Indian Institute of Technology Delhi

B.Tech Project Report

Quantum Coding Theory

Professor: Ritumoni Sharma

Students:
Akshay Reddy Vellapalem (2020MT10868)
Dayal Kumar (2020MT10797)

Submitted in partial fulfillment of the requirements for the degree of B.Tech in Mathematics and Computing

Department of Mathematics

Indian Institute of Technology Delhi

May 3, 2024

Contents

Chapter 1

Introduction to Quantum Mechanics

1.1 Postulates

1.1.1 Postulate 1: State Space

Postulate 1. *Associated to any isolated physical system is a complex vector space with inner product (that is, a Hilbert space) known as the state space of the system. The system is completely described by its state vector, which is a unit vector in the system's state space.*

Qubit: The simplest quantum mechanical system we will be working with is the *qubit* which has a two-dimensional state space with basis states $|0\rangle$ and $|1\rangle$. An arbitrary state vector $|\psi\rangle$ will be $a|0\rangle + b|1\rangle$ where $a, b \in \mathbb{C}$ and $|a|^2 + |b|^2 = 1$.

1.1.2 Postulate 2: Evolution

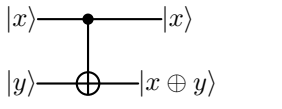
Postulate 2. *The evolution of a closed quantum system is described by a unitary transformation. That is, the state $|\psi\rangle$ of the system at time t_1 is related to the state $|\psi'\rangle$ of the system at time t_2 by a Unitary operator U which only depends on the times t_1 and t_2.*

$$|\psi'\rangle = U|\psi\rangle$$

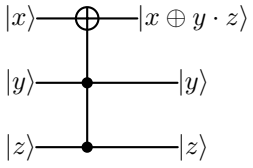
$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, $Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$, $Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ and $H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$ are some of the unitary operators describing the evolution of a single qubit, that we will be using in the further sections.

Quantum Gates: A quantum gate over n bits is defined as a unitary operation $U : \mathbb{C}^{2^n} \longrightarrow \mathbb{C}^{2^n}$ such that $U^\dagger U = I$. So, a quantum gate has to be reversible, in particular operations like AND, OR, XOR etc. cannot be implemented directly. These can be implemented by creating a unitary transformation whose one of the outputs is AND (or OR, XOR, etc.) and the other outputs help to make it reversible. Some typical quantum gates are:

$(x, y) \longrightarrow (x, x \oplus y)$:



$(x, y, z) \longrightarrow (x \oplus y \cdot z, y, z)$:



1.1.3 Postulate 3: Measurement

Postulate 3. *Quantum measurements are described by a collection $\{M_m\}$ of measurement operators. These operators act on the state space of the system being measured. The index m refers to the different measurement outcomes that may occur in the experiment. If the state of the system immediately before the measurement is $|\psi\rangle$, then the probability that result m occurs is given by*

$$p(m) = \langle\psi| M_m^\dagger M_m |\psi\rangle ,$$

and the state of the system after the measurement is

$$\frac{M_m |\psi\rangle}{\langle\psi| M_m^\dagger M_m |\psi\rangle}.$$

The measurement operators satisfy the completeness condition

$$\sum_m M_m^\dagger M_m = I.$$

The completeness equation is equivalent to the condition that the probabilities of the outcomes sum to one for every initial state.

$$\forall\psi, 1 = \sum_m p(m) = \sum_m \langle\psi| M_m^\dagger M_m |\psi\rangle \iff \sum_m M_m^\dagger M_m = I$$

Positive Operator-Valued Measure (POVM): Corresponding to a collection $\{M_m\}$ of measurement operators, the set $\{E_m | E_m = M_m^\dagger M_m\}$ is known as a *POVM* which is a set of positive operators (can be proven easily) such that $\sum_m E_m = I$ and the probability of occurence of outcome m is $p(m) = \langle\psi| E_m |\psi\rangle$. As an example of a POVM, consider a **projective measurement** described by measurement operators P_m where the P_m are projectors such that $P_m P_{m'} = \delta_{mm'} P_m$. In this instance, the POVM elements coincide with the measurement operators since $E_m = P_m^\dagger P_m = P_m$.

1.1.4 Postulate 4: Composite Systems

Postulate 4. *The state space of a composite physical system is the tensor product of the state spaces of the component physical systems. Moreover, if we have systems numbered 1 through n, and system number i is prepared in the state $|\psi_i\rangle$, then the joint state of the total system is $|\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_n\rangle$.*

Here, composite system means a system made up of more than one isolated physical systems.

1.2 Density Operator

The *density operator* gives another formulation to quantum mechanics. Suppose a quantum system is in one of a number of states $|\psi_i\rangle$ with respective probabilities p_i. $\{(p_i, |\psi_i\rangle)\}$ is called an ensemble of pure states. The density operator for the ensemble is defined as:

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

The density operator language provides a convenient means for describing quantum systems whose state is not completely known. A quantum system whose state $|\psi\rangle$ is known exactly is said to be in a pure state. In this case the density operator is simply $\rho_m = |\psi\rangle\langle\psi|$. Otherwise, ρ is in a mixed state; it is said to be a mixture of the different pure states in the ensemble for ρ. A pure state satisfies $tr(\rho^2) = 1$, while a mixed state satisfies $tr(\rho^2) < 1$.

The density operator of a quantum system present in state ρ_i (mixed or pure) with probability p_i is $\sum_i p_i \rho_i$. A proof of this is to suppose that ρ_i arises from some ensemble $\{\rho_{ij}, |\psi_{ij}\rangle\}$ (note that i is fixed) of pure states, so the probability for being in the state $|\psi_{ij}\rangle$ is $p_i p_{ij}$. The density matrix for the system is thus

$$\rho = \sum_{ij} p_i p_{ij} |\psi_{ij}\rangle \langle\psi_{ij}| = \sum_i p_i \rho_i.$$

1.2.1 Postulates

1.2.1.1 Postulate 1

Theorem 1.2.1. *An operator ρ is the density operator associated to some ensemble $\{(p_i, |\psi_i\rangle)\}$ if and only if it satisfies the conditions:*

1. ***(Trace condition)*** *ρ has trace equal to one.*

2. ***(Positivity condition)*** *ρ is a positive operator.*

Proof. Suppose $\rho = \sum_i p_i |\psi_i\rangle \langle\psi_i|$ is a density operator. Then

$$tr(\rho) = \sum_i p_i tr(|\psi_i\rangle \langle\psi_i|) = \sum_i p_i = 1,$$

so the trace condition is satisfied. Suppose $|\psi\rangle$ is an arbitrary operator in the state space. Then,

$$\begin{aligned} \langle\psi| \rho |\psi\rangle &= \sum_i p_i \langle\psi| |\psi_i\rangle \langle\psi_i| |\psi\rangle \\ &= \sum_i p_i |\langle\psi| |\psi_i\rangle|^2 \\ &\geq 0, \end{aligned}$$

so the positivity condition is satisfied.

Conversely, suppose ρ is an operator satisfying the trace and positivity conditions. Since ρ is positive, it must have a spectral decomposition

$$\rho = \sum_j \lambda_j \langle j| |j\rangle,$$

where the eigenvectors $|j\rangle$ are orthonormal, and λ_j are real and non-negative eigenvalues of ρ. From the trace condition we see that $\sum_j \lambda_j = 1$. Since all unit vectors in the state space are possible states of a physical system, a system in state $|j\rangle$ with probability λ_j has density operator ρ. □

Thus, the first postulate is reformulated as,

Postulate 1. *Associated to an isolated system is a complex vector space with inner product (i.e. a Hilbert Space) known as the state space of the system. The system is completely determined by its density operator, which is a positive operator ρ with trace one, acting on the state space of the system.*

1.2.1.2 Postulate 2

If a system was initially in an ensemble $\{(p_i, |\psi_i\rangle)\}$ of states, then the probability that the system will be in the state $|\psi\rangle$ is $\sum_i p_i 1\{|\psi\rangle = U|\psi_i\rangle\}$. Thus the final *density operator* is

$$\sum_\psi \sum_i p_i 1\{|\psi\rangle = U|\psi_i\rangle\} |\psi\rangle\langle\psi| = \sum_i p_i \sum_\psi \{|\psi\rangle = U|\psi_i\rangle\} |\psi\rangle\langle\psi| = \sum_i p_i U|\psi_i\rangle\langle\psi_i| U^\dagger = U\rho U^\dagger.$$

Thus the second postulate is reformulated as,

Postulate 2. *The evolution of a closed quantum system is described by a unitary transformation. That is, the state ρ of the system at time t_1 is related to the state at time t_2 by a unitary operator U which depends only on the times t_1 and t_2,*

$$\rho' = U\rho U^\dagger.$$

1.2.1.3 Postulate 3

Suppose we perform a measurement described by the measurement operators $\{M_m\}$. If the initial state was $|\psi\rangle$, then the probability of getting the result m is

$$p(m|i) = \langle\psi_i| M_m^\dagger M_m |\psi_i\rangle = tr(M_m^\dagger M_m |\psi_i\rangle \langle\psi|),$$

where we have used the result proved in Appendex A.1. By the law of total probability, the probability of obtaining result m is

$$\begin{aligned} p(m) &= \sum_i p(m|i)p_i \\ &= \sum_i p_i tr(M_m^\dagger M_m |\psi\rangle \langle\psi|) \\ &= tr(M_m^\dagger M_m \rho). \end{aligned}$$

If the initial state was $|\psi_i\rangle$, then the state after obtaining the result m is

$$|\psi_m\rangle = \frac{M_m |\psi_i\rangle}{\sqrt{\langle\psi_i| M_m^\dagger M_m |\psi_i\rangle}}.$$

Thus after a measurement that yields the result m, we have an ensemble of states $|\psi_i^m\rangle$ with respective probabilities $p(i|m)$, as the initial probability is complemented with the information that result m has been obtained. The corresponding density operator is therefore

$$\begin{aligned} \rho_m &= \sum_i p(i|m) |\psi_i^m\rangle \langle\psi_i^m| \\ &= \sum_i \frac{p(m|i)p_i}{p(m)} \frac{M_m |\psi_i\rangle \langle\psi_i| M_m^\dagger}{\langle\psi_i| M_m^\dagger M_m |\psi_i\rangle} \\ &= \sum_i p_i \frac{M_m |\psi_i\rangle \langle\psi_i| M_m^\dagger}{tr(M_m^\dagger M_m \rho)} \\ &= \frac{M_m \rho M_m^\dagger}{tr(M_m^\dagger M_m \rho)}. \end{aligned}$$

Postulate 3. *Quantum measurements are described by a collection $\{M_m\}$ of measurement operators. These operators act on the state space of the system being measured. The index m refers to the different measurement outcomes that may occur in the experiment. If the state of the system immediately before the measurement is ρ, then the probability that result m occurs is given by*

$$p(m) = tr(M_m^\dagger M_m \rho),$$

and the state of the system after the measurement is

$$\frac{M_m \rho M_m^\dagger}{tr(M_m^\dagger M_m \rho)}.$$

The measurement operators satisfy the completeness condition

$$\sum_m M_m^\dagger M_m = I.$$

Imagine that, for some reason, our record of the result m of the measurement was lost. We would have a quantum system in the state ρ_m with probability $p(m)$, but would no longer know the actual value of m. The state of such a quantum system would therefore be described by the density operator

$$\begin{aligned} \rho &= \sum_m p(m)\rho_m \\ &= \sum_m tr(M_m^\dagger M_m \rho) \frac{M_m \rho M_m^\dagger}{tr(M_m^\dagger M_m \rho)} \\ &= \sum_m M_m \rho M_m^\dagger, \end{aligned}$$

which can be used as the starting point for analysis of further operations on the system.

1.2.1.4 Postulate 4

Postulate 4. *The state space of a composite physical system is the tensor product of the state spaces of the component physical systems. Moreover, if we have systems numbered 1 through n, and system number i is prepared in the state ρ_i, then the joint state of the total system is $\rho_1 \otimes \rho_2 \otimes \cdots \rho_n$.*

1.2.2 Unitary Freedom

We want to understand what class of ensembles give rise to the same density matrix. For an ensemble $\{(p_i, |\psi_i\rangle)\}$, we define the set of vectors $|\tilde{\psi_i}\rangle = \sqrt{p_i}\,|\psi_i\rangle$. Thus, the set $|\tilde{\psi_i}\rangle$ generates the operator $\rho = \sum_i |\tilde{\psi_i}\rangle\langle\tilde{\psi_i}|$.

Theorem 1.2.2. *The sets $|\tilde{\psi_i}\rangle$ and $|\tilde{\phi_j}\rangle$ generate the same density matrix if and only if*

$$|\tilde{\psi_i}\rangle = \sum_j u_{ij}\,|\tilde{\phi_j}\rangle\,,$$

where u_{ij} is a unitary matrix of complex numbers, with indices i and j and we 'pad' whichever set of vectors $|\tilde{\psi_i}\rangle$ or $|\tilde{\phi_j}\rangle$ is smaller with additional zero vectors so that the two sets have the same number of elements.

Proof. Suppose $|\tilde{\psi_i}\rangle = \sum_i u_{ij}\,|\tilde{\phi_j}\rangle$ for some unitary u_{ij}. Then,

$$\begin{aligned}
\sum_i |\tilde{\psi_i}\rangle\langle\tilde{\psi_i}| &= \sum_i \left(\sum_j u_{ij}\,|\tilde{\phi_j}\rangle\right)\left(\sum_k u_{ik}\,|\tilde{\phi_k}\rangle\right)^\dagger \\
&= \sum_{jk}\left(\sum_i u_{ik}^\dagger u_{ij}\right)|\tilde{\phi_j}\rangle\langle\tilde{\phi_k}| \\
&= \sum_{jk}\delta_{jk}\,|\tilde{\phi_j}\rangle\langle\tilde{\phi_k}| \\
&= \sum_j |\tilde{\phi_j}\rangle\langle\tilde{\phi_j}|\,,
\end{aligned}$$

i.e. $|\tilde{\psi_i}\rangle$ and $|\tilde{\phi_j}\rangle$ generate the same density operator. Conversely, suppose

$$A = \sum_i |\tilde{\psi_i}\rangle\langle\tilde{\psi_i}| = \sum_j |\tilde{\phi_j}\rangle\langle\tilde{\phi_j}|\,.$$

Let $A = \sum_k \lambda_k\,|k\rangle\langle k|$ be a decomposition for A such that $|k\rangle$ are orthonormal and the λ_k are strictly positive, as A is a density operator. Then define $|\tilde{k}\rangle = \sqrt{\lambda_k}\,|k\rangle$ and let $|\psi\rangle$ be any vector orthogonal to the space spanned by $|\tilde{k}\rangle$, so $\langle\psi|\tilde{k}\rangle\langle\tilde{k}|\psi\rangle = 0$ for all k, and thus

$$0 = \langle\psi|\,A\,|\psi\rangle = \sum_i \langle\psi|\tilde{\psi_i}\rangle\langle\tilde{\psi_i}|\psi\rangle = \sum_i |\,\langle\psi|\tilde{\psi_i}\rangle\,|^2.$$

Thus, $\langle\psi|\tilde{\psi_i}\rangle = 0$ for all i and all $|\psi\rangle$ orthogonal to the space spanned by $|\tilde{k}\rangle$. It follows that each $|\tilde{\psi_i}\rangle$ belongs to the span of $|\tilde{k}\rangle$, $|\tilde{\psi_i}\rangle = \sum_k c_{ik}\,|\tilde{k}\rangle$. Since, $A = \sum_k |\tilde{k}\rangle\langle\tilde{k}| = \sum_i |\tilde{\psi_i}\rangle\langle\tilde{\psi_i}|$, we see that

$$\sum_k |\tilde{k}\rangle\langle\tilde{k}| = \sum_{kl}\left(\sum_i c_{ik}c_{il}^*\right)|\tilde{k}\rangle\langle\tilde{l}|\,.$$

The $|\tilde{k}\rangle\langle\tilde{l}|$ are linearly independent as suppose $\sum_{kl}\alpha_{kl}\,|\tilde{k}\rangle\langle\tilde{l}| = 0$ while $\alpha_{k_1l_1} \neq 0$ (say), then applying to the state $|\tilde{l_1}\rangle$ we get $\sum_k \alpha_{kl_1}\,|\tilde{k}\rangle = 0$, which is a contradiction as $|\tilde{k}\rangle$ are linearly independent. So, it must be that $c_{ik}c_{il}^* = \delta_{kl}$. That is, the columns of c are orthonormal, and by Gram-Schmidt we can extend the columns of c to obtain a unitary matrix v such that $|\tilde{\psi_i}\rangle = \sum_k v_{ik}\,|\tilde{k}\rangle$ where we have appended zero vectors to the list of $|\tilde{k}\rangle$. Similarly, we can find a unitary matrix w such that $|\tilde{\phi_j}\rangle = \sum_k w_{jk}\,|\tilde{k}\rangle$. Then v and w have the same dimensions as $|\tilde{\psi_i}\rangle$ or $|\tilde{\phi_j}\rangle$ were padded with zero vectors to make them equal in number. Thus, $|\tilde{\psi_i}\rangle = \sum_j u_{ij}\,|\tilde{\phi_j}\rangle$, where $u = vw^\dagger$ is unitary. □

1.2.3 Reduced Density Operator

Suppose we have physical systems A and B, whose state is described by a density operator ρ^{AB}. The reduced density operator for system A is defined by

$$\rho^A \equiv tr_B(\rho^{AB}),$$

where tr_B is called the partial trace over system B. The partial trace is a linear operator from the space of density operators on the composite system AB to the space of density operators on A, defined by

$$tr_B(|a_1\rangle\langle a_2| \otimes |b_1\rangle\langle b_2|) = |a_1\rangle\langle a_2| \, tr(|b_1\rangle\langle b_2|) = |a_1\rangle\langle a_2|\langle b_2|b_1\rangle,$$

where $|a_1\rangle$ and $|a_2\rangle$ are any two vectors in the state space of A, and $|b_1\rangle$ and $|b_2\rangle$ are any two vectors in the state space of B.

Theorem 1.2.3. *The partial trace operation is the unique operation which gives rise to the correct description of observable quantities for subsystems of a composite system.*

Proof. Suppose M is any observable on system A, and we have some measuring device which is capable of realizing measurements of M. Let $\tilde{M}$ denote the corresponding observable for the same measurement, performed on the composite system AB. Note that if the system AB is prepared in the state $|m\rangle|\psi\rangle$, where $|m\rangle$ is an eigenstate of M with eigenvalue m, and $|\psi\rangle$ is any state of B, then the measuring device must yield the result mfor the measurement, with probability one. Thus, if P_m is the projector onto the m eigenspace of the obeservable M, then the corresponding projector for $\tilde{M}$ is $P_m \otimes I_B$. Therefore, we have

$$\tilde{M} = \sum_m mP_m \otimes I_B = M \otimes I_B.$$

Suppose we perform a measurement on system A described by the observable M. Then, the following equation is satisfied if ρ^A the state of the system A is defined as the reduced density operator $tr_B(\rho^{AB})$.

$$tr(M\rho^A) = tr(\tilde{M}\rho^{AB}) = tr((M \otimes I_B)\rho^{AB})$$

To see the uniqueness property, let $f(\cdot)$ be any map of density operators on AB to density operators on B such that

$$tr(Mf(\rho^{AB}) = tr((M \otimes I_B)\rho^{AB}),$$

for all observables M. Let M_i be an orthonormal basis of operators for the space of Hermitian operators with respect to the Hilbert–Schmidt inner product $(X, Y) \equiv tr(XY)$. Then expanding $f(\rho^{AB})$ in this basis gives

$$f(\rho^{AB}) = \sum_i M_i tr(M_i f(\rho^{AB})) = \sum_i M_i tr((M_i \otimes I_B)\rho^{AB}).$$

It follows that f is uniquely determined. □

The following simple calculations may help us to understand the reduced density operator. First, suppose a quantum system is in the product state $\rho^{AB} = \rho \otimes \sigma$, where ρ is a density operator for system A and σ is the density operator for system B. Then

$$\rho^A = tr_B(\rho \otimes \sigma) = \rho tr(\sigma) = \rho,$$

which is consistent with the pure state of system A we started with.

A less trivial example is the Bell state $\frac{|00\rangle+|11\rangle}{\sqrt{2}}$. This has density operator

$$\begin{aligned}\rho &= \frac{|00\rangle + |11\rangle}{\sqrt{2}} \frac{\langle 00| + \langle 11|}{\sqrt{2}} \\ &= \frac{|00\rangle\langle 00| + |11\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 11|}{2}\end{aligned}$$

Tracing out the second qubit, we find the reduced density operator of the first qubit,

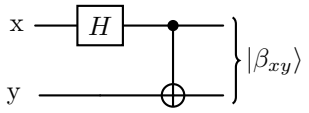
$$\begin{aligned}\rho^1 &= tr_2(\rho) \\ &= \frac{tr_2(ket00\,\langle 00| + |11\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 11|)}{2} \\ &= \frac{|0\rangle\langle 0|\langle 0|0\rangle + |1\rangle\langle 0|\langle 0|1\rangle + |0\rangle\langle 1|\langle 1|0\rangle + |1\rangle\langle 1|\langle 1|1\rangle}{2} \\ &= \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} \\ &= \frac{I}{2}.\end{aligned}$$

Notice that this state is a mixed state, since $tr(\frac{I}{2}^2) = \frac{1}{2} < 1$, while the joint system of the two qubits is in a pure state, that is, it is known exactly. This is a remarkable property exhibited by entangled states which we will discuss about in the next section.

1.3 Quantum Algorithms and Circuits

1.3.1 Quantum Entanglement

Input	Output
$\|00\rangle$	$(\|00\rangle + \|11\rangle)/\sqrt{2} = \beta_{00}$
$\|01\rangle$	$(\|01\rangle + \|10\rangle)/\sqrt{2} = \beta_{01}$
$\|10\rangle$	$(\|00\rangle - \|11\rangle)/\sqrt{2} = \beta_{10}$
$\|11\rangle$	$(\|01\rangle - \|10\rangle)/\sqrt{2} = \beta_{11}$

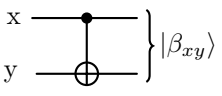


Consider the above circuit, which has a Hadamard gate followed by a CNOT gate. The adjacent table shows the output of the circuit on the four computational basis states. First, the Hadamard transform puts the top qubit in a superposition; this then acts as a control input to the CNOT, and the target gets inverted only when the control is 1. The output states are known as *Bell states* or *EPR pairs*.

The thing to note about these states is that these cannot be decomposed as $|a\rangle \otimes |b\rangle$. We say that a state of a composite system having this property (that it can't be written as a product of states of its component systems) is an *entangled state*. Entangled states play a crucial role in quantum computation and quantum information, as will be explored in following sections.

Often, we need to disentangle the qubits, i.e. change the state of the system such that a subset of the system becomes independent of the measurement or operations on the rest of the system while maintaining the measurement statistics of the concerned subset. The following circuit is used to disentangle any of the four states created above.

Input	Output
$(\|00\rangle + \|11\rangle)/\sqrt{2} = \beta_{00}$	$\frac{1}{\sqrt{2}}(\|0\rangle + \|1\rangle)\|0\rangle$
$(\|01\rangle + \|10\rangle)/\sqrt{2} = \beta_{01}$	$\frac{1}{\sqrt{2}}(\|0\rangle + \|1\rangle)\|1\rangle$
$(\|00\rangle - \|11\rangle)/\sqrt{2} = \beta_{10}$	$\frac{1}{\sqrt{2}}(\|0\rangle - \|1\rangle)\|0\rangle$
$(\|01\rangle - \|10\rangle)/\sqrt{2} = \beta_{11}$	$\frac{1}{\sqrt{2}}(\|0\rangle - \|1\rangle)\|1\rangle$

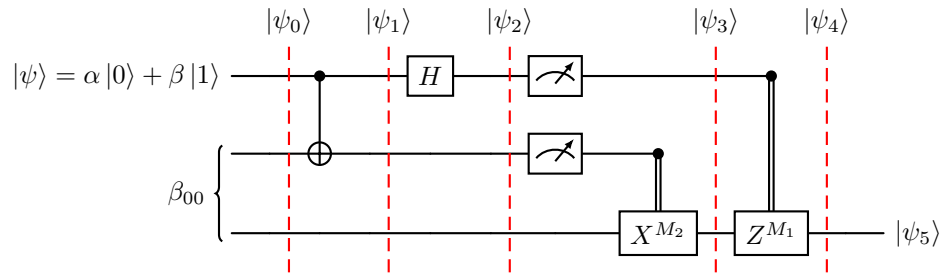


1.3.2 Quantum Teleportation

Quantum teleportation is a technique for moving quantum states around, even in the absence of a quantum communications channel linking the sender of the quantum state to the recipient.

Here's how quantum teleportation works. Alice and Bob met long ago but now live far apart. While together they generated an EPR pair, each taking one qubit of the EPR pair when they separated. Many years later, Bob is in hiding, and Alice's mission, should she choose to accept it, is to deliver a qubit $|\psi\rangle$ to Bob. She does not know the state of the qubit, and moreover can only send *classical* information to Bob.

Alice doesn't know the state $|\psi\rangle$ of the qubit she has to send to Bob, and the laws of quantum mechanics prevent her from determining the state when she only has a single copy of $|\psi\rangle$ in her possession. What's worse, even if she did know the state $|\psi\rangle$, describing it precisely takes an infinite amount of classical information since $|\psi\rangle$ takes values in a continuous space. So even if she did know $|\psi\rangle$, it would take forever for Alice to describe the state to Bob.



$$
\begin{aligned}
|\psi_0\rangle &= (\alpha|0\rangle + \beta|1\rangle)\left(\frac{|00\rangle + |11\rangle}{\sqrt{2}}\right) \\
&= \frac{\alpha|000\rangle + \alpha|011\rangle + \beta|100\rangle + \beta|111\rangle}{\sqrt{2}} \\
|\psi_1\rangle &= \frac{\alpha|000\rangle + \alpha|011\rangle + \beta|110\rangle + \beta|101\rangle}{\sqrt{2}} \\
|\psi_2\rangle &= \frac{\alpha|000\rangle + \alpha|100\rangle + \alpha|011\rangle + \alpha|111\rangle + \beta|010\rangle - \beta|110\rangle + \beta|001\rangle - \beta|101\rangle}{2} \\
|\psi_3\rangle &= \begin{cases} \frac{\alpha|000\rangle+\alpha|100\rangle+\beta|001\rangle-\beta|101\rangle}{\sqrt{2}}, & M_2 = 0 \\ \frac{\alpha|010\rangle+\alpha|110\rangle+\beta|011\rangle-\beta|111\rangle}{\sqrt{2}} & M_2 = 1 \end{cases} \\
|\psi_4\rangle &= \begin{cases} \alpha|000\rangle + \beta|001\rangle & M_1M_2 = 00 \\ \alpha|010\rangle + \beta|011\rangle & M_1M_2 = 01 \\ \alpha|100\rangle + \beta|101\rangle & M_1M_2 = 10 \\ \alpha|110\rangle + \beta|111\rangle & M_1M_2 = 11 \end{cases} \\
|\psi_5\rangle &= \alpha|0\rangle + \beta|1\rangle = |\psi\rangle
\end{aligned}
$$

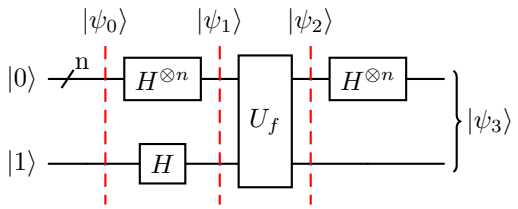
In outline, the steps of the solution are as follows: Alice interacts the qubit $|\psi\rangle$ with her half of the EPR pair, and then measures the two qubits in her possession, obtaining one of four possible classical results, 00, 01, 10, and 11. She sends this information to Bob. Depending on Alice's classical message, Bob performs one of four operations on his half of the EPR pair. As we see by the above calculations, by doing this he can recover the original state $|\psi\rangle$.

1.3.3 Deutsch–Jozsa algorithm

Deutsch's problem is described as follows: Bob has a function $f : \{0, 1, \ldots, 2^n - 1\} \to \{0, 1\}$. which is of one of two kinds; either $f(x)$ is constant for all values of x, or else $f(x)$ is balanced, that is, equal to 1 for exactly half of all the possible x, and 0 for the other half. Alice's goal is to determine with certainty whether Bob has chosen a constant or a balanced function, corresponding with him as little as possible.

In the classical case, Alice may only send Bob one value of x in each letter. At worst, Alice will need to query Bob at least $2^{n-1} + 1$ times, since she may receive $2^n/2$ 0s before finally getting a 1, telling her that Bob's function is balanced. The best deterministic classical algorithm she can use therefore requires $2^{n-1} + 1$ queries. Note that in each letter, Alice sends Bob n bits of information.

If Bob and Alice were able to exchange qubits, instead of just classical bits, and if Bob agreed to calculate $f(x)$ using a unitary transformation: $U_f : |x, y\rangle \to |x, y \oplus f(x)\rangle$, then Alice could achieve her goal in just one correspondence with Bob, using the following quantum circuit.



The first n wires here represent the query register while the last wire represents the answer register. The input state

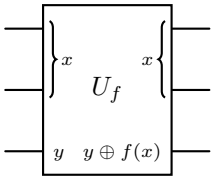
$$|\psi_0\rangle = |0\rangle^{\otimes n} |1\rangle .$$

After the Hadamard transform we have

$$|\psi_1\rangle = \prod_{i=1}^{n} \left(\frac{|0\rangle + |1\rangle}{\sqrt{2}} \right)^{\otimes n} \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right) = \sum_{x \in \{0,1\}^n} \frac{|x\rangle}{\sqrt{2^n}} \left[\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right].$$

Next, the function f is evaluated (by Bob) using $U_f : |x, y\rangle \longrightarrow |x, x \oplus f(x)\rangle$

$$\begin{aligned} |\psi_2\rangle &= \sum_x \frac{|x\rangle}{\sqrt{2^n}} \left[\frac{|0 \oplus f(x)\rangle - |1 \oplus f(x)\rangle}{\sqrt{2}} \right] \\ &= \sum_x \frac{(-1)^{f(x)} |x\rangle}{\sqrt{2^n}} \left[\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right] \end{aligned}$$



By checking the cases $x = 0$ and $x = 1$ separately we see that for a single qubit $H|x\rangle = \sum_x (-1)^{xz} |z\rangle / \sqrt{2}$. Thus,

$$H^{\otimes n} |x_1, x_2, \cdots, x_n\rangle = \frac{\sum_{z_1, z_2, \cdots, z_n} (-1)^{x_1 z_1 + x_2 z_2 + \cdots + x_n z_n} |z_1, z_2, \cdots z_n\rangle}{\sqrt{2}},$$

which can be written succinctly as

$$H^{\otimes n} |x\rangle = \frac{\sum_z (-1)^{x \cdot z} |z\rangle}{\sqrt{2^n}},$$

where $x \cdot z$ is the bitwise inner product of x and z, modulo 2. Thus,

$$|\psi_3\rangle = \sum_z \sum_x \frac{(-1)^{x \cdot z + f(x)} |z\rangle}{\sqrt{2^n}} \left[\frac{|0\rangle - |1\rangle}{\sqrt{2}} \right].$$

Observing the query registers, the amplitude of the state $|0\rangle^{\otimes n}$ is $\sum_x (-1)^{f(x)} / 2^n$. In the case where f is a constant, the amplitude for $|0\rangle^{\otimes n}$ is $+1$ or -1, depending on the constant value $f(x)$ takes. Because $|\psi_3\rangle$ is of unit length it follows that all other amplitudes must be zero and an observation will yield $0s$ for all qubits in the query register. If f is balanced then the positive and negative contributions to the amplitude for $|0\rangle^{\otimes n}$ cancel, leaving an amplitude of zero, and a measurement must yield a result other than zero on at least one qubit in the query register. Summarizing, if Alice measures all $0s$ then the function is constant; otherwise the function is balanced.

Chapter 2

Quantum Error Correction

Suppose we wish to send a bit from one location to another through a noisy classical communications channel. The effect of the noise in the channel is to flip the bit being transmitted with probability $p > 0$, while with probability $1 - p$ the bit is transmitted without error. Such a channel is known as a *binary symmetric channel*.

A simple means of protecting the bit against the effects of noise in the binary symmetric channel is to replace the bit we wish to protect with three copies of itself:

$$0 \longrightarrow 000$$
$$1 \longrightarrow 111.$$

The bit strings 000 and 111 are sometimes referred to as the logical 0 and logical 1, since they play the role of 0 and 1, respectively. We now send all three bits through the channel. At the receiver's end of the channel three bits are output, and the receiver has to decide what the value of the original bit was. Suppose 001 were output from the channel. Provided the probability p of a bit flip is not too high, it is very likely that the third bit was flipped by the channel, and that 0 was the bit that was sent.

This type of decoding is called majority voting, since the decoded output from the channel is whatever value, 0 or 1, appears more times in the actual channel output. Majority voting fails if two or more of the bits sent through the channel were flipped, and succeeds otherwise. The probability that two or more of the bits are flipped is $3p^2(1-p)+p^3$, so the probability of error is $p_e = 3p^2 - 2p^3$. Without encoding, the probability of an error was p, so the code makes the transmission more reliable provided $p_e < p$, which occurs whenever $p < 1/2$.

To protect quantum states against the effects of noise we would like to develop quantum error-correcting codes based upon similar principles. There are some important differences between classical information and quantum information that require new ideas to be introduced to make such quantum error-correcting codes possible. In particular, at a first glance we have three rather formidable difficulties to deal with:

- *No cloning:* One might try to implement the repetition code quantum mechanically by duplicating the quantum state three or more times. This is forbidden by the no-cloning theorem discussed in Appendix A.3. Even if cloning were possible, it would not be possible to measure and compare the three quantum states output from the channel.

- *Errors are continuous:* A continuum of different errors may occur on a single qubit. Determining which error occurred in order to correct it would appear to require infinite precision, and therefore infinite resources.

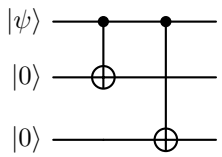
- *Measurement destroys quantum information:* In classical error-correction we observe the output from the channel, and decide what decoding procedure to adopt. Observation in quantum mechanics generally destroys the quantum state under observation, and makes recovery impossible.

2.1 Three Qubit Flip Code

Suppose we send qubits through a channel which leaves the qubits untouched with probability $1-p$, and flips the qubits with probability p. That is, with probability p the state $|\psi\rangle$ is taken to the state $X|\psi\rangle$, where X is the usual Pauli sigma x operator, or bit flip operator. This channel is called the bit flip channel, and we now explain the bit flip code, which may be used to protect qubits against the effects of noise from this channel.
Suppose we encode the single qubit state $a|0\rangle + b|1\rangle$ in three qubits as $a|000\rangle + b|111\rangle$. A convenient way to write this encoding is

$$|0\rangle \longrightarrow |0_L\rangle = |000\rangle$$
$$|1\rangle \longrightarrow |1_L\rangle = |111\rangle ,$$

where it is understood that superpositions of basis states are taken to corresponding superpositions of encoded states. A circuit performing this encoding is illustrated as:

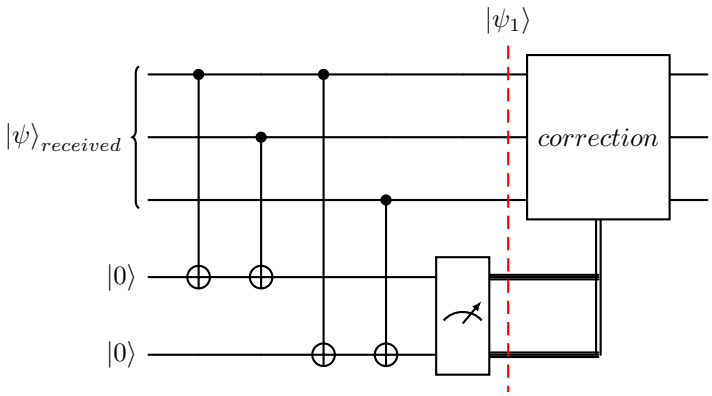


There is a simple two stage error-correction procedure which can be used to recover the correct quantum state in this case:

1. *Error-detection:* We perform a measurement which tells us what error, if any, occurred on the quantum state. The measurement result is called the error syndrome. For the bit flip channel there are four error syndromes, corresponding to the four projection operators:

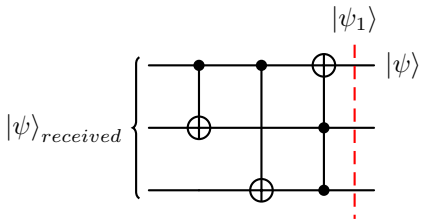
$$P_0 = |000\rangle\langle 000| + |111\rangle\langle 111| \text{ no error}$$
$$P_1 = |100\rangle\langle 100| + |011\rangle\langle 011| \text{ bit flip on first qubit}$$
$$P_2 = |010\rangle\langle 010| + |101\rangle\langle 101| \text{ bit flip on second qubit}$$
$$P_3 = |001\rangle\langle 001| + |110\rangle\langle 110| \text{ bit flip on third qubit}$$

Suppose for example that a bit flip occurs on qubit one, so the corrupted state is $a|100\rangle + b|011\rangle$. Notice that $\langle\psi| P_1 |\psi\rangle = 1$ in this case, so the outcome of the measurement result (the error syndrome) is certainly 1. Furthermore, the syndrome measurement does not cause any change to the state: it is $a|100\rangle + b|011\rangle$ both before and after syndrome measurement. Note that the syndrome contains only information about what error has occurred, and does not allow us to infer anything about the value of a or b, that is, it contains no information about the state being protected. This is a generic feature of syndrome measurements, since to obtain information about the identity of a quantum state it is in general necessary to perturb that state.
The syndrome measurement is done using a system of additional qubits known as ancilla system. We present here a procedure for this specific case while the more general proof is presented in the Appendix.



$\|\psi\rangle_{received}$	$\|\psi_1\rangle$
$\alpha\|000\rangle + \beta\|111\rangle$	$\alpha\|0\rangle\|00\rangle + \beta\|1\rangle\|00\rangle$
$\alpha\|001\rangle + \beta\|110\rangle$	$\alpha\|0\rangle\|01\rangle + \beta\|1\rangle\|01\rangle$
$\alpha\|010\rangle + \beta\|101\rangle$	$\alpha\|0\rangle\|10\rangle + \beta\|1\rangle\|10\rangle$
$\alpha\|100\rangle + \beta\|011\rangle$	$\alpha\|0\rangle\|11\rangle + \beta\|1\rangle\|11\rangle$

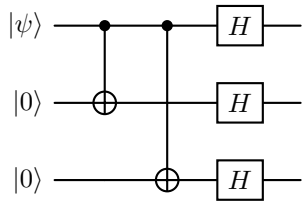
Thus, by measuring the ancilla qubits, we are able to detect the syndrome without perturbing the system. And then, the different correction transformation occurs based on the syndrome. Instead of the elaborate decoding circuit given above, a rather simpler circuit also achieves the same goal.



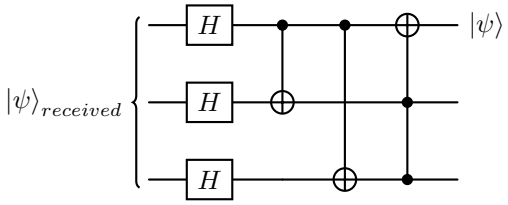
2. *Recovery:* We use the value of the error syndrome to tell us what procedure to use to recover the initial state. For example, if the error syndrome was 1, indicating a bit flip on the first qubit, then we flip that qubit again, recovering the original state $a|000\rangle + b|111\rangle$ with perfect accuracy. The four possible error syndromes and the recovery procedure in each case are: 0 (no error) – do nothing; 1 (bit flip on first qubit) – flip the first qubit again; 2 (bit flip on second qubit) – flip the second qubit again; 3 (bit flip on third qubit) – flip the third qubit again. For each value of the error syndrome it is easy to see that the original state is recovered with perfect accuracy, given that the corresponding error occurred.

2.2 Three Qubit Phase Flip Code

Another error model for noisy quantum channels is the *phase flip model*. In this error model the qubit is left alone with probability $1 - p$, and with probability p the relative phase of the $|0\rangle$ and $|1\rangle$ states is flipped. More precisely, the phase flip operator Z is applied to the qubit with probability $p > 0$, so the state $a|0\rangle + b|1\rangle$ is taken to the state $a|0\rangle - b|1\rangle$ under the phase flip. There is no classical equivalent to the phase flip channel, since classical channels don't have any property equivalent to phase. However, there is an easy way to turn the phase flip channel into a bit flip channel. Suppose we work in the qubit basis $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$. With respect to this basis the operator Z takes $|+\rangle$ to $|-\rangle$ and vice versa, that is, it acts just like a bit flip with respect to the labels $+$ and $-$. This suggests using the states $|0\rangle_L = |+++\rangle$ and $|1\rangle_L = |---\rangle$ as logical zero and one states for protection against phase flip errors. Thus the encoding circuit gets modified as:



The circuit below does correction and disentanglement to get the initial qubit in the first register.



Indeed the phase flip code is able to correct any *arbitrary* error. Following calculations illustrate that:

$$|b_1b_2b_3\rangle = a\left(\frac{|0\rangle+|1\rangle}{\sqrt{2}}\right)\left(\frac{|0\rangle+e^{i\theta}|1\rangle}{\sqrt{2}}\right)\left(\frac{|0\rangle+|1\rangle}{\sqrt{2}}\right)+b\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right)\left(\frac{|0\rangle-e^{i\theta}|1\rangle}{\sqrt{2}}\right)\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right).$$

After applying Hadamard we get

$$|b_1b_2b_3\rangle = a\frac{1+e^{i\theta}}{2}|000\rangle + a\frac{1-e^{i\theta}}{2}|010\rangle + b\frac{1-e^{i\theta}}{2}|101\rangle + b\frac{1+e^{i\theta}}{2}|111\rangle.$$

Final result is

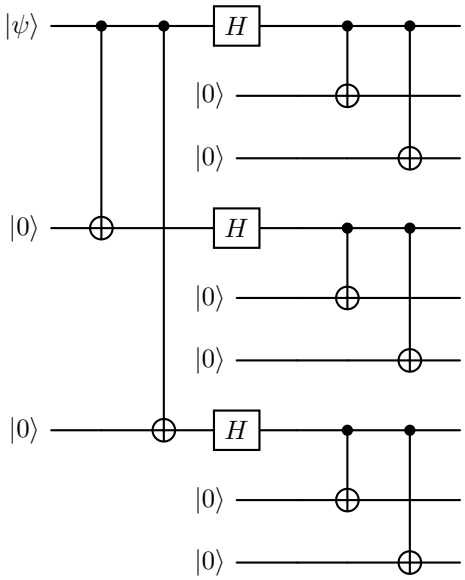
$$\begin{aligned}|b_1b_2b_3\rangle &= a\frac{1+e^{i\theta}}{2}|000\rangle + a\frac{1-e^{i\theta}}{2}|010\rangle + b\frac{1-e^{i\theta}}{2}|110\rangle + b\frac{1+e^{i\theta}}{2}|100\rangle \\ &= (a|0\rangle + b|1\rangle)\left(\frac{1+e^{i\theta}}{2}|00\rangle + \frac{1-e^{i\theta}}{2}|10\rangle\right)\end{aligned}$$

Thus the procedure outlined above corrects for arbitrary phase errors as long long as it does not occur on more than one bit. Only caveat is that we are not

2.3 Shor Code

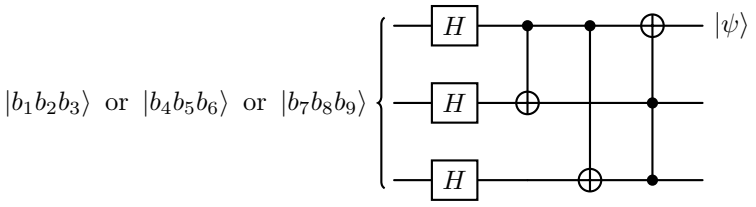
The Shor Code is a simple quantum code which can protect against the effects of an *arbitrary* error on a single qubit. The code is a combination of the three qubit phase flip and bit flip codes.

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

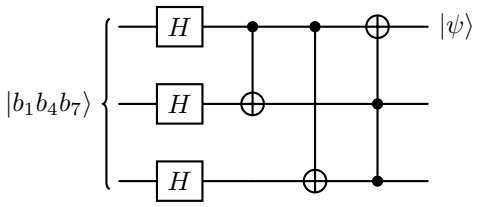


The first part of the circuit encodes the qubit using the three qubit phase flip code.The second part of the circuit encodes each of these three qubits using the bit flip code, using three copies of the bit flip code encoding circuit. This method of encoding using a hierarchy of levels is known as concatenation.

So, the message consists of 3 blocks of entangled sets of qubits. The Shor code is able to protect against phase flip and bit flip errors on any qubit. To see this, suppose a bit flip occurs on the first qubit. The circuit given below corrects any single bit flip error in an individual block.



Now the phase flip correction circuit is as follows.



Suppose both bit and phase flip errors occur on the first qubit, that is, the operator Z_1X_1 is applied to that qubit. Then it is easy to see that the procedure for detecting a bit flip error will detect a bit flip on the first qubit, and correct it, and the procedure for detecting a phase flip error will detect a phase flip on the first block of three qubits, and correct it. Thus, the Shor code also enables the correction of combined bit and phase flip errors on a single qubit.

Let's see an example of how Shor Code decoding works. Suppose the initial qubit is $|\psi\rangle = a\,|0\rangle + b\,|1\rangle$ and the encoded message undergoes bit flip and phase error of θ on the fifth qubit. The received state is therefore

$$|\psi_1\rangle = a\frac{(|000\rangle + |111\rangle)(|010\rangle + e^{i0}\,|101\rangle)(|000\rangle + |111\rangle)}{2\sqrt{2}} + b\frac{(|000\rangle - |111\rangle)(|010\rangle - e^{i\theta}\,|101\rangle)(|000\rangle - |111\rangle)}{2\sqrt{2}}.$$

After correcting the bit flip, the result is

$$|b_1b_4b_7\rangle = a\left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right)\left(\frac{|0\rangle + e^{i\theta}\,|1\rangle}{\sqrt{2}}\right)\left(\frac{|0\rangle + |1\rangle}{\sqrt{2}}\right) + b\left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right)\left(\frac{|0\rangle - e^{i\theta}\,|1\rangle}{\sqrt{2}}\right)\left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right).$$

After applying Hadamard we get

$$|b_1b_4b_7\rangle = a\frac{1+e^{i\theta}}{2}\,|000\rangle + a\frac{1-e^{i\theta}}{2}\,|010\rangle + b\frac{1-e^{i\theta}}{2}\,|101\rangle + b\frac{1+e^{i\theta}}{2}\,|111\rangle.$$

Final result is

$$\begin{aligned}|b_1b_4b_7\rangle &= a\frac{1+e^{i\theta}}{2}\,|000\rangle + a\frac{1-e^{i\theta}}{2}\,|010\rangle + b\frac{1-e^{i\theta}}{2}\,|110\rangle + b\frac{1+e^{i\theta}}{2}\,|100\rangle \\ &= (a\,|0\rangle + b\,|1\rangle)\left(\frac{1+e^{i\theta}}{2}\,|00\rangle + \frac{1-e^{i\theta}}{2}\,|10\rangle\right)\end{aligned}$$

Thus the procedure outlined above corrects for arbitrary phase errors and bit flips as long long as each of them do not occur on more than one bit.

2.4 Quantum Noise and Quantum Operations

The quantum operations formalism is a general tool for describing the evolution of quantum systems in a wide variety of circumstances, including stochastic changes to quantum states, much as Markov processes describe stochastic changes to classical states.

$$\rho' = \mathcal{E}(\rho)$$

Quantum operation is a map $\mathcal{E}(\rho)$ over the space of density operators. Two simple examples are unitary transformations and measurements for which $\mathcal{E}(\rho) = U\rho U^\dagger$ and $\mathcal{E}(\rho) = M_m \rho M_m^\dagger$, respectively. The quantum operation captures the dynamic change to a state which occurs as the result of some physical process; ρ is the initial state before the process, and $\mathcal{E}(\rho)$ is the final state after the process occurs, possibly up to some normalization factor.

Over the next two sections, we will develop two different formulations of quantum operations.

2.4.1 Environments and Quantum Operations

The dynamics of a closed quantum system are described by a unitary transform. A natural way to describe the dynamics of an open quantum system is to regard it as arising from an interaction between the system of interest, which we shall call the principal system, and an environment, which together form a closed quantum system. Then, by taking partial trace over the environment, we obtain the reduced state of the system alone:

$$\mathcal{E}(\rho) = tr_{env}\left[U(\rho \otimes \rho_{env})U^\dagger\right].$$

Of course, if U does not involve any interaction with the environment, then $\mathcal{E}(\rho) = \tilde{U}\rho\tilde{U}^\dagger$ where $U = U \otimes I$. An important assumption is made in this definition – we assume that the system and the environment start in a product state. In general, of course, this is not true. Quantum systems interact constantly with their environments, getting entangled in the process.

We present an example here. Consider a single qubit as the principal system and another qubit as the environment. U is the controlled-NOT gate and the environment is initially in the state $\rho_{env} = |0\rangle\langle 0|$ as the target qubit. It can be verified that $U = P_0 \otimes I + P_1 \otimes X$. Therefore,

$$\begin{aligned}
\mathcal{E}(\rho) &= tr_{env}\left[U(\rho \otimes \rho_{env})U^\dagger\right] \\
&= tr_{env}\left[(P_0\rho \otimes |0\rangle\langle 0| + P_1\rho \otimes |1\rangle\langle 0|)(P_0 \otimes I + P_1 \otimes X)\right] \\
&= tr_{env}\left[P_0\rho P_0 \otimes |0\rangle\langle 0| + P_0\rho P_1 \otimes |0\rangle\langle 1| + P_1\rho P_0 \otimes |1\rangle\langle 0| + P_1\rho P_1 \otimes |1\rangle\langle 1|\right] \\
&= P_0\rho P_0\langle 0|0\rangle + P_0\rho P_1\langle 1|0\rangle + P_1\rho P_0\langle 0|1\rangle + P_1\rho P_1\langle 1|1\rangle \\
&= P_0\rho P_0 + P_1\rho P_1
\end{aligned}$$

2.4.2 Operator-sum Representation

Quantum operations can be represented as the operator-sum representation where these operators act on the principal system's Hilbert space alone. Let $\{|e_k\rangle\}$ be an orthonormal basis for the state space of the environment and let $rho_{env} = |e_0\rangle\langle e_0|$ be the initial state of the environment. We assume here that the environment is described a finite dimensional Hilbert Space. Also, we assume that the environment starts in a pure state, since if it starts in a mixed state we are free to introduce an extra system purifying the environment. But, still we are making a simplifying assumption that the prinicipal system are not entangled in the beginning. Thus,

$$\begin{aligned}
\mathcal{E}(\rho) &= tr_{env}[U(\rho \otimes |e_0\rangle\langle e_0|)U^\dagger] \\
&= \sum_k (\langle e_k| U (\rho \otimes |e_0\rangle\langle e_0|) U^\dagger |e_k\rangle \\
&= \sum_k E_k \rho E_k^\dagger,
\end{aligned}$$

where $E_k = \langle e_k| U |e_0\rangle$ is an operator on the state space of the principal system where $|e_0\rangle$ and $\langle e_k|$ are actually abuse of notations which actually denote the operators $f_{1,0}$ and $f_{2,k}$ respectively:

$$\begin{aligned}
f_{1,0}(|\psi\rangle) &= |\psi\rangle \otimes |e_0\rangle \\
f_{2,k}(|\psi\rangle \otimes |\phi\rangle) &= \langle e_k|\phi\rangle |\psi\rangle,
\end{aligned}$$

where $|\psi\rangle$ and $|\phi\rangle$ belong to state spaces of the principal system and the environment respectively.

2.4.3 Freedom in the Operator-sum Representation

Theorem 2.4.1. *Suppose $\{E_1, \cdots, E_m\}$ and $\{F_1, \cdots, F_n\}$ are operation elements giving rise to qunatum operations $\mathcal{E}$ and $\mathcal{F}$, respectively. By appending zero operators to the shorter list of operation elements we may ensure that $m = n$. Then $\mathcal{E} = \mathcal{F}$ if and only if there exist com[plex numbers u_{ij} such that $E_i = \sum_j u_{ij} F_j$, and u_{ij} is an $m \times m$ unitary matrix.*

Proof. Recall from Theorem 1.2.2 that two sets of vectors $|\psi_i\rangle$ and $|\phi_j\rangle$ generate the same operator if and only if

$$|\psi_i\rangle = \sum_j u_{ij} |\phi_j\rangle ,$$

where u_{ij} is a unitary matrix of complex numbers, and we 'pad' whichever set of states $|\psi_i\rangle$ or $|\phi_j\rangle$ is smaller with additional states 0 so that the two sets have the same number of elements. Now, suppose $\{E_1, \cdots, E_m\}$ and $\{F_1, \cdots, F_n\}$ are two sets of operation elements for the same quantum operation, $\sum_i E_i \rho E_i^\dagger = \sum_j F_j \rho F_j^\dagger$ for all ρ. Let Q be the system on which the quantum operation is acting and let R be another system with the same state space. $|k_R\rangle$ and $|k_Q\rangle$ are respectively the basis sets of the state spaces of systems R and Q respectively. Define

$$|e_i\rangle = \sum_k |k_R\rangle (E_i |k_Q\rangle)$$
$$|f_j\rangle = \sum_k |k_R\rangle (F_j |k_Q\rangle)$$

Then $|e_i\rangle$ and $|f_j\rangle$ define the same density operator as

$$\begin{aligned}
\sum_i |e_i\rangle \langle e_i| &= \sum_i \left(\sum_k |k_R\rangle (E_i |k_Q\rangle) \right) \left(\sum_{k'} \langle k'_R| (\langle k'_Q| E_i^\dagger) \right) \\
&= \sum_{kk'} |k_R\rangle \langle k'_R| \sum_i E_i |k_Q\rangle \langle k'_Q| E_i^\dagger \\
&= \sum_{kk'} |k_R\rangle \langle k'_R| \sum_j F_j |k_Q\rangle \langle k'_Q| F_j^\dagger \\
&= \sum_j \left(\sum_k |k_R\rangle (F_j |k_Q\rangle) \right) \left(\sum_{k'} \langle k'_R| (\langle k'_Q| F_j^\dagger) \right) \\
&= \sum_f |f_i\rangle \langle f_j|
\end{aligned}$$

Thus there exists unitary u_{ij} such that

$$|e_i\rangle = \sum_j u_{ij} |f_j\rangle .$$

But for arbitrary $|\psi\rangle$ we have

$$\begin{aligned}
E_i |\psi\rangle &= \sum_k \psi_k E_i |k_Q\rangle \\
&= \sum_k (\psi_k^*)^* E_i |k_Q\rangle \\
&= \sum_k (\langle k_Q| |\tilde{\psi}\rangle)^* E_i |k_Q\rangle \\
&= \sum_k \langle \tilde{\psi}| k_R\rangle E_i |k_Q\rangle \\
&= \langle \tilde{\psi}| e_i\rangle \\
&= \sum_j u_{ij} \langle \tilde{\psi}| f_j\rangle \\
&= \sum_j u_{ij} F_j |\psi\rangle
\end{aligned}$$

where $|\tilde{\psi}\rangle = \sum_k \psi_k^* |k_Q\rangle$. Thus, $E_i |\psi\rangle = \sum_j u_{ij} F_j |\psi\rangle$. Conversely, suppose E_i and f_j are related by a unitary transformation of the form $E_i |\psi\rangle = \sum_j u_{ij} F_j |\psi\rangle$, then for arbitrary ρ we have

$$\begin{aligned}\mathcal{E}(\rho) &= \sum_i E_i \rho E_i^\dagger \\ &= \sum_i \left(\sum_j u_{ij} F_j \right) \rho \left(\sum_{j'} u_{ij'}^* F_{j'}^\dagger \right) \\ &= \sum_{jj'} F_j \rho F_j^\dagger \sum_i u_{ij} u_{ij'}^* \\ &= \sum_j F_j \rho F_j^\dagger = \mathcal{F}(\rho)\end{aligned}$$

because complex dot product of two different columns of a unitary matrix is 0. □

2.4.4 Depolarizing Channel

The depolarizing channel is an important type of quantum noise. Imagine we take a single qubit, and with probability p that qubit is *depolarized*. That is, it is replaced by the completely mixed state, $I/2$. With probability $1-p$ the qubit is left untouched. The state of the quantum system after this noise is

$$\mathcal{E}(\rho) = p\frac{I}{2} + (1-p)\rho.$$

The above form is not in the operator-sum representation. But, we claim that for arbitrary ρ,

$$\frac{I}{2} = \frac{\rho + X\rho X + Y\rho Y + Z\rho Z}{4}.$$

Define $\mathcal{E}(A) \equiv \frac{A+XAX+YAY+ZAZ}{4}$, and observe that $\mathcal{E}(I) = I$ and $\mathcal{E}(X) = \mathcal{E}(Y) = \mathcal{E}(Z) = 0$. Then, for arbitrary ρ we have $\rho = \frac{I+\vec{r}\cdot\vec{\sigma}}{2} = \frac{I+r_x X+r_y Y+r_z Z}{2}$, substituting which the equation is satisfied. Thus, substituting for $I/2$ in $\mathcal{E}(\rho) = p\frac{I}{2} + (1-p)\rho$, we arrive at the equation

$$\mathcal{E}(\rho) = \left(1 - \frac{3p}{4}\right)\rho + \frac{p}{4}\left(X\rho X + Y\rho Y + Z\rho Z\right),$$

showing that the depolarizing channel has operation elements $\{\sqrt{1-3p/4}I, \sqrt{p}X/2, \sqrt{p}Y/2, \sqrt{p}Z/2\}$. Note that it is frequently convenient to parametrize the depolarizing channel in different way, such as

$$\mathcal{E}(\rho) = (1-p)\rho + \frac{p}{3}(X\rho X + Y\rho Y + Z\rho Z),$$

which has the interpretation that the state ρ is left alone with probability $1-p$, and the operators X, Y and Z applied each with probability $p/3$.

2.4.5 Fidelity

A measure of distance between probability distributions, the *fidelity* of the probability distributions $\{p_x\}$ and $\{q_x\}$ indexed with the same events, is defined by

$$F(p,q) \equiv \sum_x \sqrt{p_x q_x}.$$

Analogously, fidelity is a measure of distance between quantum states. The fidelity of states ρ and σ is defined to be

$$F(\rho,\sigma) \equiv tr\sqrt{\rho^{1/2}\sigma\rho^{1/2}}.$$

There are two important special cases where it is possible to give more explicit formulae for the fidelity. The first is when ρ and σ commute, that is, are diagonal in the same basis,

$$\rho = \sum_i r_i |i\rangle\langle i|; \qquad \sigma = \sum_i s_i |i\rangle\langle i|,$$

for some orthonormal basis $|i\rangle$. In this case we see that

$$\begin{aligned} F(\rho,\sigma) &= tr\sqrt{\rho\sigma} \\ &= tr\sqrt{\sum_i r_i s_i |i\rangle\langle i|} \\ &= tr(\sum_i \sqrt{r_i s_i}\, |i\rangle\langle i|) \\ &= \sum_i \sqrt{r_i s_i} \\ &= F(r_i, s_i). \end{aligned}$$

That is, when ρ and σ commute the quantum fidelity $F(\rho,\sigma)$ reduces to the classical fidelity $F(r,s)$ between the eigenvalue distributions r_i and s_i of ρ and σ. Our second example is to calculate the fidelity between a pure state $|\psi\rangle$ and an arbitrary state ρ. As, ρ can be diagonalized with $|\psi\rangle\langle\psi|$ being the first element, they commute and hence,

$$F(|\psi\rangle, \rho) = \sqrt{\langle\psi|\rho|\psi\rangle}.$$

We present here a theorem without proof which is useful in proving some further results.

Theorem 2.4.2 (Strong concavity of the fidelity)**.** *Let p_i and q_i be the probability distributions over the same index set, and ρ_i and σ_i density operators also indexed by the same index set. Then*

$$F(\sum_i p_i\rho_i, \sum_i q_i\sigma_i) \geq \sum_i \sqrt{p_i q_i} F(\rho_i, \sigma_i).$$

2.4.6 How well does a quantum channel preserve information?

Suppose a quantum system is in the state $|\psi\rangle$ and some physical process occurs, changing the quantum system to the state $\mathcal{E}(|\psi\rangle\langle\psi|)$. We can compute the fidelity between the starting state $|\psi\rangle$ and the ending state $\mathcal{E}(|\psi\rangle\langle\psi|)$. For the case of the depolarising channel, we obtain

$$\begin{aligned} F(|\psi\rangle, \mathcal{E}(|\psi\rangle\langle\psi|)) &= \sqrt{\langle\psi|\left(p\frac{I}{2} + (1-p)|\psi\rangle\langle\psi|\right)|\psi\rangle} \\ &= \sqrt{1-\frac{p}{2}} \end{aligned}$$

But in a real quantum memory or quantum communications channel, we don't know in advance what the initial state $|\psi\rangle$ of the system will be. However, we can quantify the worst-case behavior of the system by minimizing over all possible initial states,

$$F_{min}(\mathcal{E}) \equiv \min_{|\psi\rangle} F(|\psi\rangle, \mathcal{E}(|\psi\rangle\langle\psi|)).$$

For example, for the p-depolarizing channel $F_min = \sqrt{1-p/2}$, as the fidelity of the channel is the same for all input states $|\psi\rangle$.

2.5 Theory of Quantum Error Correction

This section develops a general framework for studying quantum error-correction, including the quantum error-correction conditions, a set of equations which must be satisfied if quantum error-correction is to be possible. The basic ideas of the theory of quantum error-correction generalize in a natural way the ideas introduced by the Shor code. Quantum states are encoded by a unitary operation into a quantum error-correcting code, formally defined as a subspace C of some larger Hilbert space. After encoding the code is subjected to noise, following which a syndrome measurement is performed to diagnose the type of error which occurred, that is, the error syndrome. Once this has

been determined, a recovery operation is performed, to return the quantum system to the original state of the code. Different error syndromes correspond to undeformed and orthogonal subspaces of the total Hilbert space. The subspaces must be orthogonal, otherwise they couldn't be reliably distinguished by the syndrome measurement. Furthermore, the different subspaces must be undeformed versions of the original code space, in the sense that the errors mapping to the different subspaces must take the (orthogonal) codewords to orthogonal states, in order to be able to recover from the error. This intuitive picture is essentially the content of the quantum error-correction conditions discussed below.

We just make two very broad assumptions: the noise is described by a quantum operation E, and the complete error-correction procedure is effected by a trace-preserving quantum operation R, which we call the error-correction operation. In order for error-correction to be deemed successful, we require that for any state ρ whose support lies in the code C,

$$(\mathcal{R} \circ \mathcal{E})(\rho) \propto \rho$$

If $\mathcal{E}$ were a trace-preserving quantum operation then by taking traces of both sides of the equation we see that $\propto$ would be $=$. However, sometimes we may be interested in error-correcting non-trace-preserving operations $\mathcal{E}$, such as measurements, for which $\propto$ is appropriate. Of course, the error-correction step $\mathcal{R}$ must succeed with probability 1, which is why we required $\mathcal{R}$ to be trace-preserving.

Theorem 2.5.1 (Quantum error-correction conditions)**.** *Let C be a quantum code, and let P be the projector onto C. Suppose $\mathcal{E}$ is a quantum operation with operation elements $\{E_i\}$. A necessary and sufficient condition for the existence of an error-correction operation $\mathcal{R}$ correcting $\mathcal{E}$ on C is that*

$$PE_i^\dagger E_j P = \alpha_{ij} P,$$

for some Hermitian matrix α of complex numbers.

We call the operation elements $\{E_i\}$ for the noise $\mathcal{E}$ errors, and if such an $\mathcal{R}$ exists we say that $\{E_i\}$ constitutes a correctable set of errors.

Proof. We prove sufficiency first, by constructing an explicit error-correction operation $\mathcal{R}$ whenever the condition is satisfied. Suppose $\{E_i\}$ is a set of operation elements satisfying the quantum error-correction conditions. By assumption α is a Hermitian matrix, and thus can be diagonalized, $d = u^\dagger \alpha u$, where u is a unitary matrix and d is diagonal. Define operators $F_k \equiv \sum_i u_{ik} E_i$. Recalling Theorem 2.4.1, we see that $\{F_k\}$ is also a set of operation elements for $\mathcal{E}$. By direct substitution,

$$PF_k^\dagger F_l P = \sum_{ij} u_{kl}^\dagger u_{jl} PE_i^\dagger E_j P.$$

Substituting the given condition simplifies this to $PF_k^\dagger F_l P = \sum_{ij} u_{ki}^\dagger \alpha_{ij} u_{jl} P$ and since $d = u^\dagger \alpha u$ we obtain

$$PF_k^\dagger F_l P = d_{kl} P,$$

which can be thought of as a simplification of the quantum error-correction conditions, because d_{kl} is diagonal. Now we define the syndrome measurement. From the polar decomposition, we see that $F_k P = U_k \sqrt{PF_k^\dagger F_k P} = \sqrt{d_{kk}} U_k P$ for some unitary U_k. The effect of F_k is therefore to rotate the coding subspace into the subspace defined by the projector $P_k \equiv U_k P U_k^\dagger = F_k P U_k^\dagger / \sqrt{d_{kk}}$. These subspaces are orthogonal since when $k \neq l$,

$$P_l P_k = P_l^\dagger P_k = \frac{U_l P F_l^\dagger F_k P U_k^\dagger}{\sqrt{d_{ll} d_{kk}}} = \frac{U_l d_{kl} P U_k^\dagger}{\sqrt{d_{ll} d_{kk}}} = 0.$$

The syndrome measurement is a projective measurement defined by the projectors P_k, augmented by an additional projector if necessary to satisfy the completeness relation $\sum_k P_k = I$. Recovery is accomplished simply by applying $U_k^\dagger$. To see that this error-correction procedure works, note that the combined detection-recovery step corresponds to the quantum operation $\mathcal{R}(\sigma) = \sum_k U_k^\dagger P_k \sigma P_k U_k$. For states ρ in the code, simple algebra and the definitions show that:

$$\begin{aligned} U_k^\dagger P_k F_l P &= \frac{U_k^\dagger U_k P F_k^\dagger F_l P}{\sqrt{d_{kk}}} \\ &= \delta_{kl} \sqrt{d_{kk}} P \end{aligned}$$

Thus

$$\begin{aligned}\mathcal{R}(\mathcal{E}(\rho)) &= \sum_{kl} U_k^\dagger P_k F_l \rho F_l^\dagger P_k U_k \\ &= \sum_{kl} U_k^\dagger P_k F_l P \rho P F_l^\dagger P_k U_k \\ &= \sum_{kl} \delta_{kl} d_{kk} \rho \\ &\propto \rho,\end{aligned}$$

as required, where we have used $P\rho P = \rho$ which is true because $\rho = \sum_i |i_L\rangle \langle i_L|$ where $|i_L\rangle$ are the basis states of the code space C.

To prove necessity of the quantum error-correction conditions, suppose $\{E_i\}$ is a set of errors which is perfectly correctable by an error-correction operation $\mathcal{R}$ with operation elements $\{R_j\}$. Define a quantum operation $\mathcal{E}_C$ by $\mathcal{E}_C(\rho) \equiv \mathcal{E}(P\rho P)$. Since $P\rho P$ is in the code space for all ρ, it follows that

$$\mathcal{R}(\mathcal{E}_C(\rho)) \propto P\rho P,$$

for all states ρ. Moreover, the proportionality factor must be a constant c, bot depending on ρ, if both right and left hand sides are to be linear. Rewriting the last equation explicitly in terms of the operation elements gives

$$\sum_{ij} R_j E_i P \rho P E_i^\dagger R_j^\dagger = cP\rho P.$$

This equation holds for all ρ. It follows that the quantum operation with operation elements $\{R_j E_i\}$ is identical to the qunatum operation with a single operation element $\sqrt{c}P$. Theorem 2.4.1 implies that there exist complex numbers c_{ki} such that

$$R_k E_i P = c_{ki} P.$$

Therefore $PE_i^\dagger R_k^\dagger R_k E_j P = c_{ki}^* c_{kj} P$. But $\mathcal{R}$ is a trace-preserving operation, so $\sum_k R_k^\dagger R_k = I$. Summing the equation $PE_i^\dagger R_k^\dagger R_k E_j P = c_{ki}^* c_{kj} P$ over k we deduce that

$$PE_i^\dagger E_j P = \alpha_{ij} P,$$

where $\alpha_{ij} = \sum_k c_{ki}^* c_{kj}$ is Hermetian as $\alpha_{ij}^* = \alpha_{ji}$. □

2.5.1 Discretization of Errors

The following theorem shows that the error-correction conditions are so strong that such an $\mathcal{R}$ could be used to protect against an entire class of noise processes.

Theorem 2.5.2. *Suppose C is a quantum code and $\mathcal{R}$ is the error-correction operation constructed in the proof of Theorem 2.5.1 to recover from a noise process $\mathcal{E}$ with operation elements $\{E_i\}$. Suppose $\mathcal{F}$ is a quantum operation with operation elements $\{F_j\}$ which are linear combinations of the E_i, that is $F_j = \sum_i m_{ji} E_i$ for some matrix m_{ji} of complex numbers. Then the error-correction operation $\mathcal{R}$ also corrects for the effects of the noise process $\mathcal{F}$ on the code C.*

Proof. By Theorem 2.5.1 the operation elements $\{E_i\}$ must satisfy the quantum error-correction conditions, $PE_i E_j^\dagger P = \alpha_{ij} P$. As shown in the proof of Theorem 2.5.1, without loss of generality we may assume that the operation elements for $\mathcal{E}$ have been chosen such that $\alpha_{ij} = d_{ij}$ is diagonal with real entries (because α needs to be Hermetian as well). The error-correction operation $\mathcal{R}$ has operation elements $U_k^\dagger P_k$, where U_k and P_k are chosen such that for any ρ in the code space

$$U_k^\dagger P_k E_i P = \delta_{ki} \sqrt{d_{kk}} P.$$

Substituting $F_j = \sum_i m_{ji} E_i$ gives

$$\begin{aligned}U_k^\dagger P_k F_j P &= \sum_i m_{ji} \delta_{ki} \sqrt{d_{kk}} P \\ &= m_{jk} \sqrt{d_{kk}} P,\end{aligned}$$

and thus

$$\begin{aligned}\mathcal{R}(\mathcal{F}(\rho)) &= \sum_{kj} U_k^\dagger P_k F_j \rho F_j^\dagger P_k U_k \\ &= \sum_{kj} U_k^\dagger P_k F_j P \rho P F_j^\dagger P_k U_k \\ &= \sum_{kj} |m_{jk}|^2 d_{kk} \rho \\ &\propto \rho\end{aligned}$$

as required. □

Let's look at an example. Suppose $\mathcal{E}$ is a quantum operation acting on a single qubit. Then its operation elements $\{E_i\}$ can each be written as linear combination of the Pauli matrices σ_0, σ_1, σ_2, σ_3. Therefore, to check that the Shor code corrects against arbitrary single qubit errors on the first qubit it is sufficient to verify that the equation

$$P\sigma_i^1\sigma_j^1 = \alpha_{ij}P,$$

are satisfied, where σ_i^1 are the Pauli matrices $\{I, X, Y, Z\}$ acting on the first qubit. Thus, considering a special error channel called the *depolarising* channel, $\mathcal{E}(\rho) = (1-p)\rho + \frac{p}{3}(X\rho X + Y\rho Y + Z\rho Z)$, which has $\{I, X, Y, Z\}$ as its operation elements, the ability to error-correct the depolarising channel automatically implies the ability to error-correct an arbitrary single qubit quantum operation. Thus, we are able to *discretize* quantum errors, that to fight the continuum of errors possible on a single qubit it is sufficient merely to win the war against a finite set of errors, the four Pauli matrices.

2.5.2 Independent Error Models

If a noise process acts independently on the different qubits in the code, then provided the noise is sufficiently weak error-correction should improve the storage fidelity of the encoded state over the unencoded state. Recall that the depolarising channel on a single qubit is defined by the equation $\mathcal{E}(\rho) = (1-p)\rho + \frac{p}{3}[X\rho X + Y\rho Y + Z\rho Z]$, and can be interpreted as saying that nothing happens to the qubit with probability $1-p$, and each of the operators X, Y and Z is applied to the qubit with probability $p/3$. We showed in Section 2.4.6 that the minimum fidelity for states sent through a depolarizing channel is given by $F = \sqrt{1-2p/3} = 1 - p/3 + O(p^2)$. Suppose we encode a single qubit of information in an n qubit quantum code which corrects errors on any single qubit. Suppose the depolarizing channel with parameter p acts independently on each of the qubits, giving rise to a joint action on all n qubits of

$$\mathcal{E}^{\otimes n}(\rho) = (1-p)^n\rho + \sum_{j=1}^{n}\sum_{k=1}^{3}(1-p)^{n-1}\frac{p}{3}\sigma_k^j\rho\sigma_k^j + \cdots,$$

where the '$\cdots$' indicates higher-order terms which are positive and will drop out of the analysis. After error-correction has been performed all terms appearing in this sum will be returned to the state ρ, provided ρ was in the code originally,

$$(\mathcal{R}\circ\mathcal{E}^{\otimes n})(\rho) = [(1-p)^n + n(1-p)^{n-1}p]\rho + \cdots$$

so the fidelity satisfies

$$F \geq \sqrt{(1-p)^{n-1}(1-p+np)} = 1 - \frac{\binom{n}{2}}{2}p^2 + O(p^3)$$

. Thus, provided the probability of error p is sufficiently small, using the quantum error-correcting code leads to an improvement in the fidelity of the quantum states being protected by the code.

2.5.3 Degenerate Codes

Quantum codes can have an interesting property unknown in classical codes. Consider the effect of the errors Z_1 and Z_2 on the codewords for the Shor code. The effect of these errors is the same on both codewords. For classical error-correcting codes errors on different bits necessarily lead to different corrupted codewords. Thus, some of the proof techniques used classically to prove bounds on error-correction fall down because they can't be applied to degenerate codes. But, on the flip side, they are able to 'pack more information in' than are classical codes, because distinct errors do not necessarily have to take the code space to orthogonal spaces, and it is possible (though has not yet been shown) that this extra ability may lead to degenerate codes that can store quantum information more efficiently than any non-degenerate code.

2.5.4 Quantum Hamming Bound

In this section we develop the quantum Hamming bound, a simple bound which gives some insights into the general properties of quantum codes. Unfortunately the quantum Hamming bound only applies to non-degenerate codes, but it gives us an idea of what more general bounds may look like. Suppose a non-degenerate code is used to encode k qubits in n qubits in such a way that it can correct errors on any subset of t or fewer qubits. Suppose j errors occur, where $j \leq t$. There are $\binom{n}{j}$ sets of locations where errors may occur. With each such set of locations there are three possible errors - the three Pauli matrices X, Y, Z - that may occur in each qubit, for a total of 3^j possible errors. The total number of errors that may occur on t or fewer qubits is therefore

$$\sum_{j=0}^{t} \binom{n}{j} 3^j.$$

(Note that $j = 0$ corresponds to the case of no errors on any qubit, the 'error' I.) In order to encode k qubits in a non-degenerate way each of these errors must correspond to an orthogonal 2^k-dimensional subspace. All of these subspaces must be fitted into the total 2^n-dimensional space available to n qubits, leading to the inequality

$$\sum_{j=0}^{t} \binom{n}{j} 3^j 2^k \leq 2^n,$$

which is the quantum Hamming bound. Consider, for example, the case where we wish to encode one qubit in n qubits in such a way that errors on one qubit are tolerated. In this case the quantum Hamming bound reads:

$$2(1 + 3n) \leq 2^n.$$

Substitution shows that this inequality is not satisfied for $n \leq 4$, while it is for values of $n \geq 5$. Therefore, there is no non-degenerate code encoding one qubit in fewer than five qubits in such a way as to protect from all possible errors on a single qubit.

2.6 Calderbank-Shor-Steane codes

Our first example of a large class of quantum error-correcting codes is the *Calderbank-Shor-Steane* codes, more usually known as *CSS*.

Suppose C_1 and C_2 are $[n, k_1]$ and $[n, k2]$ classical linear codes such that $C_2 \subset C_1$ and C_1 and $C_2^\perp$ both correct t errors. We will define an $[n, k_1 - k_2]$ quantum code $CSS(C_1, C_2)$ capable of correcting errors on t qubits, the CSS code of C_1 over C_2, via the following construction. Suppose $x \in C_1$ is any codeword in the code C_1. Then we define the quantum state $|x + C_2\rangle$ by

$$|x + C_2\rangle \equiv \frac{1}{\sqrt{|C_2|}} \sum_{y \in C_2} |x + y\rangle ,$$

where $+$ is bitwise addition modulo 2. Suppose x' is an element of C_1 such that $x - x' \in C_2$. Then it is easy to see that $|x + C_2\rangle = |x' + C_2\rangle$, and thus the state $|x + C_2\rangle$ depends only upon the coset of C_1/C_2 which x is in, explaining the coset notation we have used for $|x + C_2\rangle$. Furthermore, if x and x'

belong to different cosets of C_2, then for no $y, y' \in C_2$ does $x + y = x' + y'$, and therefore $|x + c_2\rangle$ and $|x' + C_2\rangle$ are orthonormal states. The quantum code $CSS(C_1, C_2)$ is defined to be the vector space spanned by the states $|x + C_2\rangle$ for all $x \in C_1$. The number of cosets of C_2 in C_1 is $|C_1|/|C_2|$ so the dimension of $CSS(C_1, C_2)$ is $|C_1|/|C_2| = 2^{k_1 - k_2}$, and therefore $CSS(C_1, C_2)$ is an $[n, k_1 - k_2]$ quantum code.

We can exploit the classical error-correcting properties of C_1 and $C_2^\perp$ to detect and correct quantum errors! In fact, it is possible to error-correct up to t bit and phase flip errors on $CSS(C_1, C_2)$ by making use of the error-correcting properties of C_1 and $C_2^\perp$, respectively. Suppose the bit flip errors are described by an n bit vector e_1 with 1s where bit flips occured, and 0s elsewhere, and the phase flip errors are described by an n bit vector e_2 with 1s where phase flips occurred, and 0s elsewhere. If $|x + C_2\rangle$ was the original state then the corrupted state is:

$$\frac{1}{\sqrt{|C_2|}} \sum_{y \in C_2} (-1)^{(x+y)\cdot e_2} |x + y + e_1\rangle .$$

To detect where bit flips occurred it is convenient to introduce an ancilla containing sufficient qubits to the store the syndrome for the code C_1, and initially in the all zero state $|0\rangle$. We use reversible computation to apply the parity matrix H_1 for the code C_1, taking $|x + y + e_1\rangle |0\rangle$ to

$$|x + y + e_1\rangle |H_1(x + y + e_1)\rangle = |x + y + e_1\rangle |H_1 e_1\rangle ,$$

since $(x + y) \in C_1$ is annihilated by the parity check matrix. The effect of this operaton is to produce the state:

$$\frac{1}{\sqrt{|C_2|}} \sum_{y \in C_2} (-1)^{(x+y)\cdot e_2} |x + y + e_1\rangle |H_1 e_1\rangle .$$

Error-detection for the bit flip errors is completed by measureing the ancilla to obtain the result $H_1 e_1$ and discarding the sncilla, giving the state

$$\frac{1}{\sqrt{|C_2|}} \sum_{y \in C_2} (-1)^{(x+y)\cdot e_2} |x + y + e_1\rangle .$$

Knowing the error syndrome $H_1 e_1$, we can infer the error e_1 since C_1 can correct up to t error, which completes the error-detection. Recovery is performed simply by applying NOT gates to the qubits at whichever positions in the error e_1 a bit flip occurred, removing all the bit flip errors and giving the state

$$\frac{1}{\sqrt{|C_2|}} \sum_{y \in C_2} (-1)^{(x+y)\cdot e_2} |x + y\rangle .$$

To detect phase flip errors we apply Hadamard gates to each qubit, takiung the state to

$$\frac{1}{\sqrt{|C_2| 2^n}} \sum_z \sum_{y \in C_2} (-1)^{(x+y)\cdot(e_2+z)} |z\rangle ,$$

where the sum is over all possible values for n bit z. Setting $z' \equiv z + e_2$, this state may be rewritten:

$$\frac{1}{\sqrt{2^n |C_2|}} \sum_{z'} \sum_{y \in C_2} (-1)^{(x+y)\cdot z'} |z' + e_2\rangle .$$

Supposing $z' \in C_2^\perp$ it is easy to see that $\sum_{y \in C_2} (-1)^{y \cdot z'} = |C_2|$, while if $z' \notin C_2^\perp$ then $\sum_{y \in C_2} (-1)^{y \cdot z'} = 0$. Thus the state may be rewritten:

$$\frac{1}{\sqrt{2^n / |C_2|}} \sum_{z' \in C_2^\perp} (-1)^{x \cdot z'} |z' + e_2\rangle ,$$

which looks just like a bit flip error described by the vector e_2! As for the error-detection for bit flips we introduce an ancilla and reversibly apply the parity check matrix H_2 for $C_2^\perp$ to obtain $H_2 e_2$, and correct the 'bit flip error' e_2, obtaining the state

$$\frac{1}{\sqrt{2^n / |C_2|}} \sum_{z' \in C_2^\perp} (-1)^{x \cdot z'} |z'\rangle .$$

The error-correction is completed by again applying Hadamard gates to each qubit; we get

$$\frac{1}{\sqrt{|C_2|}} \sum_{y \in C_2} |x + y\rangle ,$$

which is the original encoded state!

2.6.1 Steane code

An important example of a CSS code may be constructed using the $[7, 4, 3]$ Hamming code whose parity check matrix we recall here:

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

Suppose we label this code C and define $C_1 \equiv C$ and $C_2 \equiv C^{\perp}$. To use these codes to define a CSS code we need first to check that $C_2 \subset C_1$. By definition the parity check matrix of $C_2 = C^{\perp}$ is equal to the transposed generator matrix of $C_1 = C$:

$$h[C_2] = G[C_1]^T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} .$$

Comparing with H we see that the span of the rows of $H[C_2]$ strictly contains the span of the rows of $H[C_1]$, and since the corresponding codes are the kernels of $H[C_2]$ and $H[C_1]$ we conclude that $C_2 \subset C_1$. Furthermore, $C_2^{\perp} = (C^{\perp})^{\perp} = C$, so both C_1 and $C_2^{\perp}$ are distance 3 codes which can correct errors on 1 bit. Since C_1 is a $[7, 4]$ code and C_2 is a $[7, 3]$ code it follows that $CSS(C_1, C_2)$ is a $[7, 1]$ quantum code which can correct errors on a single qubit. This $[7, 1]$ quantum code has nice properties that make it easy to work with, and will be used in many of the examples for the remainder of this chapter. It is known as the Steane code, after its inventor. The codewords of C2 are easily determined from $H[C_2]$.

$$|0_L\rangle = \frac{1}{\sqrt{8}}[|0000000\rangle + |1010101\rangle + |0110011\rangle + |1100110\rangle + |0001111\rangle + |1011010\rangle + |0111100\rangle + |1101001\rangle].$$

To determine the other logical codeword we need to find an element of C_1 that is not in C_2. An example of such an element is $(1, 1, 1, 1, 1, 1, 1)$, giving:

$$|1_L\rangle = \frac{1}{\sqrt{8}}[|1111111\rangle + |0101010\rangle + |1001100\rangle + |0011001\rangle + |1110000\rangle + |0100101\rangle + |1000011\rangle + |0010110\rangle].$$

2.7 Stabilizer Codes

Stabilizer codes are an important class of quantum codes also known as *additive* quantum codes whose construction is analogous to classical linear codes.

2.7.1 Stabilizer Formalism

Consider the EPR state of two qubits

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

It easy to verify that $X_1 X_2 |\psi\rangle = |\psi\rangle$ and $Z_1 Z_2 |\psi\rangle = |\psi\rangle$; we say that the state $|\psi\rangle$ is *stabilized* by the operators $X_1 X_2$ and $Z_1 Z_2$. We will also prove a little later that the dimension of a state described

by $n-k$ independent stabilizers is 2^k, hence $|\psi\rangle$ is the unique quantum state (upto a global phase) which is stabilized by these operators X_1X_2 and Z_1Z_2 as $k=2$ and $n=2$ in this case. So, the basic idea of stabilizer formalism is to describe quantum states by their *stabilizers* rather than working with the actual amplitudes.

Stabilizer formalism is developed here using *group theory*. The group of principal interest is the *Pauli group* G_n on n qubits. For a single qubit, the Pauli group is defined to consist of all the Pauli matrices, together with multiplicative factors $\pm 1, \pm i$: $G_1 \equiv \{\pm I, \pm iI, \pm X, \pm iX, \pm Y, \pm iY, \pm Z, \pm iZ\}$. This set of matrices is closed under matrix multiplication, and thus forms a group. The general Pauli group on n qubits is defined to consist of all n-fold tensor products of G_1.

Suppose S is a subgroup of G_n and define V_S to be the set of n qubit states which are fixed by every element of S. V_S is the vector space (which is easy to prove) stabilized by S, and S is said to be the stabilizer of the space V_S. Not just any subgroup S of the Pauli group can be used as the stabilizer for a non-trivial vector space. Two conditions are necessary and sufficient for this. We prove the necessary part here and the sufficient part at the end of this section.

Proposition 2.7.1. *A subgroup S of the Pauli group G_n stabilizes a non-trivial vector space only if:*

(a) the elements of S commute,

(b) $-I$ does not belong to S.

Proof. Suppose S stabilizes V_S which is non-trivial, that is $|\psi\rangle \in V_S$ such that $|\psi\rangle \neq 0$. Suppose M and N be elements of S which do not commute. As these are Pauli operators, they either commute or anti-commute, hence $MN = -NM$. So, we have $|\psi\rangle = MN\,|\psi\rangle = -NM\,|\psi\rangle = -\,|\psi\rangle$, which is a contradiction to $|\psi\rangle$ being non-zero.

Now, suppose $-I \in S$, then $|\psi\rangle = -I\,|\psi\rangle = |\psi\rangle$ which again leads to the contradiction. □

Next we define *check matrix*, an extremely useful way of presenting hte generators $g_1, g_2, \cdots g_l$ of S. This is an $l \times 2n$ matrix whose rows correspond to the generators g_1 through g_l; the left hand side of the matrix contains 1s to indicate which generators contain Xs, and the right hand side contains 1s to indicate which generators contain Zs; the presence of a 1 on both sides indicates a Y in the generator. More explicitly, the ith row is constructed as follows. If g_i contains an I on the jth qubit then the jth and $n+j$th column elements are 0; if it contains an X on the jth qubit then the jth column element is a 1 and the $n+j$th column element is a 0l if it contains a Z on the jth qubit then the jth column element if 0 and the $n+j$th column element if 1; if it contains a Y on the jth qubit then both the jth and $n+j$th columns are 1. In the case of the Steane seven qubit code we can read the check matrix as:

Name	Operator
g_1	$IIIXXXX$
g_2	$IXXIIXX$
g_3	$XIXIXIX$
g_4	$IIIZZZZ$
g_5	$IZZIIZZ$
g_6	$ZIZIZIZ$

$$\left[\begin{array}{ccccccc|ccccccc} 0&0&0&1&1&1&1&0&0&0&0&0&0&0\\ 0&1&1&0&0&1&1&0&0&0&0&0&0&0\\ 1&0&1&0&1&0&1&0&0&0&0&0&0&0\\ 0&0&0&0&0&0&0&0&0&0&1&1&1&1\\ 0&0&0&0&0&0&0&0&1&1&0&0&1&1\\ 0&0&0&0&0&0&0&1&0&1&0&1&0&1 \end{array}\right]$$

Note that the check matrix does not contain any information about the multiplicative factors $\pm 1, \pm i$. We define $r(g)$ as the $2n$-dimensional row vector representation of an element g of the Pauli group. Suppose we define a $2n \times 2n$ matrix Λ by

$$\Lambda = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}$$

where the I matrices on the off-diagonals are $n \times n$. Elements g and g' of the Pauli group are easily seen to commute if and only if $r(g)\Lambda r(g')^T = 0$; the formula $x\Lambda y^T$ defines a sort of *twisted* inner product between row matrices x and y expressing whether the elements of hte Pauli group corresponding to x and y commute or not.

A useful connection between independence of the generators and the check matrix is established by means of the following proposition:

Proposition 2.7.2. *Let $S = \langle g_1, \cdots g_l \rangle$ be such that $-I$ is not an element of S. The generators g_1 through g_l are independent if and only if the rows of the corresponding check matrix are linearly independent.*

Proof. We prove the contrapositive. Observe that $r(g) + r(g') = r(gg')$, which can verified taking cases. Thus the rows of the check matrix are linearly independent, $\sum_i a_i r(g_i) = 0$, and $a_j \neq 0$ for some j, if and only if $\Pi_i g_i^{a_i}$ is equal to the identity, up to an overall multiplicative factor. But $-I \notin S$ so the multiplicative factor must be 1, and the last condition corresponds to the condition $g_j = g_j^{-1} = \Pi_{i \neq j} g_i^{a_i}$, and thus $g_1, \cdots g_l$ are not independent generators. □

We will prove another very useful proposition which will used repeatedly in this chapter.

Proposition 2.7.3. *Let $S = \langle g_1, \cdots g_l \rangle$ be generated by l independent generators and satisfy $-I \notin S$. Fix i in the range $1, 2, \cdots l$. Then there exists $g \in G_n$ such that $g g_i g^\dagger = -g_i$ and $g g_j g^\dagger = g_j$ for all $j \neq i$.*

Proof. Let G be the check matrix associated to $g_1, g_2, \cdots g_l$. The rows of G are linearly independent by Proposition 2.7.2, which can be extended to an linearly independent set of size $2n$, constituting the matrix G' and vector e_i' by appending zeros to the vector e_i which is the i^{th} standard basis vector of $\mathbb{Z}^l$:

$$G' = \begin{bmatrix} G \\ \cdots \end{bmatrix}_{2n \times 2n}, \qquad e_i' = \begin{bmatrix} e_i \\ 0 \end{bmatrix}.$$

Hence there exists a $2n$-dimensional vector x such that $G'\Lambda x = e_i'$. Thus, $G\Lambda x = e_i$. Let g be such that $r(g) = x^T$. Then, by taking j^{th} row of the equation $G\Lambda x = e_i$, we have $r(g_j)\Lambda r(g)^T = 0$ for $j \neq i$ and $r(g_i)\Lambda r(g)^T = 1$, and thus $g g_i g^\dagger = g_i$ and $g g_j g^\dagger = g_j$ for $j \neq i$. □

Now, as promised in the beginning of this section, we prove that V_S is non-trivial provided S is generated by independent commuting generators and $-I \notin S$. We expect each stabilizer to cut the dimension of V_S by a factor of $\frac{1}{2}$ as there are two eigenvalues $+1$ and -1. Hence if there are $l = n - k$ generators, then it is expected that V_S is 2^k-dimensional.

Proposition 2.7.4. *Let $S = \langle g_1, \cdots, g_{n-k}$ be generated by $n-k$ independent and commuting elements from G_n, and such that $-I \notin S$. Then V_S is a 2^k-dimensional vector space.*

Proof. Let $x = (x_1, \cdots, x_{n-k})$ be a vector of $n - k$ elements of $\mathbb{Z}_2$. Define

$$P_S^x \equiv \prod_{j=1}^{n-k} \frac{I + (-1)^{x_j} g_j}{2},$$

where the addition and division operations are with respect to the vector space of operators. Since, $(I+g_j)/2$ is idempotent and stabilises the $+1$ eigenspace of g_j, it is the projector onto the $+1$ eigenspace of g_j. Thus, $P_S^{(0,0,\cdots 0)} = \prod_j (I + g_j)/2$ is the projector onto V_S as it is idempotent and stabilises V_S. By Proposition 2.7.3 for each i there exists g_i' such that $g_i' g_i g_i'^\dagger = g_i$ and $g_i' g_j g_i'^\dagger = g_j$ for $j \neq i$. And hence, for each x, we have $\hat{g}_x = \prod_{x_j=1} g_j' = g_{j_1}' g_{j_2}' \cdots g_{j_k}'$ in G_n such that:

$$\begin{aligned} \hat{g}_x P_S^{(0,0,\cdots 0)} \hat{g}_x^\dagger &= g_{j_1}' g_{j_2}' \cdots g_{j_k}' \left[\prod_j (I + g_j)/2 \right] g'^\dagger_{j_k} g'^\dagger_{j_{k-1}} \cdots g'^\dagger_{j_1} \\ &= \prod_{j=1}^{n-k} \frac{I + (-1)^{x_j} g_j}{2} \\ &= P_S^x \end{aligned}$$

Note here that g'_j for different values of j need not commute. $\hat{g}_x P_S^{(0,0,\cdots 0)} \hat{g}_x^\dagger = P_S^x$ and that $\hat{g}_x$ is unitary implies that the dimension of P_S^x is the same as the dimension of V_S. Furthermore, for $x \neq y$,

$P_S^x P_S^y$ will have a factor $(I+g_j)(I-g_j)$ if x and y differ at the j^{th} position, and hence P_S^x are orthogonal for distinct values of x. Now, observe that $I = \sum_x P_S^x$ as:

$$\begin{aligned} I &= \frac{I+g_j}{2} + \frac{I-g_j}{2} \\ &= \prod_j \left(\frac{I+g_j}{2} + \frac{I-g_j}{2} \right) \\ &= \sum_{(x_1,x_2,\cdots x_l)\in\mathbb{Z}^l} \prod_{j=1}^{l} \frac{I+(-1)^{x_j} g_j}{2} \\ &= \sum_x P_S^x. \end{aligned}$$

The left hand side is a projector onto a 2^n-dimensional space, while the right hand side is a sum over 2^{n-k} orthogonal projectors of the same dimension as V_S and thus the dimension of V_S must be 2^k. □

2.7.2 Unitary gates and the stabilizer formalism

Suppose we apply a unitary operation U to a vector space v_S stabilized by the group S. Let $|\psi\rangle$ be any element of V_S. Then for any element g of S,

$$U|\psi\rangle = Ug|\psi\rangle = UgU^\dagger U|\psi\rangle,$$

and thus the state $U|\psi\rangle$ is stabilized by $UgU^\dagger$, from which we deduce that the vector space UV_S is stabilized by the group $USU^\dagger \equiv \{UgU^\dagger | g \in S\}$. Furthermore, $g_1, g_2, \cdots g_l$ generate S, then $Ug_1U^\dagger, Ug_2U^\dagger, \cdots, Ug_lU^]\dagger$ generate $USU^\dagger$. Suppose, for example, that we apply a Hadamard gate to a single qubit. Note that

$$HXH^\dagger = Z; \quad HYH^\dagger = -Y; \quad HZH^\dagger = X.$$

As a consequence, we correctly deduce that after a Hadamard gate is applied to the quantum state stabilized by Z i.e. $|0\rangle$, the resulting state will be stabilized by X i.e. $|+\rangle$.

2.7.3 Measurement in the stabilizer formalism

Measurements in the *computational basis* may also be easliy described within the stabilizer formalism. To understand how this works, imagine we make a measurement of $g \in G_n$ (recall that g is a Hermitian operator, and can thus be regarded as an observable in the sense of projective measurements). For convenience we assume without loss of generality that g is a product of Pauli matrices with no multiplicative factor of -1 or $\pm i$ out the front. The system is assumed to be in a state $|\psi\rangle$ with stabilizer $\langle g_1, \cdots g_n\rangle$, where $g_1, \cdots g_l$ are independent (whence V_S has dimension 1 and hence a unique state upto a global phase). How does the stabilizer of the state transform under this measurement? There are two possibilities:

- g commutes with all the generators of the stabilizer.
- g anti-commutes with one or more of the generators of the stabilizer. Suppose the stabilizer has generators $g_1, \cdots h_n$ and that g anti-commutes with g_1. Without loss of generality, we may assume that g commutes with $g_2, \cdots g_n$, since if it does not commute with one of these elements (say g_2) then it is easy to verify that g does commute with g_1g_2, and we simply replace the generator g_2 by g_1g_2 in our list of generators for the stabilizer.

In the first instance, it follows that either g or $-g$ is an element of the stabilizer by the following argument. Since $g_j g|\psi\rangle = gg_j|\psi\rangle = g|\psi\rangle$ for each stabilizer generator, $g|\psi\rangle$ is in V_S and is thus a multiple of $|\psi\rangle$ as V_S has dimension 1. Because $g^2 = I$ it follows that $g|\psi\rangle = \pm|\psi\rangle$, whence wither g or $-g$ must be in the stabilizer. We assume that g is in the stabilizer, with the discurssion for $-g$ proceeding analogously. In this instance $g|\psi\rangle = |\psi\rangle$ and thus a measurement of g yields $+1$ with probability one ($tr(\frac{I+g}{2}|\psi\rangle\langle\psi|) = 1$), and the measurement does not disturb the state of the system, and thus leaves the stabilizer invariant, i.e. $\langle g_1, \cdots g_l\rangle$ stabilizes the result of the measurement.

In the second instance, g anti-commutes with g_1 and commutes with all the other generators of thwe stabilizer. Note that g has eigenvalues $\pm$ and so the projectors for the measurement outcomes ± 1 are given by $(I \pm g)/2$ respectively and thus the measurement probabilities are given by

$$p(+1) = tr\left(\frac{I+g}{2} |\psi\rangle \langle\psi|\right)$$
$$p(-1) = tr\left(\frac{I-g}{2} |\psi\rangle \langle\psi|\right)$$

Using the facts that $g_1 |\psi\rangle = |\psi\rangle$ and $gg_1 = -g_1 g$ gives

$$\begin{aligned} p(+1) &= tr\left(\frac{I+g}{2} g_1 |\psi\rangle \langle\psi|\right) \\ &= tr\left(g_1 \frac{I-g}{2} |\psi\rangle \langle\psi|\right) \\ &= tr\left(\frac{I-g}{2} |\psi\rangle \langle\psi|\right) = p(-1) \end{aligned}$$

Here we have used the cyclic property of train and $\langle\psi| g_1 = \langle\psi| g_1^\dagger = \langle\psi|$ as $g_1^\dagger = g_1$. Since $p(+1) + p(-1) = 1$, we deduce that $p(+1) = p(-1) = 1/2$. Suppose the result $+1$ occurs. In this instance hte new state of the system is $|\psi^+\rangle \equiv (I+g)|\psi\rangle / \sqrt{2}$, which is stabilized by $\langle g, g_2, \cdots g_n\rangle$ where $g, g_2, \cdots g_n$ are independent as g does not commute with g_1. Similarly, if the result -1 occurs the posterior state is stabilized by $\langle -g, g_2, \cdots g_n\rangle$.

2.7.4 Stabilizer code constructions

An $[n, k]$ stabilizer code is defined to be the vector space V_S stabilized by a subgroup S of G_n such that $-I \notin S$ and S has $n-k$ independent and commuting generators, $S = \langle g_1, \cdots, g_{n-k}\rangle$. We denote this code $C(S)$. What are the logical basis states for the code $C(S)$? In principle, given $n-k$ generators for the stabilizer S we can choose any 2^k orthonormal vectors in the code $C(S)$ to act as our logical computational basis states. In practice it makes a great deal more sense to choose the states in a more systematic way. One method is as follows. First, we choose operators $\bar{Z}_1, \cdots, \bar{Z}_k \in G_n$ such that $g_1, \cdots, g_{n-k}, \bar{Z}_1, \cdots, \bar{Z}_k$ forms an independent and commuting set. (We explain in the appendix as this can be done.) Recall that on a single qubit, the Z operator stabilizes the state $|0\rangle$ while $-Z$ stabilizes the state $|1\rangle$. The $\bar{Z}_j$ operator plays the role of a logical Pauli sigma Z operator on logical qubit number j, so the logical computational basis state $|x_1, \cdots, c_k\rangle_L$ is therefore defined to be the state with stabilizer

$$\langle g_1, \cdots, g_{n-k}, (-1)^{x_1} \bar{Z}_!, \cdots, (-1)^{x_k} \bar{Z}_k\rangle.$$

Similarly, we define $\bar{X}_j$ to be that product of Pauli matrices which anti-commutes with $\bar{Z}_j$ while commutes with all other $\bar{Z}_i$ and g_i which exists by Proposition 2.7.3. Clearly $\bar{X}_j$ has the effect of a quantum NOT gate acting on the j^{th} encoded qubit. Suppose $C(S)$ is a stabilizer code corrupted by an error $E \in G_n$. What happens to the code space when E anti-commutes with an element of the stabilizer? In this case, E takes $C(S)$ to an orthogonal subspace, and the error can in principle be detected and perhaps be corrected by performing an appropriate projective measurement. If $E \in S$ we don't need to worry since the 'error' E doesn't corrupt the space at all. But when E commutes with all the elements of S but is not actually in S, that is, $Eg = gE$ for all $g \in S$. The set of $E \in G_n$ such that $Eg = gE$ for all $g \in S$ is known as the centralizer of S in G_n and in denoted $Z(S)$. In fact, if $-I \notin S$ then the centralizer is identical to the normalizer of S, denoted $N(S)$, which is defined to consist of all elements E of G_n such that $EgE^\dagger \in S$. This is because for all $E, g \in G_n$, $EgE^\dagger$ is a scalar multiple of g where the scalar is ± 1 or $\pm i$, where it is easy to see that the scalar being one of -1 or $\pm i$ implies that $-I \in S$. We thus arrive at the quantum error-correction conditions for stabilizer codes.

Theorem 2.7.1. *Let S be the stabilizer for a stabilizer code $C(S)$. Suppose $\{E_j\}$ is a set of operators in G_n such that $E_j^\dagger E_k \notin N(S) - S$ for all j and k. Then $\{E_j\}$ is a correctable set of errors for the code $C(S)$.*

Without loss of generality we can restrict ourselves to considering errors E_j in G_n such that $E_j^\dagger = E_j$, because otherwise there is a complex scalar factor which can be removed as they have the same linear span.

Proof. Let P be the projector onto the code space $C(S)$ For given j and k there are two possibilities: either $E_j^\dagger E_k$ in S or in $G_n - N(S)$. Consider the first case. Then $PE_j^\dagger E_k P = P$ since P is invariant under multiplication by elements of S. Suppose $E_j^\dagger E_k \in G_n - N(S)$ so that $E_j^\dagger E_k$ must anticommute with some element g_1 of S. Let $g_1, \cdots, g_{n-k}$ be a set of generators of S, so that

$$P = \prod_{l=1}^{n-k} \frac{I + g_l}{2}.$$

Using the anti-commutativity gives

$$\begin{aligned} E_j^\dagger E_k P &= E_j^\dagger E_k \prod_{l=1}^{n-k} \frac{I + g_l}{2} \\ &= (I - g_1) E_j^\dagger E_k \frac{\prod_{l=2}^{n-k} (I + g_l)}{2^{n-k}} \end{aligned}$$

But $P(I - g_1) = 0$ since $(I + g_1)(I - g_1)$ and therefore $PE_j^\dagger E_k P = 0$ whenever $E_j^\dagger E_k \in G_n - N(S)$. It follows that the matrix α defined as $PE_j^\dagger E_k P = \alpha_{jk} P$ is real and symmetric and hence, the set of errors $\{E_j\}$ satisfies the quantum error-correction conditions, and thus forms a correctable set of errors. □

Above is a proof of existence of an error-correction operation when it is possible. To understand how this is achieved. suppose $g_1, \cdots, g_{n-k}$ is a set of generators for the stabilizer of an $[n, k]$ stabilizer code, and that $\{E_j\}$ is a set of correctable errors for the code. Error-detection is performed by measuring the generators of the stabilizer g_1 through g_{n-k} in turn, to obtain the error syndrome, which consists of the results of the measurements, β_1 through β_{n-k}.

2.7.5 Examples

We now give two examples of Stabilizer codes, the three qubit bit flip code and CSS codes.

The three qubit bit flip code

Consider the familiar three qubit bit flip code spanned by the states $|000\rangle$ and $|111\rangle$, with stabilizer generated by $Z_1 Z_2$ and $Z_2 Z_3$. By inspection we see that every possible product of two elements from the error set $\{I, X_1, X_2, X_3\}$ anticommutes with at least one of the generators of the stabilizer (except for I which is in S). Thus by Theorem 2.7.1 the set $\{I, X_1, X_2, X_3\}$ forms a correctable set of errors for the three qubit bit flip code with stabilizer $\langle Z_1 Z_2, Z_2 Z_3 \rangle$. Error-detection for the bit flip code is effected by measuring the stabilizer generators, $Z_1 Z_2$ and $Z_2 Z_3$. If, for example, the error X_1 occurred, then the stabilizer is transformed to $\langle -Z_1 Z_2, Z_2 Z_3 \rangle$, so the syndrome measurement gives the results -1 and $+1$. Similarly, the error X_2 gives error syndrome -1 and -1, the error X_3 gives error syndrome $+1$ and -1, and the trivial error I gives error syndrome $+1$ and $+1$. In each instance recovery is effected in the obvious way simply by applying the inverse operation to the error indicated by the error syndrome.

CSS codes and the seven qubit code

The CSS codes are an excellent example of a class of stabilizer codes, demonstrating beautifully how easy the stabilizer formalism makes it to understand quantum code construction. Suppose C_1 and C_2 are $[n, k_1]$ and $[n, k_2]$ classical linear codes such that $C_2 \subset C_1$ and C_1 and $C_2^\perp$ both correct t errors. Define a check matrix with the form

$$\left[\begin{array}{c|c} H(C_2^\perp) & 0 \\ 0 & H(C_1) \end{array}\right]$$

To see that this defines a stabilizer code, we need the check matrix to satisfy the commutativity condition $H(C_2^{\perp})H(C_1)^T = 0$. But we have $H(C_2^{\perp})H(C_1)^T = [H(C_1)G(C_2)]^T = 0$ because of the assumption $C_2 \subset C_1$. Indeed, it's an easy exercise to see that this code is exactly $CSS(C_1, C_2)$, and that it is capable of correcting arbitrary errors on any t qubits. The seven qubit Steane code is an example of a CSS code, whose check matrix we have already seen. Encoded Z and X operators may be defined for the Steane code as

$$\bar{Z} = Z_1 Z_2 Z_3 Z_4 Z_5 Z_6 Z_7; \qquad \bar{X} = X_1 X_2 X_3 X_4 X_5 X_6 X_7.$$

Appendix A

Some additional results

A.1 Trace

Trace of a square matrix A is defined to be the sum of its diagonal elements,

$$tr(A) = \sum_i A_{ii}.$$

The trace is easily seen to be cyclic $tr(AB) = tr(BA)$ and linear $tr(A + B) = tr(A) + tr(B)$, $tr(zA) = z\,tr(A)$ where A and B are arbitrary matrices and z is a complex number. Furthermore, from the cyclic property it follows that the trace of a matrix is invariant under the *unitary similarity transformation* $A \longrightarrow UAU^\dagger$ as $tr(UAU^\dagger) = tr(AU^\dagger U) = tr(A)$. So, the trace of an operator defined as the trace of its matrix representation with respect to some orthonormal basis is well-defined. As an example of the trace, suppose $|\psi\rangle$ is a unit vector and A is an arbitrary operator. To evaluate $tr(A\,|\psi\rangle\,\langle\psi|)$, use the Gram-Schmidt procedure to extend $|\psi\rangle$ to an orthonormal basis which includes $|\psi\rangle$ as its first element. Then we have,

$$tr(A\,|\psi\rangle\,\langle\psi|) = \sum_i \langle i|\,A\,|\psi\rangle\,\langle\psi|i\rangle = \langle\psi|\,A\,|\psi\rangle\,.$$

A.2 Method of Ancilla System

Suppose we have a quantum system with state space Q, and we want to perform a measurement described by measurement operators M_m on the system Q. To do this, we introduce an ancilla system, with state space M , having an orthonormal basis $|\psi\rangle$ in one-to-one correspondence with the possible outcomes of the measurement we wish to implement. This ancilla system can be regarded as merely a mathematical device appearing in the construction, or it can be interpreted physically as an extra quantum system introduced into the problem, which we assume has a state space with the required properties.

Letting $|0\rangle$ be any fixed state of M, define an operator U on products $|\psi\rangle\,|0\rangle$ of states $|\psi\rangle$ from Q with the state $|0\rangle$ by

$$U\,|\psi\rangle\,|0\rangle = \sum_m M_m\,|\psi\rangle\,|m\rangle\,.$$

Using the orthonormality of the states $|m\rangle$ and the completeness relation $\sum_m M_m^\dagger M_m = I$, we can see that U preserves inner products between states of the form $|\psi\rangle\,|0\rangle$,

$$\begin{aligned}\langle\psi|\,\langle 0|\,U^\dagger U\,|\psi\rangle\,|0\rangle &= \sum_{mm'} \langle\psi|\,M_m^\dagger M_m\,|\psi\rangle\,\langle m|m'\rangle \\ &= \sum_m \langle\psi|\,M_m^\dagger M_m\,|\psi\rangle \\ &= \langle\psi|\psi\rangle\end{aligned}$$

Theorem A.2.1. *Suppose V is a Hilbert space with a subspace W. Suppose $U : W \longrightarrow V$ is a linear operator which preserves inner products, that is, for any $|w_1\rangle$ and $|w_2\rangle$ in W,*

$$\langle w_1| U^\dagger U |w_2\rangle = \langle w_!|w_2\rangle .$$

Prove that there exists a unitary operator $U' : V \longrightarrow V$ which extends U. That is, $U'|w\rangle = U|w\rangle$ for all $|w\rangle$ in W, but U' is defined on the entire space V. Usually we omit the prime symbol $'$ and thus write U to denote the extension.

Proof. Suppose $W^\perp$ is the orthogonal complement of W. Then $V = W \oplus W^\perp$. Let $|w_i\rangle, |w'_j\rangle, |u'_j\rangle$ be orthonormal bases for W, $W^\perp$, $(\text{image}(U))^\perp$, respectively. Define $U' : V \to V$ as $U' = \sum_i |u_i\rangle\langle w_i| + \sum_j |u'_j\rangle\langle w'_j|$, where $|u_i\rangle = U|w_i\rangle$. Now

$$\begin{aligned}(U')^\dagger U' &= \left(\sum_{i=1}^{\dim W} |w_i\rangle\langle u_i| + \sum_{j=1}^{\dim W^\perp} |w'_j\rangle\langle u'_j|\right)\left(\sum_i |u_i\rangle\langle w_i| + \sum_j |u'_j\rangle\langle w'_j|\right)\\ &= \sum_i |w_i\rangle\langle w_i| + \sum_j |w'_j\rangle\langle w'_j| = I\end{aligned}$$

and

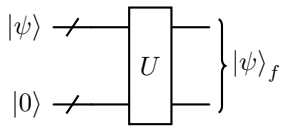
$$\begin{aligned}U'(U')^\dagger &= \left(\sum_i |u_i\rangle\langle w_i| + \sum_j |u'_j\rangle\langle w'_j|\right)\left(\sum_i |w_i\rangle\langle u_i| + \sum_j |w'_j\rangle\langle u'_j|\right)\\ &= \sum_i |u_i\rangle\langle u_i| + \sum_j |u'_j\rangle\langle u'_j| = I.\end{aligned}$$

Thus U' is an unitary operator. Moreover, for all $|w\rangle \in W$,

$$\begin{aligned}U'|w\rangle &= \left(\sum_i |u_i\rangle\langle w_i| + \sum_j |u'_j\rangle\langle w'_j|\right)|w\rangle\\ &= \sum_i |u_i\rangle\langle w_i|w\rangle + \sum_j |u'_j\rangle\langle w'_j|w\rangle\\ &= \sum_i |u_i\rangle\langle w_i|w\rangle \quad (\because |w'_j\rangle \perp |w\rangle)\\ &= \sum_i U|w_i\rangle\langle w_i|w\rangle\\ &= U|w\rangle .\end{aligned}$$

Therefore U' is an extension of U. □

Thus there exists a quantum gate which performs the transformation as:



where $|\psi\rangle_f = \sum_m M_m |\psi\rangle_m |m\rangle$. Next, suppose we perform a projective measurement on the two systems described by projectors $P_m = I_Q \otimes |m\rangle\langle m|$. Outcome m occurs with probability

$$\begin{aligned}p(m) &= \langle\psi|\langle 0| U^\dagger P_m U|\psi\rangle|0\rangle\\ &= \sum_{m',m''} \langle\psi| M^\dagger_{m'} \langle m| (I_Q \otimes |m\rangle\langle m|) M_{m''} |\psi\rangle |m\rangle\\ &= \langle\psi| M^\dagger_m M_m |\psi\rangle ,\end{aligned}$$

just as given in Postulate 3. The joint state of the system QM after measurement, conditional on result m occuring, is given by

$$\frac{P_m U |\psi\rangle |0\rangle}{\sqrt{\langle\psi| U^\dagger P_m U |\psi\rangle}} = \frac{M_m |\psi\rangle |m\rangle}{\sqrt{\langle\psi| M_m^\dagger M_m |\psi\rangle}}.$$

It follows that the state of system M after the measurement is $|m\rangle$, and the state of system Q is

$$\frac{M_m |\psi\rangle}{\sqrt{\langle\psi| M_m^\dagger M_m |\psi\rangle}},$$

just as described in Postulate 3. Thus, the method of introducing an ancilla system allows us to detect the outcome of the measurement without perturbing the system.

A.3 No Cloning Theorem

This property, that qubits cannot be copied, is known as the no-cloning theorem, and it is one of the chief differences between quantum and classical information. We describe an elementary proof of this fact that captures the essential reason this is not possible. Suppose we have a quantum machine with two slots labeled A and B. Slot A, the data slot, starts out in an unknown but pure quantum state, $|\psi\rangle$. This is the state which is to be copied into slot B, the target slot. We assume that the target slot starts out in some standard pure state, $|s\rangle$. Thus the initial state of the copying machine is

$$|\psi\rangle \otimes |s\rangle .$$

Some unitary evolution U now effects the copying procedure, ideally,

$$|\psi\rangle \otimes |s\rangle \longrightarrow U(|\psi\rangle \otimes |s\rangle) = |\psi\rangle \otimes |\psi\rangle .$$

Suppose this copying procedure works for two particular pure states, $|\psi\rangle$ and $|\phi\rangle$. Then we have

$$\begin{aligned} U(|\psi\rangle \otimes |s\rangle) &= |\psi\rangle \otimes |\psi\rangle \\ U(|\phi\rangle \otimes |s\rangle) &= |\phi\rangle \otimes |\phi\rangle . \end{aligned}$$

Taking the inner product of these two equations gives

$$\langle\psi|\phi\rangle = (\langle\psi|\phi\rangle)^2.$$

But $x = x^2$ has only two solutions, $x = 0$ and $x = 1$, so either $|\psi\rangle = |\phi\rangle$ or $|\psi\rangle$ and $|\phi\rangle$ are orthogonal. Thus a cloning device can only clone states which are orthogonal to one another, and therefore a general quantum cloning device is impossible. A potential quantum cloner cannot, for example, clone the qubit states $|\psi\rangle = |0\rangle$ and $|\phi\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, since these states are not orthogonal.

This proof rules out just one of many ways that might be possible to clone an arbitrary quantum state.

Bibliography

[1] Prof Dipan Ghosh. Quantum computation and quantum information. https://nptel.ac.in/courses/115101092, 2017.

[2] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge University Press, 2010.