q}_i$$

2 States and Qubits

In classical information processing, information is conventionally represented by a classical bit, i.e., a Bernoulli random variable that can take the value of either 0 or 1. A bit can be any two different physical states of some physical system that can be perfectly distinguished.

2.1 The Qubit

Building upon the postulates of quantum mechanics, a *qubit* describes a state in the simplest possible quantum system. A qubit can be expressed as

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

where, $a, b \in \mathbb{C}$ and $|a|^2 + |b|^2 = 1$. Noting that the absolute phase of the two states has no physical significance, we can write

$$|\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{i\varphi}\sin\left(\frac{\theta}{2}\right)|1\rangle,$$

where $\theta \in [0, \pi], \varphi \in [0, 2\pi]$. This is the Bloch-sphere representation of a qubit. Next, we introduce the *Pauli matrices*,

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

One can associate such a two-level system with an object with spin-$\frac{1}{2}$,

$$|0\rangle \equiv |\uparrow\rangle, \quad |1\rangle \equiv |\downarrow\rangle$$

In this case, the Pauli matrices are related to the spin angular momentum operator $\mathbf{S}_i$ by,

$$\mathbf{S}_i = \frac{1}{2}\hbar\sigma_i.$$

The expectation of σ_i on $|\psi\rangle$ is

$$\langle\psi|\sigma_x|\psi\rangle = \sin\theta\cos\varphi \quad \langle\psi|\sigma_y|\psi\rangle = \sin\theta\sin\varphi \quad \langle\psi|\sigma_z|\psi\rangle = \cos\theta.$$

From which we define the *Bloch vector*:

$$\vec{r} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$$

Some properties of the Bloch vector emerge,

$$|\psi\rangle\langle\psi| = \frac{1}{2}(\mathbf{I} + \vec{r}\cdot\vec{\sigma})$$

$$\|\langle\psi_1|\psi_2\rangle\|^2 = \frac{1}{2}(1 + \vec{r}_1\cdot\vec{r}_2)$$

We have seen that the angular momentum operator is the generator of rotation, i.e.

$$\mathbf{U} = e^{-i\theta\vec{r}\cdot\vec{\mathbf{S}}/\hbar} = e^{-i\frac{\theta}{2}\vec{r}\cdot\vec{\sigma}}.$$

Using the properties of the Bloch vector, we can show that $|\psi\rangle$ is indeed an eigenket of $\mathbf{U}$,

$$\mathbf{U}|\psi\rangle = e^{-i\theta/2}|\psi\rangle.$$

Following the notion that $\mathbf{U}$ acts as a rotation operator about the axis of $\vec{r}$.

2.2 The Density Operator

Given a quantum state $|\psi\rangle$, the *density operator* of $|\psi\rangle$ is given by

$$\rho = |\psi\rangle\langle\psi|. \tag{2.1}$$

For an ensemble of states $\{\psi_i\}$, the density operator is

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|.$$

This models the fact that we don't know which of the states the system is in, but we assign a probability p_i to the state $|\psi\rangle_i$. The density operator can represent uncertainty beyond the minimum required by quantum mechanics. A state of the form (2.1) is said to be a *pure state*. One that cannot be written in this form is said to be *mixed*. An example of a mixture of the two states is

$$\rho = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|).$$

However, note that it is also the same density operator for a different mixture:

$$\rho = \frac{1}{2}(|+\rangle\langle +| + |-\rangle\langle -|).$$

In either case, we note that

$$\mathrm{Tr}(\rho) = \sum_i p_i \langle\psi|\psi\rangle = 1.$$

Given a particular density operator, it is not in general possible to uniquely define the pure-state decomposition, implying some extra information that is not contained in the density operator. An important property of the density operator is that we can calculate the expected value of an observable O as

$$\langle O\rangle = \sum_i p_i \langle\psi_i|O|\psi_i\rangle = \sum_i p_i \,\mathrm{Tr}(|\psi_i\rangle\langle\psi_i|O) = \mathrm{Tr}(\sum_i p_i |\psi_i\rangle\langle\psi_i|O) = \mathrm{Tr}(\rho O).$$

The *purity* of a quantum state is defined as

$$\gamma = \mathrm{Tr}(\rho^2).$$

For pure states,

$$\mathrm{Tr}(\rho^2) = \mathrm{Tr}(\rho) = 1.$$

For mixed state, $\gamma < 1$. A *maximally entangled state* is defined as I/d, e.g.,

$$\rho = \frac{1}{4}|\psi\rangle\langle\psi| + \frac{1}{4}X|\psi\rangle\langle\psi|X^\dagger + \frac{1}{4}Y|\psi\rangle\langle\psi|Y^\dagger + \frac{1}{4}Z|\psi\rangle\langle\psi|Z^\dagger = \frac{1}{2}I.$$

A measurement along some basis $\{\phi_i\}$ has probability:

$$\sum_i p_i \langle\psi_i|\phi_i\rangle\langle\phi_i|\psi_i\rangle = \langle\phi_i|\sum_i p_i |\psi_i\rangle\langle\psi_i|\phi_i\rangle = \langle\phi_i|\rho|\phi_i\rangle.$$

2.3 Composite systems

For a joint system composed of two subsystems with Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, its Hilbert space is the tensor product $\mathcal{H}_A \otimes \mathcal{H}_B$. For $\mathcal{H}_A \simeq \mathbb{C}^{d_A}$ and $\mathcal{H}_B \simeq \mathbb{C}^{d_b}$, then

$$\mathcal{H}_A \otimes \mathcal{H}_B \simeq \mathbb{C}^{d_A d_B}$$

Consider a two-qubit state

$$|\psi\rangle = a_1|0\rangle + b_1|1\rangle, \quad |\phi\rangle = a_2|0\rangle + b_2|1\rangle,$$

We can see that

$$|\psi\rangle \otimes |\phi\rangle = a_1a_2|00\rangle + a_1b_2|01\rangle + a_2b_1|10\rangle + b_1b_2|11\rangle$$

Immediately, we can see that there are states that can not be expressed as a product form, called *entangled states*. For an n-qubit system, the Hilbert space has dimension 2^n, whereas the description of any product state grows only linearly in n. The existence of entangled states makes simulating a quantum system using a classical computer difficult.

A composition of classical systems is given by a probability distribution P_{XY},

$$\rho_{XY} = \sum_{x,y} P_{XY}(x,y)|xy\rangle\langle xy|.$$

Since the systems are independent,

$$\rho_{XY} = \sum_x P_X(x)|x\rangle\langle x| \otimes \sum_y P_Y(y)|y\rangle\langle y| = \rho_X \otimes \rho_Y.$$

The state is said to be a *product state*. For example, the maximally mixed state with

$$\rho_{XY} = \text{diag}(1/4) = \frac{1}{2}I \otimes \frac{1}{2}I.$$

On the other hand, the maximally correlated state,

$$\rho_{XY} = \frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|11\rangle\langle 11|,$$

is not a product state. In general, a state ρ_{XY} is separable if it can be written as

$$\rho_{XY} = \sum_i p_i \, \rho_X^i \otimes \rho_Y^i.$$

If not, it is entangled. A pure state is separable if and only if it is a product state.

2.4 Reduced States

Given a quantum system with associated Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ and an operator M acting on $\mathcal{H}_{AB}$. The *partial trace* of M is defined as

$$\mathrm{Tr}_B[M] = \sum_b (I_A \otimes \langle b|) M (I_A \otimes |b\rangle).$$

where $|b\rangle$ is any orthonormal basis of $\mathcal{H}_B$. The *reduced state* on system A as

$$\rho_A = \mathrm{Tr}_B[\rho_{AB}].$$

Operationally, it is equivalent to ignoring other systems. The reduced state of a classical deterministic state $\rho_{XY} = |x_0 y_0\rangle\langle x_0 y_0|$ is

$$\mathrm{Tr}_Y[\rho_{XY}] = \sum_y |x_0\rangle\langle x_0| \otimes \langle y|y_0\rangle\langle y_0|y\rangle = |x_0\rangle\langle x_0|.$$

The reduced state of a product state $\rho_{AB} = \rho_A \otimes \rho_B$ is given by

$$\mathrm{Tr}_Y[\rho_{XY}] = \sum_b \rho_A \otimes \langle b|\rho_B|b\rangle = \rho_A \otimes \mathrm{Tr}[\rho_\mathrm{B}] = \rho_A.$$

Consider the maximally entangled state

$$\rho_{AB} = \frac{1}{2}(|00\rangle\langle 00| + |11\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 11|)$$

The reduced state is

$$\mathrm{Tr}[\rho_{AB}] = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \frac{1}{2} I.$$

We recover a maximally mixed state. We see that the reduced state of a maximally entangled pure state can be mixed. This inspires the idea that a mixed state can be p̈urifiedẗo a pure state in a larger Hilbert space. Consider a pure state over systems A and B, denoted as $|\psi\rangle_{AB}$, its Schmidt decomposition is given by

$$|\psi\rangle_{AB} = \sum_{j=1}^{\min\{|\mathcal{H}_A|,|\mathcal{H}_B|\}} \sqrt{\lambda_j}|e_j\rangle \otimes |f_j\rangle.$$

Therefore, a mixed state

$$\rho_A = \sum_j \lambda_j |e_j\rangle\langle e_j|$$

can be purified into $|\psi\rangle_{AB}$. We say that $|\psi\rangle_{AB}$ is a purification of ρ_A.

$$\rho_A = \mathrm{Tr}_B[|\psi_{AB}\rangle\langle\psi_{AB}|].$$

since

$$\mathrm{Tr}_B[|\psi_{AB}\rangle\langle\psi_{AB}|] = \mathrm{Tr}_B\left[\sum_{i,j} \sqrt{\lambda_i \lambda_j}|e_i\rangle\langle e_j| \otimes |f_i\rangle\langle f_j|\right] = \sum_i \lambda_i |e_i\rangle\langle e_j|.$$

However, it is clear that a purification of ρ_A is not unique,

$$|\psi'\rangle_{AB} = (I_A \otimes U_B)|\psi\rangle_{AB}.$$

2.5 Measurement

A general measurement is described by a collection of positive semi-definite operators $\{\Pi^\omega\}_\omega$ that satisfy the completeness relation.

$$\sum_\omega \Pi^\omega = I.$$

The probability of some outcome ω is given by

$$P_\Omega(\omega) = \mathrm{Tr}[\rho\Pi^\omega].$$

The post-measurement state becomes

$$\rho \ \rightarrow \ \frac{\Pi^\omega \rho \Pi^\omega}{\mathrm{Tr}[\rho\Pi^\omega]}.$$

If we were to measure system A on the composite system AB with $\{\Pi_A^\omega\}_\omega$, we see that

$$P_\Omega(\omega) = \mathrm{Tr}[\rho_{AB}(\Pi_A^\omega \otimes I_B)] = \mathrm{Tr}[\rho_A \Pi_A^\omega].$$

This means that the measurement on system A contains no information about system B.

2.6 Evolution

In a classical channel, an input value $x \in \mathcal{X}$ is mapped to an output $y \in \mathcal{Y}$. The channel can be inherently random, i.e., given an input distribution P_X, we calculate

$$P_Y(y) = \sum_{x\in\mathcal{X}} P_X(x) P_{Y|X}(y|x).$$

A *quantum channel*, denoted as $\mathcal{C}$, transforms a given ensemble of quantum systems to another.

$$\rho \ \rightarrow \ \mathcal{C}(\rho).$$

It has some properties:

1. It preserves convex combinations.

$$\mathcal{C}(\sum_i p_i \rho_i) = \sum_i p_i \mathcal{C}(\rho_i).$$

2. It is trace-preserving.

$$\mathrm{Tr}[\mathcal{C}(\rho)] = \mathrm{Tr}[\rho].$$

3. It is positive.

$$\mathcal{C}(\rho) \succeq 0, \quad \forall \rho \succeq 0.$$

4. It is completely positive, i.e. the channel is positive when acting on just part of a larger system.

$$(\mathcal{C}_A \otimes I_B)\rho_{AB} \succeq 0.$$

Hence, a quantum channel is a linear superoperator and it is completely positive and trace-preserving(CPTP). For example, a unitary transformation is given by

$$\mathcal{C}(\rho) = U\rho U^\dagger.$$

In general, a quantum channel acts as

$$\mathcal{C}(\rho) = \sum_i M_i \rho M_i^\dagger. \tag{2.2}$$

where the operators $\{M_i\}$ satisfy the completeness relation. The word "channel" is drawn from communication theory — we are to imagine a sender who transmits the state ρ though a communication link to another party who receives the modified state $\mathcal{C}(\rho)$. Sometimes the word superoperator is used as a synonym for quantum channel, where "super" conveys that the map takes operators to operators rather than vectors to vectors. (2.2) is said to be an *operator-sum representation* of the quantum channel, and the operators $\{M_i\}$ are called *Kraus operators* of the channel.

2.6.1 Stinespring dilation

Any quantum channel $\mathcal{C}$ can be written in the form

$$\mathcal{C}(M) = \mathrm{Tr}_E[VMV^\dagger].$$

where V is an isometry from $\mathcal{H}_A$ to $\mathcal{H}_B \otimes \mathcal{H}_E$. This means that any CPTP map on a density matrix can be realized as a unitary operation on a larger system. One may view U as a purification of the quantum channel $\mathcal{C}$. We can show that the Stinespring dilation and the Kraus representation is equivalent. Let $V = \sum_i X_i \otimes |i\rangle$,

$$\mathrm{Tr}_E[VMV^\dagger] = \mathrm{Tr}_E\left[\sum_{ij} X_i M X_j^\dagger \otimes |i\rangle\langle j|\right] = \sum_i X_i M X_i^\dagger.$$

Conversely, let $X_i = (I_B \otimes \langle i|)V$,

$$\sum_i X_i M X_i^\dagger = \sum_i (I_B \otimes \langle i|)VMV^\dagger(I_B \otimes |i\rangle) = \mathrm{Tr}_E[VMV^\dagger].$$

3 No-Go Theorem

In theoretical physics, a no-go theorem is a theorem that states that a situation (either a task or a result) is not physically possible given some specific setup. An example of a no-go theorem is the Heisenberg uncertainty principle.

3.1 No-Cloning Theorem

Consider the setup of a set $\mathcal{S}$ of states that contains at least one pair of non-orthogonal states, and three systems A, B, M denoted by $|\psi\rangle|0\rangle|M_0\rangle$, where $|M_0\rangle$ is the starting state of any machines/materials required. The no-cloning theorem states that no unitary cloning process exists that achieves exact cloning for all states in $\mathcal{S}$. We prove this by considering two distinct non-orthogonal states in $\mathcal{S}$.

$$|\psi\rangle|0\rangle|M_0\rangle \to |\psi\rangle|\psi\rangle|M_\psi\rangle$$
$$|\phi\rangle|0\rangle|M_0\rangle \to |\phi\rangle|\phi\rangle|M_\psi\rangle$$

The proof is straightforward, since unitary operations preserve the inner product:

$$\langle\psi|\phi\rangle\langle 0|0\rangle\langle M_0|M_0\rangle = \langle\psi|\phi\rangle\langle\psi|\phi\rangle\langle M_\psi|M_\phi\rangle.$$

Since $|\phi\rangle \neq |\psi\rangle$, we have a contradiction. Our assumption is that this "operator" is universal and works on any arbitrary state, and that we have no knowledge of the state. If we did, we could just prepare it.

No-cloning protects the uncertainty relation in quantum mechanics, as it prevents us from performing measurements on the state and learning the observables with an arbitrarily small uncertainty. It also forbids eavesdroppers from creating copies of quantum states in quantum communication, which is vital to quantum cryptography.

Most importantly, no-cloning prevents superluminal signaling, if it were possible:

1. Alice and Bob are distinctly separated and share an EPR pair.
2. Alice sends a yes/no decision to Bob by
   - Measuring in the computational basis if the answer is yes.
   - Measuring in $\{|+\rangle, |-\rangle\}$ if the answer is no.
3. Without knowing the state, Bob clones it to create a million copies.
4. Bob now measures the state, if the answer is yes, Bob will always get $|0\rangle$ or always get $|1\rangle$ when measuring in the computational basis, and vice versa.

As a result, Bob is able to instantaneously distinguish between yes and no with an arbitrarily high probability.

3.2 No-Signaling theorem

Without cloning, Bob can measure Alice's state only once. The probability of Bob's outcome $|e_i\rangle$ when measuring with basis $\{e_0, e_1\}$ and Alice says yes is given by

$$\frac{1}{2}\|\langle 0|e_i\rangle\|^2 + \frac{1}{2}\|\langle 1|e_i\rangle\|^2 = \frac{1}{2}.$$

If Alice says no,

$$\frac{1}{2}\|\langle +|e_i\rangle\|^2 + \frac{1}{2}\|\langle -|e_i\rangle\|^2 = \frac{1}{2}$$

There is no distinction between the two.

Alice may also try to transmit information by choosing to measure or not to measure a state. If Alice did not perform measurement on a general entangled state

$$|\psi\rangle = \sum_{a,b} c_{ab}|a\rangle|b\rangle.$$

Bob will find b with probability $\sum_a |c_{ab}|^2$. Had Alice measured the state, she would get outcome a with probability

$$p_A(a) = \sum_b |c_{ab}|^2.$$

where we denote the random variable of Alice's and Bob's measurement outcome by A and B. Given this, Bob's measurement satisfies

$$p_{AB}(a,b) = p_A(a)p(b|a) = |c_{ab}|^2$$

And again

$$p_B(b) = \sum_a |c_{ab}|^2.$$

Once again, there is no distinction between Alice's actions. Algebraically, this is because the measurements are commuting operators $\Pi_a \otimes I_b$ and $I_a \otimes \Pi_b$.

4 Quantum Protocols

Communication, simply put, is the transportation of a state from system A to system B. In noiseless communication, there are three main channels:

- **Noiseless cbit channel**: Measure the state with respect to $\{|0\rangle, |1\rangle\}$ and send the collapsed state. We represent this as 1 cbit or $[c \to c]$.
- **Noiseless qubit channel**: $|\psi\rangle_A \to |\psi\rangle_B$ for all $|\psi\rangle_A \in \mathcal{H}$. We represent this as 1 qubit or $[q \to q]$.
- **Quantum entanglement**: 1 ebit or $[qq]$ is an EPR pair

  $$\frac{1}{\sqrt{2}}(|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B)$$

  shared between system A and B.

4.1 Resource Inequality

We define a resource inequality this $\geq$ that such that the resources of this can be used to do any tasks that can be done with that. Some basic resource inequalities include

- $1\,\text{qubit} \geq 1\,\text{cbit}$.
- **Holevo's bound:** It is impossible to reliably communicate more than n cbits of information in n qubits.

  $$1\,\text{qubit} \ngeq N\,\text{cbits} \quad \forall N \geq 1.$$

  The classical channel capacity of a noiseless qubit channel is 1-cbit.
- From the no-cloning theorem, finite cbits cannot perfectly represent a qubit.

  $$N\,\text{cbits} \ngeq 1\,\text{qubit} \quad \forall N \geq 1.$$

- By no-signaling theorem, entanglement alone cannot be used for communication.

  $$N\,\text{ebits} \ngeq 1\,\text{cbit}, \quad N\,\text{ebits} \ngeq 1\,\text{qubit} \quad \forall N \geq 1.$$

- Classical communication cannot be used to generate entanglement.

  $$N\,\text{cbits} \ngeq 1\,\text{ebit} \quad \forall N \geq 1$$

4.2 Entanglement Distribution

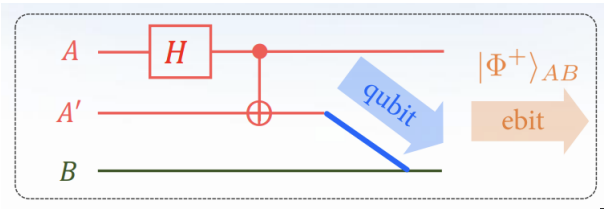


Figure 1: Entanglement Distribution

Entanglement distribution demonstrates that one ebit can be generated with the use of one qubit channel, i.e., $1\,\text{qubit} \geq 1\,\text{ebit}$. Given two parties Alice and Bob, Alice first prepares an EPR pair using on her systems A and A',

$$|00\rangle \xrightarrow{H\otimes I} \frac{1}{\sqrt{2}}(|00\rangle + |01\rangle) \xrightarrow{CNOT} \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle),$$

where H is the Hadamard gate,

$$H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}.$$

Then Alice sends system A' to Bob via a 1-qubit channel.

$$\frac{1}{\sqrt{2}}(|0\rangle_A|0\rangle_{A'} + |1\rangle_A|1\rangle_{A'}) \to \frac{1}{\sqrt{2}}(|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B)$$

Entanglement is usually implemented through photon pairs and distributed via optical fibers or, for greater distances, through satellite-to-ground downlinks.

4.3 Dense Coding

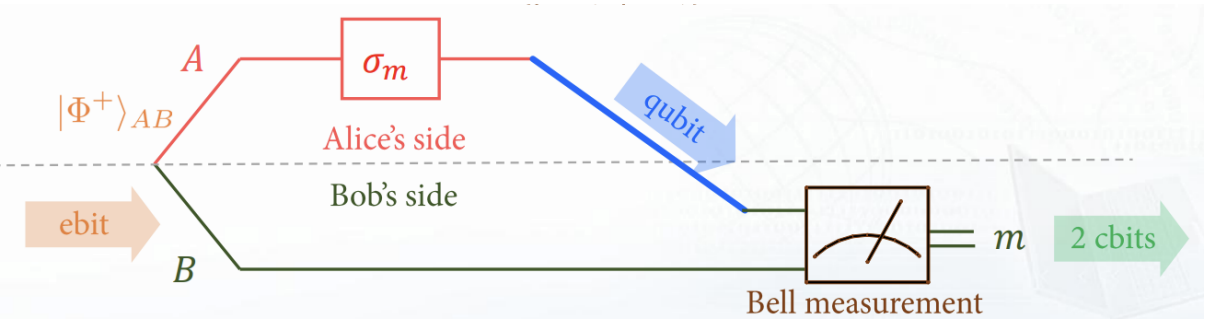


Figure 2: Dense Coding

Recall that 1 qubit can perfectly communicate 1 cbit. In dense coding, two cbits can be communicated with 1 qubit and 1 ebit,

$$1\,\text{qubit} + 1\,\text{ebit} \geq 2\,\text{cbits}\,.$$

The protocol of dense coding is as follows:

1. Alice shares an EPR pair $|\Phi\rangle_{AB}$ with Bob.
2. Alice chooses a Pauli matrix σ_i, where $\sigma_0 = I$, based on her message i and applies σ_i to her qubit.
3. Alice sends her qubit to Bob via the 1 qubit channel.
4. Bob applies the Bell measurement on both qubits to obtain Alice's message.

The result of Alice's operation is one of the four mutually orthonormal Bell states. After Bob receives one of the Bell states, he can perform the Bell measurement by applying the CNOT gate, then the Hadamard gate, which rotates the state back into the computational basis. Note that this operation is directly opposite to how one prepares a Bell state. As a result, Bob obtains Alice's message unambiguously. The dense coding protocol is secure because if Eve intercepts the transmitted 1 qubit, she will gain no information.

After Alice's action and communication		Bob's end	→ C-NOT gate	→ H gate
$i\,j$				
00	$(I \otimes I)\|\Phi_{00}\rangle$	$= \frac{1}{\sqrt{2}}(\|00\rangle + \|11\rangle)$	$= \frac{1}{\sqrt{2}}(\|00\rangle + \|10\rangle)$	$= \|00\rangle$
01	$(X \otimes I)\|\Phi_{00}\rangle$	$= \frac{1}{\sqrt{2}}(\|10\rangle + \|01\rangle)$	$= \frac{1}{\sqrt{2}}(\|11\rangle + \|01\rangle)$	$= \|01\rangle$
10	$(Z \otimes I)\|\Phi_{00}\rangle$	$= \frac{1}{\sqrt{2}}(\|00\rangle - \|11\rangle)$	$= \frac{1}{\sqrt{2}}(\|00\rangle - \|10\rangle)$	$= \|10\rangle$
11	$(XZ \otimes I)\|\Phi_{00}\rangle$	$= \frac{1}{\sqrt{2}}(\|10\rangle - \|01\rangle)$	$= \frac{1}{\sqrt{2}}(\|11\rangle - \|01\rangle)$	$= \|11\rangle$

Bell states

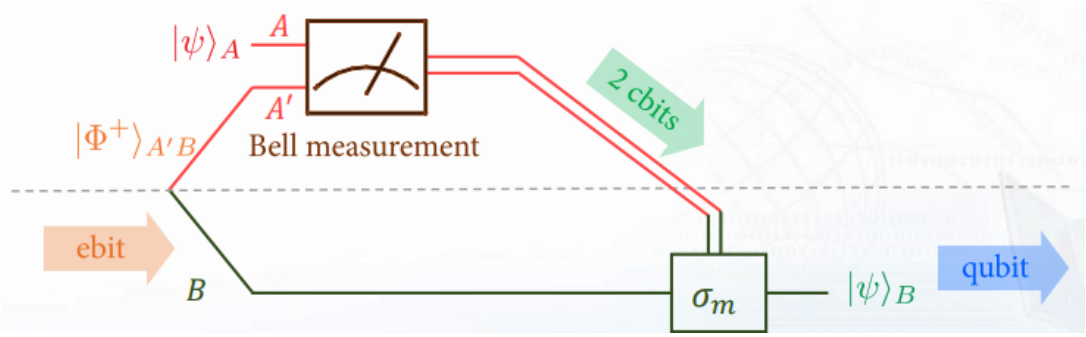


Figure 3: Quantum Teleportation

4.4 Quantum Teleportation

Suppose Alice wants to communicate an unknown qubit state $|\psi\rangle$ to Bob, but is only able to send classical bit strings. She achieves this by with the help of an ebit:

$$2\,\text{cbits} + 1\,\text{ebit} \geq 1\,\text{qubit}\,.$$

The protocol of quantum teleportation is as follows:

1. Alice shares an EPR pair $|\Phi_{00}\rangle_{A'B}$ with Bob.

$$|\psi\rangle_A \otimes |\Phi_{00}\rangle_{A'B} = (a|0\rangle_A + b|1\rangle_A) \otimes \frac{1}{\sqrt{2}}(|00\rangle_{A'B} + |11\rangle_{A'B})$$

2. Alice applies the Bell measurement (CNOT and Hadamard gate) on her state $|\psi\rangle_A$ and system A'.

$$\begin{aligned}
&|\psi\rangle_A \otimes |\Phi_{00}\rangle_{A'B} \xrightarrow{CNOT} \frac{1}{\sqrt{2}}(a|000\rangle_{AA'B} + a|011\rangle_{AA'B} + b|110\rangle_{AA'B} + b|101\rangle_{AA'B}) \\
&\xrightarrow{H} \frac{1}{\sqrt{2}}\left(a\left(\frac{|0\rangle_A + |1\rangle_A}{\sqrt{2}}\right)(|00\rangle_{A'B} + |11\rangle_{A'B}) + b\left(\frac{|0\rangle_A - |1\rangle_A}{\sqrt{2}}\right)(|10\rangle_{A'B} + |01\rangle_{A'B})\right) \\
&= \frac{1}{2}\big(|00\rangle_{AA'}(a|0\rangle + b|1\rangle) + |01\rangle_{AA'}(b|0\rangle + a|1\rangle) + |10\rangle_{AA'}(a|0\rangle - b|1\rangle) + |11\rangle_{AA'}(-b|0\rangle + a|1\rangle)\big) \\
&= \frac{1}{2}\big(|00\rangle|\psi\rangle + |01\rangle X|\psi\rangle + |10\rangle Z|\psi\rangle + |11\rangle(-Y)|\psi\rangle\big)
\end{aligned}$$

3. Alice measures her 2-qubits and gets $00, 01, 10, 11$ with equal probability. She communicates the result to Bob.

4. Bob corrects his state to recover $|\psi\rangle$.

Since each measurement outcome occurs with 25% chance, Alice learns nothing of $|\psi\rangle$. In addition, the only information communicated is the two classical bits, which contain no information about $|\psi\rangle$, so the protocol is secure.

4.5 Quantum Key Distribution

Quantum key distribution provides a method for Alice and Bob to share a secret key over public classical and quantum channels without the need for an intermediary party. It is provably secure against eavesdropping.

4.5.1 BB84

BB84 is a quantum key distribution scheme developed by Charles Bennett and Gilles Brassard in 1984. BB84 takes advantage of the *mutually unbiased bases* (MUB)

$$\mathcal{B}_0 = \{|0\rangle, |1\rangle\}, \quad \mathcal{B}_1 = \{|+\rangle, |-\rangle\}.$$

A key property of MUB is that if any state of one basis is measured in the other basis, the outcomes are equally likely. The protocol of BB84 is as follows:

1. Alice generates two uniformly random binary strings,

$$x = x_1 x_2 \cdots x_m$$

$$y = y_1 y_2 \cdots y_m,$$

   and she prepares the m qubit states $|\psi_{x_1 y_1}\rangle, |\psi_{x_2 y_2}\rangle, \cdots, |\psi_{x_m y_m}\rangle$ and sends them to Bob.

2. Bob chooses a uniformly random bit string $y' = y'_1 y'_2 \cdots y'_m$.

3. Assuming perfect transmission, Bob receives the m qubits, and measures them in basis $\mathcal{B}_{y'_i}$ to get a result x'_i.

4. Alice and Bob reveal and compare their choice of bases y and y'. If $y_i = y'_i$ then $x_i = x'_i$.

5. They discard all bits x_i and x'_i for which $y_i \neq y'_i$.

6. They use the remaining bits $\tilde{x} = \tilde{x}'$ as the key.

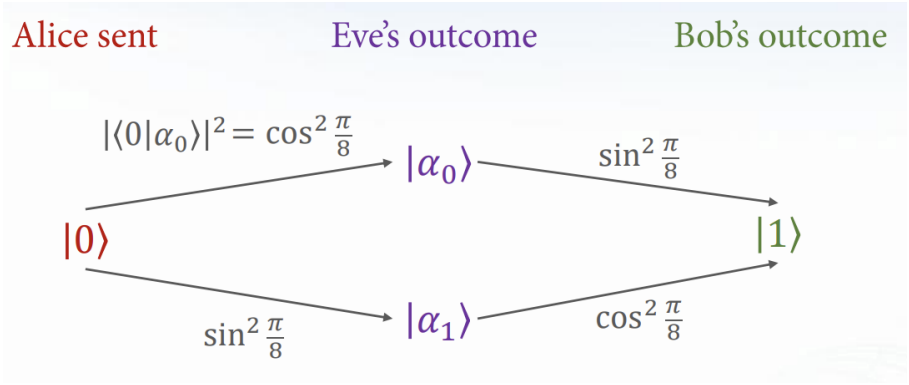


Figure 4: Analysis of the BER

Eve can attack this protocol with an *intercept-resend attack*. Eve intercepts each passing qubit and measures it in the so-called *Breidbart basis*:

$$|\alpha_0\rangle = \cos\frac{\pi}{8}|0\rangle + \sin\frac{\pi}{8}|1\rangle$$
$$|\alpha_1\rangle = -\sin\frac{\pi}{8}|0\rangle + \cos\frac{\pi}{8}|1\rangle$$

It can be shown that Eve has the highest probability of learning the correct bit with this basis. The disturbance caused by eavesdropping will account to a bit error rate of 25%. Eve will learn each bit with probability $\cos^2\frac{\pi}{8} \approx 0.85$.

5 Quantum Computation Model

A *quantum circuit* $\mathcal{C}_n$ is a collection of elementary quantum gates, which are unitary operations on qubits. A *quantum computation* or *quantum algorithm* is defined by a family of quantum circuits. Some elementary quantum gates include:

- Pauli gates X, Y, Z

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \qquad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

- Hadamard gates

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

- Controlled-Z gate

$$\begin{bmatrix} I_2 & 0 \\ 0 & Z \end{bmatrix}$$

- CNOT gate

$$\begin{bmatrix} I_2 & 0 \\ 0 & X \end{bmatrix}$$

- Toffoli gate

$$\begin{bmatrix} I_3 & 0 & 0 \\ 0 & I_3 & 0 \\ 0 & 0 & X \end{bmatrix}$$

- Swap gate $|\psi\rangle|\phi\rangle \to |\phi\rangle|\psi\rangle$

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

A key distinction between classical and quantum computation is that quantum operations are reversible since all quantum gates are unitary. However, we can still calculate any classical boolean function with $|x\rangle|y\rangle \to |x\rangle|y \otimes f(x)\rangle$.

5.1 Quantum Oracle

A query model involves a black-box function called an *oracle* that computes some function f. We define the number of queries required by an algorithm as its *query complexity*. The *quantum oracle* of any boolean function f is given by

$$U_f|x\rangle|y\rangle = |x\rangle|y \otimes f(x)\rangle$$

We require a product state input for the operation to be reversible with $U_f^{-1} = U_f$. In addition to query complexity, we are interested in the time complexity and depth of a quantum circuit. A problem is said to be BQP if there exists a polynomial time randomized quantum algorithm that gives the correct answer with probability at least 2/3. BQP is the quantum analogue of BPP.

5.2 Deutsch-Josza Algorithm

Given a black box for a Boolean function f, it is either a constant function or a balanced function, i.e., $f(x) = 0$ or 1 for exactly half of the inputs. We look to determine if f is balanced or constant. The classical solution to this problem requires at least $2^{n-1} + 1$ queries for a deterministic result. The *Deutsch-Josza Algorithm* solves this problem with only *one* query to the oracle.

1. We first prepare $n + 1$ qubits and create superpositions with the Hadamard gate,

$$|0\rangle^{\otimes n} \otimes |1\rangle \xrightarrow{H^{\otimes(n+1)}} \left(\frac{1}{\sqrt{2^n}} \sum_{x \in \mathbb{Z}_2^n} |x\rangle\right) \otimes |-\rangle.$$

2. Apply the quantum oracle U_f on the above state.

$$U_f|x\rangle\big(\frac{1}{\sqrt{2}}|0\rangle - \frac{1}{\sqrt{2}}|1\rangle\big) = \frac{1}{\sqrt{2}}|x\rangle|f(x)\rangle - \frac{1}{\sqrt{2}}|x\rangle|f(x) \oplus 1\rangle = (-1)^{f(x)}|x\rangle|-\rangle.$$

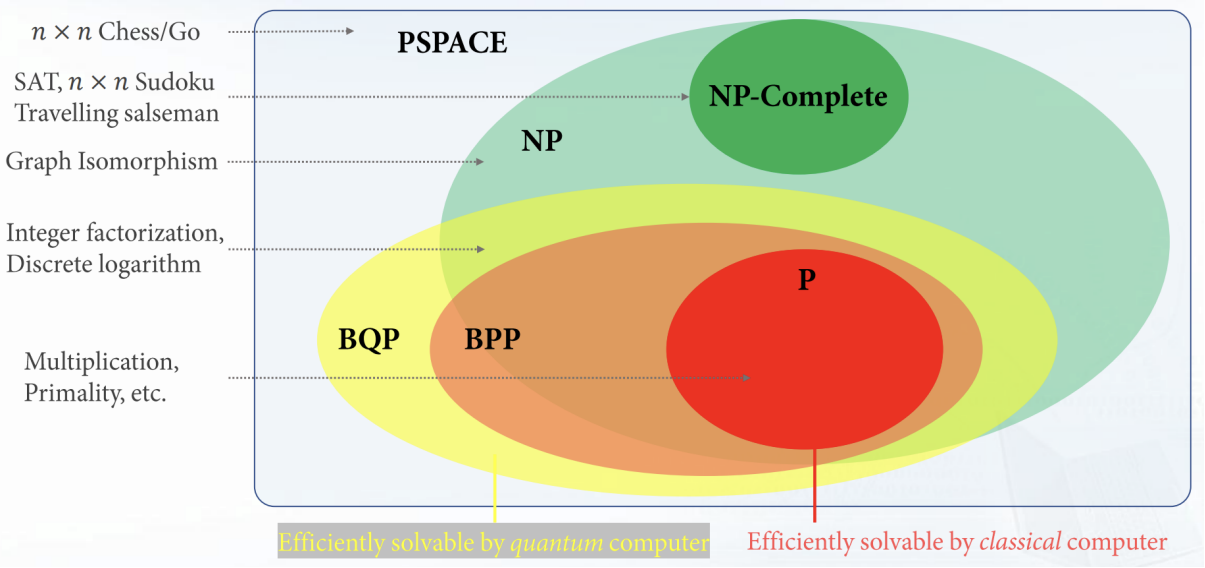


Figure 5: Complexity Classes

The resulting state is

$$|f\rangle = \left(\frac{1}{\sqrt{2^n}} \sum_{x \in \mathbb{Z}_2^n} (-1)^{f(x)} |x\rangle\right) |-\rangle.$$

If f is constant,

$$|f\rangle = \pm\left(\frac{1}{\sqrt{2^n}} \sum_{x \in \mathbb{Z}_2^n} |x\rangle\right) |-\rangle \xrightarrow{H^{\otimes n}} \pm|0^{\otimes n}\rangle.$$

We will get all zeros upon measurement. If f is balanced,

$$|f\rangle \left(\frac{1}{\sqrt{2^n}} \sum_{x \in \mathbb{Z}_2^n} |x\rangle\right) = \langle f|H^{\otimes n}|0\rangle = 0,$$

We see that $H^{\otimes n}|f\rangle$ is orthogonal to $|0\rangle^{\otimes n}$, and we will get non-zero strings with certainty. The problem is solved with one query, which is an exponential speedup compared to the classical approach. The algorithm requires $2n + 1$ Hadamard gates, one X gate, and n single-qubit measurements.

5.3 Bernstein–Vazirani Algorithm

Consider the function f_a given by

$$f_a(x) = x \cdot a = x_1 a_1 \oplus x_2 a_2 \oplus \cdots \oplus x_n a_n.$$

Assume we are provided an oracle that computes f_a, and we want to determine the n-bit string a. Classically, this would require n queries of the oracle. However, the *Bernstein–Vazirani*

Algorithm can solve the problem in *one* query. The protocol is similar to the DJ algorithm,

$$|f_a\rangle = \frac{1}{\sqrt{2^n}} \sum_{x \in \mathbb{Z}_2^n} (-1)^{f_a(x)} |x\rangle = \frac{1}{\sqrt{2^n}} \sum_{x \in \mathbb{Z}_2^n} (-1)^{x \cdot a} |x\rangle.$$

We observe that this is identical to

$$|f_a\rangle = H_n |a\rangle,$$

where $H_n = H^{\otimes n}$. It follows that, if we follow the DJ Algorithm,

$$H_n |f_a\rangle = H_n^2 |a\rangle = |a\rangle.$$

We can measure $|a\rangle$ and find a with certainty. Compared to the classical method, the Bernstein–Vazirani algorithm provides a polynomial speedup.

5.4 Simon's Algorithm

We are given an oracle of a function f, which is either one-to-one or two-to-one with $f(x) = f(y)$ if and only if $y = x \oplus \xi$, where ξ is the hidden period of the function f. We must determine which one is true.

To begin, consider $2n$ qubits with the first n comprising the input, and the last n comprising the output register for a quantum oracle U_f computing $U_f|x\rangle|y\rangle = |x\rangle|y \oplus f(x)\rangle$. We start with all zero states,

1. Apply H_n to the input register

$$H_n |0\rangle^{\otimes n} |0\rangle^{\otimes n} = \frac{1}{\sqrt{2^n}} \sum_{x \in \mathbb{Z}_2^n} |x\rangle |0\rangle^{\otimes n}$$

2. Apply U_f

$$U_f \frac{1}{\sqrt{2^n}} \sum_{x \in \mathbb{Z}_2^n} |x\rangle |0\rangle^{\otimes n} = \frac{1}{\sqrt{2^n}} \sum_{x \in \mathbb{Z}_2^n} |x\rangle |f(x)\rangle.$$

3. Measure the output register

$$\frac{1}{\sqrt{2}} (|x\rangle + |x \oplus \xi\rangle) |f(x)\rangle.$$

4. Apply H_n to the input, if f is two-to-one,

$$H_n \frac{1}{\sqrt{2}} (|x\rangle + |x \oplus \xi\rangle) = \frac{1}{\sqrt{2^{(n+1)}}} \sum_{y \in \mathbb{Z}_2^n} (-1)^{x \cdot y} + (-1)^{(x+\xi) \cdot y} |y\rangle.$$

If f is one-to-one,

$$H_n |x\rangle = \frac{1}{\sqrt{2^n}} \sum_{y \in \mathbb{Z}_2^n} (-1)^{x \cdot y} |y\rangle.$$

We observe that, regardless of whether f is one-to-one or two-to-one, the measurement outcome of $H_n|x\rangle$ will satisfy $y \cdot \xi = 0$. We can then perform this operation repeatedly to obtain a system of linear equations, which can be solved by Gaussian elimination in polynomial time to obtain ξ. If $\xi = 0$, then we know the function is one-to-one. The probability that $n-1$ bit string y are linearly independent is,

$$\prod_{k=1}^{n-1} \left(1 - \frac{2^{k-1}}{2^{n-1}}\right) \geq \frac{1}{4}.$$

We can repeat this protocol c times to achieve a success probability of at least $1 - (3/4)^c$ while maintaining a query complexity of $c(n-1) \in O(n)$.

5.5 Amplitude Amplification

The search problem involves searching through a Boolean function f for a unique input x such that $f(x) = 1$. Classically, $\Theta(2^n)$ queries are sufficient and necessary. Using quantum, *Grover's Algorithm* finds x with high probability in $O(\sqrt{N})$ queries. We define the *Grover diffusion operator* $\mathcal{G}$ as,

$$\mathcal{G} = -H_n I_{|0\rangle} H_n I_{|x_0\rangle} = I_{|\psi_0^\perp\rangle} I_{|x_0\rangle},$$

where $I_{|\phi\rangle}$ is the reflection operator in the subspace that is orthogonal to $|\phi\rangle$.

$$I_{|\phi\rangle} = I - 2|\phi\rangle\langle\phi|.$$

Initially, $|\psi_0\rangle$ is given by

$$|\psi_0\rangle = \frac{1}{\sqrt{2^n}} \sum_{x \in \mathbb{Z}_2^n} |x\rangle.$$

Then it can be shown that $\mathcal{G}$ is a rotation operator.

$$\mathcal{G} = I_{|\psi_0^\perp\rangle} I_{|x_0\rangle} = \begin{bmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix}$$

where $\sin\theta = \langle\psi_0|x_0| =\rangle 1/\sqrt{2^n}$. Thus, approximately $\pi/2\theta \approx \pi\sqrt{N}/4$ iterations will be needed to reach $|x_0\rangle$. However, since we do not know $|x_0\rangle$, we must implement $I_{|x_0\rangle}$ using phase kickback and an ancilla $|-\rangle$,

$$|x\rangle|-\rangle \xrightarrow{U_f} |x\rangle|- \oplus f(x)\rangle = (-1)^{f(x)}|x\rangle|-\rangle = I_{|x_0\rangle}|x\rangle|-\rangle.$$

To implement $I_{|\psi_0\rangle}$,

$$I_{|\psi_0\rangle} = H_n(2|0\rangle\langle 0|^{\otimes n} - I)H_n.$$

Note that $I_{|\psi_0\rangle}$ is the *inverse about mean* operator:

$$I_{|\psi_0\rangle}|\phi\rangle = 2|\psi_0\rangle\langle\psi_0|\phi\rangle - |\phi\rangle = \sum_{x \in \mathbb{Z}_2^n} (2\bar{a} - a_x)|x\rangle.$$

It can be implemented with a Multi-controlled Z gate,

$$X^{\otimes n}(\text{MCZ})X^{\otimes n}|x_1\rangle...|x_n\rangle = (-1)^{\text{if all } x_i=0}|x_1\rangle...|x_n\rangle.$$

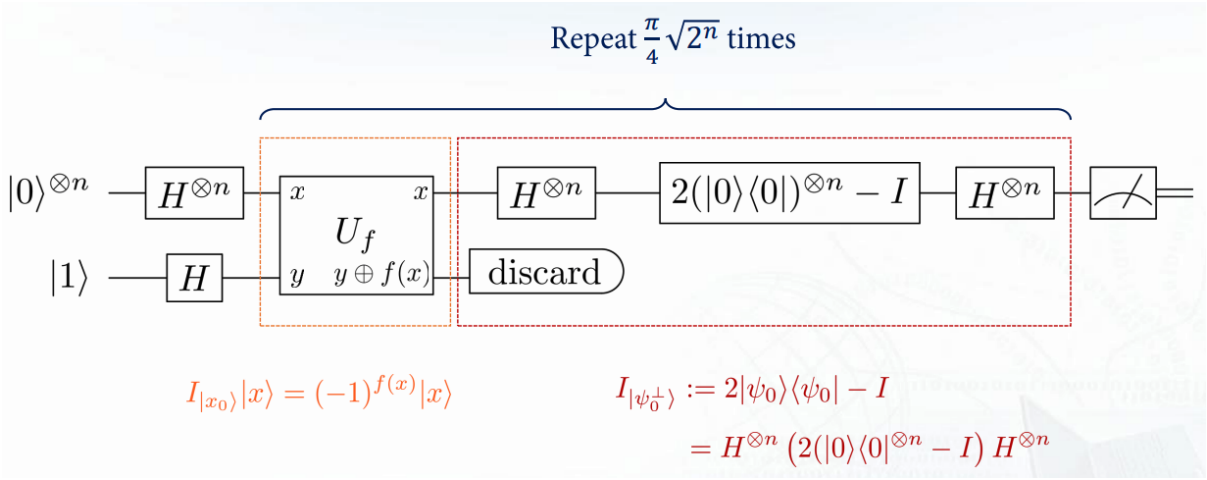


Figure 6: Circuit Diagram

6 Quantum Fourier Transform

The quantum Fourier Transform over $\mathbb{Z}_N$, denoted by Q_N, is defined as

$$Q_N|x\rangle = \frac{1}{\sqrt{N}} \sum_{y \in \mathbb{Z}_N} e^{2\pi i xy/N}|y\rangle = \frac{1}{\sqrt{N}} \sum_{y \in \mathbb{Z}_N} \omega^{xy}|y\rangle$$

The matrix entry of Q_N is given by

$$\langle j|Q_N|k\rangle = \frac{1}{\sqrt{N}}\omega^{jk}.$$

In particular,

$$Q_2 = H.$$

We can show that Q_N is unitary,

$$\langle j|Q_N^{\dagger}Q_N|k\rangle = \frac{1}{N} \sum_{y \in \mathbb{Z}_N} \omega^{-jy}\omega^{ky} = \delta_{kj}.$$

To implement the QFT, we decompose the phase ϕ into its binary form. Considering only a 1-qubit state and $\phi = (0.x_1)$,

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

Suppose we have a 2-qubit state and $\phi = (0.x_1x_2)$,

$$\begin{aligned}\frac{1}{2} \sum_{y=0}^{3} e^{2\pi i(0.x_1x_2)y}|y\rangle &= \frac{1}{2}(|0\rangle + e^{2\pi i(0.x_1x_2)}|1\rangle + e^{2\pi i(0.x_2)}|2\rangle + e^{2\pi i(0.x_2)}e^{2\pi i(0.x_1x_2)}|3\rangle) \\ &= \left(\frac{|0\rangle + e^{2\pi i(0.x_2)}|1\rangle}{\sqrt{2}}\right) \otimes \left(\frac{|0\rangle + e^{2\pi i(0.x_1x_2)}|1\rangle}{\sqrt{2}}\right).\end{aligned}$$

If $x_2 = 1$, in addition to the Hadamard gate, we need to cancel the phase difference $e^{2\pi i(0.01)}$ with the rotation Z gate:

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

Generally, a n-qubit quantum Fourier state can be rewritten as follows,

$$\frac{1}{\sqrt{2^n}} \sum_{y=0}^{2^n-1} e^{2\pi i \phi y} |y\rangle = \bigotimes_{\ell=1}^{n} \frac{1}{\sqrt{2}} (|0\rangle + e^{2\pi i (0.x_{n-\ell+1} \ldots x_n)} |1\rangle).$$

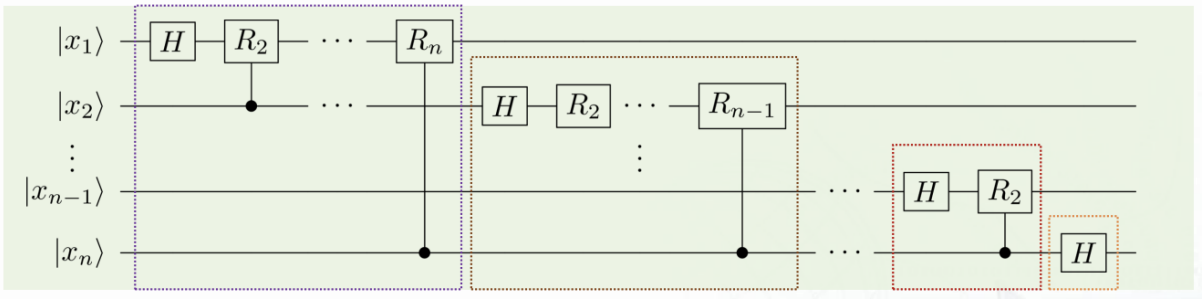


Figure 7: QFT Circuit Implenmentation

The number of gates needed is $n(n+1)/2 \in O(n^2)$. The complexity of the classical FFT is $O(n2^n)$. Since the rotation Z gates are close to identity for large k, we can remove all gates with $k \geq \log(n/\epsilon)$, which reduces the complexity to $O(n \log n)$.

6.1 Eigenvalue Estimation

We are given a unitary operator U with eigenvector $|\psi\rangle$ such that $U|\psi\rangle = e^{2\pi i \phi}|\psi\rangle$. How do we estimate the phase parameter ϕ? We use the phase estimation algorithm with the following procedure:

1. Prepare the uniform superposition state using QFT or the Hadamard gate,

$$|0\rangle^{\otimes n}|\psi\rangle \quad \xrightarrow{Q_{2^n}} \quad \frac{1}{\sqrt{2^n}} \sum_{y \in \mathbb{Z}_{2^n}} |y\rangle|\psi\rangle.$$

2. Apply the Controlled-U gate c-U^y : $|y\rangle|\psi\rangle \to |y\rangle U^y|\psi\rangle$, such that

$$|y\rangle|\psi\rangle \xrightarrow{\text{c-}U^y} e^{2\pi i \phi y}|y\rangle|\psi\rangle.$$

   Notice the phase kickback.

3. Apply the inverse quantum Fourier transform,

$$\frac{1}{\sqrt{2^n}} \sum_{y \in \mathbb{Z}_{2^n}} e^{2\pi i \phi y}|\psi\rangle \xrightarrow{Q_{2^n}^{-1}} |\phi\rangle.$$

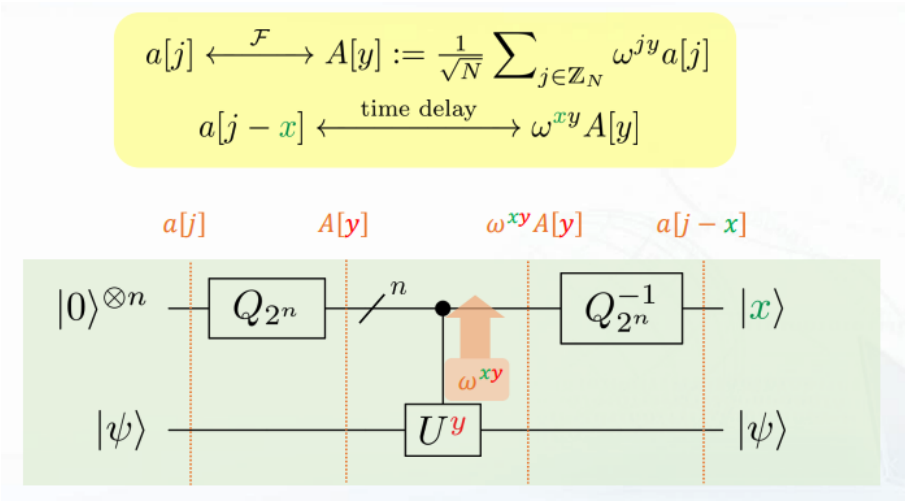


Figure 8: Circular shift property

The analogy of the phase estimation algorithm in DFT is the circular shift property.

6.2 Period-Finding Algorithm

Given a periodic function f that satisfies $f(x) = f(x+r)$ for some period r. Assume f is one-to-one in each period. Further assume r divides N exactly such that $m = N/r$. Classically, the *baby step giant step* algorithm finds the period in $O(\sqrt{N})$ queries. The quantum solution requires only $O(\log(\log N))$ queries and polynomial processing steps.

1. Construct a uniform superposition

$$\frac{1}{\sqrt{N}} \sum_{x \in \mathbb{Z}_n} |x\rangle$$

   and feed it through the oracle

$$\frac{1}{\sqrt{N}} \sum_{x \in \mathbb{Z}_n} |x\rangle |f(x)\rangle$$

2. Measure the second register to get some outcome y_0, such that the first register collapses into equal superposition of m different values $x = x_0, x_0 + r, \ldots$ and $f(x_0) = y_0$.

3. Apply QFT,

$$\frac{1}{\sqrt{m}} \sum_{j=0}^{m-1} |x_0 + jr\rangle \xrightarrow{Q_N} \frac{1}{\sqrt{mN}} \sum_{j=0}^{m-1} \sum_{y \in \mathbb{Z}_N} \omega^{(x_0+jr)y} |y\rangle.$$

   Noting that

$$\sum_{j=0}^{m-1} \omega^{jry} = \frac{1 - e^{2\pi i y}}{1 - e^{2\pi i y/m}} = \begin{cases} m, & \text{if } \ y = \ell m \\ 0 & \text{otherwise} \end{cases}$$

It follows that

$$\frac{1}{\sqrt{mN}} \sum_{j=0}^{m-1} \sum_{y \in \mathbb{Z}_N} \omega^{(x_0+jr)y} |y\rangle = \frac{1}{\sqrt{r}} \sum_{\ell=0}^{r-1} \omega^{x_0 \ell m} |\ell m\rangle.$$

4. Measure the state to obtain some outcome c that satisfies

$$\frac{c}{N} = \frac{\ell}{r}.$$

Calculate the fraction using the classical continued fraction to get the period r. It is possible that ℓ is not coprime to r, so we must verify it with $f(0) = f(r)$.

The number of integers less than r that are coprime to r grows as $O(r/\log\log r)$. Hence, if we repeat the process $O(\log\log r) \in O(\log\log N)$ times, we will obtain a coprime ℓ with an arbitrarily high probability.

It is possible that the dimension of the Hilbert space is not a multiple of the period, i.e., $2^t/r \notin \mathbb{N}$. We follow the same protocol as before, but now the result of the QFT is

$$\sum_{y \in \mathbb{Z}_{2^t}} \frac{1}{2^t m_{x_0}} \left| \frac{\sin(\pi r y m_{x_0}/2^t)}{\sin(\pi r y/2^t)} \right|^2 |y\rangle$$

It can be shown that

$$\frac{1}{2^t m_{x_0}} \left| \frac{\sin(\pi r y m_{x_0}/2^t)}{\sin(\pi r y/2^t)} \right|^2 \geq \frac{4 m_{x_0}}{2^t \pi^2} \approx \frac{4}{r\pi^2},$$

if

$$\left| \frac{r}{2^t} y - \ell \right| \leq \frac{1}{2 m_{x_0}}.$$

for the closest integer ℓ to $\frac{r}{2^t} y$.

6.3 Shor's Algorithm

An application of the period-finding algorithm is the problem of order finding. Given an integer a and a coprime number N. The *order* of a modulo N is defined as the least power of a such that $a^r = 1 \mod N$. It follows that the order-finding algorithm is equivalent to the period-finding problem if $f(x) = a^x \mod N$. With this, we can perform integer factorization on N:

1. Check if N is even.
2. Check if N is a prime power using the classical polynomial time algorithm.
3. Choose uniformly randomly $a < N$ and compute $\gcd(a, N)$. If the number is not 1, output it.
4. Find the order r of a using the period-finding algorithm. If r is even we can decompose $a^r - 1 = (a^{r/2} - 1)(a^{r/2} + 1)$. This is called *Miller's reduction*.
5. Compute $\gcd(a^{r/2} + 1, N)$, if it is not N, it is a factor of N.

It cannot be the case that $N \mid a^{r/2} - 1$ since r by definition is the smallest period of a^x. Either $N \mid a^{r/2} + 1$ or N and $a^{r/2} + 1$ share some common factor. The probability that r is even and $N \nmid a^{r/2} + 1$ is greater than $1/2$. The complexity of Shor's algorithm is as follows: We need $t = O(n)$ Hadamard gates to create a uniform superposition. We can compute $f(x)$ with $O(n^2 \log n \log\log n)$ gates via repeated squaring. QFT can be computed in $O(n^2)$ gates. The process is repeated $\log n$ times to achieve a constant level of probability. The total complexity is

$$O(\log n)(O(n) + O(n^2 \log n \log\log n) + O(n^2)) = O(n^2 \log^2 n \log\log n).$$

The runtime bottleneck of Shor's algorithm is quantum modular exponentiation, which will take about $O(n^3)$ time. It is still significantly faster than the most efficient known classical factoring algorithm, the general number field sieve, which works in sub-exponential time $O(e^{1.9n^{1/3}\log^{3/2} n})$.

7 Bell's Inequality

Hypothetical characters Alice and Bob stand in widely separated locations. Their colleague Victor prepares a pair of particles and sends one to Alice and the other to Bob. When Alice receives her particle, she chooses to perform one of two possible measurements, A_0 or A_1, the result of which is either 1 or -1. Bob performs measurements B_0 or B_1, also with binary results. Let a_0, a_1, b_0, b_1 denote the result of each measurement, respectively. Consider the quantity,

$$a_0b_0 + a_0b_1 + a_1b_0 + a_1b_1 = (a_0 + a_1)b_0 + (a_0 - a_1)b_1.$$

One of the terms in the RHS must be zero, and the other term must be ± 2, thus

$$|\langle A_0B_0\rangle + \langle A_0B_1\rangle + \langle A_1B_0\rangle - \langle A_1B_1\rangle| \leq 2.$$

No single trial can measure this quantity, because Alice and Bob can only choose one measurement each, but on the assumption that the underlying properties exist, the average value of the sum is just the sum of the averages for each term. This is a *Bell inequality*, specifically, the *CHSH inequality*.

Quantum mechanics can violate the CHSH inequality. For example, have Alice and Bob share an EPR pair,

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

And let

$$A_0 = Z,\ A_1 = X,\ B_0 = H,\ B_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ -1 & -1 \end{bmatrix}.$$

We can show that

$$\langle\psi|A_0 \otimes B_0|\psi\rangle = \langle\psi|A_0 \otimes B_1|\psi\rangle = \langle\psi|A_1 \otimes B_0|\psi\rangle = -\langle\psi|A_1 \otimes B_1|\psi\rangle = \frac{1}{\sqrt{2}}$$

and thus

$$\langle A_0B_0\rangle + \langle A_0B_1\rangle + \langle A_1B_0\rangle - \langle A_1B_1\rangle = 2\sqrt{2}.$$

In fact, the value $2\sqrt{2}$ is the largest that quantum physics permits for this combination of expectation values, according to *Tsirelson bound*.

There are many interpretations of Bell's inequality. The experiments show that

- **Hidden variables** Quantum randomness is not due to ignorance. Rather, it is intrinsic, occurring even when we have the most complete knowledge that Nature will allow. The values of each measurement are not known beforehand, instead, it is determined by the intrinsic randomness of the measurement. This rules out the hidden variable theory, in which the state of the system is predetermined by some unknown variables.

- **Nonlocality** By design, Alice and Bob are situated at locations that are far apart. However, we have shown that the result of Bob's measurement can influence the outcome of Alice's, due to quantum entanglement.

- **Counterfactuals** Don't reason about counterfactuals. When the measurements are incompatible, then if we do measurement 1, we can't speak about what would have happened if we had done measurement 2 instead.

7.1 Nonlocal Games

Consider the setting of two players, Alice and Bob, cooperatively playing a game against the referee. All communication is between the player and the referee, and no communication between players. The players answer a question given by the referee. The question is randomly chosen. Before the game, the player may meet and agree on a strategy. Afterwards, they move far apart from each other. Such a setting is called a *Nonlocal Game*.

For example, the *Clauser–Horne–Shimony–Holt Game* has the referee give out bits x and y to Alice and Bob as questions, while Alice and Bob must reply with bits a and b. The winning rule is

$$x \wedge y = a \oplus b.$$

For Alice and Bob to devise a perfect solution,

$$a_0 \oplus b_0 = a_0 \oplus b_1 = a_1 \oplus b_0 = 0, \quad a_1 \oplus b_1 = 1$$

However, we see that this is impossible when summing over each term. We can formulate this problem to an equivalent CHSH inequality with

$$A_0 = (-1)^{a_0}, \quad A_1 = (-1)^{a_1}$$
$$B_0 = (-1)^{b_0}, \quad B_1 = (-1)^{b_1}$$

where

$$\begin{aligned}
\langle A_0 B_0 \rangle &= \mathbb{E}_{a_0, b_0 \in \mathbb{Z}^2}[(-1)^{a_0 \oplus b_0}] = p_{00} - (1 - p_{00}) = 2p_{00} - 1, \\
\langle A_0 B_1 \rangle &= p_{01} - (1 - p_{01}) = 2p_{01} - 1, \\
\langle A_1 B_0 \rangle &= p_{10} - (1 - p_{10}) = 2p_{10} - 1, \\
\langle A_1 B_1 \rangle &= -p_{11} + (1 - p_{11}) = 1 - 2p_{11}.
\end{aligned}$$

and p_{xy} denotes the probability of Alice and Bob winning when the input is x and y,

$$p_{xy} = p[a \oplus b = x \wedge y].$$

It follows from the CHSH inequality that,

$$\frac{1}{4}(p_{00} + p_{01} + p_{10} + p_{11}) \leq \frac{3}{4}.$$

If the referee gives out question bits x, y uniformly, Alice and Bob cannot attain a probability of winning higher than $3/4$. Of course, if Alice and Bob share quantum entanglement, their winning probability is now bounded by the Tsirelson bound, and thus

$$\frac{1}{4}(p_{00} + p_{01} + p_{10} + p_{11}) \leq \frac{1}{2} + \frac{\sqrt{2}}{4} \approx 0.853.$$

Consider the following example, Alice and Bob share the EPR pair $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, Alice will measure the EPR pair with $\{|0\rangle, |1\rangle\}$ if she receives $x = 0$, and $\{|+\rangle, |-\rangle\}$ if she receives $x = 1$. Bob will measure along the basis with $\theta = \pi/8$ when $y = 0$, and along $\theta = -\pi/8$ when $y = 1$. Consider each input possibilities separately,

1. $x = 0, y = 0$: Assume Alice measures and get $|0\rangle$, then Bob will get $b = 0$ with probability $\cos^2(\pi/8)$.

2. $x = 0, y = 1$ and $x = 1, y = 0$: Again, since the bases are only separated by an angle of $\pi/8$, we have again a winning probability of $\cos^2(\pi/8)$.

3. $x = 1, y = 1$: If Alice gets $a = 0$, Bob will get $b = 1$ with probability $\cos^2(\pi/8)$, but since $a \oplus b = x \wedge y$, the winning probability is again $\cos^2(\pi/8)$.

With this strategy, we have shown that this indeed achieves a success probability of

$$\cos^2(\pi/8) = \frac{1}{2} + \frac{\sqrt{2}}{4} \approx 0.853.$$

In agreement with the Tsirelson bound.

8 Quantum Fidelity

The *trace distance* between two state ρ, σ is defined as

$$T(\rho, \sigma) = \frac{1}{2}\|\rho - \sigma\|_1,$$

where $\|\cdot\|_1$ is the *trace norm*,

$$\|M\|_1 = \mathrm{Tr}\left[\sqrt{M^\dagger M}\right].$$

The trace distance has the following important properties,

- $T(\rho, \sigma) = 0$ if and only if $\rho = \sigma$.
- $0 \leq T(\rho, \sigma) \leq 1$.
- For pure states $\rho = |\psi\rangle\langle\psi|$ and $\sigma = |\phi\rangle\langle\phi|$, we have

$$T(\rho, \sigma) = \sqrt{1 - |\langle\psi|\phi\rangle|^2}.$$

- Variational formula:

$$T(\rho, \sigma) = \sup_{0 \leq \Pi \leq I} \mathrm{Tr}[(\rho - \sigma)\Pi].$$

- It is invariant under unitary transformation.
- It is invariant under ancilla.
- It is convex.
- It is contractive under CPTP: $T(\mathcal{C}(\rho), \mathcal{C}(\sigma)) \leq T(\rho, \sigma)$.

A operational interpretation of the trace distance is through the Holevo-Helstöm theorem. Consider the discrimination problem of two states with equal prior, the success probability is given by

$$P_S = \frac{1}{2}\,\mathrm{Tr}[\rho\Pi_0] + \frac{1}{2}\,\mathrm{Tr}[\sigma\Pi_1] = \frac{1}{2}(1 + \mathrm{Tr}[(\rho - \sigma)\Pi_1]).$$

Using the variational formula, the maximal success probability is

$$P_S^* = \frac{1}{2}(1 + \mathrm{Tr}(\rho, \sigma)).$$

For pure states, we recover

$$P_S^* = \frac{1}{2}\left(1 + \sqrt{1 - |\langle\psi|\phi\rangle|^2}\right).$$

The *fidelity* of two states ρ, σ is defined as

$$F(\rho, \sigma) = \left(\mathrm{Tr}\left[\sqrt{\sqrt{\sigma}\rho\sqrt{\sigma}}\right]\right)^2.$$

It follows similar properties to the trace distance,

- $F(\rho,\sigma)=1$ if and only if $\rho=\sigma$.
- $0\le F(\rho,\sigma)\le 1$.
- For pure states $\rho=|\psi\rangle\langle\psi|$ and $\sigma=|\phi\rangle\langle\phi|$, we have

$$F(\rho,\sigma)=|\langle\psi|\phi\rangle|^2.$$

- It is invariant under unitary transformation.
- It is invariant under ancilla.
- It is convex.
- It is increasing under CPTP: $T(\mathcal{C}(\rho),\mathcal{C}(\sigma))\ge T(\rho,\sigma)$.
- **Uhlmann's Theorem** Let $|\psi\rangle,|\phi\rangle$ be purifications of ρ,σ, then

$$F(\rho,\sigma)=\max\{|\langle\psi|\phi\rangle|^2\}.$$

Proof. A standard purification of ρ,σ can be written in the form

$$|\psi\rangle=(\sqrt{\rho}\otimes U)\sum_i|i\rangle\otimes|i\rangle,\quad |\phi\rangle=(\sqrt{\sigma}\otimes V)\sum_i|i\rangle\otimes|i\rangle$$

then

$$|\langle\psi|\phi\rangle|=\sum_{i,j}(\langle i|\otimes\langle i|)(\sqrt{\rho}\sqrt{\sigma}\otimes UV)(|j\rangle\otimes|j\rangle)=\mathrm{Tr}[\sqrt{\rho}\sqrt{\sigma}V^TU^T].$$

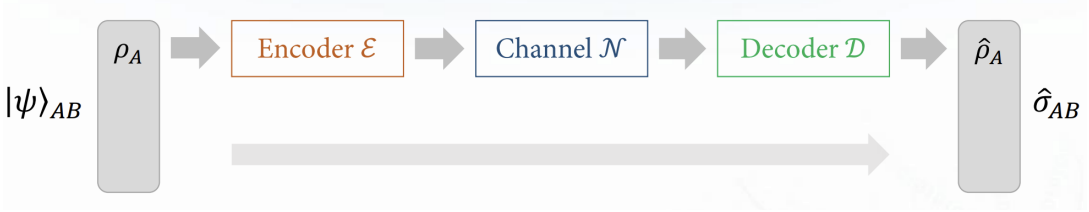
It follows from the variational formula that

$$\mathrm{Tr}[\sqrt{\rho}\sqrt{\sigma}]=\max\{|\langle\psi|\phi\rangle|^2\}.$$

□

- A corollary of Uhlmann's Theorem is the monotonicity of fidelity.

$$F(\rho_{AB},\sigma_{AB})\le F(\rho_A,\sigma_A)$$



A general framework of quantum information processing involves an encoder $\mathcal{E}$, a channel $\mathcal{N}$, and a decoder $\mathcal{D}$. There are two benchmarks for evaluating the fidelity of a system.

- Evaluating a single system:

$$F(\rho_A, \mathcal{D} \circ \mathcal{N} \circ \mathcal{E}(\rho_A))$$

  If the quantum source emits a distribution of quantum states ρ_A^x, then we can calculate the average fidelity as

$$\bar{F} = \sum_{x \in X} p_X(x) F(\rho_A^x, \mathcal{D} \circ \mathcal{N} \circ \mathcal{E}(\rho_A^x)).$$

- Evaluating on a joint system AB,

$$F(\sigma_{AB}, \mathcal{D} \circ \mathcal{N} \circ \mathcal{E}(\sigma_{AB})).$$

Thus, given a channel $\Lambda_{A \to A}$ and a state ρ_A, we define the *channel fidelity* as

$$F_{\text{ch}}(\Lambda_{A \to A}, \rho_A) = \inf\{F\left(\sigma_{AB}, (\Lambda_{A \to A} \otimes \text{Id}_B)\sigma_{AB}\right), \forall \sigma_{AB}, \sigma_A = \rho_A\}.$$

Given a channel $\Lambda_{A \to A}$ and a state ρ_A with any decomposition $\sum_x p_X(x)\rho_A^x = \rho_A$. Then,

$$F_{\text{ch}}(\Lambda_{A \to A}, \rho_A) \leq \bar{F} = \sum_{x \in X} p_X(x) F(\rho_A^x, \Lambda(\rho_A^x)).$$

This means that transfer of a single entangled purification of AB ensures transfer of every state in the ensemble of A.

9 Quantum Information Theory

9.1 Classical Compression

Consider a string of symbols $\{x_i\}$, each with its associated i.i.d random variables X_i. A particular sampling of $x_1 x_2 ... x_n$ occurs with probability

$$p_{X_1 ... X_n}(x_1, ..., x_n) = \prod_{i=1}^{n} p_{X_i}(x_i).$$

If the symbols are binary, we assign a probability p to outcome 1 of the distribution X. We call strings with np 1s and $n(1-p)$ 0s a *typical strings*, the number of distinct typical strings is approximately,

$$\log \binom{n}{np} \approx (n \log n - n) - (np \log(np) - np) - \big(n(1-p)\log(n(1-p)) - n(1-p)\big) = nH(p).$$

where H is the entropy,

$$H(p) = -p \log p - (1-p)\log(1-p).$$

There are of order $2^{nH(p)}$ typical strings, each with probability

$$\log(p^{np}(1-p)^{n(1-p)}) \approx 2^{-nH_n(p)}.$$

Formally, a typical set is defined as

$$\left|\frac{1}{n}\log p_{X_1...X_n}(x_1,...,x_n)+H(X)\right|\leq\delta$$

or

$$2^{-n(H+\delta)}\leq p_{X_1...X_n}(x_1,...,x_n)\leq 2^{-n(H-\delta)}$$

We enforce a lower bound of the probability of typical sequences to be $1-\epsilon$, such that

$$2^{n(H+\delta)}\geq N_{\text{typ}}\geq(1-\epsilon)2^{n(H-\delta)}.$$

Where X is not necessarily a Bernoulli random variable. To convey essentially all the information carried by a string of n bits, it suffices to choose a block code that assigns a nonnegative integer to each of the typical strings. The block code shortens the message for all $p\neq 1/2$. The probability that the actual message is atypical becomes negligible asymptotically due to the law of large numbers.

The number of bits per letter encoding the compressed message is called the *rate* R of the compression code. Note that if we were to compress the messages further into, say, $H(X)-\delta'$ bits. The success probability decreases to

$$2^{n(H-\delta')}2^{-n(H-\delta)}=2^{-n(\delta-\delta')}$$

which decreases exponentially with δ'. In addition, for large n the probability becomes very small. This is Shannon's source coding theorem.

9.2 Quantum Compression

We define a *quantum source* as a probability distribution over an ensemble of pure states with a density matrix ρ. The goal of quantum compression is to compress the system ρ into $\lfloor nR\rfloor$ bits, and then later recover it with high fidelity. Consider a binary source A with states $\{|0\rangle,|+\rangle\}$ with uniform probabilities, i.e.,

$$\rho=\frac{1}{2}|0\rangle\langle 0|+\frac{1}{2}|+\rangle\langle +|.$$

We define the eigenbasis of ρ as $|0\rangle'$ and $|1\rangle'$, where

$$|0'\rangle=\cos\left(\frac{\pi}{8}\right)|0\rangle+\sin\left(\frac{\pi}{8}\right)|1\rangle,\quad |1'\rangle=\sin\left(\frac{\pi}{8}\right)|0\rangle-\cos\left(\frac{\pi}{8}\right)|1\rangle.$$

and

$$\rho=\cos^2\left(\frac{\pi}{8}\right)|0'\rangle\langle 0'|+\sin^2\left(\frac{\pi}{8}\right)|1'\rangle\langle 1'|.$$

Then for any source state $|\psi\rangle=|\psi_1\rangle|\psi_2\rangle...|\psi_n\rangle$, where each qubit is either $|0\rangle$ or $|+\rangle$, we have

$$|\psi\rangle=\sum_{x\in\{0',1'\}^n}\pm\left(\sin\frac{\pi}{8}\right)^k\left(\cos\frac{\pi}{8}\right)^{n-k}|\phi\rangle.$$

Therefore, the projection of any source state $|\psi\rangle$ on a basis vector $|\phi\rangle$ is

$$|\langle\psi|\phi\rangle|^2 = \left(\sin^2\frac{\pi}{8}\right)^k \left(\cos^2\frac{\pi}{8}\right)^{n-k},$$

which is precisely the probability of a classical binary source with k 1s. This means that there exists $\sim 2^{nH(\rho)}$ bases that span a typical subspace with the property that the source states are projected into it with probability close to 1. This suggests a simple encoding scheme that projects An into the typical subspace, thus reliably describing the source using $nH(\rho)$ qubits per block, resulting in a compression rate of $R \sim H(\rho)$. We denote the subspace spanned by these typical eigenstates by Λ, and call it the *typical subspace*. Following that, the compression scheme work as follows:

1. Make a measurement that project the input state into Λ or $\Lambda^\perp$.
2. If the measurement indicates the state was projected into $\Lambda^\perp$, substitute a predefined state $|0'0'...0'\rangle \in \Lambda$.
3. Apply a unitary operator U such that

   $$|\phi\rangle \xrightarrow{U} |\psi\rangle|0_{\text{rem}}.\rangle,$$

   where the state is projected into a smaller subspace with $nH(\rho)$ qubits.
4. The decoder receives the compressed bits and adds the $|0_{\text{rem}}\rangle$ ancilla bits.
5. Apply $U^\dagger$ to retrieve $|\phi\rangle$.

The decoder's output density matrix is then

$$\rho' = \Pi_\Lambda|\psi\rangle\langle\psi|\Pi_\Lambda + \langle\psi|I - \Pi_\Lambda|\psi\rangle|0'...0'\rangle\langle 0'...0'|.$$

And the fidelity is

$$F(\rho,\rho') = |\langle\psi|\Pi_\Lambda|\psi\rangle|^2 + |\langle\psi|I - \Pi_\Lambda|\psi\rangle||\langle\psi|0'...0'\rangle|^2.$$

10 Quantum Error Correction

Typically, a practical quantum algorithm using n qubits and M elementary steps has $nM \gg 10^10$. However, the best known fidelity among physical implementations is $\sim 0.1\%$ error for single-qubit gates and $\sim 5\%$ for two-qubit gates, which means that the logical error for each logical operation should be much less than 10^{-10}. Classical error-correcting codes use repetition encoding against bit flip errors. Consider a bit-flip channel with probability p, we can clone the bit twice to produce three bits, e.g., $0 \to 000$. We can effectively reduce the decoding error,

$$3p^2(1-p) + p^3 = 3p^2 - 2p^3 < p, \quad \text{for } p \in (0, 1/2).$$

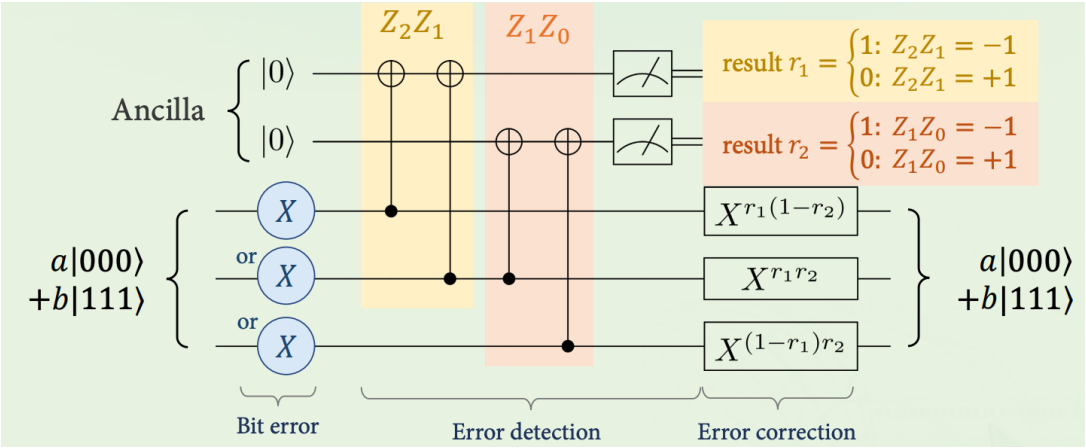
On the other hand, a quantum bit-flip channel has the form

$$\mathcal{N}(\rho) = (1-p)\rho + pX\rho X.$$

Unfortunately, repetition codes are not possible for quantum states due to the no-cloning theorem. Instead, we encode a qubit with the following,

$$|\psi\rangle = a|0\rangle + b|1\rangle \;\rightarrow\; a|000\rangle + b|111\rangle$$

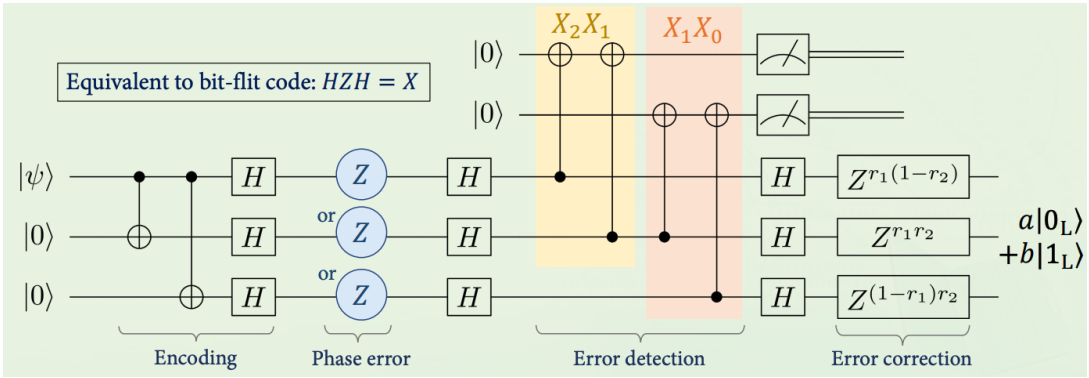
We can implement this with two CNOT gates on the second and third qubits. To detect a bit flip, we measure with operators Z_2Z_1 and Z_1Z_0, such that if the two qubits are not the same, we will measure -1 since Z has eigenvalues ± 1 for $|0\rangle, |1\rangle$ respectively. We implement the error correction circuit for bit flips as the following circuit.



To protect against phase flips, we use the following encoding,

$$\begin{cases} |0\rangle \;\rightarrow\; \frac{1}{2\sqrt{2}}(|0\rangle + |1\rangle)(|0\rangle + |1\rangle)(|0\rangle + |1\rangle) \\ |1\rangle \;\rightarrow\; \frac{1}{2\sqrt{2}}(|0\rangle - |1\rangle)(|0\rangle - |1\rangle)(|0\rangle - |1\rangle) \end{cases}$$

We measure the qubits with X_2X_1 and X_1X_0 to detect phase flips. The circuit diagram is as follows,



Note that this is similar to the bit-flip correction codes since $HZH = X$.

10.1 Shor's [9,1,3] Code

The [9,1,3] in Shor's code indicates 9 qubits for a 1 bit of information with code distance 3. Consider the concatenation of the previous two error-correction codes,

$$\begin{cases} |0\rangle & \to & \frac{1}{2\sqrt{2}}(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle) \\ |1\rangle & \to & \frac{1}{2\sqrt{2}}(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle) \end{cases}$$

Note that Y-errors are taken care of since $Y = -iZX$. Shor's [9,1,3] code requires 8 measurement operators,

$$Z_8Z_7, Z_7Z_6, Z_5Z_4, Z_4Z_3, Z_2Z_1, Z_1Z_0, X_8X_7X_6X_5X_4X_3, X_5X_4X_3X_2X_1X_0.$$

The X-errors only need to be measured within each block, so there are 6 total operators. We can then correct the incorrect qubit i with X_i. If Z-error is detected, we can correct it by applying Z to any qubit in this block. There are 256 possible measurement results from the 8 measurements. However, there are at most $1 + 3 \times 9 = 28$ types of errors, so Shor's [9,1,3] code is not optimal. There exists a [5,1,3] code.

References

[1] Hao-Chung Cheng. Quantum information and computation, Spring 2025.

[2] John Preskill. Ph 219: Quantum computation. `https://preskill.caltech.edu/ph219/`.