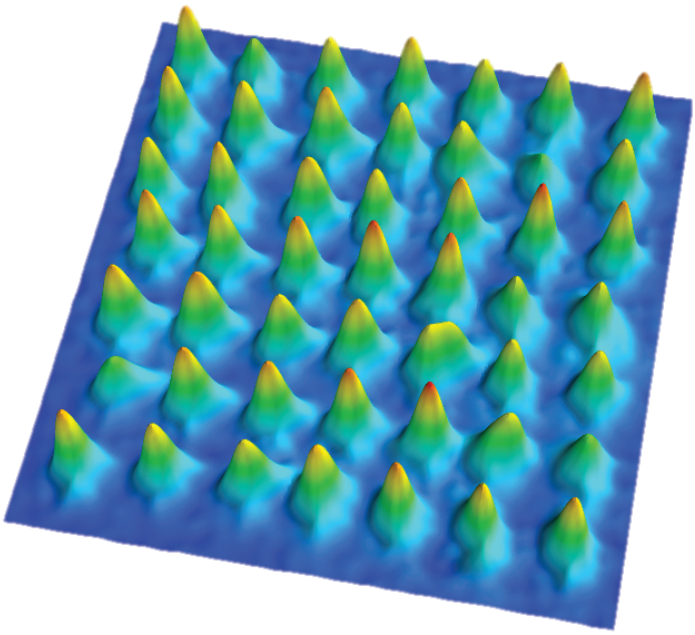


A little bit of quantum information

©M. Saffman
Department of Physics, University of Wisconsin-Madison

February 10, 2020

Contents

Chapter 1

Circuit model quantum computing

The framework of "circuit" model quantum computing is we encode an initial state $|\psi\rangle$ in n qubits. The qubits then evolve in time due to the application of various unitary operators. This leads to an output state of n qubits $|\psi'\rangle = \mathbf{U}|\psi\rangle$ as illustrated in Fig. 1.1. The result of the computation is then found from measurements on the output states. To perform arbitrary computations we need to be able to apply arbitrary unitary transformations $\mathbf{U}$ on the n qubits. Fortunately it turns out that such arbitrary transformations can be performed with a universal gate set consisting of a finite number of one- and two-qubit gates. Thus we only need to implement a handful of different operations in order to create a general quantum computing device. This is analogous to classical computing for which a few logic gates are universal. For example the combination of NOT and AND gates can implement any Boolean logic functions. Just a NAND gate is also sufficient.

The circuit model is only one form of quantum computing. There is also adiabatic quantum computing which formulates the answer to a desired problem in terms of finding the ground state of a Hamiltonian, which is prepared by adiabatic variation of parameters[1]. Another approach is "one-way" quantum computing[2]. This involves preparing a highly entangled so-called "cluster" state which is then measured. Depending on the results of the measurements single qubit gate operations are performed followed by more measurements and more gates. There is no requirement for additional two-qubit gates that create entanglement. Remarkably any circuit model quantum algorithm can be mapped onto adiabatic evolution or one way quantum computing, hence these different approaches are closely related in their computational power. On the other hand the necessary physical resources

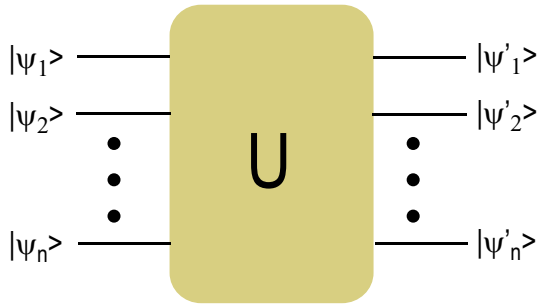


Figure 1.1: Circuit model quantum computation. Time flows from left to right.

may be very different which makes different approaches better suited for different hardware platforms. We will mainly discuss circuit model computation in these notes.

An obvious question is the following: why is the model of computation depicted in Fig. 1.1 more powerful than classical computation? This important question has not yet been fully answered, although, as we will see later, there are a number of quantum algorithms with a provable speed up compared to a classical computing device. One place where we can see a separation between classical and quantum data processing is in the use of qubits instead of classical binary bits. Given n binary bits we can represent one of 2^n different data values. A qubit can be in a superposition of states $|0\rangle$ and $|1\rangle$, as will be described in the next section, and the product of n qubits can simultaneously represent a superposition of all 2^n distinct binary states. This is expressed mathematically as

$$|\psi\rangle = \bigotimes_{j=1}^{n} \frac{|0\rangle_j + |1\rangle_j}{\sqrt{2}} = \frac{1}{2^{n/2}} \sum_{x=0}^{2^n-1} |x\rangle \tag{1.1}$$

where $|x\rangle$ are n digit binary strings. If we think of the qubits as providing input data for a computation then we see that all possible values, limited only by the size of the data register, can be input in parallel. From a naive perspective this gives a parallel processing speedup compared to a classical machine which requires that all data values are entered sequentially.

On the other hand the phenomenon of superposition is not unique to quantum mechanics. Any linear wave equation allows for superposition of wave solutions. Quantum superposition can therefore be simulated with linear waves. This was realized early on and there have been several demonstrations of quantum algorithms using linear wave superposition and interference[3, 4, 5]. However, a careful evaluation of such classical simulations always reveals that computational resource requirements in space or time scale faster than true quantum implementations.

If superposition alone does not explain quantum speedup another suspect is entanglement[6]. Entanglement was emphasized by Schrödinger as the decisive characteristic distinguishing quantum from classical phenomena[7, 8, 9]. The presence of entanglement inside a quantum computer implies that a classical description is not adequate to explain the inner workings, and it is therefore perhaps not surprising that the computational power may exceed what is possible in a machine that allows a purely classical description. On the other hand the presence of entanglement alone does not guarantee computational speedup and indeed the dynamics of some highly entangled states can be efficiently simulated on classical computers[10]. Furthermore there are counter examples of quantum machines with computational speedup yet vanishingly small entanglement[11]. Recent work has suggested that the important ingredient is contextuality, another type of quantum correlation[12]. Suffice it to say that a full understanding of the underlying reasons for quantum speedup remains a topic of ongoing research.

Irrespective of the reasons for the power of a quantum computer we know that classical computers are inadequate if we wish to understand the behavior of multi-particle quantum systems from first principles. The unitary evolution of a quantum state described by a density operator $\hat{\rho}$ is found by solving the equation

$$i\hbar \frac{\partial \hat{\rho}}{\partial t} = [\hat{\mathcal{H}}, \hat{\rho}]$$

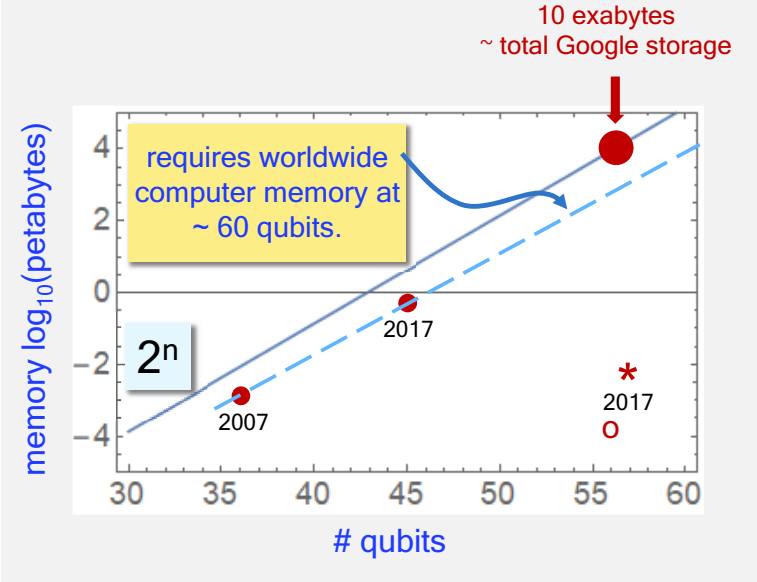


Figure 1.2: Memory requirements for classical simulation of a n qubit circuit. The indicated points are: 2007- [15], 2017- [16], * - [17], and O - [18].

with $\hat{\mathcal{H}}$ the system Hamiltonian. If the system of interest is comprised of n spin 1/2 particles, e.g. electrons, the density operator has dimensions $2^n \times 2^n$. Not all of the components of the density operator are independent since it is Hermitian and has unit trace. The trace condition gives a single constraint and Hermiticity implies that $\rho_{ij} = \rho^*_{ji}$ which gives

$$\frac{2^{2n} - 2^n}{2} 2$$

real constraints. The number of independent real degrees of freedom is therefore

$$2^n + 2(2^{2n} - 2^n) - 1 - \left(2^{2n} - 2^n\right) = 2^{2n} - 1.$$

For n large this is approximately 2^{2n} and for $n = 100$ this is approximately 10^{87} which is larger than the estimated number of atoms in the known universe! This simple estimate shows that calculating the full evolution of a system containing less than 100 particles is in general not possible on a classical computer. One solution is to think of simulating the unknown evolution by observing the dynamics of a controllable, analogous quantum system, a quantum simulator. This idea was proposed by Feynman in 1982[13]. The dynamics can also be studied by programming a quantum circuit machine containing n qubits to simulate the effect of the Hamiltonian $\hat{\mathcal{H}}$[14].

As of early 2018, when this material is being written, there is a great deal of excitement in the research community about the development of quantum computers. It is also the case that development is no longer the sole province of university and government research labs, but is being vigorously pursued in commercial companies who hope to reap large economic benefits from building and selling quantum computers. Although quantum computing research has been ongoing for about 30 years the shift to commercial activities is relatively new. Why the interest at this time? The remarkable fact in 2018 is that we are now on the

verge of having quantum hardware available that exceeds the capability of classical machines and can therefore demonstrate a quantum advantage.

To see this recall that the storage capacity needed to represent an arbitrary n qubit state (1.1) on a classical computer grows exponentially. The required storage as a function of n, making the conservative estimate of only requiring one byte per expansion coefficient, is shown in Fig. 1.2. At about 60 qubits the storage requirement exceeds the total worldwide computer memory presently available. Also shown in the figure are some recent calculations that evade the storage requirement at the cost of an exponential increase in computational time. Since quantum hardware is now emerging at the scale of 50 qubits, it is a good bet that we will see demonstrations of a quantum computational advantage in the near future. This will usher in a new era in humankinds ability to compute and understand nature. The rest of this course aims to provide a quantitative understanding of how a quantum computer works.

1.1 Qubits and Gates

Our building blocks for creating quantum circuits are qubits and gates.

1.1.1 Qubits

Any two-state quantum object can represent a qubit. In general the state of a qubit is

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

where a, b are complex numbers. There is one constraint, since $|a|^2+|b|^2 = 1$, and the common phase of a and b has no physical consequences. This leaves us with two real numbers which we can parameterize in terms of angles on the Bloch sphere

$$|\psi\rangle \;=\; \cos(\theta/2)|0\rangle + \sin(\theta/2)e^{\imath\phi}|1\rangle. \tag{1.2}$$

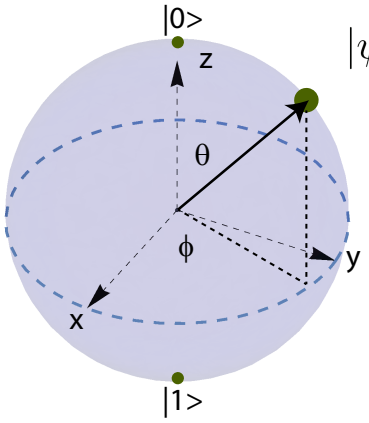
Here $0 \le \theta \le \pi$ and $0 \le \phi \le 2\pi$. Note that $|0\rangle$ (not $|1\rangle$) is at the north pole of the Bloch sphere. With the convention $|0\rangle = \begin{pmatrix}1\\0\end{pmatrix}$, $|1\rangle = \begin{pmatrix}0\\1\end{pmatrix}$ we can write

$$|\psi\rangle = \begin{pmatrix}\cos(\theta/2)\\ \sin(\theta/2)e^{\imath\phi}\end{pmatrix}.$$

This is shown graphically in Fig. 1.3.

The states $|0\rangle, |1\rangle$ are referred to as the computational basis for one qubit. Several different notations can be used for the basis states. We have

$$\begin{aligned} |0\rangle &= |0\rangle_z = |\mathbf{z}\rangle = |\uparrow\rangle = |1/2\rangle \\ |1\rangle &= |1\rangle_z = |-\mathbf{z}\rangle = |\downarrow\rangle = |-1/2\rangle. \end{aligned}$$



$$|\psi\rangle = \cos(\theta/2)|0\rangle + \sin(\theta/2)e^{\imath\phi}|1\rangle$$

Figure 1.3: Bloch sphere representation of a qubit.

We will use the convention that $|0\rangle = |0\rangle_z, |1\rangle = |1\rangle_z$ with the subscript z understood. The basis states in other directions are found from (1.2) as follows[1]

$$\begin{aligned}
\mathbf{x}: &\quad \theta = \pi/2, \phi = 0,\ |\mathbf{x}\rangle = |0\rangle_x = \frac{|0\rangle + |1\rangle}{\sqrt{2}},\\
-\mathbf{x}: &\quad \theta = \pi/2, \phi = \pi,\ |-\mathbf{x}\rangle = |1\rangle_x = \frac{|0\rangle - |1\rangle}{\sqrt{2}},\\
\mathbf{y}: &\quad \theta = \pi/2, \phi = \pi/2,\ |\mathbf{y}\rangle = |0\rangle_y = \frac{|0\rangle + i|1\rangle}{\sqrt{2}},\\
-\mathbf{y}: &\quad \theta = \pi/2, \phi = 3\pi/2,\ |-\mathbf{y}\rangle = |1\rangle_y = \frac{|0\rangle - i|1\rangle}{\sqrt{2}}.
\end{aligned}$$

Note that basis states along orthogonal axes are not orthogonal states, i.e. $\langle\mathbf{y}|\mathbf{x}\rangle \neq 0$ whereas antipodal states are orthogonal since

$$\begin{aligned}
\langle -\boldsymbol{\psi}|\boldsymbol{\psi}\rangle &= \left[\cos(\pi/2-\theta/2)\langle 0| + \sin(\pi/2-\theta/2)e^{-\imath(\pi+\phi)}\langle 1|\right]\left[\cos(\theta/2)|0\rangle + \sin(\theta/2)e^{\imath\phi}|1\rangle\right]\\
&= \cos(\pi/2-\theta/2)\cos(\theta/2) + \sin(\pi/2-\theta/2)e^{-\imath(\pi+\phi)}\sin(\theta/2)e^{\imath\phi}\\
&= 0.
\end{aligned}$$

The computational basis for two qubits is $|00\rangle, |01\rangle, |10\rangle, |11\rangle$. This has an obvious generalization to n qubits, for which there are 2^n basis states. For two qubits an alternative basis is the Bell basis

$$|\beta_{xy}\rangle = \frac{|0, y\rangle + (-1)^x|1, 1\oplus y\rangle}{\sqrt{2}}, \quad x, y = 0, 1$$

[1]It is important to remember that the state $|-\mathbf{x}\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}}$, whereas $-|\mathbf{x}\rangle = -\frac{|0\rangle+|1\rangle}{\sqrt{2}}$ are different quantum states.

where $\oplus$ denotes addition modulo 2. Explicitly

$$|\beta_{00}\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}, \quad |\beta_{01}\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}},$$
$$|\beta_{10}\rangle = \frac{|00\rangle - |11\rangle}{\sqrt{2}}, \quad |\beta_{11}\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}.$$

1.1.2 One-qubit Gates

Gates are used to change the state of a qubit. It is useful to know about some standard sets of gates. A basis for single qubit gates is provided by the Pauli group $\{I, \sigma_x, \sigma_y, \sigma_z\}$:

$$I \equiv \sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \ \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \ \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.3}$$

The operators have been expressed in the basis $\left\{|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right\}$. They satisfy $\sigma_j^2 = -i\sigma_x\sigma_y\sigma_z = I$ and the commutation relations

$$[\sigma_j, \sigma_k] = 2i \sum_\ell \epsilon_{jk\ell}\sigma_\ell, \quad \{\sigma_j, \sigma_k\} = 2\delta_{jk} I.$$

Any single qubit operator can be expressed as a linear combination of Pauli group operators. That is to say that an arbitrary operator $\mathbf{O}$ can be written as

$$\mathbf{O} = c_0 I + c_x\sigma_x + c_y\sigma_y + c_z\sigma_z$$

with c_0, c_x, c_y, c_z c numbers. This decomposition will be central to methods of tomographic reconstruction as discussed in chapters 10,11.

Note that

$$\sigma_z|0\rangle = +|0\rangle, \quad \sigma_z|1\rangle = -|1\rangle.$$

We see that the state $|0\rangle$ is what we normally think of as spin up (positive eigenvalue of σ_z) while $|1\rangle$ corresponds to spin down. This is consistent with the geometrical picture of $|0\rangle$ at the top of the Bloch sphere.

It is common in the quantum information literature to refer to $\sigma_x, \sigma_y, \sigma_z$ as X, Y, Z. The action of the Pauli operators on the basis states is

$$X|0\rangle = |1\rangle, \quad Y|0\rangle = i|1\rangle, \quad Z|0\rangle = |0\rangle$$
$$X|1\rangle = |0\rangle, \quad Y|1\rangle = -i|0\rangle, \quad Z|1\rangle = -|1\rangle.$$

Using the identity

$$e^{\imath\theta\mathbf{A}} = \cos(\theta)\mathbf{I} + i\sin(\theta)\mathbf{A}, \tag{1.4}$$

which holds provided $\mathbf{A}\cdot\mathbf{A} = \mathbf{I}$, we can express the rotation operators about the x, y, z axes as

$$R_x(\theta) = e^{-\imath\theta\sigma_x/2} = \cos\frac{\theta}{2}I - i\sin\frac{\theta}{2}\sigma_x = \begin{pmatrix} \cos\theta/2 & -i\sin\theta/2 \\ -i\sin\theta/2 & \cos\theta/2 \end{pmatrix},$$
$$R_y(\theta) = e^{-\imath\theta\sigma_y/2} = \cos\frac{\theta}{2}I - i\sin\frac{\theta}{2}\sigma_y = \begin{pmatrix} \cos\theta/2 & -\sin\theta/2 \\ \sin\theta/2 & \cos\theta/2 \end{pmatrix},$$
$$R_z(\theta) = e^{-\imath\theta\sigma_z/2} = \cos\frac{\theta}{2}I - i\sin\frac{\theta}{2}\sigma_z = \begin{pmatrix} e^{-\imath\theta/2} & 0 \\ 0 & e^{\imath\theta/2} \end{pmatrix} = e^{-\imath\theta/2}\begin{pmatrix} 1 & 0 \\ 0 & e^{\imath\theta} \end{pmatrix}.$$

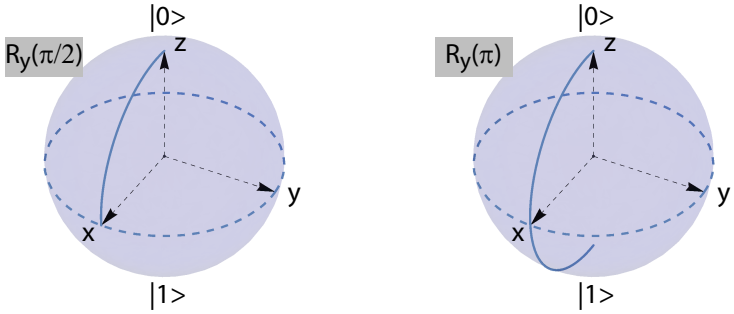


Figure 1.4: Rotations about the y axis on the Bloch sphere starting in state $|0\rangle$.

Thus

$$X = \sigma_x = iR_x(\pi), \quad Y = \sigma_y = iR_y(\pi), \quad Z = \sigma_z = iR_z(\pi).$$

Using (1.4) we can define a rotation about an arbitrary direction $\mathbf{a}$ as

$$R_{\mathbf{a}}(\theta) = e^{-\imath\theta\mathbf{a}\cdot\boldsymbol{\sigma}/2} = \cos(\theta/2)\mathbf{I} - i\sin(\theta/2)\mathbf{a}\cdot\boldsymbol{\sigma}.$$

An example of some basic rotations is shown in Fig. 1.4.

The X, Y, Z gates can be transformed into each other using $\pi/2$ rotations since

$$\begin{aligned}
X &= R_y^\dagger(-\pi/2)ZR_y(-\pi/2) = R_z^\dagger(\pi/2)YR_z(\pi/2) = R_z(-\pi/2)YR_z(\pi/2),\\
Y &= R_z^\dagger(-\pi/2)XR_z(-\pi/2) = R_x^\dagger(\pi/2)ZR_x(\pi/2) = R_x(-\pi/2)ZR_x(\pi/2),\\
Z &= R_x^\dagger(-\pi/2)YR_x(-\pi/2) = R_y^\dagger(\pi/2)XR_y(\pi/2) = R_y(-\pi/2)XR_y(\pi/2).
\end{aligned}$$

It is also useful to express the X, Y, Z operators as products of the other two using $\sigma_j = -i\epsilon_{jkl}\sigma_k\sigma_l$ which yields the explicit relations

$$\begin{aligned}
X &= -iYZ = iZY,\\
Y &= iXZ = -iZX,\\
Z &= -iXY = iYX.
\end{aligned}$$

In some cases we can approximate the action of a Y gate by an X gate preceded by a $\pi/2$ rotation about Z. Such a gate is

$$Y' = XR_z(-\pi/2) = \begin{pmatrix} 0 & e^{-\imath\pi/4} \\ e^{\imath\pi/4} & 0 \end{pmatrix} = e^{-\imath\pi/4}\begin{pmatrix} 0 & 1 \\ i & 0 \end{pmatrix}.$$

The action of this gate on a generic state $|\psi\rangle = a|0\rangle + b|1\rangle$ is

$$|\psi'\rangle = Y'|\psi\rangle = e^{-\imath\pi/4}\left(b|0\rangle + ia|1\rangle\right)$$

whereas

$$Y|\psi\rangle = -ib|0\rangle + ia|1\rangle.$$

The probability of measuring $|0\rangle$ or $|1\rangle$ is the same for these two states, but they have different phases.

Some often used gates are the Hadamard gate

$$H = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix} = \frac{\sigma_x + \sigma_z}{\sqrt{2}},$$

the phase gate

$$S = e^{\imath\pi/4}R_z(\pi/2) = \begin{pmatrix}1 & 0\\ 0 & e^{\imath\pi/2}\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & i\end{pmatrix} = \frac{(1+i)I + (1-i)\sigma_z}{2},$$

and the $\pi/8$ gate

$$T = e^{\imath\pi/8}R_z(\pi/4) = \begin{pmatrix}1 & 0\\ 0 & e^{\imath\pi/4}\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & \frac{1+i}{\sqrt{2}}\end{pmatrix} = \frac{(1+\sqrt{2}+i)I + (2^{3/2}-\sqrt{2}-1-i)\sigma_z}{2^{3/2}}.$$

Note the relations $S = T^2$, $Z = S^2 = T^4$.

The action of the one-qubit gates on the computational basis states is

$$\begin{array}{lll}
X|0\rangle = |1\rangle & X|1\rangle = |0\rangle & X^2 = I\\
Y|0\rangle = i|1\rangle & Y|1\rangle = -i|0\rangle & Y^2 = I\\
Z|0\rangle = |0\rangle & Z|1\rangle = -|1\rangle & Z^2 = I\\
H|0\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}} & H|1\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}} & H^2 = I\\
S|0\rangle = |0\rangle & S|1\rangle = i|1\rangle & S^2 = Z\\
T|0\rangle = |0\rangle & T|1\rangle = e^{\imath\pi/4}|1\rangle & T^2 = S
\end{array}$$

Note that the Hadamard gate serves to switch between the z and x bases since

$$\begin{array}{ll}
H|0\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}} = |0\rangle_x, & H|1\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}} = |1\rangle_x\\
H|0\rangle_x = |0\rangle, & H|1\rangle_x = |1\rangle
\end{array}$$

An arbitrary 1-qubit gate U can be decomposed into rotations as

$$U = e^{\imath\alpha}R_z(\beta)R_y(\gamma)R_z(\delta)$$

where $\alpha, \beta, \gamma, \delta$ are real numbers. The operator ordering is read from right to left, so

$$\begin{aligned}
U|\psi\rangle &= e^{\imath\alpha}R_z(\beta)R_y(\gamma)[R_z(\delta)|\psi\rangle]\\
&= e^{\imath\alpha}R_z(\beta)[R_y(\gamma)[R_z(\delta)|\psi\rangle]]\\
&= e^{\imath\alpha}[R_z(\beta)[R_y(\gamma)[R_z(\delta)|\psi\rangle]]].
\end{aligned}$$

There is nothing special about the y, z axes. Any one-qubit operator can be decomposed into rotations about any two non-parallel axes on the Bloch sphere.

An additional useful result is

$$U = e^{\imath\alpha}AXBXC \tag{1.5}$$

where A, B, C are unitary and satisfy $ABC = I$ and $X = \sigma_x$.

One-qubit gates can be implemented experimentally using Rabi oscillations. Suppose the qubit levels are coupled by a resonant driving field with a Rabi frequency Ω. The unitary transformation of a qubit after time t is

$$U(t,\phi) = \begin{pmatrix}\cos(|\Omega|t/2) & ie^{-\imath\phi}\sin(|\Omega|t/2)\\ ie^{\imath\phi}\sin(|\Omega|t/2) & \cos(|\Omega|t/2)\end{pmatrix}$$

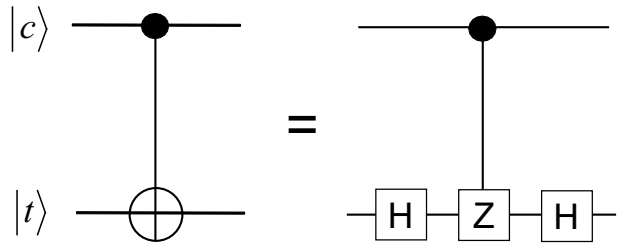


Figure 1.5: CNOT gate implemented as $CNOT_{ct} = H_t C_{Z,ct} H_t$.

where $\Omega = |\Omega| e^{\imath\phi}$. Setting $t = \theta/|\Omega|$ gives

$$U(\theta, \phi) = \begin{pmatrix} \cos(\theta/2) & ie^{-\imath\phi}\sin(\theta/2) \\ ie^{\imath\phi}\sin(\theta/2) & \cos(\theta/2) \end{pmatrix}.$$

Setting $\phi = 0$ then gives $U(\theta, 0) = R_x(-\theta)$ and setting $\phi = \pi/2$ gives $U(\theta, \pi/2) = R_y(-\theta)$. Any single qubit gate can then be composed from combinations of x and y rotations as described above.

1.1.3 Two-qubit Gates

The standard two-qubit gate is the controlled-not

$$CNOT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

The operation of the $CNOT_{xy}$ gate with x the control bit and y the target bit can be expressed as $|x\rangle|y\rangle \rightarrow |x\rangle|y \oplus x\rangle$. The gate can also be written as

$$CNOT = C_X = (P_0 \otimes I) + (P_1 \otimes X)$$

where $P_0 = |0\rangle\langle 0|, P_1 = |1\rangle\langle 1|$ are projectors onto the computational basis. The CNOT gate can be implemented with a controlled phase gate

$$C_Z = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \tag{1.6}$$

and Hadamard gates as

$$CNOT = (I \otimes H) C_Z (I \otimes H).$$

The circuit diagram for this construction is shown in Fig. 1.5. It follows from (1.6) that $C_{Z,xy} = C_{Z,yx}$ so this gate is often drawn in a symmetric fashion as shown in Fig. 1.6.

We can implement a C_Y gate with a similar construction. Using $Y = R_x(-\pi/2) Z R_x(\pi/2)$ we find

$$\begin{aligned} C_Y &= (P_0 \otimes I) + (P_1 \otimes Y) \\ &= (I \otimes R_x(-\pi/2)) C_Z (I \otimes R_x(\pi/2)) \\ &= (I \otimes H)[(I \otimes R_x(\pi/2)) C_Z (I \otimes R_x(-\pi/2))](I \otimes H). \end{aligned}$$

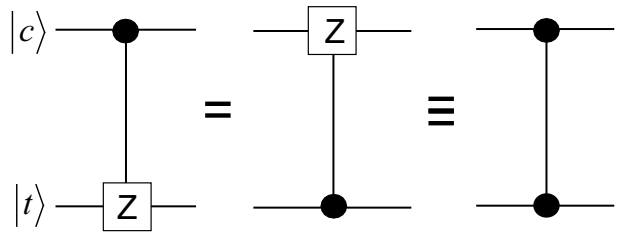


Figure 1.6: The C_Z gate is symmetric with respect to control and target qubits.

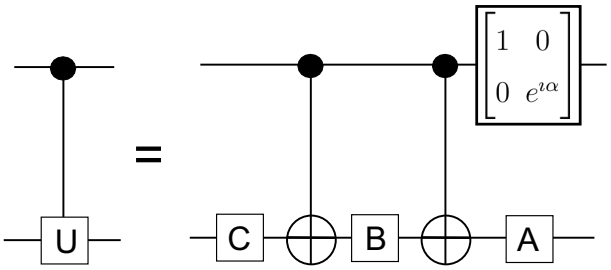


Figure 1.7: C_U gate construction.

More generally we may wish to implement the controlled U gate shown in Fig. 1.7 with U some arbitrary operator. This can be done using the result (1.5) $U = e^{\imath\alpha}AXBXC$.

1.1.4 Phase kickback

Consider the effect of a CNOT gate due to a control qubit $|c\rangle = c_0|0\rangle + c_1|1\rangle$ acting on a target qubit $|t\rangle = t_0|0\rangle + t_1|1\rangle$. Let us rewrite the target qubit state using

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

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

so

$$|t\rangle = \frac{t_0 + t_1}{\sqrt{2}}\frac{|0\rangle + |1\rangle}{\sqrt{2}} + \frac{t_0 - t_1}{\sqrt{2}}\frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

Recall that

$$CNOT: \ |x\rangle\frac{|0\rangle + |1\rangle}{\sqrt{2}} = |x\rangle\frac{|0\rangle + |1\rangle}{\sqrt{2}}$$

but

$$CNOT: \ |x\rangle\frac{|0\rangle - |1\rangle}{\sqrt{2}} = -1^x|x\rangle\frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

to obtain

$$CNOT: \ c_1|1\rangle|t\rangle \ = \ c_1|1\rangle\left(\frac{t_0 + t_1}{\sqrt{2}}\frac{|0\rangle + |1\rangle}{\sqrt{2}} - \frac{t_0 - t_1}{\sqrt{2}}\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right).$$

Therefore

$$\begin{aligned} CNOT:\ |c\rangle|t\rangle &= c_0|0\rangle\left(\frac{t_0+t_1}{\sqrt{2}}\frac{|0\rangle+|1\rangle}{\sqrt{2}}+\frac{t_0-t_1}{\sqrt{2}}\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right) \\ &+ c_1|1\rangle\left(\frac{t_0+t_1}{\sqrt{2}}\frac{|0\rangle+|1\rangle}{\sqrt{2}}-\frac{t_0-t_1}{\sqrt{2}}\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right). \end{aligned} \tag{1.7}$$

This is of course just a complicated way of writing

$$CNOT:\ |c\rangle|t\rangle = c_0t_0|00\rangle + c_0t_1|01\rangle + c_1t_1|10\rangle + c_1t_0|11\rangle.$$

Nevertheless (1.7) shows that if $t_0 = 1/\sqrt{2} = -t_1$ so $|t\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}}$, which is an eigenvector of X with eigenvalue -1, we get

$$\begin{aligned} CNOT:\ |c\rangle|t\rangle &= (c_0|0\rangle - c_1|1\rangle)\frac{|0\rangle-|1\rangle}{\sqrt{2}} \\ &= (c_0|0\rangle - c_1|1\rangle))|t\rangle \\ &= (Z\otimes I)|c\rangle|t\rangle. \end{aligned} \tag{1.8}$$

The $CNOT = C_X$ gate leaves the control qubit unaffected. Nevertheless, when the target qubit is an eigenstate of X with eigenvalue -1 the action of the gate can be thought of as imprinting a Z phase shift on the control qubit. This behavior plays a central role in the phase estimation algorithm described in Sec. 3.3 and is frequently used in the construction of error correcting codes.

1.2 Subroutines

There are a few circuit constructions which appear repeatedly in algorithms. Starting with an n qubit register in the fiducial state $|0\rangle^{\otimes n}$ we can create a uniform superposition of all $|\mathbf{z}\rangle$ where $\mathbf{z}$ is an $n-$bit binary string using n one-qubit Hadamard gates

$$H^{\otimes n}|0\rangle^{\otimes n} = \frac{1}{2^{n/2}}\sum_{\mathbf{z}\in\{0,1\}^n}|\mathbf{z}\rangle.$$

More generally

$$H^{\otimes n}|\mathbf{x}\rangle = \frac{1}{2^{n/2}}\sum_{\mathbf{z}\in\{0,1\}^n}(-1)^{\mathbf{x}\cdot\mathbf{z}}|\mathbf{z}\rangle.$$

Here $\mathbf{x}\cdot\mathbf{z}$ is the bitwise inner product modulo 2. These relations follow easily from the definition of the Hadamard gate.

The quantum Fourier transform at order $N = 2^n$ is

$$QFT_N|x\rangle = \frac{1}{\sqrt{N}}\sum_{y=0}^{N-1}e^{\imath 2\pi\frac{xy}{N}}|y\rangle$$

and the inverse transform is

$$QFT_N^{-1}|x\rangle = \frac{1}{\sqrt{N}}\sum_{y=0}^{N-1}e^{-\imath 2\pi\frac{xy}{N}}|y\rangle.$$

Here x, y are integers ranging from 0 to $N-1$ and $|x\rangle, |y\rangle$ are shorthand for $|\mathbf{x}\rangle, |\mathbf{y}\rangle$ which are $n-$qubit encodings of the corresponding binary strings $\mathbf{x}, \mathbf{y}$.

Note that for $N = 2$ the QFT operations are

$$\begin{aligned} QFT_2|x\rangle &= \frac{1}{\sqrt{2}}\sum_{y=0}^{1} e^{\imath\pi xy}|y\rangle = \frac{|0\rangle + (-1)^x|1\rangle}{\sqrt{2}} = H|x\rangle \\ QFT_2^{-1}|x\rangle &= \frac{1}{\sqrt{2}}\sum_{y=0}^{1} e^{-\imath\pi xy}|y\rangle = \frac{|0\rangle + (-1)^x|1\rangle}{\sqrt{2}} = H|x\rangle. \end{aligned}$$

The QFT acting on a single qubit is a Hadamard gate.

The QFT is an essential ingredient of several quantum algorithms, including the factoring algorithm. How do we implement the QFT in a circuit and what are the resource requirements? The QFT of a n qubit state $|x\rangle$ is

$$QFT|x\rangle = \frac{1}{\sqrt{2^n}}\sum_{y=0}^{2^n-1} e^{\imath 2\pi\frac{xy}{2^n}}|y\rangle$$

In this expression y is shorthand for the n binary digits $y_1y_2...y_n$ which represent the number $y = \sum_{j=1}^{n} y_j 2^{n-j} = y_1 2^{n-1} + y_2 2^{n-2} + ... + y_n$ and $|y\rangle = |\mathbf{y}\rangle = |y_1y_2...y_n\rangle$. We can therefore write the QFT as

$$QFT|x\rangle = \frac{1}{\sqrt{2^n}}\sum_{y_1y_2...y_n\in\{0,1\}} e^{\frac{\imath 2\pi x}{2^n}\sum_{k=1}^{n} 2^{n-k}y_k}|y_1y_2...y_n\rangle$$

The exponential of a sum can be written as a product of exponentials so

$$\begin{aligned} QFT|x\rangle &= \frac{1}{\sqrt{2^n}}\sum_{y_1,y_2...y_n\in\{0,1\}}\bigotimes_{k=1}^{n} e^{\frac{\imath 2\pi x}{2^n}2^{n-k}y_k}|y_k\rangle \\ &= \frac{1}{\sqrt{2^n}}\bigotimes_{k=1}^{n}\sum_{y_k\in\{0,1\}} e^{\frac{\imath 2\pi x}{2^n}2^{n-k}y_k}|y_k\rangle \\ &= \frac{1}{\sqrt{2^n}}\bigotimes_{k=1}^{n}\left(|0\rangle_k + e^{\frac{\imath 2\pi x}{2^n}2^{n-k}}|1\rangle_k\right). \end{aligned} \tag{1.9}$$

We have succeeded in expressing the result of the QFT as a product state of the n qubits. We can simplify further using

$$\begin{aligned} e^{\frac{\imath 2\pi x}{2^n}2^{n-k}} &= e^{\frac{\imath 2\pi\sum_{j=1}^{n} 2^{n-j}x_j}{2^n}2^{n-k}} \\ &= e^{\imath 2\pi\sum_{j=1}^{n} 2^{n-j-k}x_j} \\ &= e^{\imath 2\pi\left(2^{n-1-k}x_1 + 2^{n-2-k}x_2 + ...2^{-k}x_n\right)}. \end{aligned}$$

Then note that for $k = 1$ the right hand side of the above expression is equal to

$$e^{\imath 2\pi 2^{-1}x_n} = (-1)^{x_n},$$

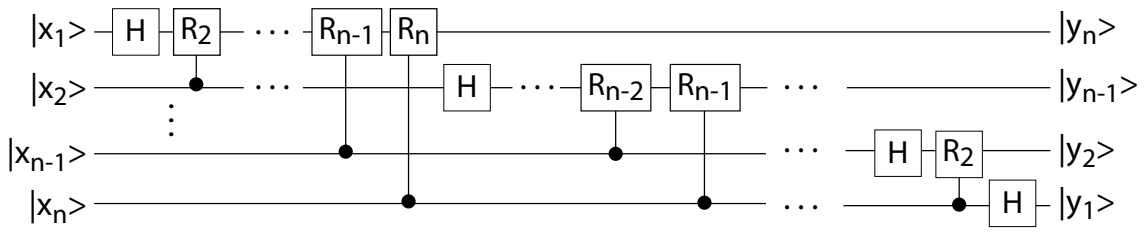


Figure 1.8: QFT circuit.

for $k = 2$

$$e^{\imath 2\pi\left(2^{-1}x_n + 2^{-2}x_{n-1}\right)} = (-1)^{x_n} e^{\imath 2\pi 2^{-2}x_{n-1}},$$

for $k = 3$

$$e^{\imath 2\pi\left(2^{-1}x_n + 2^{-2}x_{n-1} + 2^{-3}x_{n-2}\right)} = (-1)^{x_n} e^{\imath 2\pi\left(2^{-2}x_{n-1} + 2^{-3}x_{n-2}\right)},$$

and for $k = n$

$$e^{\imath 2\pi \sum_{j=1}^{n} 2^{-j}x_j}.$$

Thus the first qubit of the output only depends on the last qubit of the input, the second qubit only depends on the last two input qubits, and so on, until the last qubit of the output depends on all the input qubits. Furthermore the first qubit of the output is in the state

$$|0\rangle + (-1)^{x_n}|1\rangle = H|x_n\rangle$$

which is just a Hadamard acting on the last qubit of the input. The phase factors acting on the other output qubits can be written as the effect of conditional rotation operators

$$R_k = \begin{pmatrix} 1 & 0 \\ 0 & e^{\imath 2\pi/2^k} \end{pmatrix} \tag{1.10}$$

which are controlled by the appropriate input qubits. Combining these elements gives the circuit in Fig. 1.8. The output register is in inverted order. This can be corrected with swap gates or the qubits can simply be relabeled. The gate count needed for the full QFT on n qubits is n Hadamard gates and $\sum_{k=1}^{n} k - 1 = \frac{n(n-1)}{2}$ controlled rotation gates. The resource scaling is thus $O(n^2)$ which is polynomial in the number of qubits. This can be compared with the classical Fast Fourier Transform (FFT) algorithm which requires $O(N \log N) = O(n2^n)$ operations[19]. Although it might appear that we could use the QFT as a more efficient replacement for the FFT this is not the case. For starters there is no efficient method of preparing a general superposition state with desired amplitudes. Even if we could, the result of the QFT is encoded in the amplitudes of the output state and we have no way of extracting those amplitudes. Making a measurement, or many measurements, only tells us about the relative probabilities of different state components. Nevertheless the exponential speedup of quantum algorithms is in many cases a consequence of the efficiency of the QFT, as we will see.

The controlled rotations can be implemented with the construction shown in Fig. 1.7. This general construction requires two CNOT gates and four single qubit gates. For the particular case of a controlled rotation operation the general approach can be simplified to a single CNOT and three single qubit rotations[20]. The QFT is of interest because it is an

integral part of quantum algorithms demonstrating an exponential speedup such as the phase estimation algorithm and the factoring algorithm. It turns out that for these applications only the probabilities, not the amplitudes are needed after the QFT is applied. This allows the controlled rotation gates to be replaced by rotation operations that are executed, or not, depending on the result of a measurement[21]. This "semi-classical" QFT eliminates the need for CNOT gates in implementing the rotations and considerably simplifies the circuit implementation[22].

1.3 Universal gate sets

It was realized in the early days of quantum circuit design that a finite set of one- and two-qubit gates is sufficient for universal qauntum computation[23, 24, 25]. One set of universal gates is $\{H, S, T, CNOT\}$. This is referred to by Nielsen & Chuang as the standard set. The T gate is special. It is needed for universal quantum computation, but it cannot be implemented transversally, which complicates the construction of fault-tolerant error correcting codes. The T gate can be dispensed with if we include the three qubit Toffoli gate (a CNOT type gate with two control qubits) giving the set $\{H, S, CNOT, \text{Toffoli}\}$. As we will see later on it is not possible to implement a universal gate set transversally[26, 27].

Even if we have a universal gate set available it is not generally obvious how to construct a quantum circuit to perform a desired mathematical transformation. One approach which may not be optimal, but is useful, is to first find or devise a classical circuit with Boolean logic elements that implements the desired function. Classical Boolean logic that uses, for example, AND gates is not reversible. However, all Boolean logic circuits can be implemented with reversible hardware using three terminal gates (Toffoli or Fredkin gates). The reversible classical circuit can then be implemented with the corresponding quantum gates. An example of this type of design methodology appears in the construction of circuits for modular exponentiation which is used in the factoring algorithm[28].

1.4 Approximating gates

We can in principle implement arbitrary gates by combining rotations to generate any 1-qubit gate and using the C_U construction to generate any 2-qubit gate. However, as we will learn later on, only a limited set of gates can be readily implemented in a way which is compatible with fault tolerant error correction. That limited set could be the standard set of the Clifford group plus the T gate. This seems problematic as for some algorithms we need more general gates. For example $\pi/2^k$ rotations appear in Shor's algorithm.

The Solovay-Kitaev theorem[29] shows that we can efficiently approximate any desired gate by concatenating a finite number of gates from a universal set. The theorem says that we can always approximate a desired gate with error at most ϵ using a finite sequence of gates of length

$$\mathcal{O}(\log^c(1/\epsilon))$$

where c is a positive constant. Thus to approximate a circuit containing m gates to accuracy ϵ requires each gate to have error $\mathcal{O}(\epsilon/m)$ so we need $\mathcal{O}(m \log^c(m/\epsilon))$ gates which is only a "polylogarithmic" increase in circuit size.

index	x axis	y axis	z axis	U	index	x axis	y axis	z axis	U
1	I	I	I	$\begin{pmatrix}1&0\\0&1\end{pmatrix}$	2	I	I	$\pi/2$	$e^{-\imath\pi/4}\begin{pmatrix}1&0\\0&i\end{pmatrix}$
3	I	I	π	$-i\begin{pmatrix}1&0\\0&-1\end{pmatrix}$	4	I	I	$-\pi/2$	$e^{\imath\pi/4}\begin{pmatrix}1&0\\0&-i\end{pmatrix}$
5	I	π	I	$-1\begin{pmatrix}0&1\\-1&0\end{pmatrix}$	6	I	π	$\pi/2$	$-e^{\imath\pi/4}\begin{pmatrix}0&1\\i&0\end{pmatrix}$
7	π	I	I	$-i\begin{pmatrix}0&1\\1&0\end{pmatrix}$	8	π	I	$\pi/2$	$e^{-\imath\pi/4}\begin{pmatrix}0&1\\-i&0\end{pmatrix}$
9	π	$\pi/2$	I	$\frac{-i}{\sqrt{2}}\begin{pmatrix}1&1\\1&-1\end{pmatrix}$	10	I	$-\pi/2$	I	$\frac{1}{\sqrt{2}}\begin{pmatrix}1&1\\-1&1\end{pmatrix}$
11	$\pi/2$	I	$\pi/2$	$\frac{e^{-\imath\pi/4}}{\sqrt{2}}\begin{pmatrix}1&1\\-i&i\end{pmatrix}$	12	$\pi/2$	π	$\pi/2$	$-\frac{e^{\imath\pi/4}}{\sqrt{2}}\begin{pmatrix}1&1\\i&-i\end{pmatrix}$
13	π	$-\pi/2$	I	$\frac{i}{\sqrt{2}}\begin{pmatrix}1&-1\\-1&-1\end{pmatrix}$	14	$-\pi/2$	I	$\pi/2$	$\frac{e^{-\imath\pi/4}}{\sqrt{2}}\begin{pmatrix}1&-1\\i&i\end{pmatrix}$
15	I	$\pi/2$	I	$\frac{1}{\sqrt{2}}\begin{pmatrix}1&-1\\1&1\end{pmatrix}$	16	$-\pi/2$	π	$\pi/2$	$\frac{e^{\imath\pi/4}}{\sqrt{2}}\begin{pmatrix}1&-1\\-i&-i\end{pmatrix}$
17	$-\pi/2$	$-\pi/2$	I	$\frac{e^{-\imath\pi/4}}{\sqrt{2}}\begin{pmatrix}1&i\\-1&i\end{pmatrix}$	18	$-\pi/2$	$\pi/2$	I	$\frac{e^{\imath\pi/4}}{\sqrt{2}}\begin{pmatrix}1&i\\1&-i\end{pmatrix}$
19	$-\pi/2$	π	I	$\frac{i}{\sqrt{2}}\begin{pmatrix}1&i\\-i&-1\end{pmatrix}$	20	$-\pi/2$	I	I	$\frac{1}{\sqrt{2}}\begin{pmatrix}1&i\\i&1\end{pmatrix}$
21	$\pi/2$	$-\pi/2$	I	$\frac{e^{\imath\pi/4}}{\sqrt{2}}\begin{pmatrix}1&-i\\-1&-i\end{pmatrix}$	22	$\pi/2$	I	I	$\frac{1}{\sqrt{2}}\begin{pmatrix}1&-i\\-i&1\end{pmatrix}$
23	$\pi/2$	π	I	$\frac{-i}{\sqrt{2}}\begin{pmatrix}1&-i\\i&-1\end{pmatrix}$	24	$\pi/2$	$\pi/2$	I	$\frac{e^{-\imath\pi/4}}{\sqrt{2}}\begin{pmatrix}1&-i\\1&i\end{pmatrix}$

Table 1.1: Elements of the Clifford group $\mathcal{C}_1$ for a single qubit. The notation π_j is shorthand for $R_j(\pi) = -i\sigma_j$ and $\pm\pi_j/2$ is shorthand for $R_j(\pm\pi/2)$.The gates are applied along z, y, then x axes, i.e. reading right to left along each row. The operators have been grouped according to the top row of the matrix, and have been factored so that the first nonzero element in the top row is unity.

1.5 Clifford Group

According to the Gottesman-Knill theorem[30] the action of combinations of Clifford gates can be efficiently simulated (in polynomial time) on a classical computer. The power of quantum computers requires non Clifford gates.

The Clifford group consists of those operators which map elements of the Pauli group onto elements of the Pauli group. The Pauli group $\mathcal{P}_1$ for one qubit is the set of Pauli operators $\{I, X, Y, Z\}$ multiplied by ± 1 and $\pm i$, i.e.

$$\mathcal{P}_1 = \{\pm I, \pm iI, \pm X, \pm iX, \pm Y, \pm iY, \pm Z, \pm iZ\}.$$

The number of elements in the Pauli group acting on n qubits is $|\mathcal{P}_n| = 4^{n+1}$.

For one qubit the Clifford group $\mathcal{C}_1$ is generated by the operations $I, R_j(\pm\pi/2), R_j(\pi)$ about axes $j = x, y, z$. We include $R_j(\pm\pi/2)$ since these are distinct operators but do

not need $R_j(\pm\pi)$ since $R_j(\pi) = -R_j(-\pi)$. Forming all permutations of the four operators $I, R_j(\pi/2), R_j(-\pi/2), R_j(\pi)$ acting on axes $j = x, y, z$ gives $4^3 = 64$ possible operators. We eliminate operators that are the same up to a global phase leaving the 24 distinct Clifford group elements shown in Table 12.2.

This method can be extended to generate the 11520 elements of $\mathcal{C}_2$ acting on two qubits. For two qubits the Clifford group is generated by the set $\{H, S, CNOT\}$. The operators are $I_a, R_{aj}(\pm\pi/2), R_{aj}(\pi)$ on each of two qubits. In addition to the single qubit operators we allow for three additional two-qubit operators $CNOT_{12}, CNOT_{21}, CNOT_{12}CNOT_{21}$ where $CNOT_{ab}$ is a CNOT gate with qubit a the control and qubit b the target. This gives $4(4^3)^2 = 16384$ possible operators. A large number of these are the same up to a global phase. The number of distinct elements in the Clifford group acting on n qubits is [M. Ozols 2008]

$$|\mathcal{C}_n| = 2^{n^2+2n} \prod_{j=1}^{n} \left(4^j - 1\right),$$

which gives $|\mathcal{C}_2| = 11520$, $|\mathcal{C}_3| = 92\,897\,280$, and $|\mathcal{C}_4| = 12\,128\,668\,876\,800$.

1.6 Gottesman-Knill theorem

The Gottesman-Knill theorem[30] says:

Suppose a quantum computation is performed which involves only the following elements: state preparations in the computational basis, Clifford group gates, and measurements of observables in the Pauli group (which includes measurement in the computational basis as a special case). Such a computation may be efficiently simulated on a classical computer.

The proof of this theorem is most easily presented using the stabilizer formalism - but we have not discussed stabilizer operators yet. Aaronson and Gottesman have provided a computer program which can be used for such simulations[10].

The implication of this theorem is that the Clifford group is not sufficient for universal quantum computation that can efficiently solve classically hard problems. Nevertheless Clifford group operators are important for error correction, can be used to create entangled states, and are important for randomized benchmarking tests of gate fidelity.

1.7 No cloning theorem

Any desired quantum state $|\psi\rangle$ can in principle be prepared by starting with $|0\rangle$ and applying a unitary operator U such that $|\psi\rangle = U|0\rangle$. If we prepare n qubits in $|0\rangle$ and apply the same unitary to each one we will create n copies of the state $|\psi\rangle$. In other words we can in this way clone $|\psi\rangle$. This construction applies equally well to states of more than one qubit.

However, the above construction only works for a known $|\psi\rangle$. If we are provided with an unknown state $|\psi\rangle$ it is not possible to copy the state. The proof of this important result known as the no cloning theorem [31, 32] is a simple exercise in quantum mechanics. The physics behind it is the assumption that quantum mechanics is a linear theory. This assumption[33] can be put to experimental test by postulating nonlinear extensions to quan-

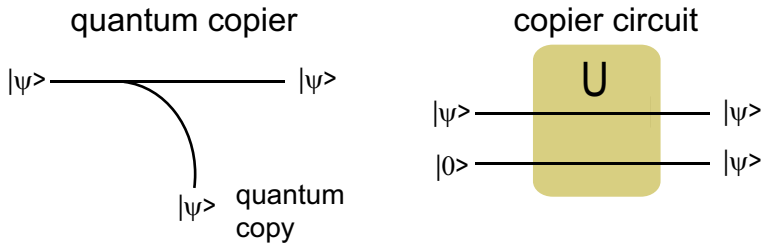


Figure 1.9: Quantum copying machine. The quantum copier shown on the left and its' circuit implementation on the right.

tum mechanics and looking for experimental evidence of nonlinearity. So far no experiment has shown any deviation from the assumed linearity of the quantum theory[34].

The proof of the theorem goes as follows. Suppose we have a quantum copying machine is shown in Fig. 1.9. Then

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

This must be true for any $|\psi\rangle$ so

$$U|\alpha\rangle|0\rangle = |\alpha\rangle|\alpha\rangle.$$

and

$$U|\beta\rangle|0\rangle = |\beta\rangle|\beta\rangle.$$

Since quantum mechanics is linear it must also be possible to copy $|\gamma\rangle = \frac{|\alpha\rangle+|\beta\rangle}{\sqrt{2}}$. However, this doesn't work since

$$\begin{aligned} U|\gamma\rangle|0\rangle &= \frac{U|\alpha\rangle|0\rangle}{\sqrt{2}} + \frac{U|\beta\rangle|0\rangle}{\sqrt{2}} \\ &= \frac{|\alpha\rangle|\alpha\rangle}{\sqrt{2}} + \frac{|\beta\rangle|\beta\rangle}{\sqrt{2}}. \end{aligned}$$

On the other hand

$$|\gamma\rangle|\gamma\rangle = \frac{|\alpha\rangle|\alpha\rangle + |\beta\rangle|\beta\rangle + 2|\alpha\rangle|\beta\rangle}{2}$$

and we see that $U|\gamma\rangle|0\rangle \neq |\gamma\rangle|\gamma\rangle$. We can teleport an unknown state but must erase the input when doing so. The impossibility of copying unknown quantum states lies behind the security of quantum key distribution, a topic we will discuss in more detail later on.

1.8 Qudits

We can generalize the two-level qubit to a $d-$ level quantum object called a qudit. For $d = 3$ these are referred to as qutrits. Clearly more information can be stored in a qudit than a qubit, but it is also more demanding to perform gate operations on qudits. Nevertheless there is evidence that for some quantum operations it may be more efficient to encode data in qudits. An example is the discovery that the fault tolerance threshold of a quantum error correcting code can be raised by using a qudit encoding.

Defining the $d-$dimensional qudit basis as $\{|0\rangle, |1\rangle, ...|d-1\rangle$ qudit generalizations of elementary X, Z gates are[35]

$$\begin{aligned} X|j\rangle &= |j+1(\text{mod d})\rangle, \\ Z|j\rangle &= (\omega)^j|j\rangle, \end{aligned}$$

with $\omega = e^{\imath 2\pi/d}$ the d^{th} root of unity. Also $XZ = \frac{1}{\omega}ZX$ so the commutator is $[X, Z] = XZ - ZX = (\frac{1}{\omega} - 1)ZX$ which provides a definition of Y gates. The Z gate is closely related to a discrete Fourier transform operation. The non-Clifford $\pi/8$ gate can be defined as a diagonal operator [36].

A two-qudit C_Z gate can be defined as

$$C_Z|j\rangle|k\rangle = \omega^{jk}|j\rangle|k\rangle.$$

Implementation requires phases ω^{jk} applied to all d^2 pairs of states. An alternative is the qudit generalization of the C_X gate which is called a SUM gate and is defined as[35]

$$SUM|j\rangle|k\rangle = |j\rangle|k+j\rangle$$

where the addition is understood to be modulo d.

Actual implementation of these multi-dimensional gates will be highly dependent on the chosen physical platform.

Chapter 2

Measurements

2.1 Projective measurements

Consider an observable $\hat{A}$ with eigenspectrum $|n\rangle$ satisfying $\hat{A}|n\rangle = a_n|n\rangle$. An arbitrary state $|\psi\rangle$ can be expanded as $|\psi\rangle = \sum_{n=0}^{\infty} c_n|n\rangle$. The result of a measurement of the observable A is

$$\langle\hat{A}\rangle = \langle\psi|\hat{A}|\psi\rangle = \sum_m c_m^*\langle m|\hat{A}\sum_n c_n|n\rangle = \sum_m\sum_n c_m^* c_n a_n\langle m|n\rangle = \sum_n |c_n|^2 a_n. \tag{2.1}$$

The probability of the measurement result being the eigenvalue a_n is

$$\Pi(a_n) = |c_n|^2 = |\langle n|\psi\rangle|^2 = \langle\psi|n\rangle\langle n|\psi\rangle = ||\hat{P}_n|\psi\rangle||^2 = \langle\psi|\hat{P}_n|\psi\rangle \tag{2.2}$$

where we have introduced the projector onto state $|n\rangle$, $\hat{P}_n = |n\rangle\langle n|$. After $\hat{A}$ has been measured giving result a_n the new state of the system is

$$|\psi'\rangle = \frac{\hat{P}_n|\psi\rangle}{||\hat{P}_n|\psi\rangle||} = \frac{c_n}{|c_n|}|n\rangle. \tag{2.3}$$

A subsequent measurement of $\langle\hat{A}\rangle$ will return a_n with unit probability. The theory of quantum measurements can be extended beyond projective measurements to a more general situation described by a positive operator valued measure (POVM). See [37] for details.

The uncertainty in the measurement of an operator depends on the state being measured. Recall the definition of the variance

$$\langle(\Delta\hat{A})^2\rangle = \langle\hat{A}^2\rangle - \langle\hat{A}\rangle^2.$$

Let us assume $|\psi\rangle = a|0\rangle + b|1\rangle$ describes a two-state particle and that $\hat{A}$ has eigenvalues $\hat{A}|0\rangle = \lambda_0|0\rangle, \hat{A}|1\rangle = \lambda_1|1\rangle$. The variance of $\hat{A}$ is

$$\begin{aligned}\langle(\Delta\hat{A})^2\rangle &= \lambda_0^2|a|^2 + \lambda_1^2|b|^2 - (\lambda_0|a|^2 + \lambda_1|b|^2)^2 \\ &= \lambda_0^2(|a|^2 - |a|^4) + \lambda_1^2(|b|^2 - |b|^4) - 2\lambda_0\lambda_1|a|^2|b|^2.\end{aligned}$$

If $a = 0$ or $b = 0$ we are in an eigenstate of $\hat{A}$ and $\langle(\Delta\hat{A})^2\rangle = 0$. If $a = b = 1/\sqrt{2}$ then $\langle(\Delta\hat{A})^2\rangle = \lambda_0^2/4 + \lambda_1^2/4 - 2\lambda_0\lambda_1/4$. This variance is often referred to as quantum projection

noise. Variation of the measurement variance as the state changes is a signature of a quantum limited measurement.

An important aspect of projective measurements is that the results of sequential measurements of non-commuting operators depend on the order of the measurements. Suppose we prepare the state $|\psi\rangle = |1\rangle_x = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$. A measurement of $\hat{X}$ followed by a measurement of $\hat{Z}$ will give the results $+1_x$ followed by 0_z or 1_z with 50% probability. Alternatively a measurement of $\hat{Z}$ followed by a measurement of $\hat{X}$ will give the results 0_z or 1_z with 50% probability followed by 0_x or 1_x with 50% probability. Clearly the sequence of measurement results may be different in the two cases.

2.2 Measurement in the z basis

Given a state $|\psi\rangle = c_0|0\rangle + c_1|1\rangle$ the probabilities of observing $|0\rangle$ or $|1\rangle$ are found from the expectation values of the projectors $\hat{P}_0 = |0\rangle\langle 0|$, $\hat{P}_1 = |1\rangle\langle 1|$ as

$$\Pi_{|0\rangle} = \langle\psi|\hat{P}_0|\psi\rangle = |\langle 0|\psi\rangle|^2 = |c_0|^2, \quad \Pi_{|1\rangle} = \langle\psi|\hat{P}_1|\psi\rangle = |\langle 1|\psi\rangle|^2 = |c_1|^2.$$

The expected value of the measurement result assigning $+1$ to $|0\rangle$ and -1 to $|1\rangle$ is

$$E = \Pi_{|0\rangle} - \Pi_{|1\rangle} = |c_0|^2 - |c_1|^2.$$

This can also be written as the expectation value of Z since

$$\langle\psi|Z|\psi\rangle = (c_0^*, c_1^*)\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}\begin{pmatrix}c_0\\ c_1\end{pmatrix} = |c_0|^2 - |c_1|^2 = E.$$

2.3 Measurement in the x basis

In many cases measurements in the z basis are directly implemented. Effective measurements in other bases are implemented by preceding a z basis measurement with an appropriate rotation operation.

For example suppose we wish to measure a qubit state in the $\pm x$ basis. Recall that

$$|0\rangle_x = \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \quad |1\rangle_x = \frac{|0\rangle - |1\rangle}{\sqrt{2}},$$

and

$$|0\rangle = \frac{|0\rangle_x + |1\rangle_x}{\sqrt{2}}, \quad |1\rangle = \frac{|0\rangle_x - |1\rangle_x}{\sqrt{2}}.$$

Consider a qubit state

$$\begin{aligned}|\psi\rangle &= \alpha_0|0\rangle + \alpha_1|1\rangle\\ &= \alpha_0\frac{|0\rangle_x + |1\rangle_x}{\sqrt{2}} + \alpha_1\frac{|0\rangle_x - |1\rangle_x}{\sqrt{2}}\\ &= \frac{\alpha_0+\alpha_1}{\sqrt{2}}|0\rangle_x + \frac{\alpha_0-\alpha_1}{\sqrt{2}}|1\rangle_x.\end{aligned}$$

The probabilities of observing states $|0\rangle_x, |1\rangle_x$ are

$$\Pi_\psi(|0\rangle_x) = |_x\langle 0|\psi\rangle|^2 = \frac{|\alpha_0+\alpha_1|^2}{2}, \quad \Pi_\psi(|1\rangle_x) = |_x\langle 1|\psi\rangle|^2 = \frac{|\alpha_0-\alpha_1|^2}{2}.$$

Acting with $R_y(\pi/2)$ on $|\psi\rangle$ gives

$$|\psi'\rangle = R_y(\pi/2)|\psi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}\begin{pmatrix}\alpha_0\\ \alpha_1\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}\alpha_0-\alpha_1\\ \alpha_0+\alpha_1\end{pmatrix}.$$

Therefore

$$\Pi_\psi(|0\rangle_x) = \Pi_{\psi'}(|1\rangle_z), \quad \Pi_\psi(|1\rangle_x) = \Pi_{\psi'}(|0\rangle_z),$$

so we can perform effective x measurements by rotating the state followed by z basis measurements.

2.4 Measurement in any basis

Suppose we prepare the state $|\psi\rangle$ corresponding to a direction $\boldsymbol{\psi}$ on the Bloch sphere specified by polar angles θ_ψ, ϕ_ψ and wish to measure the projection in the direction $\mathbf{u}$ specified by polar angles θ_u, ϕ_u. We can perform this measurement either by rotating the Pauli operators into the $\mathbf{u}$ basis or by rotating the state.

In the first case the rotated operator is

$$\begin{aligned}\sigma_{\mathbf{u}} &= \mathbf{u}\cdot\boldsymbol{\sigma}\\ &= u_x\sigma_x + u_y\sigma_y + u_z\sigma_z\\ &= \begin{pmatrix}\cos(\theta_u) & \sin(\theta_u)e^{-\imath\phi_u}\\ \sin(\theta_u)e^{\imath\phi_u} & -\cos(\theta_u)\end{pmatrix}.\end{aligned} \tag{2.4}$$

The probability of a projective measurement along $\mathbf{u}$ resulting in the state $|\mathbf{u}\rangle$ is

$$\Pi_{|\mathbf{u}\rangle} = \langle\psi|\hat{P}_{|\mathbf{u}\rangle}|\psi\rangle = |\langle\mathbf{u}|\psi\rangle|^2 = \frac{1+\mathbf{u}\cdot\boldsymbol{\psi}}{2}$$

and

$$\Pi_{|-\mathbf{u}\rangle} = \langle\psi|\hat{P}_{|-\mathbf{u}\rangle}|\psi\rangle = |\langle-\mathbf{u}|\psi\rangle|^2 = \frac{1-\mathbf{u}\cdot\boldsymbol{\psi}}{2}.$$

The expected value of the qubit projection is

$$\begin{aligned}E &= \langle\psi|\sigma_{\mathbf{u}}|\psi\rangle\\ &= \Pi_{|\mathbf{u}\rangle} - \Pi_{|-\mathbf{u}\rangle}\\ &= \cos(\theta_\psi)\cos(\theta_u) + \sin(\theta_\psi)\sin(\theta_u)\cos(\phi_\psi-\phi_u)\\ &= \mathbf{u}\cdot\boldsymbol{\psi}.\end{aligned}$$

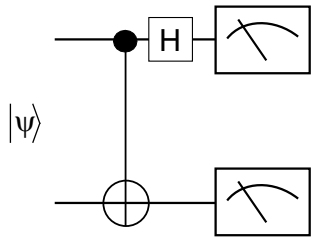


Figure 2.1: Circuit for measuring in the Bell basis. When the input state is one of the Bell states the outputs from the z measurements label the Bell states.

2.5 Measurement in the Bell basis

We often need to make measurements in the basis of Bell states $|\beta_{xy}\rangle$. This can be done with the circuit shown in Fig. 2.1. The circuit results in the transformation

$$\begin{aligned}
|\beta_{xy}\rangle &\rightarrow (H \otimes I) C_X |\beta_{xy}\rangle \\
&= (H \otimes I) C_X \frac{|0,y\rangle + (-1)^x |1, 1 \oplus y\rangle}{\sqrt{2}} \\
&= (H \otimes I) \frac{|0,y\rangle + (-1)^x |1,y\rangle}{\sqrt{2}} \\
&= \frac{|0,y\rangle + |1,y\rangle + (-1)^x \left(|0,y\rangle - |1,y\rangle\right)}{2} \\
&= \frac{\left(1 + (-1)^x\right)|0,y\rangle + \left(1 - (-1)^x\right)|1,y\rangle}{2}.
\end{aligned}$$

Therefore the Bell states are transformed as $|\beta_{xy}\rangle \rightarrow |xy\rangle$, i.e.

$$|\beta_{00}\rangle \rightarrow |00\rangle, \quad |\beta_{01}\rangle \rightarrow |01\rangle, \quad |\beta_{10}\rangle \rightarrow |10\rangle, \quad |\beta_{11}\rangle \rightarrow |11\rangle$$

and the measurement results immediately reveal which Bell state was input. Running this circuit backwards gives the standard design for preparing Bell states, starting from computational basis states.

Chapter 3

Algorithms

In this chapter we describe some basic quantum algorithms.

3.1 Deutsch-Jozsa

The Deutsch-Jozsa algorithm demonstrates a quantum speed up for a computational task. The two-qubit version of the algorithm was first proposed by Deutsch[38]. This was later extended to n bits by Deutsch and Jozsa[39].

Consider the following problem. The function f takes a one bit input with value 0 or 1 and maps it to a one bit output, 0 or 1. There are four possibilities given in Table 3.1. The problem is to determine $f(0) \oplus f(1)$ which tells us whether the function is constant or balanced.

Classically this requires two evaluations of the function f. Using a quantum circuit we need only a single evaluation. The function evaluation can be expressed as a unitary operator

$$|x\rangle|y\rangle \stackrel{U_f}{\rightarrow} |x\rangle|y \oplus f(x)\rangle.$$

The four possible versions of f can be implemented with the circuits shown in Fig. 3.1. The operator U_f corresponds to one of these four circuits.

To determine if the function is constant or balanced with a single evaluation put the first qubit in the state $\frac{|0\rangle+|1\rangle}{\sqrt{2}}$ so the input to the circuit is

$$\frac{|0\rangle + |1\rangle}{\sqrt{2}}|0\rangle = \frac{|00\rangle + |10\rangle}{\sqrt{2}}.$$

The output is

$$\begin{aligned} U_f \frac{|00\rangle + |10\rangle}{\sqrt{2}} &= \frac{1}{\sqrt{2}} U_f|00\rangle + \frac{1}{\sqrt{2}} U_f|10\rangle \\ &= \frac{1}{\sqrt{2}}|0\rangle|0 \oplus f(0)\rangle + \frac{1}{\sqrt{2}}|1\rangle|0 \oplus f(1)\rangle \\ &= \frac{1}{\sqrt{2}}|0\rangle|f(0)\rangle + \frac{1}{\sqrt{2}}|1\rangle|f(1)\rangle. \end{aligned}$$

$f(0)$	$f(1)$	$f(0) \oplus f(1)$	type of function
0	0	0	constant
0	1	1	balanced
1	0	1	balanced
1	1	0	constant

Table 3.1: Truth table for a function f. We refer to $f(0) = f(1)$ as a constant function and $f(0) \neq f(1)$ as balanced.

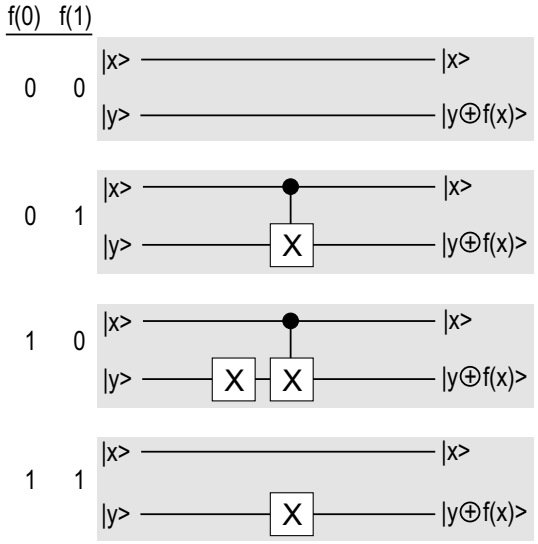


Figure 3.1: Circuit diagrams implementing $|x\rangle|y\rangle \xrightarrow{U_f} |x\rangle|y \oplus f(x)\rangle$ for the four possible functions f.

We have computed both $f(0)$ and $f(1)$ in a single function call. Of course a single measurement will only tell us one of the answers. For the Deutsch problem we are interested in whether f is constant or balanced. We can determine this by interference with the circuit shown in Fig. 3.2. Note that when the function is balanced, $f(0) \neq f(1)$ the output state is entangled. The presence of entanglement plays a crucial role in the computation.

Following the state evolution through the circuit we have

$$\begin{aligned}
|\psi\rangle_{\text{in}} &= |0\rangle|0\rangle \\
|\psi\rangle_1 &= (I \otimes X)|\psi\rangle_{\text{in}} = |0\rangle|1\rangle \\
|\psi\rangle_2 &= (H \otimes H)|\psi\rangle_1 \\
&= \frac{|0\rangle + |1\rangle}{\sqrt{2}} \frac{|0\rangle - |1\rangle}{\sqrt{2}} \\
&= \frac{1}{\sqrt{2}}|0\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}} + \frac{1}{\sqrt{2}}|1\rangle \frac{|0\rangle - |1\rangle}{\sqrt{2}} \\
|\psi\rangle_3 &= U_f|\psi\rangle_2.
\end{aligned}$$

At this point there is an interesting effect called "phase kickback" as explained in Sec. 1.1.4. We note that

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

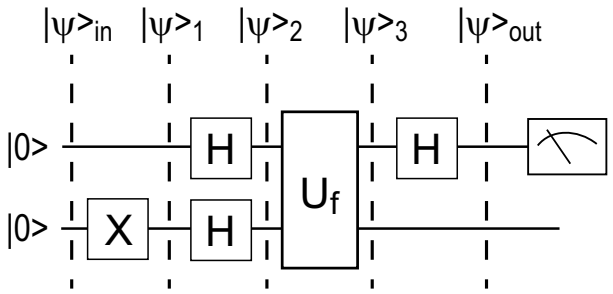


Figure 3.2: Circuit diagram for Deutsch algorithm.

We see that the state $\frac{|0\rangle-|1\rangle}{\sqrt{2}}$ is an eigenstate of U_f and the eigenvalue $(-1)^{f(x)}$ effectively gets imprinted on the first qubit.

Using phase kickback we see that

$$\begin{aligned}
|\psi\rangle_3 &= U_f|\psi\rangle_2 \\
&= \frac{(-1)^{f(0)}}{\sqrt{2}}|0\rangle\frac{|0\rangle-|1\rangle}{\sqrt{2}} + \frac{(-1)^{f(1)}}{\sqrt{2}}|1\rangle\frac{|0\rangle-|1\rangle}{\sqrt{2}} \\
&= \left[\frac{(-1)^{f(0)}|0\rangle+(-1)^{f(1)}|1\rangle}{\sqrt{2}}\right]\left[\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right] \\
&= (-1)^{f(0)}\left[\frac{|0\rangle+(-1)^{f(0)\oplus f(1)}|1\rangle}{\sqrt{2}}\right]\left[\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right]
\end{aligned}$$

and the final Hadamard gate gives

$$\begin{aligned}
|\psi\rangle_{\text{out}} &= (H\otimes I)|\psi\rangle_3 \\
&= (-1)^{f(0)}\left[\frac{|0\rangle+|1\rangle+(-1)^{f(0)\oplus f(1)}(|0\rangle-|1\rangle)}{\sqrt{2}}\right]\left[\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right] \\
&= (-1)^{f(0)}\left[\frac{|0\rangle+|1\rangle+(-1)^{f(0)\oplus f(1)}(|0\rangle-|1\rangle)}{\sqrt{2}}\right]\left[\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right].
\end{aligned}$$

If f is constant, $f(0)=f(1)$ so $(-1)^{f(0)\oplus f(1)}=1$ and

$$|\psi\rangle_{\text{out}} = (-1)^{f(0)}|0\rangle(|0\rangle-|1\rangle).$$

If f is balanced, $f(0)\neq f(1)$ so $(-1)^{f(0)\oplus f(1)}=-1$ and

$$|\psi\rangle_{\text{out}} = (-1)^{f(0)}|1\rangle(|0\rangle-|1\rangle).$$

We see that measuring the first qubit, the one that is not changed by U_f, reveals whether f is constant or balanced. Note the essential ingredients of the circuit: superposition (the input is put in a superposition state with the first H gate), interference (the second H gate distinguishes between $|0\rangle+|1\rangle$ and $|0\rangle-|1\rangle$), and entanglement which appears during the computation. The algorithm has been demonstrated in several experiments: with trapped ions[40], with superconducting qubits[41], with nuclear magnetic resonance[42], with optical fields[43], and with photonic one-way quantum computation[44].

3.2 Quantum search

The quantum search algorithm is due to Grover[45]. Finding an entry in an unsorted database containing N elements requires, on average, $N/2$ queries. The quantum search algorithm, which uses amplitude amplification, requires $O(\sqrt{N})$ queries where $N = 2^n$ and n is the number of qubits. It can be shown that this is optimal[46].

We can think of the algorithm in terms of an "oracle" to which we pass a quantum register containing n qubits in a superposition of $N = 2^n$ states. The oracle tags the correct state we are searching for by inverting the phase of that state. Using Grover's algorithm it takes only $\mathcal{O}(\sqrt{n})$ queries of the oracle to find the correct state.

In general there is no recipe for constructing such an oracle. If we imagine the search problem as finding a name in a phone book that is not alphabetically sorted it seems that we need to know the answer in order to construct the oracle, so the algorithm is not useful. The utility is more apparent if we think of the oracle as evaluating a function. For example we could apply Grover's algorithm to an NP problem such as factoring. The oracle stores a large composite number and the register contains a superposition of possible factors. The oracle can efficiently check the possible factors without knowing the correct answer in advance. For this particular problem Shor's algorithm provides a larger speedup, exponential instead of quadratic, but the same approach could be applied to other problems such as finding an input string that optimizes a given function relative to a threshold.

An intuitive description of Grover's algorithm goes as follows. We prepare the n qubit register in a superposition of all possible values. The register is input to the oracle and the "correct" component has its amplitude inverted. This corresponds to making the replacement

$$|\psi\rangle = \sum_{j=1}^{2^n} c_j |x_j\rangle \rightarrow |\psi'\rangle = \sum_{j=1}^{2^n} c_j (-1)^{f(x_j)} |x_j\rangle$$

where $f(x_k) = 1$ and $f(x_{j\neq k}) = 0$ where x_k is the target state.

The amplitudes of all components of the input state are then inverted about the mean. This corresponds to making the replacement

$$|\psi\rangle = \sum_{j=1}^{2^n} c_j |x_j\rangle \rightarrow |\psi'\rangle = \sum_{j=1}^{2^n} c'_j |x_j\rangle \tag{3.1}$$

where $c'_j = \mu - (c_j - \mu) = 2\mu - c_j$ with $\mu = \frac{1}{2^n}\sum_j c_j$. We are here using the fact that all the c_j are initially real and therefore stay real during the computation. The combination of amplitude inversion and inversion about the mean constitutes a Grover iteration. After a number of iterations that scales as $\sqrt{N}$ the amplitude of the correct component of the state vector becomes very close to one. An example of this procedure is shown in Fig. 3.3.

The question is then how do we convert this computation into a quantum circuit? The circuit diagram is shown in Fig. 3.4. Note that for the case of $n = 2$ qubits the circuit can be simplified to that shown in Fig. 3.5. This circuit has equivalent functionality to that in Fig. 3.4 after swapping of the Oracle rotations ($X \rightarrow I, I \rightarrow X$).

Let's analyze the operation of this circuit. The input state is

$$|\psi\rangle = H^{\otimes n}|0\rangle^{\otimes n} \equiv \vec{H}|\vec{0}\rangle.$$

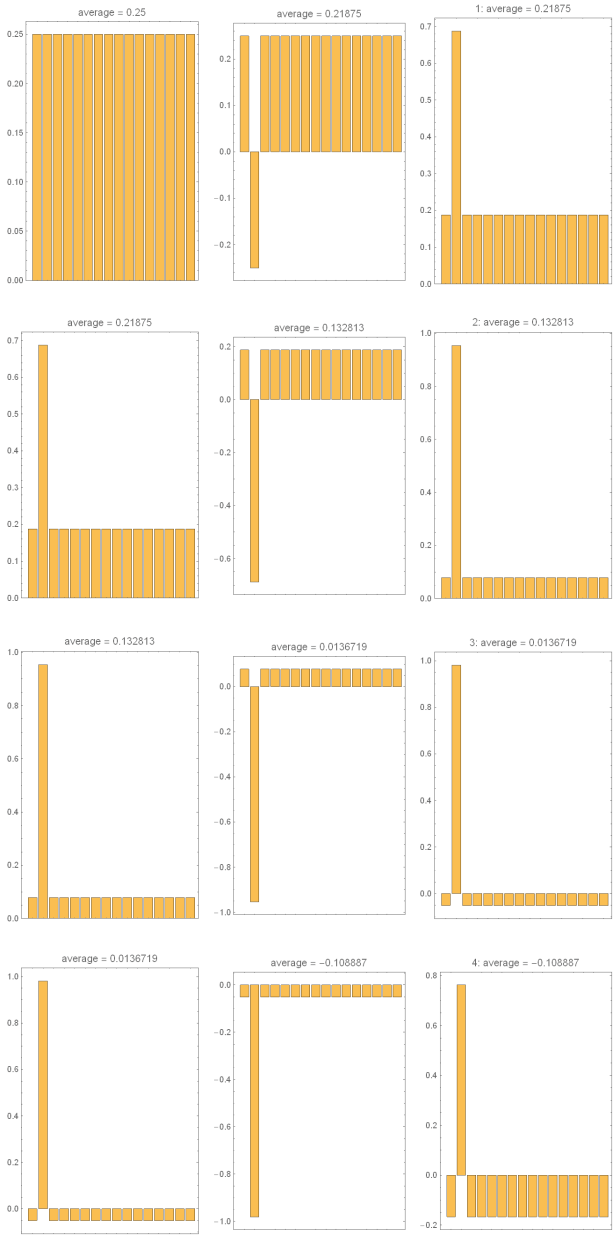


Figure 3.3: From top to bottom the rows show the results of amplitude amplification for $n = 4, N = 16$. Each row shows the initial amplitudes, amplitudes after sign inversion, and after inversion about the mean. Iteration 3 gives the highest success probability.

One Grover iteration corresponds to the operator G with $G = \vec{H}\tilde{U}_\perp\vec{H}U_f$ where the sign inversion is

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

and the inversion about the mean is accomplished by $\vec{H}\tilde{U}_\perp\vec{H}$ with

$$\tilde{U}_\perp|0\rangle = |0\rangle, \qquad \tilde{U}_\perp|x \neq 0\rangle = -|x \neq 0\rangle.$$

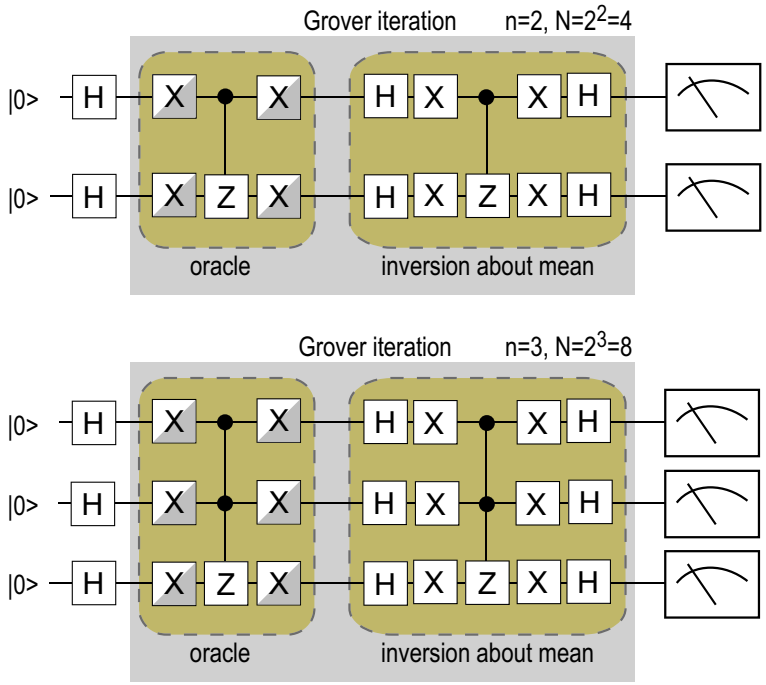


Figure 3.4: Circuit diagrams for quantum search with $n = 2, 3$ qubits. The combination of the oracle query and inversion about the mean defines a Grover iteration which is repeated K times, after which the result is read out. The half grayed X gates in the oracle are implemented or not to select the target string. If the corresponding bit of the target string is 0(1) the X gate in the oracle is(is not) implemented. The circuit for n qubits is the same as for $n = 3$, but with $n - 1$ upper rows.

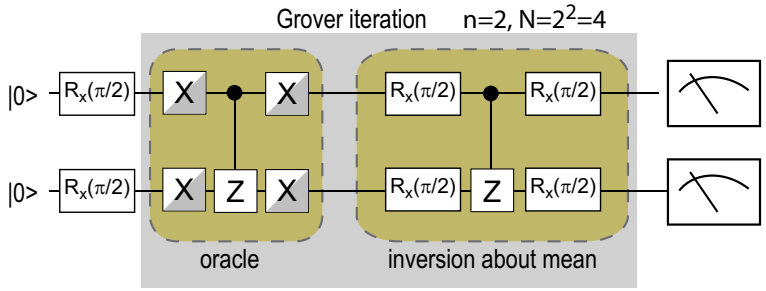


Figure 3.5: Simplified circuit diagrams for quantum search with $n = 2$ qubits.

Observe the following relations:

$$\vec{H}\tilde{U}_\perp\vec{H}|\psi\rangle = \vec{H}\tilde{U}_\perp|0\rangle = |\psi\rangle$$

so $|\psi\rangle$, the uniform superposition state, is an eigenvector of $\vec{H}\tilde{U}_\perp\vec{H}$ with eigenvalue $+1$.

The state $|\psi\rangle$ is one state out of the 2^n different n qubit states. There is an orthogonal space $V_\perp$ which is spanned by $\vec{H}|x \neq 0\rangle$. In other words states in $V_\perp$ are $|\psi_\perp\rangle = \vec{H}|x \neq 0\rangle$. These orthogonal states satisfy

$$\vec{H}\tilde{U}_\perp\vec{H}|\psi_\perp\rangle = \vec{H}\tilde{U}_\perp|x \neq 0\rangle = -|\psi_\perp\rangle$$

so $|\psi_\perp\rangle$ is an eigenvector of $\vec{H}\tilde{U}_\perp\vec{H}$ with eigenvalue -1.

To proceed let's clean up the notation by defining $U_\perp = \vec{H}\tilde{U}_\perp\vec{H}$ so the operator of a Grover iteration is $G = U_\perp U_f$ and $U_\perp|\psi\rangle = |\psi\rangle, U_\perp|\psi_\perp\rangle = -|\psi_\perp\rangle$. It can be shown that

$$U_\perp = 2|\psi\rangle\langle\psi| - I = |\psi\rangle\langle\psi| - (I - |\psi\rangle\langle\psi|)$$

which corresponds to inversion about the mean, the second part of the Grover iteration. This is inversion about the mean because if we consider an arbitrary state $|\phi\rangle = \sum_j c_j|x_j\rangle$ the mean amplitude is $\mu = \frac{1}{2^n}\sum_j c_j$ and

$$U_\perp|\phi\rangle = \sum_j (2\mu - c_j)|x_j\rangle$$

which is the same as Eq. (3.1). We see that the inversion about the mean is also a conditional sign change, for all states orthogonal to $|\psi\rangle$.

Using these definitions we are now ready to analyze the circuit behavior. The uniform state $|\psi\rangle$ is a superposition of all possible bit strings. Divide the components of $|\psi\rangle$ into two parts $\{x_{\rm bad}\}$ for which $f(x_{\rm bad}) = 0$ and the target string $|t\rangle$ for which $f(t) = 1$. Assuming there is only one valid target string we can write

$$|\psi\rangle = \frac{1}{\sqrt{N}}|t\rangle + \sqrt{\frac{N-1}{N}}|\psi_{\rm bad}\rangle,$$

where $|\psi_{\rm bad}\rangle = \sum_j c_j|x_{\rm bad}\rangle$ is a superposition state and $f(x_{\rm bad}) = 0$. The orthogonal state is

$$|\psi_\perp\rangle = \sqrt{\frac{N-1}{N}}|t\rangle - \frac{1}{\sqrt{N}}|\psi_{\rm bad}\rangle.$$

With these definitions $\langle\psi_\perp|\psi\rangle = 0$. Then define $\sin(\theta) = 1/\sqrt{N}$, $\cos(\theta) = \sqrt{(N-1)/N}$ so

$$\begin{aligned} |t\rangle &= \sin(\theta)|\psi\rangle + \cos(\theta)|\psi_\perp\rangle, \\ |\psi_{\rm bad}\rangle &= \cos(\theta)|\psi\rangle - \sin(\theta)|\psi_\perp\rangle. \end{aligned}$$

It is then not hard to show that

$$\begin{aligned} G|\psi\rangle &= U_\perp U_f|\psi\rangle \\ &= \sin(3\theta)|t\rangle + \cos(3\theta)|\psi_{\rm bad}\rangle. \end{aligned} \tag{3.2}$$

After k Grover iterations we find

$$G^k|\psi\rangle = \sin[(2k+1)\theta]|t\rangle + \cos[(2k+1)\theta]|\psi_{\rm bad}\rangle. \tag{3.3}$$

If we stop at $(2k+1)\theta \simeq \pi/2$ the amplitude of the desired target state will be approximately unity. The optimal number of iterations for large N is therefore

$$k = \frac{\pi/2 - \theta}{2\theta} \simeq \frac{\pi}{4}\sqrt{N}.$$

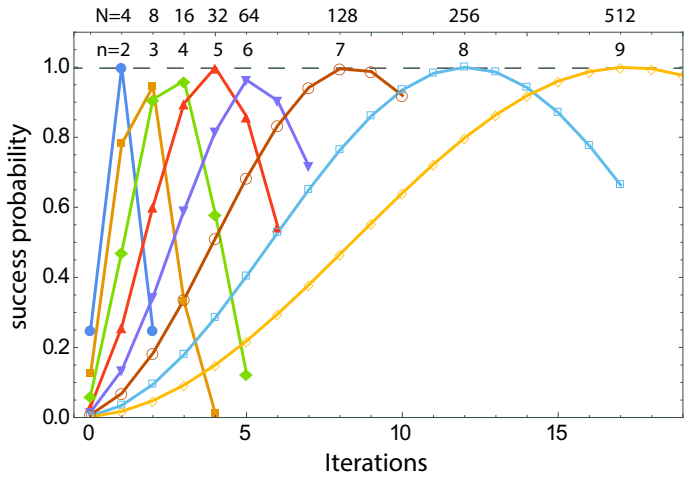


Figure 3.6: Success probability of Grover algorithm for $n = 2 - 9$ qubits as a function of the number of iterations.

This demonstrates the $\sqrt{N}$ scaling of the algorithm.

For $n = 2$ ($N = 4$) a single iteration results in finding the correct string with unit probability. For $n = 3$ ($N = 8$) we find the correct string with 94.5% probability after two iterations. The success probability versus number of iterations is shown in Fig. 3.6. The algorithm has been simulated for n up to 10 using a Mathematica code running on a desktop computer. Simulations up to $n = 40$ have been reported using specialized codes[47]. For $n = 2$ the algorithm works perfectly at finding the marked element out of four possibilities after a single iteration. The average number of tries $\langle k \rangle$ needed to find the answer classically is

$$\langle k \rangle = 1 \times 1/4 + 2 \times (3/4)(1/3) + 3 \times (3/4)(2/3) = 9/4.$$

For $n > 2$ the success probability after the optimal number of iterations is high but slightly less than 1. For each value of n the iterations have been stopped after the first maximum to make the figure readable. Nevertheless it can be helpful to run more iterations. For example for $n = 4$ the success probability is 0.961 after 3 iterations and 0.992 after 9 iterations. It is also possible to get perfect success probability by using a modified phase shift in the inversion about the mean part of the circuit[48].

The execution time scales with the number of bits n and the number of iterations K. Here we give some plausible numbers based on a neutral atom implementation with microwave single qubit rotations and Rydberg two-qubit gates[49]. We need $2K + 1$ global Hadamards which can be done with microwaves. We need $2K$ global X gates which can also be done with microwaves. We need $2Kn$ site selected X gates which can be done with microwaves and a Stark shifting beam as in [49]. In total we need $2K(n + 1)$ X gates and $2K + 1$ Hadamards. The $C_{n-1,\,Z}$ gates can be done with $2n$ Rydberg π pulses as described in [50] . Taking the pulse times as $T_{\pi,\mu\mathrm{w}} = 50$ μs for a microwave π pulse and $T_{\pi,\mathrm{Ryd}} = 0.5$ μs for a Rydberg π pulse we get a time of

$$T = 2K(n + 1)T_{\pi,\mu\mathrm{w}} + (K + 1/2)T_{\pi,\mu\mathrm{w}} + 4nT_{\pi,\mathrm{Ryd}}.$$

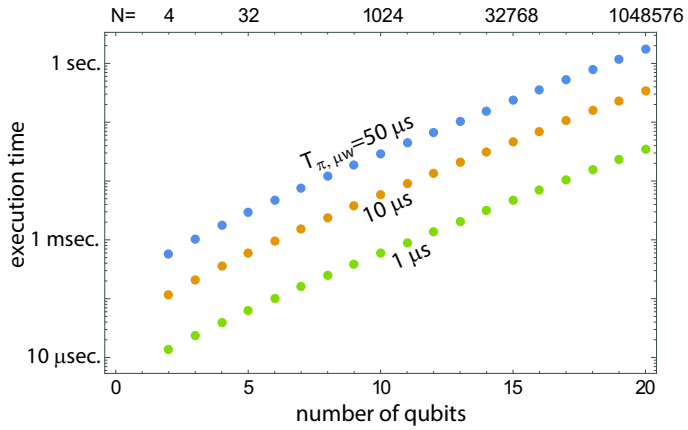


Figure 3.7: Execution time for n qubit Grover search with, from top to bottom, $T_{\pi,\mu\mathrm{w}} = 50, 10, 1$ μs and $T_{\pi,\mathrm{Ryd}} = 0.5$ μs.

Putting $K = \frac{\pi}{4}N^{1/2} = \frac{\pi}{4}2^{n/2}$ we get

$$T = \left[\frac{\pi(n+3/2)}{2}2^{n/2} + 1/2\right] T_{\pi,\mu\mathrm{w}} + 4nT_{\pi,\mathrm{Ryd}}.$$

The execution time for several different π pulse times is shown in Fig. 3.7. These execution time estimates assume the efficient $C_{n,Z}$ implementation described in [50]. This implementation does not scale to arbitrarily large n. As described in [51] a sub-register architecture can be used to extend n with multi-bit gates each of which only act on a subset of the qubits. Alternatively one can decompose $C_{n,Z}$ gates into multiple $C_{2,Z}$ (Toffoli type) gates. The decomposition given in [52] requires $2n-7$ Toffoli type gates and $n-3$ scratch bits to implement $C_{n-1,Z}$. Since implementation of $C_{2,Z}$ requires 6 Rydberg π pulses the $C_{n-1,Z}$ gate would require $6\times(2n-7) = 12n-42$ Rydberg π pulses. If implemented directly as a Rydberg $C_{n-1,Z}$ gate we would only need $2n$ Rydberg π pulses, so the overhead in the multi-qubit gate decomposition is a factor of 6 in Rydberg pulses plus scratch bits. This time overhead does not appreciably change the execution times shown in Fig. 3.7 which are dominated by the single qubit gates needed for the oracle.

3.3 Phase estimation

The phase estimation algorithm can be used to find the eigenvalue of a unitary operator. It plays a central role in the factoring algorithm which is discussed in the next section. Consider a unitary operator U with eigenket $|u\rangle$ satisfying $U|u\rangle = \lambda|u\rangle$. The eigenvalues of unitary operators have unit modulus so we can write $\lambda = e^{\imath 2\pi\phi}$ with $0 \le \phi \le 1$. The goal of the algorithm is to find ϕ. We will use the fact that $U^{2^j}|u\rangle = e^{\imath 2\pi 2^j\phi}|u\rangle$ for $j \ge 0$.

Consider a controlled U gate C_U with control qubit in the state $H|0\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$ and target qubit in the eigenstate $|u\rangle$. The gate results in the output state

$$C_U\frac{|0\rangle+|1\rangle}{\sqrt{2}}|u\rangle = \frac{|0\rangle|u\rangle + |1\rangle U|u\rangle}{\sqrt{2}} = \frac{|0\rangle|u\rangle + e^{\imath 2\pi\phi}|1\rangle|u\rangle}{\sqrt{2}} = \frac{|0\rangle + e^{\imath 2\pi\phi}|1\rangle}{\sqrt{2}}|u\rangle.$$

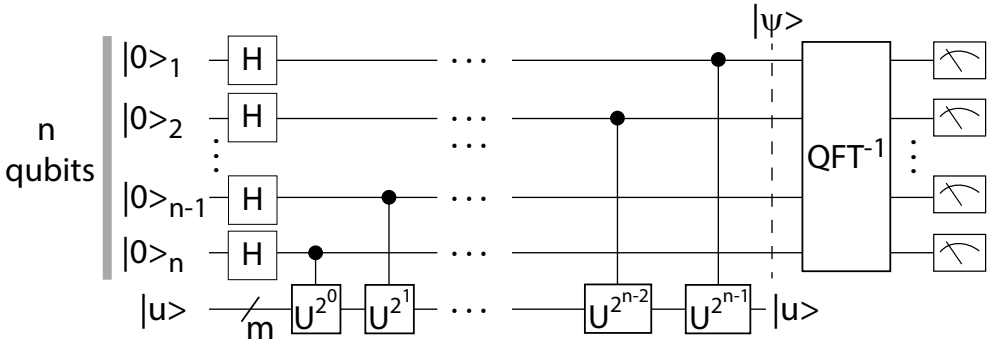


Figure 3.8: Circuit for finding the phase of the unitary operator U. The notation $/_m$ means a bundle of m qubits. The size m of the second register is chosen to be large enough to hold $|u\rangle$.

Similarly

$$C_{U^{2^j}} \frac{|0\rangle + |1\rangle}{\sqrt{2}} |u\rangle = \frac{|0\rangle + e^{\imath 2\pi 2^j \phi}|1\rangle}{\sqrt{2}} |u\rangle.$$

The circuit shown in Fig. 3.8 uses an n qubit register that controls C_U raised to sequentially higher powers. The control qubits are all prepared in the state $\frac{|0\rangle+|1\rangle}{\sqrt{2}}$ so that after the C_U operations the circuit state is

$$|\psi\rangle \;=\; \frac{1}{\sqrt{2^n}} \bigotimes_{j=1}^{n} \left(|0\rangle_j + e^{\imath 2\pi 2^{n-j}\phi} |1\rangle_j \right) \otimes |u\rangle$$

If ϕ has a value that can be exactly represented with n bits, which is to say $\phi = a/2^n$ where a is an n digit binary number, then

$$|\psi\rangle = \frac{1}{\sqrt{2^n}} \bigotimes_{j=1}^{n} \left(|0\rangle_j + e^{\frac{\imath 2\pi}{2^n} 2^{n-j} a} |1\rangle_j \right) \otimes |u\rangle.$$

Comparing with Eq. (1.9) we see that

$$|\psi\rangle = QFT|a\rangle \otimes |u\rangle,$$

so taking the inverse QFT as shown in Fig. 3.8 gives the state

$$QFT^{-1}|\psi\rangle = |a\rangle \otimes |u\rangle.$$

Measuring the first n qubits of the output determines a from which we find $\phi = a/2^n$.

When the phase does not have an exact n bit binary representation the phase algorithm does not return the exact value of the phase but provides a good estimate. The uncertainty of the phase can be estimated as follows. When ϕ is not an integer multiple of $1/2^n$ the output state after the QFT^{-1} operation will be

$$|\chi\rangle = \sum_x c_x(\phi)|x\rangle.$$

Here $|\chi\rangle$ is not a computational basis state but is a sum of basis states $|x\rangle$. The probability of observing $|x\rangle$ is $|\langle x|\chi\rangle|^2 = |c_x(\phi)|^2$. It can be shown that if $\frac{a}{2^n} < \phi < \frac{a+1}{2^n}$ and ϕ is exactly in between these bounds then with probability $\Pi \geq 8/\pi^2 = 0.81$ the observed $|x\rangle$ will be $|a\rangle$ or $|a+1\rangle$. Furthermore if $\frac{a}{2^n}$ is the multiple of $1/2^n$ that is closest to ϕ, then with probability $\Pi \geq 4/\pi^2 = 0.405$ the observed $|x\rangle$ will be $|a\rangle$. Importantly the probability of obtaining a good estimate for ϕ stays constant as n increases and the uncertainty $1/2^n$ decreases.

3.4 Factoring

The most famous quantum algorithm is Shor's factoring algorithm[53, 54]. Factoring is believed to be a hard problem on a classical computer. The best known classical algorithm is the number field sieve. A lucid presentation of the development of this algorithm can be found in [55]. This algorithm has run time

$$t \sim e^{\left(\frac{64}{9}\ln N\right)^{1/3}(\ln\ln N)^{2/3}}$$

where N is the number to be factored. Say we have enough computer power to factor a number with 500 binary digits in 30 minutes, a not unreasonable time to wait to decrypt a message. Then the time for factoring a number with 1000 binary digits would be 370 years. In fact the 1061 bit composite number $2^{1061} - 1$ was factored between early 2011 and 4 August 2012 using about 300 CPU-years of sieving (`http://en.wikipedia.org/wiki/Integer_factorization_records`). Nevertheless increasing the number of digits by an additional factor of two would require 350 billion CPU-years, which would render factoring completely infeasible.

It is current belief, but has not been proven, that a polynomial time classical factoring algorithm does not exist. Shor's algorithm provides us with a solution that runs in polynomial time on quantum hardware. The resource requirements in terms of number of qubits and gate depth depend on the particular implementation of the algorithm. One recent example[56] presents an architecture which requires $9n+2$ qubits and has a depth of about $2000n^2$ where $n = \log_2(N)$ is the number of bits in the binary representation of the number. Thus to factor a 1000 bit integer would require about 9000 qubits, and a circuit depth of $\sim 10^9$ gates. At a clock speed of 1 MHz this would take about 30 min. execution time. A 2000 bit number would require just over two hours, a vastly different scaling than for the classical algorithm.

Since gate error rates of 10^{-9} are not available, and may never be, quantum error correction is required which will boost the qubit count by a factor of 100 - 1000, depending on the code used, and the execution time by an even larger factor. Nevertheless the polynomial scaling of the algorithm implies that for sufficiently large numbers a future scalable and error corrected quantum computer will succeed where classical computers fail.

As of early 2018 the largest numbers to be factored with Shor's algorithm on a quantum computer were 15 with trapped ions[57] and 21 with photons[58]. Other, quantum related demonstrations, using adiabatic and annealing approaches include factoring 143 using liquid state nuclear magnetic resonance[59] and a much larger number, 200099 on a D-Wave 2X processor[60]. A distinction should be made between the demonstrations of Shor's algorithm which has a proven polynomial scaling, and approaches based on annealing for which it is unknown how the spectral gap, and therefore the run time, scales with problem size.

3.4.1 Shor's algorithm

This is a physics course and the problem of factoring does not appear to have a lot to do with physics. However, the existence of the factoring algorithm has played a central role in the high level of interest in and rapid development of quantum computing in the last two decades. The reason being that the difficulty of factoring on a classical computer is central to the Rivest-Shamir-Adleman (RSA) public encryption scheme[61] which underpins much of internet commerce. It therefore seems relevant to understand in more detail what all the fuss is about.

All composite numbers N can be expressed in terms of their prime factorization as

$$N = p_1^{m_1} p_2^{m_2} p_3^{m_3} ... p_k^{m_k}. \tag{3.4}$$

Here the p_i are the prime factors, $m_i = 1, 2, ...$ are the multiplicities of each factor, and k is the number of distinct factors. The number of digits in the binary representation of N is $n = \lceil \log_2(N) \rceil$ (the notation $\lceil x \rceil$ means the smallest integer larger than or equal to x and $\lfloor x \rfloor$ means the largest integer less than or equal to x).

New algorithms are invented at unexpected junctures. For example Agrawal, Kayal, and Saxena (AKS) found a deterministic polynomial time algorithm for deciding whether a number is prime or composite in 2004[62]. Unfortunately AKS does not help to find the prime factors when the number is composite but it can be used to check whether or not a number can be factored before trying Shor's algorithm. Even before checking with AKS there are of course a few simple cases for which we know a number is not prime. When expressed in decimal form the number is even if the last digit is $0, 2, 4, 6$, or 8. The number is divisible by 3 if the sum of the digits is divisible by 3. The number is divisible by 5 if the last digit is 5.

If we have run the AKS algorithm and are sure the number is not prime we know the number is of the form (3.4). We can first check if $N = p^m$, $m \geq 2$. In this case $\log_2 N = m \log_2 p$ so $m = \frac{\log_2 N}{\log_2 p} \leq \log_2 N$ and there is an efficient algorithm for this case. We might also choose a random number $1 < x < N - 1$ and calculate $\gcd(x, N)$ (gcd is the greatest common divisor) using Euclid's algorithm which is efficient. If $\gcd(x, N) \neq 1$ we have found a factor. We could of course do this for all possible x, but such a procedure is exponential in n.

Having checked that we cannot efficiently factor the number of interest with classical tools we may turn to Shor's factoring algorithm which has a classical and a quantum part. First the problem of factoring is reduced to finding the order of a number, a calculation that is classically hard. The quantum part provides an efficient method of finding the order which, with high probability, provides one of the factors of the original number.

Why does finding the order help us find a prime factor? For integers x and N with no common factors the order of x modulo N is the least positive integer r such that

$$x^r = 1 \bmod N. \tag{3.5}$$

N is the number we wish to factor and x is a randomly chosen number. The length of the number is $L = \log_2 N$. There is no known classical algorithm for finding r which is polynomial in L, but we can efficiently find r with Shor's algorithm.

If r is even Eq. (3.5) can be written as

$$(x^{r/2}+1)(x^{r/2}-1) = 0 \bmod N. \tag{3.6}$$

This equation implies that N has a common factor with $x^{r/2}+1$ or $x^{r/2}-1$. There are two trivial solutions to (3.6): $x^{r/2} = 1 (\bmod N)$ and $x^{r/2} = -1 (\bmod N)$. Suppose the r we find is not one of these trivial solutions. Then a factor of N is $\gcd(x^{r/2}+1, N)$ or $\gcd(x^{r/2}-1, N)$ and the gcd can be computed efficiently with Euclid's algorithm.

The algorithm is not guaranteed to produce a factor, but the probability of it doing so is not small. The probability that r is even as well as $x^{r/2} \neq 1$ and $x^{r/2} \neq -1$ is at least $1/2$ provided N is not a prime power, a case which we have already dealt with. Therefore if we found the period r correctly we will get a factor with probability $\sim 1/2$. There is also a finite probability that the eigenvalue finding part of Shor's algorithm actually returns a correct value for r. The probability for this is ~ 0.4. Thus the probability of finding a factor from a single run of the algorithm is ~ 0.2. The probability of not having found a factor after running the algorithm 10 times is $\sim (0.8)^{10} = 0.11$. In other words we have a high probability of about 90% of finding a factor after a modest number of tries. The important point is that the success probability does not become smaller as N increases. Indeed for large N it can be shown that the probability of finding a factor given r reaches 0.95 for large N and r[63]. In such a limit the success probability for one run becomes $\sim 0.4 \times 0.95 = 0.38$ and the probability of finding a factor after only 5 tries is better than 99%.

To make the procedure clearer let's consider some examples. Suppose we want to factor $N = 91$. Choose as a test number the smallest prime greater than 1, i.e. $x = 2$. Then $x^a \bmod N$ gives

$$\begin{aligned} a &: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, ... \\ 2^a &: 1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048, ... \\ 2^a \bmod 91 &: 1, 2, 4, 8, 16, 32, 64, 37, 74, 57, 23, 46, 1, 2, 4, ... \end{aligned} \tag{3.7}$$

We see that $2^{12} \bmod 91 = 1$ so $r = 12$. Then calculate $2^{12/2} - 1 = 63$ and $2^{12/2} + 1 = 65$. We see that $63 \times 65 = 4095 = 0 \bmod 91$. Therefore $\gcd(63, 91) = 7$ and $\gcd(65, 91) = 13$ are factors of 91. This is correct since $7 \times 13 = 91$.

As another example let's factor $N = 161$. Choosing $x = 2$ we find

$$\begin{aligned} a &: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, ... \\ 2^a &: 1, 2, 4, 8, 16, 32, 64, 128, 256, 512, 1024, 2048, ... \\ 2^a \bmod 161 &: 1, 2, 4, 8, 16, 32, 64, 128, 95, 29, 58, 116, 71, 142, 123, 85, 9, 18, 36, 72, \\ &\quad 144, 127, 93, 25, 50, 100, 39, 78, 156, 151, 141, 121, 81, 1, 2, 4, ... \end{aligned}$$

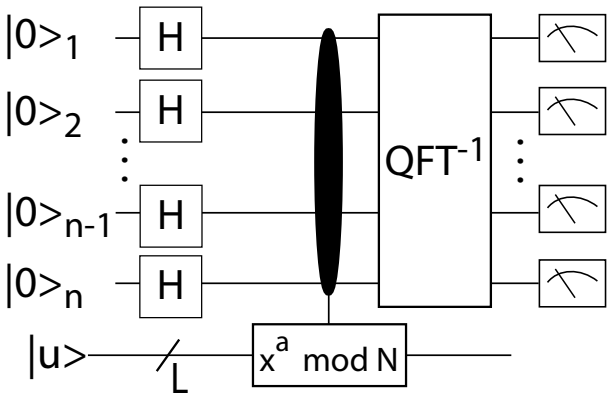


Figure 3.9: Circuit for order finding. The input register has n bits and the modular exponentiation operation acts on $L = \lceil \log(N) \rceil$ qubits with N the number to be factored.

We see that $r = 33$ and the algorithm fails. We then try $x = 3$ and find

$$\begin{aligned}
a &: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, ... \\
3^a &: 1, 3, 9, 27, 81, 243, 729, 2187, 6561, 19683, 59049, 177147, ... \\
3^a \bmod 161 &: 1, 3, 9, 27, 81, 82, 85, 94, 121, 41, 123, 47, 141, 101, 142, 104, \\
&\quad 151, 131, 71, 52, 156, 146, 116, 26, 78, 73, 58, 13, 39, 117, 29, 87, \\
&\quad 100, 139, 95, 124, 50, 150, 128, 62, 25, 75, 64, 31, 93, 118, 32, 96, \\
&\quad 127, 59, 16, 48, 144, 110, 8, 24, 72, 55, 4, 12, 36, 108, 2, 6, 18, \\
&\quad 54, 1, 3, 9, 27, ...
\end{aligned}$$

We see that $3^{66} \bmod 161 = 1$ so $r = 66$. Then calculate $3^{66/2} - 1 = 5559060566555522 = y_1$ and $3^{66/2} + 1 = 5559060566555524 = y_2$. We see that
$y_1 \times y_2 = 30903154382632612361920641803528 = 0 \bmod 161$. Therefore $\gcd(y_1, 161) = 23$ and $\gcd(y_2, 161) = 7$ are factors of 161. This is correct since $7 \times 23 = 161$.

There is no known classical algorithm that can efficiently find the order of a number. Instead we can use Shor's polynomial time quantum algorithm to find the order, and thereby find a factor. Interestingly there are also other quantum algorithms for rapid factoring. Factoring can be recast as an optimization problem which can be solved by adiabatic quantum computing[64].

The basic idea of the factoring algorithm is to use phase estimation, as presented in the previous section, to find the order r of x modulo N where x is a random number we choose in polynomial time. The circuit implementation for the quantum part of Shor's algorithm is therefore the same as Fig. 3.8 except U is now the operation of modular exponentiation, to be explained in the following, and the eigenket $|u\rangle$ is replaced by $|1\rangle$. The circuit is shown in Fig. 3.9.

The modular multiplication operator U_x has the property

$$U_x|y\rangle = |xy \,(\mathrm{mod}\, N)\rangle, \quad \text{for } 0 \leq y \leq N - 1.$$

For $N \leq y \leq 2^L - 1$ we define U_x to have the property $U_x|y\rangle = |y\rangle$. Since $x^r = 1 (\mathrm{mod}\, N)$

with r the order, we see that

$$U_x^r|y\rangle = |x^r y \,(\mathrm{mod}\, N)\rangle = |y\rangle$$

so U_x is an r^{th} root of the identity operator. The operator U_x is unitary and has eigenstates $|u_s\rangle$. The eigenstates for $0 \leq s \leq r-1$ are

$$|u_s\rangle = \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} e^{-\imath 2\pi s k/r} |x^k (\mathrm{mod}\, N)\rangle.$$

To verify this we see that

$$\begin{aligned}
U_x|u_s\rangle &= \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} e^{-\imath 2\pi s k/r} |x^{k+1} (\mathrm{mod}\, N)\rangle \\
&= e^{\imath 2\pi s/r} \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} e^{-\imath 2\pi s (k+1)/r} |x^{k+1} (\mathrm{mod}\, N)\rangle \\
&= e^{\imath 2\pi s/r} \frac{1}{\sqrt{r}} \sum_{k=1}^{r} e^{-\imath 2\pi s k/r} |x^k (\mathrm{mod}\, N)\rangle \\
&= e^{\imath 2\pi s/r} \frac{1}{\sqrt{r}} \left[\sum_{k=0}^{r-1} e^{-\imath 2\pi s k/r} |x^k (\mathrm{mod}\, N)\rangle - |1 (\mathrm{mod}\, N)\rangle + e^{-\imath 2\pi s} |x^r (\mathrm{mod}\, N)\rangle \right] \\
&= e^{\imath 2\pi s/r} |u_s\rangle
\end{aligned}$$

which shows that the eigenvalue of $|u_s\rangle$ is $e^{\imath 2\pi s/r}$. Using U_x as the controlled operator in the phase estimation algorithm, as shown in Fig. 3.9 we can measure the phase s/r and thereby determine r since we know s. The phase estimation algorithm involves evaluating powers of the modular multiplication operator $U_x^{2^j}$, hence the name modular exponentiation. The actual circuit that implements $U_x^{2^j}$ for phase estimation depends on the choice of the number x so the circuit has to be changed to accommodate different values of x. To complete the phase estimation circuit we need to implement the controlled operations $C_{U_x^{2^j}}$ for $j = 0, n-1$. The most obvious way of doing this is to repeat C_{U_x}, 2^j times at each step, which would seem to imply an exponential number of operations. However, multiplying by $x (\mathrm{mod}\, N)$, 2^j times gives the same result as multiplying by $x^{2^j} (\mathrm{mod}\, N)$ from which it follows that $U_x^{2^j} = U_{x^{2^j}}$. We therefore only need n modular multiplication circuits, each of which implements modular multiplication of numbers with L digits which has a resource cost of L^2 using the standard method of multiplying two numbers together. The total computational cost is therefore $O(nL^3)$ and using $n \sim L$ we find the modular exponentiation part of the circuit has a cost scaling as L^3. There are classical reversible logic circuits that implement modular multiplication as a permutation matrix. These can be implemented as quantum circuits using only swap gates which can be constructed with CNOT operations. The quantum circuit depth for the modular exponentiation part of order finding therefore scales as $O(L^3)$.

There is an obvious problem in that without knowing the order in advance we cannot prepare an eigenstate $|u_s\rangle$ of the modular multiplication operator. However, we can use the fact that U_x is a unitary and normal operator so its eigenvectors form a complete basis. We

are therefore guaranteed that a state which it is easy to prepare such as $|1\rangle$ can be expressed as a superposition of eigenstates of U_x. In fact, it is not hard to show that

$$\frac{1}{\sqrt{r}}\sum_{s=0}^{r-1}|u_s\rangle = |1\rangle.$$

The output of the phase estimation algorithm using the lower register prepared in the state $|1\rangle$ is

$$|\psi\rangle = \frac{1}{\sqrt{r}}\sum_{s=0}^{r-1}|\widetilde{s/r}\rangle|u_s\rangle$$

where $|\widetilde{s/r}\rangle$ is the n bit binary state closest to the fraction s/r. This follows from the linearity of quantum operations. The state of the first register is therefore a uniform superposition of states $|\widetilde{s/r}\rangle$. A measurement will give an estimate of s/r and using an efficient algorithm for continued fractions the values of s and r can be determined with high probability.

The gate count needed for the QFT on n qubits is n Hadamard gates and $\sum_{k=1}^{n} k-1 = \frac{n(n-1)}{2}$ controlled rotation gates. Some shortcuts can be taken. The final QFT is followed by a measurement so the phases are not important. Therefore the QFT can be replaced by a semi-classical version which only requires one-qubit gates[21, 22]. On the other hand Eq. (1.10) shows that implementation of the QFT on an n qubit register requires small phase rotations with an exponential precision of $\theta \sim 2\pi/2^n$, which for large n will be unrealistic to implement with imperfect hardware. This would seem to negate the usefulness of the QFT and the factoring algorithm. Fortunately this turns out not to be a limiting problem since excluding rotations smaller than some threshold from the QFT has a minimal impact on the output state[65]. The performance of the factoring algorithm with this approximate QFT substituted for the full QFT verifies the scalability[20]. Another concern is the difficulty of performing CNOT gates between separated qubits in a large register. It has also been shown that the factoring algorithm retains its performance with a banded QFT that only implements the controlled rotations on qubits with indices that are closer than some threshold[66].

Taking $n = L$ we see that the resource cost of the QFT scales as L^2 whereas the modular exponentiation requires resources scaling as L^3. Modular exponentiation is thus the computational bottleneck so much work has gone into designing optimized circuits[28, 67, 56]. For an early detailed design see the paper from the Preskill group[28]. A recent example is Ref. [56] where an architecture is presented which requires $9n + 2$ qubits and has a depth of about $2000n^2$.

Unfortunately the above estimates for computational speedup are completely unrealistic due to the need to include error correction to mitigate errors and loss of coherence during the computation. The overhead from error correction will likely be many orders of magnitude in both circuit size and depth. We will discuss error correction later in the course.

Chapter 4

Entanglement

The presence of entanglement separates classical from quantum states. Schrödinger referred to entanglement as the defining property of quantum systems already in the 1930s[7, 8, 9]. Nevertheless, it is not fully understood to what extent entanglement is required for the computational speedup of quantum systems. From a practical point of view it is useful to have metrics which allow us to verify the quantum character of some processor we are developing. Methods to detect the presence of entanglement are therefore important. This is a solved problem for two-qubit (bipartite) systems. For more than two qubits there remain open questions about the detection and classification of entanglement.

4.1 Density matrix theory

When we have a statistical mixture of states the state vector $|\psi\rangle$ is equal to one of the set $\{|\psi_i\rangle\}$ with probability P_i. If the state is described by any one of the $|\psi_i\rangle$ the quantum mechanical expectation value of an operator $\hat{O}$ is simply $\langle\hat{O}\rangle_i = \langle\psi_i|\hat{O}|\psi_i\rangle$. When we have a statistical mixture we define the expectation value of an operator as

$$\langle\hat{O}\rangle = \sum_i P_i\langle\hat{O}\rangle_i = \sum_i P_i\langle\psi_i|\hat{O}|\psi_i\rangle. \tag{4.1}$$

It should be emphasized that the probabilistic nature of the expectation value $\langle\hat{O}\rangle$ is not quantum mechanical in origin but arises from our imperfect knowledge of the state $|\psi\rangle$. This could for example be due to imperfect preparation of the state. In addition there is the quantum mechanical uncertainty due to the probabilistic interpretation of $\langle\hat{O}\rangle_i$. The probabilities satisfy $0 \leq P_i \leq 1$, $\sum_i P_i = 1$, and $\sum_i P_i^2 \leq 1$. A pure state refers to the situation where only one $P_i = 1$ and all the other P_i vanish. In this case $\langle\hat{O}\rangle = \langle\hat{O}\rangle_i$ and we recover our usual quantum mechanical result. If this is not the case we refer to the state as a mixed state.

In order to deal with situations where we only have statistical knowledge of the quantum state we introduce the density matrix defined by

$$\hat{\rho} = \sum_i P_i|\psi_i\rangle\langle\psi_i|.$$

For a pure state this reduces to $\hat{\rho} = |\psi\rangle\langle\psi|$. We now introduce a complete set of orthonormal basis states $|n\rangle$ using which we can write

$$|\psi_i\rangle = \hat{I}|\psi_i\rangle = \left(\sum_n |n\rangle\langle n|\right)|\psi_i\rangle = \sum_n \langle n|\psi_i\rangle|n\rangle = \sum_n c_{in}|n\rangle$$

with $c_{in} = \langle n|\psi_i\rangle$. The trace of the density matrix is the sum of the diagonal components which is

$$\begin{aligned}
\mathrm{Tr}[\hat{\rho}] &= \sum_n \langle n|\hat{\rho}|n\rangle \\
&= \sum_n \langle n| \sum_i P_i|\psi_i\rangle\langle\psi_i|n\rangle \\
&= \sum_i P_i \sum_{n,n',n''} c_{in'} c^*_{in''} \langle n|n'\rangle\langle n''|n\rangle \\
&= \sum_i P_i \sum_n c_{in} c^*_{in} \\
&= \sum_i P_i = 1. \qquad (4.2)
\end{aligned}$$

which corresponds to conservation of probability. Some important properties of density matrices are $\hat{\rho} = \hat{\rho}^\dagger$, and $\langle j|\hat{\rho}^2|i\rangle \leq \langle j|\hat{\rho}|i\rangle$. The equality holds for pure states with $\hat{\rho}^2 = \hat{\rho}$ in which case the density matrix is referred to as idempotent.

The expectation value of an arbitrary operator $\hat{O}$ is given by

$$\begin{aligned}
\langle\hat{O}\rangle &= \sum_i P_i \langle\psi_i|\hat{O}|\psi_i\rangle \\
&= \sum_i P_i \sum_{n',n''} \langle\psi_i|(|n'\rangle\langle n'|)\hat{O}(|n''\rangle\langle n''|)|\psi_i\rangle \\
&= \sum_i P_i \sum_{n',n''} \langle\psi_i|n'\rangle\langle n'|\hat{O}|n''\rangle\langle n''|\psi_i\rangle \\
&= \sum_i P_i \sum_{n',n''} \langle n''|\psi_i\rangle\langle\psi_i|n'\rangle\langle n'|\hat{O}|n''\rangle \\
&= \sum_{n''} \langle n''| \left(\sum_i P_i|\psi_i\rangle\langle\psi_i|\right) \sum_{n'} |n'\rangle\langle n'|\hat{O}|n''\rangle \\
&= \sum_{n''} \langle n''|\hat{\rho} \sum_{n'} |n'\rangle\langle n'|\hat{O}|n''\rangle \\
&= \sum_{n''} \langle n''|\hat{\rho}\hat{O}|n''\rangle \\
&= \mathrm{Tr}\left[\hat{\rho}\hat{O}\right].
\end{aligned}$$

It can be shown that the trace is unchanged with cyclic reordering of the operators so for any operator

$$\langle\hat{O}\rangle = \mathrm{Tr}\left[\hat{\rho}\hat{O}\right] = \mathrm{Tr}\left[\hat{O}\hat{\rho}\right]. \qquad (4.3)$$

Equation (4.3) defines the expectation value of an operator when the state is only known statistically. The equation of motion for the density matrix is

$$\frac{d\hat{\rho}}{dt} = \frac{i}{\hbar}[\hat{\rho}, \hat{\mathcal{H}}]. \tag{4.4}$$

This can be derived by plugging the definition of $\hat{\rho}$ into the Schrödinger equation. Comparing with the equation for the time evolution of a Heisenberg frame operator it can be seen that Eq. (4.4) is of the same form but differs by a minus sign.

A physically admissible density matrix must satisfy three conditions

$$\mathrm{Tr}[\rho] = 1, \tag{4.5a}$$
$$\rho^\dagger = \rho, \tag{4.5b}$$
$$\langle\psi|\rho|\psi\rangle \geq 0 \;\forall\; |\psi\rangle. \tag{4.5c}$$

The first condition is conservation of probability, the second says ρ is Hermitian which is apparent from the definition of a mixed state density operator $\rho = \sum_i P_i|\psi_i\rangle\langle\psi_i|$, with the P_i non-negative real numbers, and the third condition says that ρ is a positive operator. This last condition follows from

$$\langle\psi|\rho|\psi\rangle = \langle\psi|\left(\sum_i P_i|\psi_i\rangle\langle\psi_i|\right)|\psi\rangle = \sum_i P_i\langle\psi|\psi_i\rangle\langle\psi|\psi_i\rangle^* = \sum_i P_i|\langle\psi|\psi_i\rangle|^2 \geq 0.$$

A density operator describing the quantum state in a D dimensional basis can be represented as a $D \times D$ matrix with D^2 elements. There are D real entries on the diagonal and $D^2 - D$ off-diagonal complex coherences giving a total of $D + 2(D^2 - D)$ real values. There is one constraint from $\mathrm{Tr}[\rho] = 1$ and $2(D^2 - D)\frac{1}{2} = D^2 - D$ constraints from $\rho^\dagger = \rho$. The density matrix can thus be described by $D + 2(D^2 - D) - 1 - (D^2 - D) = D - 1 + D^2 - D = D^2 - 1$ independent real parameters. If the quantum system of interest is N spin 1/2 particles then $D = 2^N$ and the density matrix has $2^{2N} - 1$ independent real parameters. An arbitrary bipartite density matrix of two spins can thus be described by $2^4 - 1 = 15$ real parameters.

Let's now look at a few examples. Consider a single atom with two levels in a pure state. The density matrix is $\hat{\rho} = |\psi\rangle\langle\psi|$ and using $|\psi\rangle = c_g|g\rangle + c_e|e\rangle$ gives

$$\hat{\rho} = \begin{pmatrix} |c_g|^2 & c_g c_e * \\ (c_g c_e^*)^* & |c_e|^2 \end{pmatrix}.$$

The maximum possible value of the off-diagonal entries occurs for $|c_g| = |c_e| = 1/\sqrt{2}$ giving an off-diagonal entry with magnitude 1/2. We note that there is a useful geometric parameterization of the density operator for a single qubit in the form

$$\hat{\rho} = \frac{\hat{I} + \mathbf{v}\cdot\hat{\boldsymbol{\sigma}}}{2} \tag{4.6}$$

with

$$v_x = 2\mathrm{Re}(c_g c_e^*), \quad v_y = -2\mathrm{Im}(c_g c_e^*), \quad v_z = |c_g|^2 - |c_e|^2.$$

If the atom is in a coherent superposition of ground and excited states $|\psi\rangle = (1/\sqrt{2})(|g\rangle + |e\rangle)$, and the density matrix is

$$\hat{\rho} = \sum_i P_i |\psi_i\rangle\langle\psi_i| = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

The nonzero off-diagonal elements are referred to as coherences. They show that the expectation value of for example the dipole operator $\hat{\mathbf{d}} = e\hat{\mathbf{r}}$ will be nonzero. On the other hand an atom that is prepared in an incoherent mixture of ground and excited states has density matrix

$$\hat{\rho} = \sum_i P_i |\psi_i\rangle\langle\psi_i| = \frac{1}{2}|g\rangle\langle g| + \frac{1}{2}|e\rangle\langle e| = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}.$$

The off-diagonal elements are zero and there is no coherence. Note that measurements of the probabilities of the ground and excited states do not distinguish between the two cases.

As a second example of a mixed state consider a two-level atom in thermal equilibrium at temperature T. The probability of occupation of state $|j\rangle$ is proportional to $\langle e^{-\hat{\mathcal{H}}/k_BT}\rangle = e^{-E_j/k_BT}$. Taking the energy of the ground state to be 0 and the energy of the excited state to be $\hbar\omega_a$ the correctly normalized density operator is

$$\begin{aligned} \hat{\rho} &= \frac{e^{-\hat{\mathcal{H}}/k_BT}}{\mathrm{Tr}\left[e^{-\hat{\mathcal{H}}/k_BT}\right]} \\ &= \frac{1}{1+e^{-\hbar\omega_a/k_BT}} \begin{pmatrix} 1 & 0 \\ 0 & e^{-\hbar\omega_a/k_BT} \end{pmatrix} \\ &= \frac{1}{2\cosh(\hbar\omega_a/2k_BT)} \begin{pmatrix} e^{\hbar\omega_a/2k_BT} & 0 \\ 0 & e^{-\hbar\omega_a/2k_BT} \end{pmatrix}. \end{aligned}$$

In the case of a sample of N atoms the wavefunction can be written as

$$|\psi\rangle = c_0|1_g 2_g ... N_g\rangle + (c_1|1_e 2_g ... N_g\rangle + ...) + C_{N+2}|1_e 2_e ... N_g\rangle + ... + C_{2^N-1}|1_e 2_e ... N_e\rangle.$$

When N is large this is a very complicated wavefunction with 2^N coefficients. When the atoms are uncorrelated, that is to say the state of any atom is not dependent on the states of the other atoms, the total density matrix is just the product of the density matrices of the individual atoms,

$$\rho = \prod_i \rho^{(i)} = \rho^{(1)} \otimes \rho^{(2)} \otimes .... \otimes \rho^{(N)}.$$

A system of N two-level atoms is described by a composite density matrix of dimensions $2^N \times 2^N$. For N large we have approximately $2^{2N}/2$ independent complex quantities which highlights the difficulty of solving quantum dynamical many body problems.

4.1.1 Composite systems

Even when there is no classical uncertainty density matrices play an essential role in describing the properties of composite systems. Consider a qubit in a pure state $|\psi\rangle = a_0|0\rangle + a_1|1\rangle$.

The expectation value of an arbitrary Hermitian operator U is

$$\begin{aligned}\langle U\rangle &= \langle\psi|U|\psi\rangle \\ &= \begin{pmatrix}a_0^* a_1^*\end{pmatrix}\begin{pmatrix}U_{00} & U_{01}\\ U_{10} & U_{11}\end{pmatrix}\begin{pmatrix}a_0\\ a_1\end{pmatrix} \\ &= |a_0|^2 U_{00} + |a_1|^2 U_{11} + a_0 a_1^* U_{10} + a_0^* a_1 U_{01}.\end{aligned}$$

Since $U_{01} = U_{10}^*$ this simplifies to

$$\langle U\rangle_1 = |a_0|^2 U_{00} + |a_1|^2 U_{11} + 2\mathrm{Re}\left(a_0 a_1^* U_{10}\right). \tag{4.7}$$

When there are two qubits the most general pure state is

$$|\psi\rangle = a_{00}|00\rangle + a_{01}|01\rangle + a_{10}|10\rangle + a_{11}|11\rangle.$$

The expectation value of an operator U acting on the first qubit is

$$\begin{aligned}\langle U\rangle_2 &= \langle\psi|U\otimes I|\psi\rangle \\ &= \left(|a_{00}|^2 + |a_{01}|^2\right)U_{00} + \left(|a_{10}|^2 + |a_{11}|^2\right)U_{11} + 2\mathrm{Re}\left[\left(a_{00}a_{10}^* + a_{01}a_{11}^*\right)U_{10}\right].\end{aligned} \tag{4.8}$$

Clearly Eqs. (4.7, 4.8) are not the same. In general we need to use the second expression to find the expectation value, the exception being when the two qubits are in a separable state. As we will show below the expectation value (4.8) can be written as

$$\langle U\rangle_2 = \mathrm{Tr}\left[U\rho_A\right]$$

where

$$\rho_A = \mathrm{Tr}_B\left[\rho\right]$$

is the reduced density matrix of the first subsystem.

Consider two subsystems A, B in the product state

$$|\psi\rangle = |\phi\rangle_A \otimes |\chi\rangle_B.$$

Here $\otimes$ denotes the tensor product. We will often omit the $\otimes$ symbol and write

$$|\psi\rangle = |\phi\rangle_A|\chi\rangle_B \quad \text{or} \quad |\psi\rangle = |\phi_A\chi_B\rangle \quad \text{or} \quad |\psi\rangle = |\phi\chi\rangle.$$

In this last version the subsystems are identified by their ordering in the ket. The corresponding bra is

$$\langle\psi| = |\psi\rangle^\dagger = \langle\phi|_A\langle\chi|_B \quad \text{or} \quad \langle\psi| = \langle\phi_A, \chi_B| \quad \text{or} \quad \langle\psi| = \langle\phi\chi|.$$

Unfortunately the other convention is also in use for which

$$\langle\psi| = \langle\chi|_B\langle\phi|_A = \langle\chi\phi|.$$

In many cases the ordering will be clear from the context. If not, then it is a good idea to be explicit by using subscripts to indicate the subsystems.

Using the partial trace the density matrix of the joint state ρ_{AB} can be reduced to give

$$\rho_A = \mathrm{Tr}_B(\rho_{AB}) \quad \text{or} \quad \rho_B = \mathrm{Tr}_A(\rho_{AB}).$$

The reduced density matrices ρ_A or ρ_B encapsulate what we know about the subsystems individually. In many applications of density matrices the composite system may consist of a material object A such as an electron or an atom or a piece of condensed matter containing many atoms. The second half of the system B may be the set of oscillator modes describing a radiation field that is couped to the matter. We are often interested in the state and the dynamics of the matter, but have only statistical knowledge of the radiation field B. In this case the result of measurements performed on the matter are found using the reduced density matrix $\hat{\rho}_A$.

As an example of a basic composite system consider the two-particle Bell state

$$|\psi\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}}. \tag{4.9}$$

If we measure one of the particles individually we will get the answer 0 or 1 with 50% probability. However, we will also know that a measurement of the other particle would reveal the opposite state. This type of behavior, where there is no local certainty, but strong two-particle correlations is characteristic of entanglement, and is revealed by contrasting the full density matrix with the reduced density matrices. We find

$$\begin{aligned}
\rho_{AB} &= |\psi\rangle\langle\psi| \\
&= \frac{|01\rangle + |10\rangle}{\sqrt{2}} \frac{\langle 01| + \langle 10|}{\sqrt{2}} \\
&= \frac{|01\rangle\langle 01| + |01\rangle\langle 10| + |10\rangle\langle 01| + |10\rangle\langle 10|}{2} \\
&= \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}
\end{aligned}$$

which displays the maximum possible coherence. Indeed the Bell state (4.9) is a maximally entangled state of two, two-level objects.

On the other hand

$$\begin{aligned}
\rho_A &= \mathrm{Tr}_B(\rho_{AB}) \\
&= \sum_x {}_B\langle x| \frac{|01\rangle\langle 01| + |01\rangle\langle 10| + |10\rangle\langle 01| + |10\rangle\langle 10|}{2} |x\rangle_B \\
&= \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} \\
&= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
\end{aligned}$$

is a mixed state with no coherence.

For reference let's calculate the partial trace for an arbitrary mixed state of two spin 1/2 objects. The most general 4×4 matrix is

$$\rho_{AB} = \begin{pmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \\ m & n & o & p \end{pmatrix}.$$

Of course the coefficients have to satisfy the conditions (4.5) for this to be a valid density matrix. The reduced density matrices are

$$\rho_A = \mathrm{Tr}_B(\rho_{AB}) = \begin{pmatrix} a+f & c+h \\ i+n & k+p \end{pmatrix}$$

and

$$\rho_B = \mathrm{Tr}_A(\rho_{AB}) = \begin{pmatrix} a+k & b+l \\ e+o & f+p \end{pmatrix}.$$

4.2 Density matrix properties

The most general bipartite density matrix composed from subsystems A and B, each having a two-dimensional basis, can be written in the form

$$\rho = \begin{pmatrix} P_{00} & B_1 & A_1 & C_1 \\ B_1^* & P_{01} & C_2 & A_2 \\ A_1^* & C_2^* & P_{10} & B_2 \\ C_1^* & A_2^* & B_2^* & P_{11} \end{pmatrix}. \tag{4.10}$$

Here P_{ij} are real populations, and the $A_1, A_2, B_1, B_2, C_1, C_2$ are complex coherences. We have written ρ in the computational basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ so A_1, A_2 correspond to coherences of subsystem A, B_1, B_2 are coherences of subsystem B, and C_1, C_2 are coherences that involve both subsystems.

A physically admissible density matrix must satisfy three conditions

$$\begin{aligned} &\mathrm{Tr}[\rho] = 1 \\ &\rho^\dagger = \rho \\ &\langle\psi|\rho|\psi\rangle \geq 0 \ \forall\ |\psi\rangle. \end{aligned} \tag{4.11}$$

The first condition is conservation of probability, the second says ρ is Hermitian which is apparent from the definition of a mixed state density operator $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$, with the p_i non-negative real numbers, and the third condition says that ρ is a positive operator. This last condition follows from

$$\langle\psi|\rho|\psi\rangle = \langle\psi| \left(\sum_i p_i |\psi_i\rangle\langle\psi_i| \right) |\psi\rangle = \sum_i p_i (\langle\psi|\psi_i\rangle)(\langle\psi|\psi_i\rangle)^* = \sum_i p_i |\langle\psi|\psi_i\rangle|^2 \geq 0.$$

An arbitrary bipartite density matrix can thus be described by $2^4 - 1 = 15$ real parameters.

4.3 Definition of entanglement

In simple cases it is apparent whether or not a state is entangled. The two-qubit state

$$|\psi\rangle = |00\rangle$$

is clearly a separable product state. The Bell state

$$|\psi\rangle = \frac{|01\rangle + |10\rangle}{\sqrt{2}}$$

is clearly a nonseparable state. However for other states such as

$$|\psi\rangle = a|00\rangle + b\frac{|01\rangle + |10\rangle}{\sqrt{2}}$$

with $|a|^2 + |b|^2 = 1$ it is not immediately obvious if $|\psi\rangle$ is separable or not. In this section we give a definition of entanglement that can be used for both pure and mixed bipartite states.

If the density matrix can be decomposed as a convex sum

$$\rho = \sum_i p_i \rho_A^i \otimes \rho_B^i \tag{4.12}$$

with $\sum_i p_i = 1$ then ρ is not an entangled state, but it may exhibit nonclassical correlations[68]. When no representation of the form of Eq. (4.12) is possible then the two-qubit state is entangled. Another way of thinking about this is that states of the form (4.12) can be formed by local operations and classical communication (LOCC). Party A prepares $\rho_A^1, \rho_A^2, ..$ and communicates to party B to prepare $\rho_B^1, \rho_B^2, ...$ and the two systems are combined with probabilities $p_1, p_2, .. .$ Conversely entangled states require more than LOCC for their creation, there must be a two-qubit quantum gate involved.

To emphasize the ability of classical correlations to mimic entanglement consider the following example. We prepare the product state

$$|\psi\rangle = \frac{|0\rangle_A + i|1\rangle_A}{\sqrt{2}}|0\rangle_B$$

and run it through a CNOT gate to create

$$|\psi\rangle \overset{CNOT}{\rightarrow} |\psi\rangle_{\text{ent}} = \frac{|00\rangle + i|11\rangle}{\sqrt{2}}$$

which is an entangled state. The corresponding density matrix is

$$\rho_{\text{ent}} = \begin{pmatrix} 1/2 & 0 & 0 & -i/2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ i/2 & 0 & 0 & 1/2 \end{pmatrix}. \tag{4.13}$$

If we measure the populations we get $P_{00} = 1/2, P_{01} = 0, P_{10} = 0, P_{11} = 1/2$ which looks like an entangled state. However a two-qubit density matrix with the same populations but no coherence can be decomposed in the form of Eq. (4.12) as

$$\rho_{\rm sep} = \frac{1}{2}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}_A \otimes \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}_B + \frac{1}{2}\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}_A \otimes \begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}_B = \begin{pmatrix}1/2 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 1/2\end{pmatrix}$$

which is a classically correlated mixture of both atoms being in $|0\rangle$ with 50% probability or both atoms being in $|1\rangle$ with 50% probability. This is a separable state which cannot be distinguished from the entangled state (4.13) by measuring populations alone.

4.4 Quantum correlations

There are states that are not entangled but have non-classical, i.e. quantum, correlations. Such states have the decomposition

$$\rho = \sum_{ij} p_{ij}|a_i\rangle\langle a_i| \otimes |b_j\rangle\langle b_j| = \sum_i |a_i\rangle\langle a_i| \otimes \sum_j p_{ij}|b_j\rangle\langle b_j|. \tag{4.14}$$

We may think of Eq. (4.14) as implying that the $i^{\rm th}$ realization of subsystem A may be correlated with multiple realizations of B in different bases.

Here is a simple example

$$\rho = \frac{1}{2}|0\rangle_a\langle 0| \otimes |0\rangle_b\langle 0| + \frac{1}{2}|1\rangle_a\langle 1| \otimes |+\rangle_b\langle +|$$

where $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$. This is a separable state which can be expanded to

$$\rho = \frac{1}{2}|0\rangle_a\langle 0| \otimes |0\rangle_b\langle 0| + \frac{1}{4}|1\rangle_a\langle 1| \otimes (|0\rangle_b\langle 0| + |0\rangle_b\langle 1| + |1\rangle_b\langle 0| + |1\rangle_b\langle 1|)\,.$$

The density matrix in the basis $(00, 01, 10, 11)$ is

$$\rho = \frac{1}{4}\begin{pmatrix}2 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 1\\ 0 & 0 & 1 & 1\end{pmatrix}.$$

It is unclear to what extent these states are important or not for understanding the computational power of quantum systems.

Chapter 5

Entanglement measures

Entanglement can be detected for bipartite systems in several different ways. We describe some standard methods below. See the review articles by Gühne and Toth[69] and the Horodeckis[70] for more details and other methods.

5.1 Positive partial Transpose

A known method is based on the partial transpose which is positive definite for separable states[71, 72]. Let the density matrix be

$$\rho = \sum_{ij\mu\nu} p_{ij\mu\nu} |i\rangle\langle j| \otimes |\mu\rangle\langle\nu|$$

where i, j refer to subsystem A and μ, ν refer to subsystem B. The partial transpose is

$$\rho^{T_B} = I \otimes T_B(\rho) = \sum_{ij\mu\nu} p_{ij\mu\nu} |i\rangle\langle j| \otimes (|\mu\rangle\langle\nu|)^T = \sum_{ij\mu\nu} p_{ij\mu\nu} |i\rangle\langle j| \otimes (|\nu\rangle\langle\mu|).$$

We can also write this as $\rho^{T_B} = (\hat{I} \otimes \hat{T})\rho(\hat{I} \otimes \hat{T})^\dagger$ where $\hat{T}$ represents a transpose operation.

For a 2×2 system the density operator is 4×4

$$\rho = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

where A_{ij} are 2×2 matrices. Then the partial transpose is

$$\rho^{T_B} = \begin{pmatrix} A_{11}^T & A_{12}^T \\ A_{21}^T & A_{22}^T \end{pmatrix}.$$

If ρ is separable then ρ^{T_B} is guaranteed to have non-negative eigenvalues. Thus if ρ^{T_B} has a negative eigenvalue then ρ is guaranteed to be entangled. The partial transpose can be taken on either subsystem since $\rho^{T_A} = (\rho^{T_B})^T$. This test is sufficient for two qubits or two qutrits, (2×2 or 2×3 dimensional problems). In higher dimensions the results are inconclusive. For a proof of this result see[71, 72].

5.2 von Neumann entropy

The von Neumann entropy of a density matrix ρ is

$$S = -\mathrm{Tr}[\rho \log_2 \rho] = -\sum_i \lambda_i \log \lambda_i$$

where λ_i are the eigenvalues of ρ. The entropy is nonnegative, $S \geq 0$.

The conditional entropy of two systems A, B is

$$S(A|B) = S(A,B) - S(B),$$

where $S(A,B)$ is the joint entropy of the density matrix of both systems ρ_{AB} and $S(B)$ is the entropy of the reduced density matrix $\rho_B = \mathrm{Tr}_A[\rho_{AB}]$. If the conditional entropy is negative the systems are entangled.

5.3 Entanglement of formation

For a pure state ψ the entanglement of a bipartite system is defined as

$$E(\psi) = -\mathrm{Tr}[\rho_A \log_2 \rho_A] = -\mathrm{Tr}[\rho_B \log_2 \rho_B]$$

where ρ_A, ρ_B are the partial traces of $\rho = |\psi\rangle\langle\psi|$ over the individual subsystems. A mixed state has density matrix

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

The entanglement of formation of a mixed state ρ is defined as

$$E_f(\rho) = \min \sum_i p_i E(\psi_i).$$

The entanglement of formation may be interpreted as the minimal number of maximally entangled singlets that is required to build a single copy of the state.

If we have access to all elements of ρ then we can find the eigenvalues and compute the entanglement of formation(Wootters 1998, 2001)[73]. Consider a bipartite system with density matrix ρ. The quantity $R = \sqrt{\sqrt{\rho}\tilde{\rho}\sqrt{\rho}}$, has eigenvalues $\lambda_1..\lambda_4$ ordered such that $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \lambda_4$. Here $\tilde{\rho} = (Y \otimes Y)\rho^*(Y \otimes Y)$.

The concurrence is defined as

$$\mathcal{C} = \max(0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4)$$

and the entanglement of formation is

$$E_f(\rho) = \mathcal{E}(C(\rho))$$

where

$$\mathcal{E}(C) = h\left(\frac{1+\sqrt{1-C^2}}{2}\right)$$

and

$$h(x) = -x \log_2(x) - (1-x)\log_2(1-x).$$

The function $\mathcal{E}(C)$ increases monotonically for $0 \leq C \leq 1$. A positive $\mathcal{E}(C)$ is a necessary and sufficient condition for the presence of entanglement.

5.4 Coherence criterion

There are entanglement measures which do not require knowledge of the full density matrix. For a bipartite system it is sufficient to measure the diagonal populations and the coherence C_1(Sackett 2000)[74]. An arbitrary separable pure state of two-atoms can be written as

$$|\psi\rangle = (a_0|0\rangle + a_1|1\rangle)_A \otimes (b_0|0\rangle + b_1|1\rangle)_B \,.$$

Equivalently an arbitrary separable bipartite density matrix describing a pure state can be written as

$$\rho = |\psi\rangle\langle\psi| = \begin{pmatrix} |a_0|^2|b_0|^2 & ... & ... & a_0 b_0 a_1^* b_1^* \\ ... & |a_0|^2|b_1|^2 & a_0 b_1 a_1^* b_0^* & ... \\ ... & a_0^* b_1^* a_1 b_0 & |a_1|^2|b_0|^2 & ... \\ a_0^* b_0^* a_1 b_1 & ... & ... & |a_1|^2|b_1|^2 \end{pmatrix}.$$

Elements written as ... are not needed in the rest of this argument.

The normalization conditions are $|a_0|^2 + |a_1|^2 = 1$ and $|b_0|^2 + |b_1|^2 = 1$ which can be combined to give

$$(|a_0| - |b_0|)^2 + (|a_1| - |b_1|)^2 + 2|a_0||b_0| + 2|a_1||b_1| = 2. \tag{5.1}$$

This implies

$$|a_0||b_0| + |a_1||b_1| \leq 1$$

or

$$(|a_0||b_0| + |a_1||b_1|)^2 \leq 1$$

which can be written as

$$|a_0|^2|b_0|^2 + |a_1|^2|b_1|^2 + 2|a_0||b_0||a_1||b_1| \leq 1.$$

Rewriting this inequality in the notation of Eq. (4.10) gives

$$P_{00} + P_{11} + 2|C_1| \leq 1. \tag{5.2}$$

Defining

$$F = \frac{1}{2}(P_{00} + P_{11}) + |C_1|$$

a necessary condition for the density matrix to represent a separable state is

$$F \leq 1/2. \tag{5.3}$$

Conversely if we create a density matrix and measure $F > 1/2$ then we know that we have a non-separable or entangled state. It is also possible that a state with $F \leq 1/2$ is entangled since (5.3) is a necessary but not sufficient condition for separability. It is important to recognize that if we only measure diagonal elements of the density matrix, and do not determine the coherence, then all we can say is that F is at most $1/2$ which does not prove entanglement. It is necessary to measure the coherence.

The extension of this argument to a classically correlated mixed state as in (4.12) is trivial. Writing $\rho = \sum_i p_i \rho^i$ we have from (5.2) that $P^i_{00} + P^i_{11} + 2|C^i_1| \leq 1$ and summing over the probabilities gives

$$\sum_i p_i \left[P^i_{00} + P^i_{11} + 2|C^i_1|\right] = P_{00} + P_{11} + 2|C_1| \leq \sum_i p_i = 1$$

or

$$P_{00} + P_{11} + 2|C_1| \leq 1 \tag{5.4}$$

where the variables in Eq. (5.4) now refer to the general separable mixed state density matrix.

For completeness note that we could have used, instead of (5.1), the equality

$$(|a_0| - |b_1|)^2 + (|a_1| - |b_0|)^2 + 2|a_0||b_1| + 2|a_1||b_0| = 2.$$

This implies

$$(|a_0||b_1| + |a_1||b_0|)^2 \leq 1$$

which can be written as

$$|a_0|^2|b_1|^2 + |a_1|^2|b_0|^2 + 2|a_0||b_0||a_1||b_1| \leq 1.$$

Rewriting this inequality in the notation of Eq. (4.10) gives

$$P_{01} + P_{10} + 2|C_2| \leq 1. \tag{5.5}$$

Thus we can equivalently define the fidelity as

$$F = \frac{P_{01} + P_{10}}{2} + |C_2|.$$

To summarize, for both pure and mixed states, we can distinguish between separable and entangled bipartite states on the basis of the 'Fidelity' F as follows:

$$\begin{aligned} F &\leq 1/2 \quad \text{separable or entangled state,} \\ 1/2 < F &\leq 1 \quad \text{entangled state.} \end{aligned}$$

For any physically allowed density matrix $P_{00} + P_{11} \leq 1$ and $C_1 \leq 1/2$ so it is necessary to have some contribution from both populations and coherence to cross the entanglement boundary.

5.4.1 Parity oscillations

The coherence can be measured by observing parity oscillations. Assume we apply a unitary rotation to each of the subsystems A, B of the form

$$R_0(\theta, \phi) = \begin{pmatrix} \cos(\theta/2) & ie^{-\imath\phi}\sin(\theta/2) \\ ie^{\imath\phi}\sin(\theta/2) & \cos(\theta/2) \end{pmatrix}.$$

This corresponds to a Rabi pulse with pulse area θ and phase ϕ. The density matrix evolves according to

$$\rho \to \rho' = R\rho R^\dagger = \begin{pmatrix} P'_{00} & B'_1 & A'_1 & C'_1 \\ B_1^{*'} & P'_{01} & C'_2 & A'_2 \\ A_1^{*'} & C_2^{*'} & P'_{10} & B'_2 \\ C_1^{*'} & A_2^{*'} & B_2^{*'} & P'_{11} \end{pmatrix}$$

where $R = R_{0A} \otimes R_{0B}$. Define a parity signal

$$P = P_{00} + P_{11} - P_{01} - P_{10}.$$

The parity of the transformed density matrix is

$$\begin{aligned} P' &= P'_{00} + P'_{11} - P'_{01} - P'_{10} \\ &= P\cos^2(\theta) - 2\mathrm{Re}[C_1 e^{2\imath\phi}]\sin^2(\theta) + 2\mathrm{Re}[C_2]\sin^2(\theta) + \mathrm{Im}[(A_1 - A_2 + B_1 - B_2)\, e^{\imath\phi}]\sin(2\theta). \end{aligned}$$

If we use $\theta = \pi/2$ then

$$P' = 2\mathrm{Re}[C_2] - 2\mathrm{Re}[C_1 e^{2\imath\phi}] = 2\mathrm{Re}[C_2] - 2|C_1|\cos(2\phi + \xi)$$

where ξ is the phase of the coherence C_1 ($C_1 = |C_1|e^{\imath\xi}$). We see that the parity oscillates with peak to peak amplitude given by $4|C_1|$. Measuring the amplitude of the parity oscillation thus determines $|C_1|$ which together with Eq. (5.2) can be used to verify entanglement. This approach was followed in ion trap experiments(Turchette 1998, Sackett2000)[75, 74].

The parity oscillations are due to all four populations oscillating. In the simplest case where $P_{01} = P_{10} = A_1 = A_2 = B_1 = B_2 = C_2 = 0$, $P_{11} = P_{00}$, and $C_1 = P_{00}$ the populations oscillate as

$$\begin{aligned} P'_{00} = P'_{11} = P_{00}\sin^2(\phi), \\ P'_{01} = P'_{10} = P_{00}\cos^2(\phi). \end{aligned}$$

Alternatively we might keep ϕ fixed and vary θ. This gives a more complicated parity signal that depends on θ even when C_1, C_2 vanish. If we, for example, set $\phi = 0$ then

$$P' = P\cos^2(\theta) - 2\mathrm{Re}[C_1 - C_2]\sin^2(\theta) + \mathrm{Im}[A_1 - A_2 + B_1 - B_2]\sin(2\theta). \tag{5.6}$$

This approach does not readily lend itself to unambiguous extraction of the coherence C_1.

5.5 Continuous variables

The methods described so far are suitable for bipartite, finite dimensional systems. There is another important situation that arises in entanglement of beams of light that are described by continuous variable wavefunctions. A suitable entanglement criterion for this case was formulated in [76, 77].

Chapter 6

Cats, EPR, and Bell

Already in the early years after the development of the quantum theory it became apparent that the interpretation of the quantum state was troublesome. For a careful discussion of this topic the book by Peres is an excellent resource[78]. Here we give a brief summary of some of the main points.

6.1 Schrödinger's cat

Schrödinger identified entanglement as the essential feature separating classical from quantum mechanics[7, 8, 9] and gave his famous example of a cat to illustrate the unreasonableness of quantum mechanics. He envisioned a box containing a cat described by a quantum state $|A\rangle$ for alive or $|D\rangle$ for dead. Inside the box was a vial of poison gas which would be broken open by a mechanism triggered by the decay of an excited atom. This construction led to the principle possibility of a superposition of either a live cat and an excited atom $|A, e\rangle$ or a dead cat and an atom in the ground state $|D, g\rangle$.

The quantum state should then be written as

$$|\psi\rangle \sim |A, e\rangle + |D, g\rangle.$$

Needless to say noone has ever observed cats, or other beings, in a superposition of alive and dead. The point of the thought experiment is that quantum mechanics, when applied to macroscopic objects, leads to seemingly absurd predictions. One way out of this conundrum is to subscribe to the belief that quantum mechanics does not apply to sufficiently large or complicated objects, and that there is a quantum-classical boundary at some scale. It is not known where such a boundary might exist and experimental advances keep revealing quantum phenomena in larger and larger objects. An example of this is the observation of interference fringes for buckyballs sent through a two-slit apparatus[79].

6.2 EPR

Einstein, Podolsky, and Rosen, in what is now known as the EPR paper[80] raised other objections. The considered a system with parts $1, 2$ and considered the position and momentum

operators of the two parts, $\hat{x}_{1,2}, \hat{p}_{1,2}$. Suppose the following joint state is prepared

$$|\psi\rangle \sim \delta(\hat{x}_1 - \hat{x}_2)\delta(\hat{p}_1 + \hat{p}_2).$$

The two parts are then separated and the position of particle 1 is measured resulting in a value a. The momentum of particle 2 is measured resulting in a value b. Quantum mechanics then predicts that a subsequent measurement of the momentum of particle 1, if it were performed, would yield the result $-b$. We see that in principle we thereby can know both the position and momentum of particle 1 without any uncertainty, which would seem to violate the Heisenberg uncertainty principle which is a cornerstone of quantum mechanics.

While not denying the success of quantum theory at predicting experimentally measured quantities with high accuracy, EPR characterized this paradox as being due to the incompleteness of quantum mechanics. Either, one should give up the notion of the wave function containing "elements of reality", or one should give up the notion of observations at one point in space and time only depending on phenomena occurring in a causally connected region. This may be referred to as Einstein locality. EPR proferred a possible solution to the paradox, while retaining Einstein locality, by introducing the idea of hidden variables which define the real physical state of a particle, but are not accounted for in the incomplete quantum theory. Another response to this "paradox" is to observe that it rests on assumptions about measurements that could be performed, but are not actually performed. As Peres has pointed out[81], *Unperformed experiments have no results.*

6.3 Bell

For many years these issues were relegated to discussions of a more philosophical than physics bent. In the last 50 years these questions have been brought into sharper focus both theoretically and experimentally such that the validity of classical vs. quantum descriptions of reality can now be studied in a quantitative fashion. In the 1950s Bohm reformulated the EPR argument in terms of dichotomic (binary) measurements of a spin 1/2 particle[82].

In the 1960s Bell, in a paper completed while visiting Madison[83], introduced his famous inequalities which provide a means of testing for the possible existence of hidden variables. Innumerable experiments have since shown conclusively that quantum phenomena violate the Bell inequalities thereby excluding the possibility of local hidden variables. It should be mentioned that until recently there were experimental loopholes related to fair sampling, space like separation, and "free will" that put the Bell inequality violations in doubt. The spacelike separation and fair sampling loopholes were first closed in separate experiments using photons in 1998[84] and trapped ions in 2000[85]. Only in 2015 were both loopholes closed simultaneously[86, 87, 88]. Interestingly all of these loophole free Bell experiments used random numbers generated by an optoelectronic device based on spontaneous emission[89].

The "free will" loophole relates to the possibility that the detector settings in a Bell test are somehow influenced by events prior to running the experiment that control the experimentalists choices. To a certain degree this has been closed separately, but not together with the other loopholes[90]. It has been proposed to definitively close this loophole using astrophysical photons from causally disconnected cosmic sources to determine the detector settings[91]. Such an experiment may yet be performed but, in my opinion, will be unlikely to deliver any surprises.

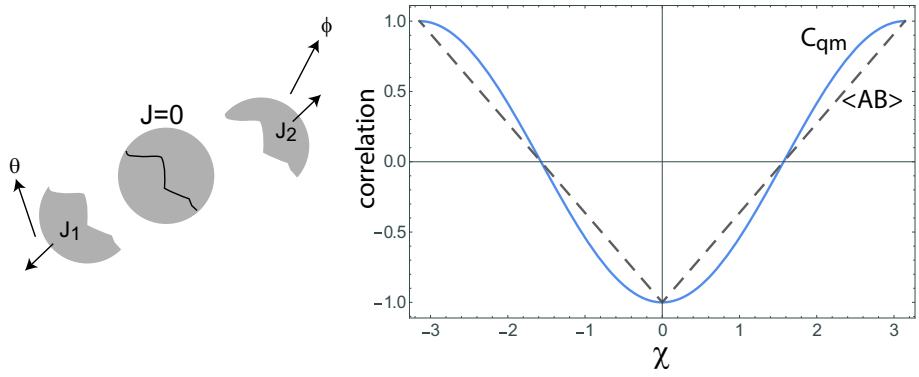


Figure 6.1: Comparison of quantum spin correlation from Eq. (6.1) and a classical model of correlated measurements on a particle splitting into two parts.

To proceed let's first calculate the correlation coefficient for measurements of two spin 1/2 particles in a singlet state $|S\rangle = \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}}$. This state can be produced using a quantum circuit, as we have discussed, or by a physical process such as parametric down conversion of photons or two-photon decay of an excited atom. Let us then imagine that particle 1 is measured by observer A and particle 2 is measured by observer B. Many copies of the singlet state are prepared and the observers measure the spin projection along directions $\boldsymbol{\theta}$ for observer A and $\boldsymbol{\phi}$ for observer B, with the directions assumed to be coplanar. Thus, given the joint state $\hat{\rho}$, they measure $A = \langle \hat{\sigma}_{\boldsymbol{\theta}} \rangle$ and $B = \langle \hat{\sigma}_{\boldsymbol{\phi}} \rangle$ where $\boldsymbol{\theta}, \boldsymbol{\phi}$ are unit vectors in the corresponding directions. Each measurement gives a dichotomic value $+1$ or -1. The expected value of the joint value of these measurements is, according to quantum mechanics,

$$C_{\rm qm} = \langle \hat{\sigma}_{\boldsymbol{\theta}} \hat{\sigma}_{\boldsymbol{\phi}} \rangle = -\cos(\theta - \phi). \tag{6.1}$$

This correlation is stronger than we would expect from a classical model. Following [81] consider an object, initially at rest with zero angular momentum, that breaks up into two parts carrying angular momenta $\mathbf{J}_1$ and $\mathbf{J}_2$. These two parts separate and the projection of the angular momenta are measured along directions $\boldsymbol{\theta}, \boldsymbol{\phi}$ respectively. The measurement results are assigned the values ± 1 according to $\text{sign}(\boldsymbol{\theta} \cdot \mathbf{J}_1)$ and $\text{sign}(\boldsymbol{\phi} \cdot \mathbf{J}_2)$. If the experiment is repeated N times and the angular momenta are randomly distributed in space, the expected values of these quantities tend to zero

$$\langle A \rangle = \frac{\sum_j \text{sign}(\boldsymbol{\theta} \cdot \mathbf{J}_{1j})}{N} \to 0, \quad \langle B \rangle = \frac{\sum_j \text{sign}(\boldsymbol{\phi} \cdot \mathbf{J}_{2j})}{N} \to 0.$$

However, the correlation need not be zero. If we take $\boldsymbol{\theta} = \boldsymbol{\phi}$ then $\langle AB \rangle = -1$. A geometrical argument[81] shows that if χ is the angle between $\boldsymbol{\theta}, \boldsymbol{\phi}$ then $\langle AB \rangle = -1 + 2|\chi|/\pi$. This correlation is compared with the quantum result $C_{\rm qm}$ in Fig. 6.1. We see that the magnitude of the quantum correlation is always larger than the classical correlation. This goes against the superstition that quantum phenomena are more uncertain or random than classical phenomena. In fact the quantum correlations are stronger.

The above comparison of quantum and classical correlations used a specific model. The importance of the Bell inequalities is that they are formulated in terms of a very general

model of hidden variables λ. Let us now see what correlation we predict using classical theory augmented by some hidden variables λ. In the presence of hidden variables λ

$$A = A(\boldsymbol{\theta}, \lambda), \qquad B = B(\boldsymbol{\phi}, \lambda).$$

We assume locality and therefore exclude a dependence $A = A(\boldsymbol{\theta}, \boldsymbol{\phi}, \lambda)$ or $B = B(\boldsymbol{\phi}, \boldsymbol{\theta}, \lambda)$. Let the hidden variables satisfy some normalized probability distribution

$$\int d\lambda\, P(\lambda) = 1.$$

Then the predicted value of the correlation coefficient, allowing for hidden variables, is

$$C_{\rm hv}(\boldsymbol{\theta}, \boldsymbol{\phi}) = \int d\lambda\, P(\lambda) A(\boldsymbol{\theta}, \lambda) B(\boldsymbol{\phi}, \lambda).$$

Remember that the individual measurement results A, B are all ± 1. Let us consider four measurement results corresponding to different settings of the angles: $A = A(\boldsymbol{\theta}, \lambda)$, $A' = A(\boldsymbol{\theta}', \lambda)$, $B = B(\boldsymbol{\phi}, \lambda)$, $B' = B(\boldsymbol{\phi}', \lambda)$. It is then an algebraic identity that

$$S = AB + AB' + A'B - A'B' = A(B + B') + A'(B - B') = \pm 2.$$

Therefore it must be true that $-2 \le S_{\rm hv} \le 2$ or

$$-2 \le C_{\rm hv}(\boldsymbol{\theta}, \boldsymbol{\phi}) + C_{\rm hv}(\boldsymbol{\theta}, \boldsymbol{\phi}') + C_{\rm hv}(\boldsymbol{\theta}', \boldsymbol{\phi}) - C_{\rm hv}(\boldsymbol{\theta}', \boldsymbol{\phi}') \le 2,$$

for any values of $\boldsymbol{\theta}, \boldsymbol{\theta}', \boldsymbol{\phi}, \boldsymbol{\phi}'$. This form of the Bell inequality was introduced by CHSH[92].

According to the predictions of quantum mechanics this inequality can be violated in an experiment. Using $C_{\rm qm} = -\cos(\theta - \phi)$, which we calculated above, and the angles $\theta = 0, \theta' = \pi/2, \phi = \pi/4, \phi' = -\pi/4$ we find

$$S_{\rm qm} = -2\sqrt{2} \simeq -2.83 < -2.$$

A popular exposition of such a situation is given in the quantum baking paper of Kwiat and Hardy[93]. The value $|S_{\rm qm}| = 2\sqrt{2}$ is the maximum possible violation of the CHSH inequality assuming standard quantum mechanics. This is known as the Cirel'son bound[94].

The CHSH form of the Bell inequality can be violated by an entangled state, but entanglement alone is not sufficient. There are states that are entangled but have $|S_{\rm qm}| < 2$. Consider the state

$$\begin{aligned} \hat{\rho} &= \alpha |S\rangle\langle S| + \frac{1-\alpha}{2} \left(|\uparrow\uparrow\rangle\langle\uparrow\uparrow| + |\downarrow\downarrow\rangle\langle\downarrow\downarrow| \right) \\ &= \frac{1}{2} \begin{pmatrix} 1-\alpha & 0 & 0 & 0 \\ 0 & \alpha & -\alpha & 0 \\ 0 & -\alpha & \alpha & 0 \\ 0 & 0 & 0 & 1-\alpha \end{pmatrix}. \end{aligned} \tag{6.2}$$

For $\alpha = 1$ this is a maximally entangled singlet state which saturates the Cirel'son bound, while for $\alpha = 0$ this is a non-entangled but classically correlated state of the form (4.12).

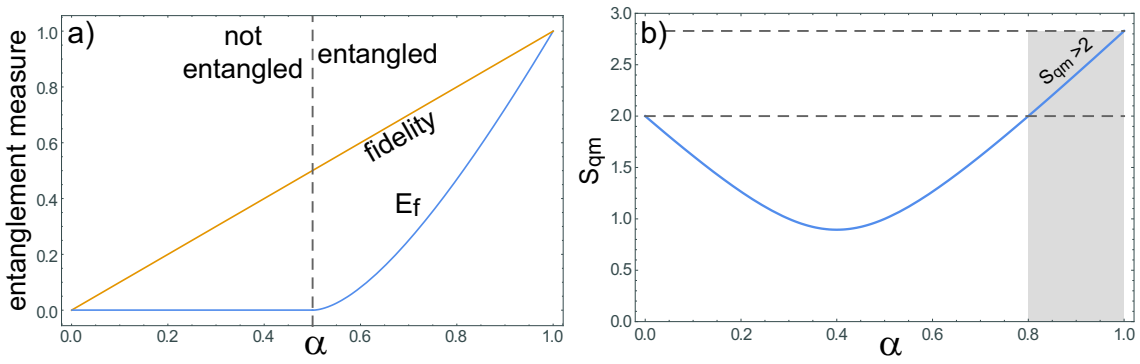


Figure 6.2: Entanglement of formation and fidelity a) and $S_{\rm qm}$ parameter b) for the state (6.2) as a function of α. The $S_{\rm qm}$ parameter was maximized over all possible settings of $\theta, \theta', \phi, \phi'$ at each value of α.

The entanglement as a function of α is shown in Fig. 6.2a. We see that the state is entangled for $\alpha > 1/2$. The spin correlation is

$$C_{\rm qm} = \langle \hat{\sigma}_{\boldsymbol{\theta}} \hat{\sigma}_{\boldsymbol{\phi}} \rangle = (1-2\alpha)\cos(\theta)\cos(\phi) - \alpha \sin(\theta)\sin(\phi).$$

Maximizing $S_{\rm qm}$ over angles we obtain the curve shown in Fig. 6.2b. Although the state is entangled for $\alpha > 1/2$ the Bell inequality is only violated for $\alpha > 0.8$. The converse is readily shown: any state that violates the Bell inequality is entangled. This follows immediately from the fact that for an unentangled state $C_{\rm qm}$ factors into $\sum_j p_j \langle \hat{\sigma}_{\boldsymbol{\theta}} \rangle_j \langle \hat{\sigma}_{\boldsymbol{\phi}} \rangle_j$. Each term in the sum is the product of quantities that are bounded by ± 1 and we get a CHSH parameter that is bounded by the classical limit of ± 2. Therefore violation of a Bell inequality implies that the state was entangled. Such a violation is a stronger condition than entanglement, and has practical uses for ensuring the security of quantum key distribution[95].

Chapter 7

Multiqubit entangled states

There are several important classes of multipartite entangled states. These include the N particle states

$$\begin{aligned} |GHZ\rangle_N &= \frac{|00...0\rangle + |11...1\rangle}{\sqrt{2}}, \\ |W\rangle_N &= \frac{1}{\sqrt{N}}\sum_{j=1}^{N}|01_j...0\rangle = \frac{1}{\sqrt{N}}\sum_{j=1}^{N}|\bar{1}_j\rangle_N. \end{aligned}$$

In the last expression we introduce the notation $|\bar{1}_j\rangle_N$ for the N particle ket with only element j in state $|1\rangle$. We will denote the state with all N particles in $|x\rangle$ by $|x\rangle_N$ and $|x_j\rangle$ without a subscript exterior to the ket will refer to the state of particle j. The GHZ state has maximal phase sensitivity and is important for atomic clocks. The W state can be readily prepared using Rydberg blockade. These are inequivalent multiparticle entangled states that cannot be transformed into each other with local operations and classical communication (LOCC)[96]. While the GHZ state has maximal phase sensitivity it is also extremely fragile. Tracing over a single particle leaves

$$\begin{aligned} \hat{\rho}_{N-1} &= \mathrm{Tr}_1[\hat{\rho}_N^{(\mathrm{GHZ})}] \\ &= \frac{1}{2}\,{}_1\langle 0|\left(|0^{\otimes N}\rangle\langle 0^{\otimes N}| + |0^{\otimes N}\rangle\langle 1^{\otimes N}| + |1^{\otimes N}\rangle\langle 0^{\otimes N}| + |1^{\otimes N}\rangle\langle 1^{\otimes N}|\right)|0\rangle_1 \\ &+ \frac{1}{2}\,{}_1\langle 1|\left(|0^{\otimes N}\rangle\langle 0^{\otimes N}| + |0^{\otimes N}\rangle\langle 1^{\otimes N}| + |1^{\otimes N}\rangle\langle 0^{\otimes N}| + |1^{\otimes N}\rangle\langle 1^{\otimes N}|\right)|1\rangle_1 \\ &= \frac{1}{2}|0^{\otimes N-1}\rangle\langle 0^{\otimes N-1}| + \frac{1}{2}|1^{\otimes N-1}\rangle\langle 1^{\otimes N-1}|. \end{aligned}$$

This is a completely unentangled mixed state of N particles.

The W state behaves very differently under particle loss. Tracing over the last particle the reduced density matrix is

$$\begin{aligned} \hat{\rho}_{N-1} &= \mathrm{Tr}_N[\hat{\rho}_N^{(\mathrm{W})}] \\ &= \frac{1}{N}\langle 0_N|\left(\sum_{j=1}^{N}|\bar{1}_j\rangle_N \sum_{k=1}^{N}{}_N\langle\bar{1}_k|\right)|0_N\rangle + \frac{1}{N}\langle 1_N|\left(\sum_{j=1}^{N}|\bar{1}_j\rangle_N \sum_{k=1}^{N}{}_N\langle\bar{1}_k|\right)|1_N\rangle. \end{aligned}$$

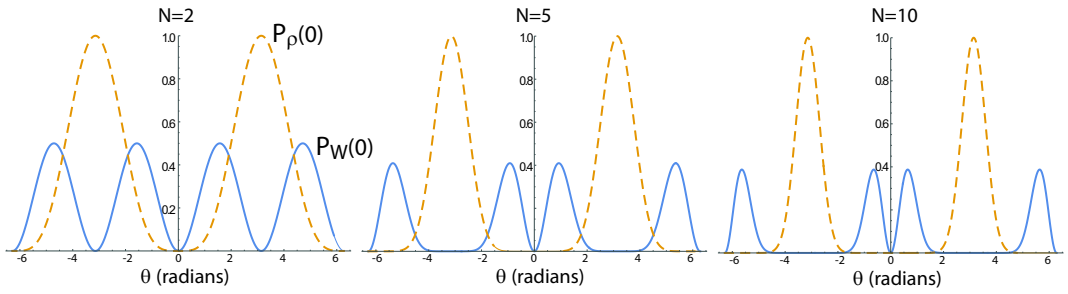


Figure 7.1: Probability of measuring $|\bar{0}\rangle_N$ after rotating the $|W\rangle$ state and the nonentangled singly excited state for $N = 2, 5, 10$ qubits.

The expression in parentheses can be written as

$$\begin{aligned}
\sum_{j=1}^{N} |\bar{1}_j\rangle_N \sum_{k=1}^{N} {}_N\langle\bar{1}_k| &= \left(\sum_{j=1}^{N-1} |\bar{1}_j\rangle_{N-1} \sum_{k=1}^{N-1} {}_{N-1}\langle\bar{1}_k|\right) |0_N\rangle\langle 0_N| + \left(\sum_{j=1}^{N-1} |\bar{1}_j\rangle_{N-1}\, {}_{N-1}\langle\bar{0}|\right) |0_N\rangle\langle 1_N| \\
&+ \left(|\bar{0}\rangle_{N-1} \sum_{j=1}^{N-1} {}_{N-1}\langle\bar{1}_j|\right) |1_N\rangle\langle 0_N| + \left(|\bar{0}\rangle_{N-1}\, {}_{N-1}\langle\bar{0}|\right) |1_N\rangle\langle 1_N| \\
&= (N-1)\hat{\rho}_{N-1}^{(W)} |0_N\rangle\langle 0_N| + \left(\sum_{j=1}^{N-1} |\bar{1}_j\rangle_{N-1}\, {}_{N-1}\langle\bar{0}|\right) |0_N\rangle\langle 1_N| \\
&+ \left(|\bar{0}\rangle_{N-1} \sum_{j=1}^{N-1} {}_{N-1}\langle\bar{1}_j|\right) |1_N\rangle\langle 0_N| + \left(|\bar{0}\rangle_{N-1}\, {}_{N-1}\langle\bar{0}|\right) |1_N\rangle\langle 1_N|
\end{aligned}$$

so

$$\hat{\rho}_{N-1} = \frac{N-1}{N}\hat{\rho}_{N-1}^{(W)} + \frac{1}{N}|\bar{0}\rangle_{N-1}\, {}_{N-1}\langle\bar{0}|. \tag{7.1}$$

With probability $\frac{N-1}{N}$ we have a W state of $N-1$ particles and with probability $1/N$ we have a product state. The state (7.1) can be converted to a state arbitrarily close to $\hat{\rho}_{N-1}^{(W)}$ by a filtering procedure [97]. Verifying entanglement for states close to these states can be done using entanglement witnesses.

7.1 W state tomography

It is interesting to look at the tomographic signature of a W state. Suppose we prepare $|W\rangle$ and then apply $R_x(\theta)^{\otimes N}$. The probability of measuring $|\bar{0}\rangle_N = |00...0\rangle$ before applying the rotation operator is clearly 0. After rotating we find

$$P_{|W\rangle}(0) = \left|{}_N\langle\bar{0}|R_x(\theta)^{\otimes N}|W\rangle_N\right|^2 = N\cos^{2N-2}(\theta/2)\sin^2(\theta/2).$$

This can be contrasted with the probability of measuring $|\bar{0}\rangle_N$ when we start with the non-entangled, singly excited state state

$$\rho_{\text{incoh}} = \frac{1}{N}\sum_j \rho_j,$$

with $\rho_j = |\bar{1}_j\rangle\langle\bar{1}_j|$. For this state the probability of measuring $|\bar{0}\rangle_N$ after applying the rotation operator is

$$P_{\rho_{\rm incoh}}(0) = {\rm Tr}\left[R_x(\theta)^{\otimes N}\rho_{\rm incoh}\left(R_x(\theta)^{\otimes N}\right)^\dagger|\bar{0}\rangle_N\langle\bar{0}|\right] = \sin^{2N}(\theta/2).$$

The probability of observing $|\bar{0}\rangle_N$ is compared for these two states in Fig. 7.1. There is clearly a qualitative difference in the curves for the entangled and nonentangled states. This has been used in [98] to verify entanglement of $|W\rangle$ states.

It is also interesting to look at the probability of measuring a single excitation after rotating the states. The matrix element between $|\bar{1}\rangle_j$ and $R_x^{\otimes N}|\bar{1}\rangle_k$ is

$$\begin{aligned} j \neq k \quad &: \quad {}_j\langle\bar{1}|R_x^{\otimes N}(\theta)|\bar{1}\rangle_k = -\cos^{N-2}(\theta/2)\sin^2(\theta/2) \\ j = k \quad &: \quad {}_j\langle\bar{1}|R_x^{\otimes N}(\theta)|\bar{1}\rangle_j = \cos^{2N}(\theta/2). \end{aligned}$$

Therefore the probability of observing a single excitation upon rotating $|\bar{1}\rangle_k$ is

$$P(1) = (N-1)\cos^{2(N-2)}(\theta/2)\sin^4(\theta/2) + \cos^{4N}(\theta/2). \tag{7.2}$$

The probability of observing a single excitation upon rotating $\rho_{\rm incoh}$ is $P_{\rho_{\rm incoh}}(1) = N \times \frac{1}{N}P(1) = P(1)$ as given by Eq. (7.2).

For the $|W\rangle$ state we find

$$\begin{aligned} {}_j\langle\bar{1}|R_x^{\otimes N}(\theta)|W\rangle_N &= \frac{1}{\sqrt{N}}\sum_k {}_j\langle\bar{1}|R_x^{\otimes N}(\theta)|\bar{1}\rangle_k \\ &= -\frac{N-1}{\sqrt{N}}\cos^{N-2}(\theta/2)\sin^2(\theta/2) + \frac{1}{\sqrt{N}}\cos^{2N}(\theta/2). \end{aligned}$$

The probability of observing a single excitation is therefore

$$\begin{aligned} P_{|W\rangle}(1) &= N\left|{}_j\langle\bar{1}|R_x^{\otimes N}(\theta)|W\rangle_N\right|^2 \\ &= \left[(N-1)\cos^{N-2}(\theta/2)\sin^2(\theta/2) - \cos^{2N}(\theta/2)\right]^2 \\ &= P_{\rho_{\rm incoh}}(1) + (N-1)(N-2)\cos^{2(N-2)}(\theta/2)\sin^4(\theta/2) - 2(N-1)\cos^{3N-2}(\theta/2)\sin^2(\theta/2). \end{aligned}$$

Finally the probability of observing one excitation starting in $|\bar{0}\rangle_N$ is

$$\begin{aligned} P_{|\bar{0}\rangle}(1) &= N\left|{}_j\langle\bar{1}|R_x^{\otimes N}(\theta)|\bar{0}\rangle_N\right|^2 \\ &= N\cos^{2N-2}(\theta/2)\sin^2(\theta/2). \end{aligned}$$

This expression has a maximum at $\sin^2(\theta/2) = 1/(N+1)$ so for large N, $\theta_{\rm max} \simeq 2/\sqrt{N}$ and $P_{|\bar{0}\rangle}(1) \to 1/e \simeq 0.37$. This implies that a QND measurement of the singly excited state can be used to probabilistically prepare a W state with at best 0.37 probability for large N[98].

The probability of observing no excitations is

$$\begin{aligned} P_{|\bar{0}\rangle}(0) &= \left|\langle\bar{0}|R_x^{\otimes N}(\theta)|\bar{0}\rangle_N\right|^2 \\ &= \cos^{2N}(\theta/2) \end{aligned}$$

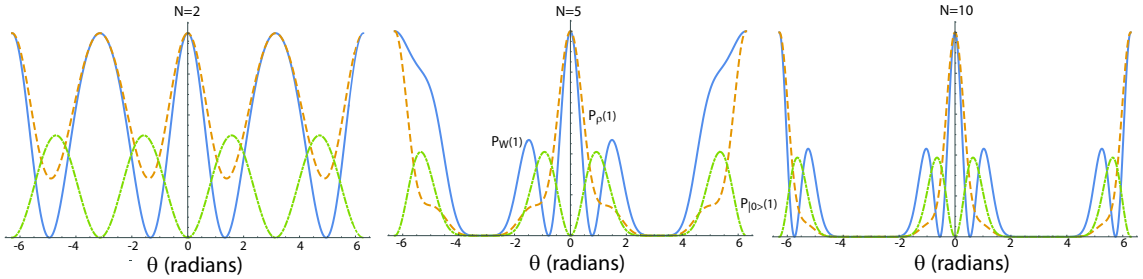


Figure 7.2: Probability of measuring a single excitation after rotating the $|W\rangle$ state (solid blue), the nonentangled singly excited state (dashed yellow), and the $|\bar{0}\rangle$ state (dash-dot green) for $N = 2, 5, 10$ qubits.

so that the probability of observing two or more excitations is

$$\begin{aligned} P_{|\bar{0}\rangle}(k>1) &= 1 - P_{|\bar{0}\rangle}(0) - P_{|\bar{0}\rangle}(1) \\ &= 1 - \cos^{2N}(\theta/2) - N\cos^{2N-2}(\theta/2)\sin^2(\theta/2). \end{aligned}$$

The probability of observing a single excitation for these different states is shown in Fig. 7.2. For large N the single excitation probability $P_{|\bar{0}\rangle}(1)$ has its maximum at small θ for which

$$\begin{aligned} P_{|\bar{0}\rangle}(1) &\simeq N\left(1-\frac{\theta^2}{8}\right)^{2(N-1)}\frac{\theta^2}{4} \\ &\simeq N\frac{\theta^2}{4}\left(1-\frac{(N-1)\theta^2}{4}\right) \end{aligned} \tag{7.3}$$

7.2 GHZ state tomography

Suppose we start with $|\bar{0}\rangle_N$ and rotate it. The overlap with the GHZ state is

$$\begin{aligned} {}_N\langle GHZ|R_x^{\otimes N}(\theta)|\bar{0}\rangle_N &= \frac{1}{\sqrt{2}}\,{}_N\langle\bar{0}|R_x^{\otimes N}(\theta)|\bar{0}\rangle_N + \frac{1}{\sqrt{2}}\,{}_N\langle\bar{1}|R_x^{\otimes N}(\theta)|\bar{0}\rangle_N \\ &= \frac{\cos^N(\theta/2) + i^N\sin^N(\theta/2)}{\sqrt{2}}. \end{aligned}$$

The overlap probability never exceeds the initial value of $1/2$ at $\theta = 0$ and a QND measurement analogous to that used for the W state does not prepare a GHZ state.

Chapter 8

Quantum communication

Just as quantum mechanics changes what is computable, it also results in new capabilities for the transmission of information. In this chapter we will start with a brief introduction to classical notions of information, and then look at how much information can be transmitted per unit time with a classical channel. This will then be compared with the information transmitting capacity of a quantum channel. Quantum channels are more powerful than classical channels in at least two different ways. They can be used to transmit more than one bit of classical information with a single quantum bit. This is known as dense coding. They can also be used to transmit a qubit by sending two classical bits. This is known as teleportation, which can also be applied to operations, not just states. We will then explore an application of quantum channels which is the secure distribution of cryptography keys. Quantum secured key distribution is known as QKD, which has been one of the first practical successes of quantum information. Very recently QKD was extended to intercontinental distances with the successful sharing of secure keys via satellite between China and Austria[99].

8.1 Classical information

Information can be transmitted in various ways, as analog signals modulated in amplitude, frequency, or phase and as binary bits. From a user perpsective we often wish to maximize the rate of reliable information transmission over a nonideal channel that adds noise. In order to compare different approaches we need to establish a metric for quantifying the amount of information that has been transmitted.

The first step is to agree on a definition for information. Let x be a random discrete variable that can take on X different values. This defines an alphabet. The probability of observing the value x_i is $p_i = \langle x_i \rangle$ and $\sum_i p_i = 1$. If all of the x_i are equally likely then $p_i = 1/X$. Knowledge of all the probabilities $\{p_i\}$ encapsulates complete information about the random variable x.

Suppose Alice sends a single x_i to Bob. How much information does he gain? The answer depends on Bob's prior knowledge. If he already knows all $\{p_i\}$ then receiving x_i does not add to his knowledge. If Bob initially has no information about the $\{p_i\}$ then the amount

of information he possesses will increase. A function which captures the information gain is

$$H(p) = -\sum_{i=1}^{X} p_i \log(p_i).$$

If the log function is base 2 then H is measured in bits. If the log is a natural log then H is measured in nats. Note that the same function is used in physics to define entropy (with a prefactor of k_B).

Some properties of H are

1) $H(p) \geq 0$ and $H(p) = 0$ only when there is a single p_i that is nonzero. This can be expressed colloquially as "You can always learn something new unless you already know everything".

2) There is an upper bound $C(X)$ on the information gain, $H(p) \leq C(X)$ and $H(p) = C(X)$ when $p_i = 1/X$. If $X' > X$ then $C(X') > C(X)$. A single value can carry more information if the alphabet is larger.

3) If we have two alphabets x, y with independent probabilities $p_i = \langle x_i \rangle, q_i = \langle y_i \rangle$. The information gain is

$$H(p, q) = H(p) + H(q).$$

This corresponds to the statement that the information in independent events is the sum of the information in the individual events.

Let's now consider a more complicated situation where not a singel sample, but a string of N samples $x_1, x_2 ... x_N$ is transmitted. The probability to observe a particular string is

$$p(x_1, x_2, ...x_N) = \prod_{i=1}^{N} p(x_i) = \prod_{i=1}^{X} p_i^{N_i}$$

where $\sum_i N_i = N$ and N_i is the number of times symbol x_i appears in the string.

The information in the string is found from the relation

$$\begin{aligned} -\frac{1}{N} \log \left[p(x_1, x_2, ...x_N) \right] &= -\frac{1}{N} \log \left[\prod_{i=1}^{X} p_i^{N_i} \right] \\ &= -\frac{1}{N} \sum_{i=1}^{X} N_i \log p_i \\ &= -\sum_{i=1}^{X} \frac{N_i}{N} \log p_i \\ &\simeq -\sum_{i=1}^{X} p_i \log p_i \\ &= H(p). \end{aligned}$$

Here we have made the approximation $Ni/N \simeq p_i$ which is accurate for large N. We can then write the string probability as

$$p(x_1, x_2, ...x_N) \simeq 2^{-NH(p)}$$

which shows that the probability of observing a particular string decreases exponentially with the length N, but is independent of the particular elements of the string. In other words $2^{-NH(p)}$ is the probability of observing any string of length N.

We can write the probability of observing a string as the inverse of the effective number of strings N_{eff} so

$$N_{\text{eff}} = \frac{1}{2^{-NH(p)}} = 2^{NH(p)} \leq 2^{NC(X)} = X^N.$$

The number of different strings of length N is X^N which in general is larger than N_{eff}. The difference between N_{eff} and X^N makes data compression possible. An average string can be transmitted using only $NH(p)$ bits which is less than $N\log_2(X)$ bits. This result is known as Shannon's coding theorem. A rare string can cause the coding to fail. The probability of this happening can be made arbitrarily small by using a sufficiently long code. This is known as the Shannon-McMillan theorem.

Coding schemes that occasionally fail are called lossy codes. This is acceptable for music or pictures and is widely used in for example JPEG and MPEG. A lossy code is clearly unacceptable for some applications such as transferring bank data.

Here is a simple example of a code that achieves data compression. Alice sends an alphabet to Bob with $X = 4$ symbols: a,b,c,d. The probabilities are

$$p(a) = 1/2, \quad p(b) = 1/8, \quad p(c) = 1/4, \quad p(d) = 1/8.$$

We can encode the alphabet using two bits per symbol

$$a \to 00, \quad b \to 01, \quad c \to 10, \quad d \to 11.$$

This is a lossless encoding.

We could gain efficiency by using a shorter codeword for more likley symbols and a longer codeword for less likely symbols. For example

$$a \to 0, \quad b \to 110, \quad c \to 10, \quad d \to 111.$$

This is still a lossless encoding since any string can be uniquely decoded.

The average number of bits per symbol needed to encode an average string is

$$1p(a) + 3p(b) + 2p(c) + 3p(d) = 7/4 < 2.$$

This code is therefore more efficient than using two bits for each symbol and works perfectly as long as there is no noise in the channel.

8.2 Channel capacity

We saw in the previous section that an intelligent choice of the encoding can increase the data rate. This leads to the question of what the maximum possible rate is for a noiseless channel.

Consider an alphabet described by $p(x)$ and a string of N symbols $x_1, x_2 ... x_N$. Let the ouput of the channel be described by $p(y)$. The conditional output probability distribution is $p(y|x)$. From Bayes' rule, if x, y are independent distributions then

$$p(x, y) = p(x|y)p(y) = p(y|x)p(x).$$

Then use this relation for the probabilities to find relations between the information of x and y. It is not hard to show that

$$\begin{aligned} H(x,y) &= -\sum_x \sum_y p(x,y)\log[p(x,y)] \\ &= -\sum_x \sum_y p(x,y)\log[p(x|y)] - \sum_y p(y)\log[p(y)] \\ &= H(x|y) + H(y) \\ &= H(y|x) + H(x). \end{aligned}$$

where $H(x|y)$ is the information in x given y. We then introduce a quantity $I(x,y)$ called the mutual information which is the information of x and y considered separately minus the information in x and y when they are taken together. We can write

$$H(x,y) = H(x) + H(y) - I(x,y)$$

so

$$I(x,y) = H(x) + H(y) - H(x,y) = H(y) - H(y|x).$$

When the variables are independent $I(x,y) = 0$. If the variables are completely dependent $I(x,y) = H(x)$. The channel capacity is defined as

$$C = \max_{p(x)} \left[I(x,y) \right].$$

8.2.1 Noisy Gaussian channel

Let's calculate the capacity of a noisy Gaussian channel. Think of the channel as

$$\{x\} \rightarrow \quad \text{channel} \quad \rightarrow \{y\}$$

with $y = x + \eta$ and η is the noise. Assume η is Gaussian distributed with probability distribution

$$p(\eta) = \frac{1}{\sqrt{2\pi}\sigma_\eta} e^{-\eta^2/2\sigma_\eta^2}.$$

Let the input also have Gaussian statistics

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-(x-\mu_x)^2/2\sigma_x^2}$$

with mean $\mu_x = \langle x \rangle$.

If the channel has bandwidth Δf, signal power $S = \langle x^2 \rangle$, and noise power $N = \langle \eta^2 \rangle$ it can be shown that the capacity is

$$C = \Delta f \log\left(1 + \frac{S}{N_0 \Delta f}\right) \quad \text{bits/sec}$$

where N_0 is the one sided noise power spectral density.

8.3 Quantum channel capacity

The information capacity of a quantum channel that transmits qubits is determined by the mutual information which is $I(x,y) \leq N$ when the channel transmits N qubits. Using codewords composed of tensor products of states ρ_i with the probability of ρ_i being used given by p_i the Holevo bound is[100]

$$C = H\left[\sum_i p_i \rho_i\right] - \sum_i p_i H(\rho_i)$$

where $H(\rho) = -\mathrm{Tr}\left[\rho \log(\rho)\right]$ is the von Neumann entropy. This says that if we transmit one qubit we can at most gain one bit of classical information.

8.3.1 Superdense coding

Superdense coding uses entanglement to, in some sense, surpass the Holevo bound. The caveat is that we must first prepare the channel by sharing an entangled pair. We then send one quantum bit over the channel to transmit two classical bits. The protocol is as follows[101].

1) Alice and Bob share the Bell state

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

2) Alice operates with I, X, Y, or Z on her qubit. Bob does nothing. The shared state is then

$U_A \otimes U_B$	resulting state	classical data
$I \otimes I$	$\|\beta_{00}\rangle = \frac{\|00\rangle+\|11\rangle}{\sqrt{2}}$	00
$X \otimes I$	$\|\beta_{01}\rangle = \frac{\|10\rangle+\|01\rangle}{\sqrt{2}}$	01
$Z \otimes I$	$\|\beta_{10}\rangle = \frac{\|00\rangle-\|11\rangle}{\sqrt{2}}$	10
$Y \otimes I$	$i\|\beta_{11}\rangle = i\frac{\|10\rangle-\|01\rangle}{\sqrt{2}}$	11

3) Alice sends her qubit to Bob.

4) Bob measures both qubits in the Bell basis, thereby extracting two classical bits of information.

In actuality two quantum bits were sent to Bob, but only one bit was sent after Alice decided which classical information to transmit. Experimental demonstrations have been made[102].

8.3.2 State Teleportation

This is complementary to superdense coding in that by sending two classical bits one quantum state is transmitted. The original state $|\psi\rangle$ held by Alice is erased in the process so the non-cloning theorem is not violated. The protocol is as follows[103].

1) Alice has three qubits. The first is in the unknown state $|\psi\rangle$ that is to be teleported. The other two qubits are prepared in the Bell state $|\beta_{00}\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$. The three qubit state is therefore

$$|\psi\rangle|\beta_{00}\rangle = \frac{|\beta_{00}\rangle|\psi\rangle + |\beta_{01}\rangle X|\psi\rangle + |\beta_{10}\rangle Z|\psi\rangle + |\beta_{11}\rangle XZ|\psi\rangle}{2}.$$

2) The third qubit is sent to Bob.

3) Alice measures the first two qubits in the Bell basis.

measurement result	state of third qubit
00	$\|\psi\rangle$
01	$X\|\psi\rangle$
10	$Z\|\psi\rangle$
11	$XZ\|\psi\rangle$

4) The classical results of Alice's measurement are sent to Bob who rotates his qubit accordingly.

measurement result	Bob's operation	final state of third qubit
00	I	$\|\psi\rangle$
01	X	$\|\psi\rangle$
10	Z	$\|\psi\rangle$
11	ZX	$\|\psi\rangle$

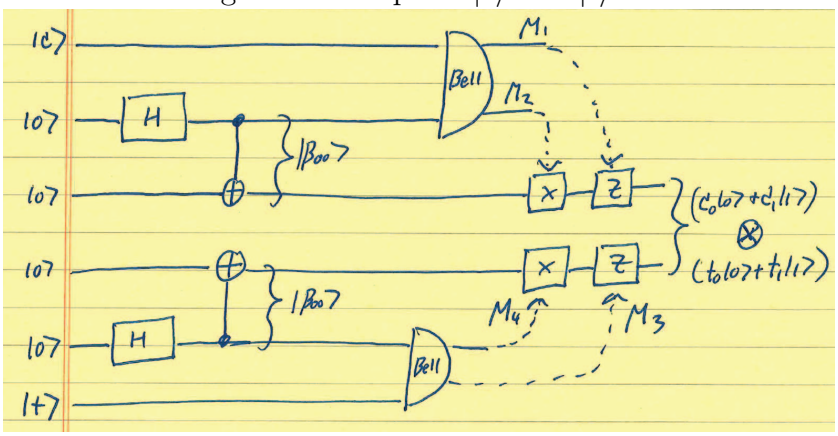
8.3.3 Gate teleportation

In addition to teleporting quantum states we can teleport quantum operations. This is very useful for some approaches to quantum computation for which two-qubit gates can only be performed probabilistically. Using a resource state $|r\rangle$, which also can only be prepared probabilistically, but can be distilled to high fidelity, a gate operation can be teleported onto a pair of control $|c\rangle$ and target $|t\rangle$ qubits. This architecture can be used for universal computation[104, 105].

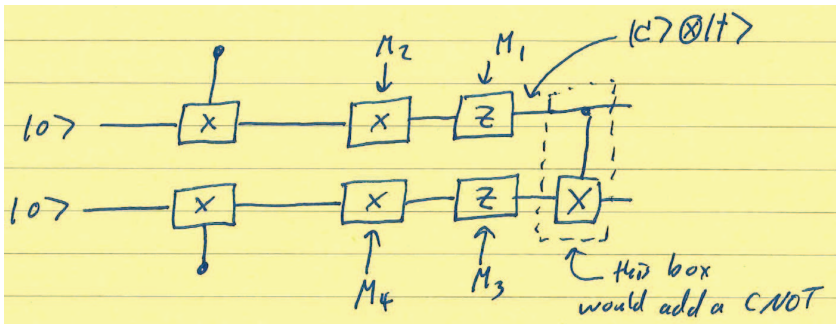
To understand how gate teleportation works consider the state that results from applicaiton of a CNOT gate on control qubit $|c\rangle = c_0|0\rangle + c_1|1\rangle$ and target qubit $|t\rangle = t_0|0\rangle + t_1|1\rangle$. The resulting state is

$$|\psi\rangle = c_0t_0|00\rangle + c_0t_1|01\rangle + c_1t_1|10\rangle + c_1t_0|11\rangle.$$

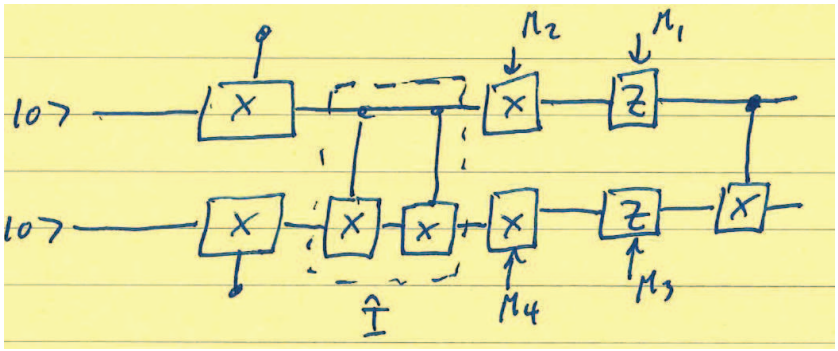
The following circuit teleports $|c\rangle$ and $|t\rangle$ but does not implement the CNOT operation:



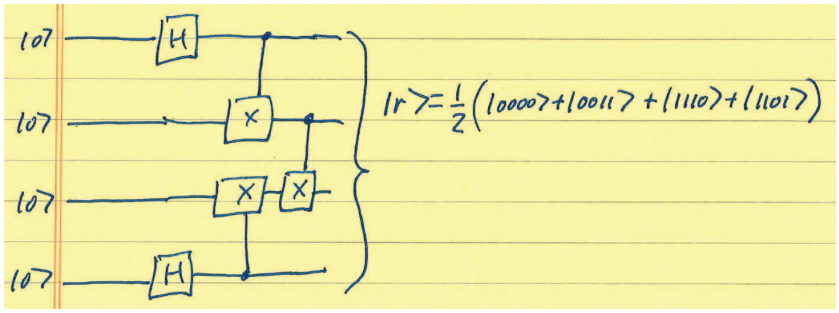
Let's look at the center two rows



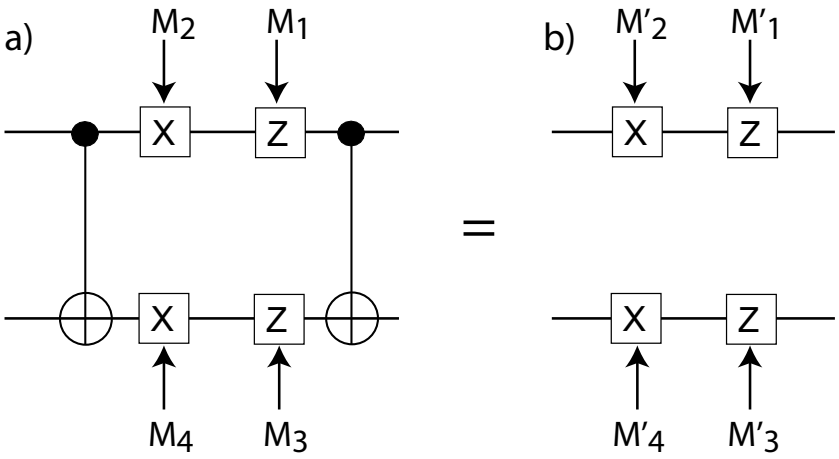
This is the same as



Regrouping the extra CNOT at the front of the circuit gives



If we then replace the rest of the circuit with controlled single qubit operations such that



we will obtain the desired $|\psi\rangle$ at the output of the circuit.

The complete circuit for gate teleportation is shown in Fig. 8.1. By changing the classical logic part of the circuit we can teleport any two-qubit operation. Of course the teleportation procedure does not remove the requirement for entangling gates since they are needed to prepare the resource state $|r\rangle$. The advantage is nevertheless that $|r\rangle$ can be prepared "off-line", independently of $|c\rangle$ and $|t\rangle$, and can be prepared with high fidelity using a distillation procedure[106], even if the available two-qubit gate is not of high fidelity.

8.4 QKD

to be added

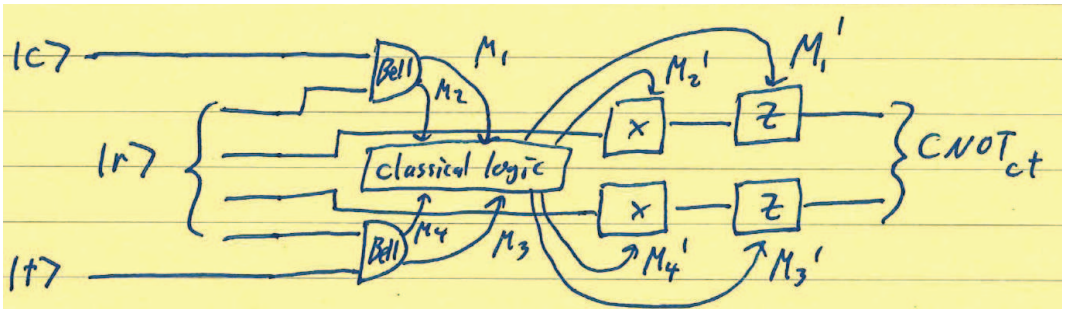


Figure 8.1: Circuit for gate teleportation.

Chapter 9

Quantum processes

A quantum evolution or quantum process acting on a pure state maps the state vector to a new state vector,

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

If we are hoping that $|\psi'\rangle$ is close to target state $|\psi_t\rangle$ the success probability can be quantified in terms of the overlap

$$\mathcal{O} = |\langle\psi_t|\psi'\rangle|^2.$$

The density matrix corresponding to a pure state evolves according to

$$\rho = |\psi\rangle\langle\psi| \rightarrow \rho' = |\psi'\rangle\langle\psi'| = (U|\psi\rangle)(\langle\psi|U^\dagger) = U\rho U^\dagger.$$

This transformation is a quantum process defined as

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

with

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

Also for mixed states we can write

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

A general quantum process can be expanded in an operator sum representation as

$$\mathcal{E}(\rho) = \sum_i A_i \rho A_i^\dagger. \tag{9.1}$$

The operation elements A_i are also known as Kraus operators. Since the density matrix must be normalized we have

$$1 \;=\; \mathrm{Tr}\left[\mathcal{E}(\rho)\right] = \mathrm{Tr}\left[\sum_i A_i \rho A_i^\dagger\right] = \mathrm{Tr}\left[\sum_i \rho A_i^\dagger A_i\right] = \mathrm{Tr}\left[\rho \sum_i A_i^\dagger A_i\right].$$

This is true for any ρ which implies the completeness condition

$$\sum_i A_i^\dagger A_i = I.$$

When we consider measurements and include both the system and an environment in the density matrix the measurement process is not in general trace preserving. In that case a normalized output state is

$$\rho' = \frac{\mathcal{E}(\rho)}{\mathrm{Tr}[\mathcal{E}(\rho)]}.$$

We will continue by introducing T_1 and T_2 coherence times and giving a geometric picture of a qubit subject to a noisy environment in Sec. 9.1. There are two more general complementary approaches to describing the dynamics of density operators. The first is based on solving a differential equation for the time evolution with terms that account for the coupling of the quantum system of interest to a larger environment. This is often referred to as a master equation which will be derived in Sec. 9.2. The second approach derives Kraus operators for specific quantum channels, examples of which will be given starting in Sec. 9.3.

9.1 A qubit in a noisy environment

If we represent the qubit as a point on the surface of the Bloch sphere we can parameterize the state with two angles as

$$|\psi\rangle = \cos(\theta/2)|0\rangle + \sin(\theta/2)e^{\imath\phi}|1\rangle.$$

Noise acting on the qubit will change both θ and ϕ. If we prepare the qubit in state $|0\rangle$ and wait a time t the probability to be in the initial state will decay. The $1/e$ time constant for the population decay is called the T_1 time, or longitudinal decay time. This can be determined by measuring the population at different delay times after the qubit is prepared.

Similarly if we prepare the qubit with a particular value of ϕ and wait a time t the state will change due to the presence of noise. The value of ϕ cannot be determined directly from a measurement in the computational or Z basis which is only sensitive to θ. To determine ϕ we need to either measure in the X or Y basis or use interference to make a Z basis measurement sensitive to ϕ. The latter approach was introduced by Ramsey[107, 108] and is commonly referred to as Ramsey interferometry based on the pulse sequence shown in Fig. 9.1. The characteristic time for the decay of the phase is called the T_2 time or transverse decay time. The notation of T_1 and T_2 times stems from the field of nuclear magnetic resonance. When talking about phase decay there is an additional subtlety in that two different times T_2

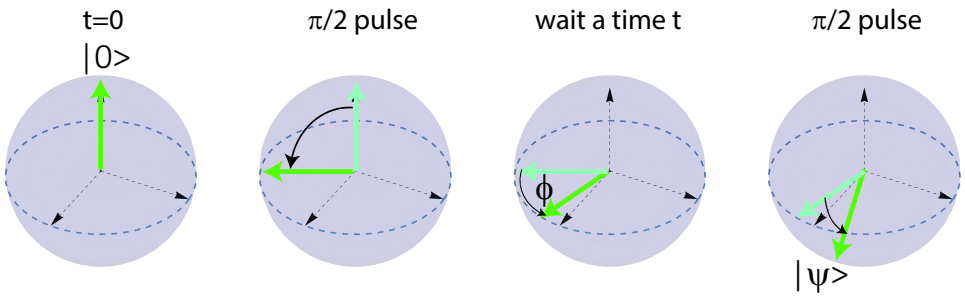


Figure 9.1: Pulse sequence for Ramsey interferometry.

and T_2^* are often reported. The time T_2^* is dependent on the noise characteristics of the perturbing fields and can often be lengthened using pulse sequences that average out the effects of noise. Doing so reveals a longer decay time T_2 that is intrinsic to the qubit itself. The Ramsey interferometry method used to determine the phase forms the basis for modern atomic clocks[109].

Let's calculate the output state after the sequence of Fig. 9.1. We start with the qubit in state $|\psi\rangle = |0\rangle$. We then perform a rotation $R_y(\pi/2)$ giving $|\psi\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$. After waiting a time t the state becomes $|\psi\rangle = \frac{|0\rangle+e^{\imath\phi}|1\rangle}{\sqrt{2}}$. The final state after the second $\pi/2$ pulse is

$$\begin{aligned} |\psi\rangle &= R_y(\pi/2)\frac{1}{\sqrt{2}}\begin{pmatrix}1\\ e^{\imath\phi}\end{pmatrix} \\ &= e^{\imath\phi/2}\begin{pmatrix}-i\sin(\phi/2)\\ \cos(\phi/2)\end{pmatrix}. \end{aligned} \tag{9.2}$$

The probability of observing $|1\rangle$ as the final state is $\cos^2(\phi/2)$ which allows the phase ϕ to be determined. If $\phi(t)$ is a linear function of t we will observe sinusoidal oscillation of the output probability as a function of the delay time t. If ϕ has a stochastic noise component the population oscillations will decay. The time after which the amplitude of the oscillations has decayed to $1/e$ of the initial amplitude defines T_2^*.

To make these notions more concrete let's calculate qubit decoherence in the presence of noise. Consider a qubit with levels $|0\rangle, |1\rangle$ subject to a "Rabi" field that couples the levels with frequency Ω as shown in Fig. 9.2. The Rabi field is turned on a for a time t which results in a $R_x(\theta)$ rotation of the qubit with $\theta = \Omega t$. We have the equality

$$R_x(\theta) = e^{-\imath\theta\sigma_x/2} = e^{-i\hat{\mathcal{H}}_x t/\hbar}$$

where $\hat{\mathcal{H}}_x = \hbar\Omega\sigma_x/2$ is the effective Hamiltonian due to the Rabi drive. The actual value of Ω depends on the strength of the driving field and the matrix element connecting the qubit levels.

In addition we will assume there are some background noise fields that change the energy separation of the qubit levels. Let's define the zero of energy as midway between the levels so

$$\begin{aligned} \hat{\mathcal{H}}_z|0\rangle &= \left(U_{\text{zero}} + \frac{\hbar\omega_q}{2}\right)|0\rangle \\ \hat{\mathcal{H}}_z|1\rangle &= \left(U_{\text{zero}} - \frac{\hbar\omega_q}{2}\right)|1\rangle \end{aligned}$$

with ω_q the unperturbed qubit frequency. We can write $\hat{\mathcal{H}}_z$ as

$$\hat{\mathcal{H}}_z = U_{\text{zero}}\hat{I} + \frac{\hbar\omega_q}{2}\sigma_z.$$

Due to additional electric or magnetic background fields $\omega_q \rightarrow \omega_q + \Delta(E, B)$. Constant terms in the Hamiltonian can be ignored and working in a rotating frame at frequency ω_q we can write

$$\hat{\mathcal{H}}_z = \frac{\hbar\Delta}{2}\sigma_z.$$

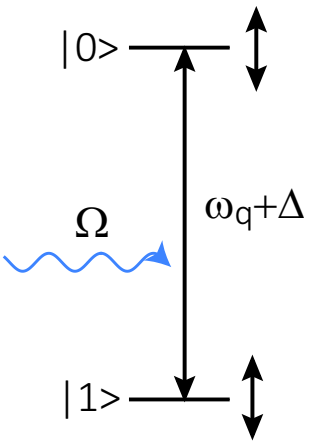


Figure 9.2: A qubit with Rabi drive Ω and perturbation due to external fields giving frequency shift Δ.

Combining the Rabi drive and the background fields we get the total Hamiltonian acting on the qubit

$$\begin{aligned}\hat{\mathcal{H}} &= \frac{\hbar}{2}\left(\Omega\sigma_x + \Delta\sigma_z\right) \\ &= \frac{\hbar}{2}\begin{pmatrix}\Delta & \Omega \\ \Omega & -\Delta\end{pmatrix}\end{aligned} \tag{9.3}$$

Equation (9.3) has eigenvalues and eigenvectors

$$\begin{aligned}\lambda_\pm &= \pm\frac{\hbar}{2}\sqrt{\Omega^2+\Delta^2} \\ v_+ &= \begin{pmatrix}\frac{\Delta}{\Omega} + \frac{\sqrt{\Omega^2+\Delta^2}}{\Omega} \\ 1\end{pmatrix} \\ v_- &= \begin{pmatrix}1 \\ \frac{\Delta}{\Omega} - \frac{\sqrt{\Omega^2+\Delta^2}}{\Omega}\end{pmatrix}.\end{aligned}$$

We then introduce a generalized Rabi frequency $\Omega' = \sqrt{\Omega^2+\Delta^2}$ giving the normzlized eigensystem

$$\begin{aligned}\lambda_\pm &= \pm\frac{\hbar\Omega'}{2} \\ v_+ &= \frac{1}{\left[\Omega^2 + (\Omega'+\Delta)^2\right]^{1/2}}\begin{pmatrix}\Omega'+\Delta \\ \Omega\end{pmatrix} \\ v_- &= \frac{1}{\left[\Omega^2 + (\Omega'-\Delta)^2\right]^{1/2}}\begin{pmatrix}\Omega \\ -\Omega'+\Delta\end{pmatrix}.\end{aligned}$$

One way to solve for the time evolution is to express the initial state in terms of the eigenstates and then evolve in time with eigenvalue dependent phase factors.

A more convenient approach is based on returning to the Hamiltonian and developing a geometrical picture of the evolution. We can write

$$\hat{\mathcal{H}} = \frac{\hbar\Omega'}{2}\left[\cos(\xi)\sigma_x + \sin(\xi)\sigma_z\right]$$

with $\xi = \tan^{-1}(\Delta/\Omega)$, $\cos(\xi) = \Omega/\Omega'$, $\sin(\xi) = \Delta/\Omega'$. The Hamiltonian is now in the same form as that describing a spin 1/2 particle in an effective magnetic field $\mathbf{B}$

$$\hat{\mathcal{H}} = \frac{\hbar}{2}\mathbf{B}\cdot\boldsymbol{\sigma} = \frac{\hbar}{2}\left(B_x\sigma_x + B_z\sigma_z\right)$$

with

$$B_x = \Omega'\cos(\xi) = \Omega, \quad B_z = \Omega'\sin(\xi) = \Delta.$$

The effective driving field is in the $x-z$ plane and makes an angle ξ with respect to the x axis.

We are now ready to calculate the qubit dynamics due to fields $\Omega(t)$ and $\Delta(t)$. Let the qubit be in an initial state $|\psi\rangle = a_0|0\rangle + a_1|1\rangle$. As shown in Eq. (4.6) the density operator can be written as

$$\rho = \frac{I + \mathbf{M}\cdot\boldsymbol{\sigma}}{2} = \frac{I + M_x\sigma_x + M_y\sigma_y + M_z\sigma_z}{2}$$

with

$$M_x = 2\mathrm{Re}(a_0 a_1^*), \quad M_y = -2\mathrm{Im}(a_0 a_1^*), \quad M_z = |a_0|^2 - |a_1|^2.$$

The equation of motion $\frac{d\rho}{dt} = \frac{i}{\hbar}[\rho, \hat{\mathcal{H}}]$ then takes the form

$$\frac{d\mathbf{M}}{dt} = \mathbf{B}\times\mathbf{M}. \tag{9.4}$$

This is just the equation for the precession of a magnetic moment $\mathbf{M}$ about a magnetic field $\mathbf{B}$. We see that the quantum dynamics of a qubit (any two-level system) can be recast in the form of an equation for the classical motion of a magnet subjected to a torquing field. This geometrical correspondence is useful for solving problems in magnetism and in laser dynamics[110].

Let's use Eq. (9.4) to study qubit decoherence. Assume the effective field $\mathbf{B}$ is composed of a control field $\mathbf{B}_0(t)$ that may be time dependent, plus a noise field $\mathbf{b}(t)$. We will further assume that $\mathbf{b}$ is parallel to $\mathbf{B}_0$. This "longitudinal" noise causes dephasing of the qubit as we proceed to show. With these definitions $\mathbf{B} = \mathbf{B}_0 + \mathbf{b}$ has a constant direction in space but changes in magnitude. The magnetization vector $\mathbf{M}$ precesses about $\mathbf{B}$ at a rate that is proportional to $|\mathbf{B}|$.

In addition the magnitude of the magnetization, averaged over $\mathbf{b}(t)$ will decay. To see this put $\mathbf{M} = \mathbf{M}_\parallel + \mathbf{M}_\perp$ where $\mathbf{M}_\parallel, \mathbf{M}_\perp$ are the components parallel and perpendicular to $\mathbf{B}$ respectively. The equation of motion (9.4) then takes the form

$$\begin{aligned}\frac{d\mathbf{M}}{dt} &= \frac{d\mathbf{M}_\parallel}{dt} + \frac{d\mathbf{M}_\perp}{dt} \\ &= \mathbf{B}\times(\mathbf{M}_\parallel + \mathbf{M}_\perp) \\ &= \mathbf{B}\times\mathbf{M}_\perp. \end{aligned} \tag{9.5}$$

An immediate consequence is $\frac{d\mathbf{M}_\parallel}{dt} = 0$ so $\mathbf{M}_\parallel$ is a constant and $\frac{d\mathbf{M}_\perp}{dt} = \mathbf{B} \times \mathbf{M}_\perp$. To proceed we define a coordinate system with $\hat{z}$ along $\mathbf{B}$, transverse coordinates x, y and complex magnetization amplitude $M_+ = M_x + iM_y$. The equations of motion for the components are

$$\begin{aligned} \frac{dM_x}{dt} &= (\mathbf{B} \times \mathbf{M})_x = -BM_y \\ \frac{dM_y}{dt} &= (\mathbf{B} \times \mathbf{M})_y = BM_x \end{aligned}$$

so

$$\begin{aligned} \frac{dM_+}{dt} &= \frac{d}{dt}(M_x + iM_y) \\ &= iB(iM_y + M_x) \\ &= iBM_+ \end{aligned}$$

which is solved by

$$M_+(t) = e^{\imath\phi(t)} M_+(0)$$

with $\phi(t) = \int_0^t dt'\, B(t')$. We then divide ϕ into a deterministic and random part as $\phi = B_0 t + \chi(t)$ with $\chi(t) = \int_0^t dt'\, b(t')$. The time averaged value of M_+ is then

$$\begin{aligned} \langle M_+(t) \rangle &= \left\langle e^{\imath\phi(t)} M_+(0) \right\rangle \\ &= \left\langle e^{\imath B_0 t} e^{\imath\chi(t)} M_+(0) \right\rangle \\ &= e^{\imath B_0 t} M_+(0) \left\langle e^{\imath\chi(t)} \right\rangle . \end{aligned}$$

The decay of the magnetization vector can be found if we can calculate the average value $\left\langle e^{\imath\chi(t)} \right\rangle$. One way to proceed is to assume a probability distribution $P(\chi)$ that is Gaussian

$$P(\chi) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\chi^2/2\sigma^2}$$

with $\sigma = \sqrt{\langle \chi^2 \rangle}$ the RMS fluctuation level. Using $P(\chi)$ we find

$$\left\langle e^{\imath\chi(t)} \right\rangle = \int_{-\infty}^{\infty} d\chi e^{\imath\chi} P(\chi) = e^{-\sigma^2/2}$$

and

$$\langle M_+(t) \rangle = e^{\imath B_0 t} M_+(0) e^{-\sigma^2/2}.$$

The justification for the assumption of Gaussian statistics is that if the time t at which we are interested in knowing $\langle M_+(t) \rangle$ is long compared to the noise correlation time then $\chi(t)$ will be the sum of many random parts and the central limit theorem says $\chi(t)$ has Gaussian statistics

To finish the calculation we need to relate σ to the noise $b(t)$. We have

$$\begin{aligned} \sigma^2 &= \left\langle \chi^2 \right\rangle \\ &= \langle \chi(t)\chi(t) \rangle \\ &= \left\langle \int_0^t dt_1\, b(t_1) \int_0^t dt_2\, b(t_2) \right\rangle \\ &= \int_0^t dt_1 \int_0^t dt_2\, \langle b(t_1) b(t_2) \rangle . \end{aligned}$$

We then use the he Wiener-Khinchin theorem which says that the correlation function $\langle b(t_1)b(t_2)\rangle$ is the Fourier transform of the power spectrum $|b(\omega)|^2$. Explicitly

$$\langle b(t_1)b(t_2)\rangle = \int_{-\infty}^{\infty} d\omega\, |b(\omega)|^2 e^{\imath\omega(t_2-t_1)}$$

and

$$\begin{aligned}
\sigma^2 &= \int_0^t dt_1 \int_0^t dt_2 \int d\omega\, |b(\omega)|^2 e^{\imath\omega(t_2-t_1)} \\
&= \int d\omega\, |b(\omega)|^2 \int_0^t dt_1 \int_0^t dt_2\, e^{\imath\omega(t_2-t_1)} \\
&= \int d\omega\, |b(\omega)|^2 \frac{\sin^2(\omega t/2)}{(\omega/2)^2}. \qquad (9.6)
\end{aligned}$$

Thus the transverse vector decays with time due to the noise $b(t)$. The exact form of the decay depends on $\sigma^2(t)$ which is a function of the noise spectrum $b(\omega)$. We can estimate T_2^* using

$$\lim_{t\to\infty} \frac{1}{2\pi t} \frac{\sin^2(\omega t/2)}{(\omega/2)^2} = \delta(\omega).$$

Thus for long times

$$\frac{\sin^2(\omega t/2)}{(\omega/2)^2} \approx 2\pi t \delta(\omega)$$

and

$$\sigma^2(t) \approx \int d\omega\, |b(\omega)|^2 2\pi t\delta(\omega) = 2\pi t |b(0)|^2. \qquad (9.7)$$

The magnetization therefore decays as

$$\begin{aligned}
|\langle M_+(t)\rangle| &= \left| e^{\imath B_0 t} M_+(0) e^{-\pi t |b(0)|^2} \right| \\
&= M_+(0) e^{-\pi t |b(0)|^2}.
\end{aligned}$$

Defining T_2^* by $|M_+(t)/M_+(0)| = e^{-t/T_2^*}$ we get

$$T_2^* = \frac{1}{\pi |b(0)|^2}. \qquad (9.8)$$

The stronger the noise is, the larger $|b(0)|^2$, and the shorter the coherence time. Note that for some types of noise the result (9.8) is ambiguous. For example for $1/f$ noise $|b(0)|^2$ diverges and $T_2^* \to 0$. For a careful treatment of this case see [111]. Note that we found a finite T_2 but no longitudinal decay. If we had made other assumptions about the direction of $\mathbf{b}(t)$ relative to $\mathbf{B}_0$ finite T_1 and T_2 would have resulted.

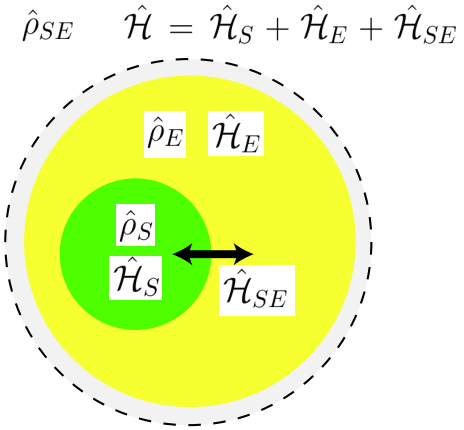


Figure 9.3: Quantum system and environment. The total Hamiltonian $\hat{\mathcal{H}}$ and the combined density operator $\hat{\rho}_{SE}$ consist of parts describing the system, the environment, and their interaction.

9.2 Time evolution of open quantum systems

The unitary evolution of the density operator is described by

$$\frac{d\hat{\rho}}{dt} = \frac{i}{\hbar}[\hat{\rho}, \hat{\mathcal{H}}] \tag{9.9}$$

with $\hat{\mathcal{H}}$ the Hamiltonian. Since $\hat{\mathcal{H}}$ is a hermitian operator the solution for $\hat{\rho}(t)$ is described by a unitary transformation

$$\hat{\rho}(t) = U(t, t_0)\hat{\rho}(t_0)U^\dagger(t, t0)$$

with U the unitary time evolution operator. Unitary time evolution is characteristic of "closed" quantum systems that are isolated from the environment.

On the other hand we often deal with "open" quantum systems that interact with the environment. This is depicted qualitatively in Fig. 9.3. In such cases we may observe non-unitary evolution. A familiar example is the decay of an excited atomic level. The atom decays to a lower level and emits a photon, but is never observed to spontaneously transition from a lower to a higher level in the absence of some source of energy in the environment. The dynamics is therefore non-unitary since independent of the initial conditions the atom is always observed in the lower level at long times.

This apparent departure from unitary evolution can be understood from different points of view. If we believe that quantum mechanics applies to everything then the non-unitary dynamics is simple an expression of our lack of knowledge of the environment. If we had a complete description of the environment, i.e. knew the Hamiltonian, there should be a unitary description of the combined dynamics of the system and the environment. In the example of the decaying atom the emitted photon, after rattling around in the environment, should eventually return to the atom and have the possibility of exciting it to a higher level, thereby restoring unitary dynamics.

An alternative viewpoint is that quantum mechanics only applies to systems of a finite size, and that at some point we have to give up the quantum description and use classical mechanics. Otherwise we would predict quantum behavior in macroscopic objects, which is never observed. This is the famous Schrödinger's cat paradox. A problem with this point of view is that there is no obvious scale at which to insert a boundary between quantum and classical behavior. Indeed experimental advances of the last few decades have succeeded in observing quantum behavior in larger and larger objects.

Whatever our point of view about these questions, it is true that we often wish to predict the evolution of a quantum system without full knowledge of the environment. This leads to a genaralization of Eq. (9.9) to describe effectively non-unitary evolution. Let's now make these ideas more precise[1]. Consider a quantum system S and an environment E. The density matrix describing system and environment is $\hat{\rho}_{SE}$ and the Hamiltonian is $\hat{\mathcal{H}} = \hat{\mathcal{H}}_S + \hat{\mathcal{H}}_E + \hat{\mathcal{H}}_{SE}$. Here $\hat{\mathcal{H}}_S$ is the system Hamiltonian, $\hat{\mathcal{H}}_E$ is the Hamiltonian of the environment, and $\hat{\mathcal{H}}_{SE}$ describes the coupling between the system and the environment.

Let's assume that the system and the environment are initially in a separable state, $\hat{\rho}_{SE}(t_0) = \hat{\rho}_S(t_0) \otimes \hat{\rho}_E(t_0)$. At a later time t we have

$$\hat{\rho}_{SE}(t) = U\hat{\rho}_{SE}(t_0)U^\dagger$$

where the time evolution operator U is defined by the total Hamiltonian $\hat{\mathcal{H}}$. The state of the system alone at a later time is

$$\hat{\rho}_S(t) = \mathrm{Tr}_E[\hat{\rho}_{SE}] = \mathrm{Tr}_E[U\hat{\rho}_{SE}(t_0)U^\dagger] = \mathrm{Tr}_E[U\hat{\rho}_S(t_0) \otimes \hat{\rho}_E(t_0)U^\dagger].$$

The evolution of the system density operator can be formally described as due to the action of a superoperator $\mathcal{E}$,

$$\hat{\rho}_S(t) = \mathcal{E}[\hat{\rho}_S(t_0)].$$

The map $\mathcal{E}$ is referred to as a quantum process, or as a superoperator, since it maps operators to operators.

Define a basis of environment states $|E_j\rangle$ then

$$\begin{aligned}\hat{\rho}_S(t) &= \mathrm{Tr}_E[U\hat{\rho}_S(t_0) \otimes \hat{\rho}_E(t_0)U^\dagger] \\ &= \sum_j \langle E_j|U\hat{\rho}_S(t_0) \otimes \hat{\rho}_E(t_0)U^\dagger|E_j\rangle.\end{aligned}$$

Let the initial state of the environment be $\rho_E(t_0) = |E_0\rangle\langle E_0|$ and the initial system state be a pure state $\rho_S(t_0) = |S_0\rangle\langle S_0|$ then

$$\begin{aligned}\hat{\rho}_S(t) &= \sum_j \langle E_j|U|S_0\rangle\langle S_0| \otimes |E_0\rangle\langle E_0|U^\dagger|E_j\rangle \\ &= \sum_j \langle E_j|U|S_0\rangle|E_0\rangle \otimes \langle S_0|\langle E_0|U^\dagger|E_j\rangle \\ &= \sum_j \langle E_j|U|E_0\rangle|S_0\rangle \otimes \langle S_0|\langle E_0|U^\dagger|E_j\rangle.\end{aligned}$$

[1] We primarily draw on the treatment in B. Schumacher and M. Westmoreland, *Quantum processes systems, & information*, Cambridge University Press, Cambridge (2010).

If we now define an operator $\hat{A}_j$ by

$$\hat{A}_j|S_0\rangle = \langle E_j|U|E_0\rangle|S_0\rangle$$

then

$$\hat{\rho}_S(t) \;=\; \sum_j \hat{A}_j|S_0\rangle\langle S_0|\hat{A}_j^\dagger.$$

Since any density operator can be written as a sum of pure state density operators we arrive at

$$\hat{\rho}_S(t) = \mathcal{E}[\hat{\rho}_S(t_0)] = \sum_j \hat{A}_j\hat{\rho}_S(t_0)\hat{A}_j^\dagger. \tag{9.10}$$

This result, which justifies Eq. (9.1), provides an expression for the evolution of the system called the operator sum representation of the process $\mathcal{E}$. The $\hat{A}_j$ are often called Kraus operators. There are two important things to note about Eq. (9.10). First, we have found a representation for the time evolution in terms of operators $\hat{A}_j$ that act only on the system. Second, the operators are not unitary and the temporal dynamics of the system density matrix is therefore not unitary. The act of tracing over the environment effectively leads to non-unitary system evolution.

The Kraus operators satisfy a normalization condition since

$$1 = \mathrm{Tr}[\hat{\rho}_S(t)] = \mathrm{Tr}[\sum_j \hat{A}_j|S_0\rangle\langle S_0|\hat{A}_j^\dagger] = \mathrm{Tr}[\langle S_0|\sum_j \hat{A}_j^\dagger\hat{A}_j|S_0\rangle] = \langle S_0|\sum_j \hat{A}_j^\dagger\hat{A}_j|S_0\rangle$$

must hold for any pure system state $|S_0\rangle$. Therefore

$$\sum_j \hat{A}_j^\dagger\hat{A}_j = \hat{I}.$$

We can now use the operator sum representation to find a general form for the time evolution of the system density operator when coupled to an environment. Consider an infinitesimal time evolution

$$\mathcal{E}[\hat{\rho}] = \hat{\rho} + \delta\hat{\rho} = \sum_j \hat{A}_j\hat{\rho}\hat{A}_j^\dagger \tag{9.11}$$

with $\delta\hat{\rho} \sim \delta t$. Here we have dropped the subscript S with the understanding that $\hat{\rho}$ is the reduced density operator of the system. Assume one of the Kraus operators $\hat{A}_0$ is dominant and is given by

$$\hat{A}_0 = \hat{I} + \hat{L}_0\delta t$$

and the other operators are

$$\hat{A}_j = \hat{L}_j\sqrt{\delta t}, \quad j \neq 0.$$

Then

$$\hat{A}_0\hat{\rho}\hat{A}_0^\dagger = \hat{\rho} + (\hat{L}_0\hat{\rho} + \hat{\rho}\hat{L}_0^\dagger)\delta t + \mathcal{O}(\delta t^2)$$

and

$$\hat{A}_j\hat{\rho}\hat{A}_j^\dagger = \hat{L}_j\hat{\rho}\hat{L}_j^\dagger\delta t.$$

To order δt (9.11) becomes

$$\hat{\rho} + \delta\hat{\rho} = \hat{\rho} + \left(\hat{L}_0\hat{\rho} + \hat{\rho}\hat{L}_0^\dagger + \sum_{j\neq 0} \hat{L}_j\hat{\rho}\hat{L}_j^\dagger \right) \delta t.$$

This implies the differential equation

$$\frac{d\hat{\rho}}{dt} = \hat{L}_0\hat{\rho} + \hat{\rho}\hat{L}_0^\dagger + \sum_{j\neq 0} \hat{L}_j\hat{\rho}\hat{L}_j^\dagger.$$

If we compare with (9.9) we see that in order to recover the Hamiltonian part of the dynamics we must put $\hat{L}_0 = -i\hat{\mathcal{H}}/\hbar + L_0'$ so that

$$\frac{d\hat{\rho}}{dt} = \frac{i}{\hbar}[\hat{\rho}, \hat{\mathcal{H}}] + \hat{L}_0'\hat{\rho} + \hat{\rho}\hat{L_0'}^\dagger + \sum_{j\neq 0} \hat{L}_j\hat{\rho}\hat{L}_j^\dagger.$$

Since $\mathrm{Tr}[\hat{\rho}] = 1$ at all times we require $\mathrm{Tr}[d\hat{\rho}/dt] = 0$ which results in the condition $\hat{L}_0' = -\frac{1}{2}\sum_j \hat{L}_j^\dagger\hat{L}_j$. Plugging in we get

$$\frac{d\hat{\rho}}{dt} = \frac{i}{\hbar}[\hat{\rho}, \hat{\mathcal{H}}] + \sum_j \hat{L}_j\hat{\rho}\hat{L}_j^\dagger - \frac{1}{2}\hat{L}_j^\dagger\hat{L}_j\hat{\rho} - \frac{1}{2}\hat{\rho}\hat{L}_j^\dagger\hat{L}_j.$$

This is known as the Lindblad equation[112] and is widely used to study open quantum system dynamics. The specific form of the Lindblad operators $\hat{L}_j$ depends on the problem being treated.

9.3 Kraus operators

Kraus operators are used as a basis to define a quantum process as in Eq. (9.1). It will be useful to establish the Kraus operators for several quantum processes.

9.3.1 Bit flip channel

A bit flip with probability p gives the process

$$\rho \to \rho' = (1-p)\rho + pX\rho X.$$

The Kraus operators are therefore

$$A_0 = \sqrt{1-p}I, \quad A_x = \sqrt{p}X.$$

9.3.2 Phase flip channel

A phase flip with probability p gives the process

$$\rho \to \rho' = (1-p)\rho + pZ\rho Z.$$

The Kraus operators are therefore

$$A_0 = \sqrt{1-p}I, \quad A_z = \sqrt{p}Z.$$

9.3.3 Combined bit and phase flip channel

The quantum process due to X and Z operations is

$$\rho \to \rho' = (1-p)\rho + pXZ\rho(XZ)^\dagger.$$

The Kraus operators are therefore

$$A_0 = \sqrt{1-p}I, \quad A_{xz} = \sqrt{p}XZ.$$

Since $XZ = -iY$ we can also write this as

$$\rho \to \rho' = (1-p)\rho + p(-iY)\rho(iY) = (1-p)\rho + pY\rho Y.$$

The Kraus operators are therefore

$$A_0 = \sqrt{1-p}I, \quad A_y = \sqrt{p}Y.$$

9.3.4 Depolarizing channel

The depolarizing channel for a single qubit is the quantum process for an X, Y or Z operation to occur with equal probabilities. This can be written as

$$\rho \to \rho' = (1-p)\rho + \frac{p}{3}X\rho X + \frac{p}{3}Y\rho Y + \frac{p}{3}Z\rho Z.$$

The Kraus operators are

$$A_0 = \sqrt{1-p}I, \quad A_1 = \sqrt{p/3}\,X, \quad A_2 = \sqrt{p/3}\,Y, \quad A_3 = \sqrt{p/3}\,Z. \tag{9.12}$$

It is not hard to show that for any ρ

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

so that

$$X\rho X + Y\rho Y + Z\rho Z = 2I - \rho.$$

The depolarizing channel can therefore be written as

$$\rho \to \rho' = \left(1 - \frac{4p}{3}\right)\rho + \frac{4p}{3}\frac{I}{2}.$$

This says that with probability $4p/3$ the qubit state ρ is depolarized and is replaced by the completely mixed state $I/2$, hence the name depolarizing channel. Note that when the elementary operations X, Y, Z each occur with probability $p/3$ the depolarization probability is $4p/3$.

Alternatively if the probability of depolarizing towards $I/2$ is p the depolarizing channel can be written as

$$\begin{aligned}\rho \to \rho' &= (1-p)\rho + p\frac{I}{2} \\ &= \left(1 - \frac{3p}{4}\right)\rho + \frac{p}{4}(X\rho X + Y\rho Y + Z\rho Z).\end{aligned}$$

With this convention the Kraus operators are

$$A_0 = \sqrt{1 - \frac{3p}{4}} I, \quad A_1 = \frac{\sqrt{p}}{2} X, \quad A_2 = \frac{\sqrt{p}}{2} Y, \quad A_3 = \frac{\sqrt{p}}{2} Z. \tag{9.13}$$

We see that each elementary X, Y or Z operation occurs with probability $p/4$. Both conventions (9.12, 9.13) are encountered in the literature.

9.3.5 Amplitude damping channel

To be added.

9.3.6 Phase damping channel

To be added.

9.3.7 Decoherence described by T_1 and T_2

Qubit decoherence is often modeled as an exponential decay with time constants T_1 for the population and T_2 for the coherence. It is often the case that

$$\frac{1}{T_2} = \frac{1}{2T_1} + \frac{1}{T_\phi}$$

where $1/T_\phi$ is a dephasing rate that leads to loss of coherence without changing the population. The evolution of the density matrix can be written as

$$\rho = \begin{pmatrix} 1-\rho_{11} & \rho_{01} \\ \rho_{01}^* & \rho_{11} \end{pmatrix} \rightarrow \rho' = \begin{pmatrix} 1-\rho_{11}e^{-t/T_1} & \rho_{01} e^{-t\left(\frac{1}{2T_1}+\frac{1}{T_\phi}\right)} \\ \rho_{01}^* e^{-t\left(\frac{1}{2T_1}+\frac{1}{T_\phi}\right)} & \rho_{11}e^{-t/T_1} \end{pmatrix}.$$

The Krauss operators corresponding to this process are[113]

$$A_1 = \frac{1+\sqrt{1-a-b}}{2} I + \frac{1-\sqrt{1-a-b}}{2} Z, \quad A_2 = \frac{\sqrt{a}}{2} X + \frac{i\sqrt{a}}{2} Y, \quad A_3 = \frac{\sqrt{b}}{2} I - \frac{\sqrt{b}}{2} Z,$$

with $a = 1 - e^{-t/T_1}$, $b = e^{-t/T_1}\left(1 - e^{-2t/T_\phi}\right)$.

9.3.8 Decoherence described by $T_{1\uparrow}$, $T_{1\downarrow}$

In some situations the population decay rate is bidirectional and asymmetric. For example there may be a rate $1/T_{1\uparrow}$ to make transitions from $|0\rangle \rightarrow |1\rangle$ and a different rate $1/T_{1\downarrow}$ to make transitions from $|1\rangle \rightarrow |0\rangle$. This occurs for example when measuring atomic qubits using light scattering or due to population redistribution from light scattering in an optical lattice if the two qubit levels couple differently to the trapping light. Putting $1/T_\phi = 0$ for simplicity we model this as

$$\rho = \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{01}^* & \rho_{11} \end{pmatrix} \rightarrow \rho' = \begin{pmatrix} \rho_{00}e^{-t/T_{1\uparrow}} + \rho_{11}\left(1 - e^{-t/T_{1\downarrow}}\right) & \rho_{01} e^{-t\left(\frac{1}{2T_{1\downarrow}}+\frac{1}{2T_{1\uparrow}}\right)} \\ \rho_{01}^* e^{-t\left(\frac{1}{2T_{1\downarrow}}+\frac{1}{2T_{1\uparrow}}\right)} & \rho_{11}e^{-t/T_{1\downarrow}} + \rho_{00}\left(1 - e^{-t/T_{1\uparrow}}\right) \end{pmatrix}.$$

Using $\rho_{00} = 1 - \rho_{11}$ we can rewrite this as

$$\rho' = \begin{pmatrix} e^{-t/T_{1\uparrow}} + \rho_{11}\left(1 - e^{-t/T_{1\downarrow}} - e^{-t/T_{1\uparrow}}\right) & \rho_{01} e^{-t\left(\frac{1}{2T_{1\downarrow}} + \frac{1}{2T_{1\uparrow}}\right)} \\ \rho_{01}^* e^{-t\left(\frac{1}{2T_{1\downarrow}} + \frac{1}{2T_{1\uparrow}}\right)} & 1 - e^{-t/T_{1\uparrow}} + \rho_{11}\left(e^{-t/T_{1\downarrow}} + e^{-t/T_{1\uparrow}} - 1\right) \end{pmatrix}.$$

Defining $a = 1 - e^{-t/T_{1\downarrow}} - e^{-t/T_{1\uparrow}}$, $b = e^{-t\left(\frac{1}{T_{1\downarrow}} + \frac{1}{T_{1\uparrow}}\right)}$, this can be expressed as

$$\rho' = \begin{pmatrix} e^{-t/T_{1\uparrow}} + \rho_{11} a & \rho_{01}\sqrt{b} \\ \rho_{01}^*\sqrt{b} & 1 - \left(e^{-t/T_{1\uparrow}} + \rho_{11} a\right) \end{pmatrix}.$$

The Kraus operators are ????

9.3.9 C_Z gate errors

The C_Z gate is a standard two-qubit primitive that provides universal computation. The error channels depend on teh physical realization of the C_Z. Reference [113] provides an example of a relatively simple error model. The ideal C_Z gate is

$$C_Z = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

An imperfect gate with amplitude switching error $ae^{\imath\phi}$ and phase error b can be written as

$$U = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \sqrt{1-a} & \sqrt{a}e^{\imath\phi} & 0 \\ 0 & -\sqrt{a}e^{-\imath\phi} & \sqrt{1-a} & 0 \\ 0 & 0 & 0 & -e^{\imath b} \end{pmatrix}.$$

The operator sum representation for an imperfect C_Z in the Pauli basis can be implemented as $\rho' = C_Z \rho C_Z$ followed by Pauli errors that give the output state

$$\rho_{\text{out}} = \sum_{j=1}^{7} p_j A_j \rho' A_j^\dagger$$

with

$$\begin{aligned} A_1 &= I \otimes I, \quad A_2 = Z \otimes I, \quad A_3 = I \otimes Z, \\ A_4 &= X \otimes X, \quad A_5 = Y \otimes Y, \quad A_6 = X \otimes Y \\ A_7 &= Y \otimes X, \quad A_8 = Z \otimes Z, \end{aligned}$$

and

$$\begin{aligned} p_1 &= \left|\frac{1 + 2\sqrt{1-a} + e^{\imath b}}{4}\right|^2, \quad p_2 = p_3 = \left|\frac{1 - e^{\imath b}}{4}\right|^2 \\ p_4 = p_5 &= \left|\frac{\sqrt{a}\sin(\phi)}{2}\right|^2, \quad p_6 = p_7 = \left|\frac{\sqrt{a}\cos(\phi)}{2}\right|^2 \\ p_8 &= \left|\frac{1 - 2\sqrt{1-a} + e^{\imath b}}{4}\right|^2. \end{aligned}$$

The error parameters can be related to a measured gate fidelity $\mathcal{F}$, averaged over input states, as

$$1 - \mathcal{F} = \frac{2}{5}a + \frac{3}{20}b^2.$$

To determine the relative magnitude of a and b and the phase ϕ additional measurements beyond the average fidelity are needed. A simple assumption would be to set $\frac{2}{5}a = \frac{3}{20}b^2$ and $\phi = \pi/4$.

9.4 Distance measures

An important figure of merit for a quantum process is how well the process output ρ approximates a desired, or target output σ. There is no universal answer to this question - it depends on what ρ will be used for. Nevertheless it is useful to have a single number to characterize the performance of a quantum process. There are several different measures in use for characterizing processes. The trace overlap or fidelity is defined as

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

Clearly $F(\sigma, \sigma) = 1$. The square of the fidelity

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

is also used. Unfortunately both F and F_{sq} are referred to as the fidelity in the literature. The fidelity $F(\rho, \sigma)$ is used in Nielsen & Chuang although Nielsen and others have argued[114] for the use of F_{sq}, since it corresponds to the probability of observing the desired target state. Although it is not obvious from the definition the fidelity is symmetric in its arguments[115] $F(\rho, \sigma) = F(\sigma, \rho)$.

Another useful measure is the trace distance

$$D(\rho, \sigma) = \frac{1}{2}\mathrm{Tr}\left[\sqrt{(\rho - \sigma)^\dagger (\rho - \sigma)}\right].$$

One minus the trace distance can be used to quantify the similarity of two states. The trace distance is an alternative that is sensitive to phase errors which may not be picked up by the fidelity[116]. Using Eq. (4.6) we can interpret the trace distance geometrically for single qubit states as

$$\begin{aligned} D(\rho, \sigma) &= \frac{1}{2}\mathrm{Tr}\left[\sqrt{\left(\frac{\hat{I} + \mathbf{r}\cdot\hat{\boldsymbol{\sigma}}}{2} - \frac{\hat{I} + \mathbf{s}\cdot\hat{\boldsymbol{\sigma}}}{2}\right)^\dagger \left(\frac{\hat{I} + \mathbf{r}\cdot\hat{\boldsymbol{\sigma}}}{2} - \frac{\hat{I} + \mathbf{s}\cdot\hat{\boldsymbol{\sigma}}}{2}\right)}\right] \\ &= \frac{1}{4}\mathrm{Tr}\left[\sqrt{[(\mathbf{r} - \mathbf{s})\cdot\hat{\boldsymbol{\sigma}}]^\dagger [(\mathbf{r} - \mathbf{s})\cdot\hat{\boldsymbol{\sigma}}]}\right] \end{aligned}$$

where $\mathbf{r}, \mathbf{s}$ are the vectors associated with the one-qubit density operators.

Chapter 10

Quantum state tomography

The density matrix can be reconstructed by tomographic measurements. Here we explain explicitly how this can be done for one or two qubits. An extensive reference for this topic is [117].

An arbitrary $2^n \times 2^n$ matrix A acting on an n qubit Hilbert space can be decomposed as a sum of orthogonal operators O_i satisfying $\mathrm{Tr}[O_i O_j] = \delta_{ij}$. The decomposition is

$$A = \sum_{i=1}^{2^{2n}} \lambda_i O_i$$

where the coefficients λ_i are found from

$$\lambda_i = \mathrm{Tr}[A O_i].$$

This is a consistent definition since

$$A = \sum_i \lambda_i O_i = \sum_i \mathrm{Tr}[A O_i] O_i = \sum_i \mathrm{Tr}\left[\sum_j \lambda_j O_j O_i\right] O_i = \sum_i \lambda_i O_i = A.$$

As we have discussed earlier the density operator for N qubits depends on $2^{2N} - 1$ independent real parameters. The number of measurements needed for tomographic state reconstruction therefore scales exponentially with the number of qubits which renders the technique infeasible for large systems.

10.1 One qubit

For a single qubit we may use $O_i = \frac{1}{\sqrt{2}}\sigma_i$ with $i = 0, x, y, z$ corresponding to the four basis states $i = 1, 2, 3, 4$. The first and last expansion coefficients are

$$\begin{aligned} \lambda_1 &= \frac{1}{\sqrt{2}}\mathrm{Tr}[\rho I] = \frac{1}{\sqrt{2}}(\rho_{00} + \rho_{11}) = \frac{1}{\sqrt{2}}(P_0 + P_1) \\ \lambda_4 &= \frac{1}{\sqrt{2}}\mathrm{Tr}[\rho \sigma_z] = \frac{1}{\sqrt{2}}(\rho_{00} - \rho_{11}) = \frac{1}{\sqrt{2}}(P_0 - P_1) \end{aligned}$$

coefficient	operation on qubit	measurement	density matrix elements
λ_1	I	$\frac{1}{\sqrt{2}}(P_0+P_1)$	$\frac{1}{\sqrt{2}}(\rho_{00}+\rho_{11})$
λ_2	$R_y(-\pi/2)$	$\frac{1}{\sqrt{2}}(P_0-P_1)$	$\frac{1}{\sqrt{2}}(\rho_{01}+\rho_{10})$
λ_3	$R_x(\pi/2)$	$\frac{1}{\sqrt{2}}(P_0-P_1)$	$\frac{i}{\sqrt{2}}(\rho_{01}-\rho_{10})$
λ_4	I	$\frac{1}{\sqrt{2}}(P_0-P_1)$	$\frac{1}{\sqrt{2}}(\rho_{00}-\rho_{11})$

Table 10.1: Tomography measurements needed to determine the density matrix of one qubit.

where P_0, P_1 are the probabilities of measuring states $|0\rangle$ or $|1\rangle$ in the z basis. The other two coefficients

$$\begin{aligned}\lambda_2 &= \frac{1}{\sqrt{2}}\mathrm{Tr}[\rho\sigma_x]\\ \lambda_3 &= \frac{1}{\sqrt{2}}\mathrm{Tr}[\rho\sigma_y]\end{aligned}$$

cannot be directly obtained from measurements in the z basis. To extract them we need effective x or y basis measurements which are obtained by rotating the qubit before making a z measurement. This is accomplished with the operator identities

$$\begin{aligned}\sigma_x &= R_y^\dagger(-\pi/2)\sigma_z R_y(-\pi/2),\\ \sigma_y &= R_x^\dagger(\pi/2)\sigma_z R_x(\pi/2).\end{aligned}$$

There is an obvious geometrical interpretation of these relations: rotating about the y axis by $-\pi/2$ converts a vector pointing along $\pm z$ into a vector along $\pm x$ and rotating about the x axis by $+\pi/2$ converts a vector pointing along $\pm z$ into a vector along $\pm y$. This geometrical analogy is valid for spin 1/2 but not for higher spins.

Using these identities together with the cyclic property of the trace we can write

$$\mathrm{Tr}[\sigma_x\rho] = \mathrm{Tr}\left[\left(R_y^\dagger(-\pi/2)\sigma_z R_y(-\pi/2)\right)\rho\right] = \mathrm{Tr}\left[\sigma_z\left(R_y(-\pi/2)\rho R_y^\dagger(-\pi/2)\right)\right] \tag{10.1}$$

and

$$\mathrm{Tr}[\sigma_y\rho] = \mathrm{Tr}\left[\left(R_x^\dagger(\pi/2)\sigma_z R_x(\pi/2)\right)\rho\right] = \mathrm{Tr}\left[\sigma_z\left(R_x(\pi/2)\rho R_x^\dagger(\pi/2)\right)\right]. \tag{10.2}$$

Thus

$$\begin{aligned}\lambda_2 &= \frac{1}{\sqrt{2}}\mathrm{Tr}[\sigma_x\rho]\\ &= \frac{1}{\sqrt{2}}\mathrm{Tr}\left[\sigma_z\left(R_y(-\pi/2)\rho R_y^\dagger(-\pi/2)\right)\right]\\ &= \frac{1}{\sqrt{2}}(\rho_{01}+\rho_{10})\end{aligned} \tag{10.3}$$

and

$$\begin{aligned}\lambda_3 &= \frac{1}{\sqrt{2}}\mathrm{Tr}[\sigma_y\rho]\\ &= \frac{1}{\sqrt{2}}\mathrm{Tr}\left[\sigma_z\left(R_x(\pi/2)\rho R_x^\dagger(\pi/2)\right)\right]\\ &= \frac{i}{\sqrt{2}}(\rho_{01}-\rho_{10}).\end{aligned} \tag{10.4}$$

The analysis pulses and measurements needed are given in Table 10.1. We see that we need only 3 distinct analysis pulse settings and two measurements required for each setting, giving a total of 6 measurements. We can easily check that the reconstruction works since

$$\lambda_1 O_1 + \lambda_2 O_2 + \lambda_3 O_3 + \lambda_4 O_4 = \rho.$$

A physically valid density matrix for one qubit can be parameterized by three real parameters t_1, t_2, t_3 as

$$\rho = \begin{pmatrix} t_1 & t_2 + it_3 \\ t_2 - it_3 & 1 - t_1 \end{pmatrix}.$$

It may therefore seem inefficient that we need to make six measurements to perform tomography on one qubit. The six measurements are P_0, P_1 for operations of $I, R_y(-\pi/2), R_x(\pi/2)$ applied to the qubit. The reason for making twice as many measurements as should be needed is that due to measurement errors we can not necessarily assume that $P_1 = 1 - P_0$. If we made this assumption then it would suffice to only measure P_0 for each of the three operations. Those three measurements are the minimum number needed to reconstruct the three real parameters t_1, t_2, t_3. Even making the full six measurements we will generally observe that $P_0 + P_1 \neq 1$ due to experimental uncertainties, or loss of population to unmeasured states. In order to reconstruct a physically valid density matrix we can use a maximum likelihood estimator described below in Sec. 10.3.

10.2 Two qubits

One possible choice for the O_i are $O_i = \frac{1}{2}\sigma_{i_A} \otimes \sigma_{i_B}$ with $i_A(i_B) = 0, x, y,$ or z. These definitions provide a set of 16 O_i:

$$O_1 = \frac{1}{2}\sigma_0 \otimes \sigma_0 = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad O_2 = \frac{1}{2}\sigma_0 \otimes \sigma_x = \frac{1}{2}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$O_3 = \frac{1}{2}\sigma_0 \otimes \sigma_y = \frac{1}{2}\begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \end{pmatrix} \quad O_4 = \frac{1}{2}\sigma_0 \otimes \sigma_z = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

$$O_5 = \frac{1}{2}\sigma_x \otimes \sigma_0 = \frac{1}{2}\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad O_6 = \frac{1}{2}\sigma_x \otimes \sigma_x = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$$O_7 = \frac{1}{2}\sigma_x \otimes \sigma_y = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix} \quad O_8 = \frac{1}{2}\sigma_x \otimes \sigma_z = \frac{1}{2}\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{pmatrix}$$

$$O_9 = \frac{1}{2}\sigma_y \otimes \sigma_0 = \frac{1}{2}\begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \end{pmatrix} \quad O_{10} = \frac{1}{2}\sigma_y \otimes \sigma_x = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix}$$

$$O_{11} = \frac{1}{2}\sigma_y \otimes \sigma_y = \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} \quad O_{12} = \frac{1}{2}\sigma_y \otimes \sigma_z = \frac{1}{2}\begin{pmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \\ i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{pmatrix}$$

$$O_{13} = \frac{1}{2}\sigma_z \otimes \sigma_0 = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad O_{14} = \frac{1}{2}\sigma_z \otimes \sigma_x = \frac{1}{2}\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

$$O_{15} = \frac{1}{2}\sigma_z \otimes \sigma_y = \frac{1}{2}\begin{pmatrix} 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \end{pmatrix} \quad O_{16} = \frac{1}{2}\sigma_z \otimes \sigma_z = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (10.5)$$

Given a density matrix $\rho = \sum_i \lambda_i O_i$ we determine the λ_i by measuring the expectation value of the operator O_i, i.e. we evaluate

$$\lambda_i = \mathrm{Tr}[\rho O_i].$$

For example $\lambda_1 = \mathrm{Tr}[\rho O_1] = \frac{1}{2}(P_{00} + P_{01} + P_{10} + P_{11})$ and $P_{00}, etc.$ are measured in the computational basis. The measurement of λ_1 therefore involves 4 physical measurements. This is not as bad as it sounds since these four measurements determine $\lambda_1, \lambda_4, \lambda_{13}$, and λ_{16}. Explicitly we have

$$\begin{aligned} \lambda_1 &= \frac{1}{2}(P_{00} + P_{01} + P_{10} + P_{11}) \\ \lambda_4 &= \frac{1}{2}(P_{00} - P_{01} + P_{10} - P_{11}) \\ \lambda_{13} &= \frac{1}{2}(P_{00} + P_{01} - P_{10} - P_{11}) \\ \lambda_{16} &= \frac{1}{2}(P_{00} - P_{01} - P_{10} + P_{11}). \end{aligned}$$

What about the other coefficients? They involve combinations of σ_x or σ_y which we do not directly measure. We therefore have to rotate the basis to measure them. This is accomplished in the same way as for the tomography of a single qubit. Thus, for example,

$$\begin{aligned} \lambda_3 &= \mathrm{Tr}[O_3\rho] = \frac{1}{2}\mathrm{Tr}[\sigma_{0A} \otimes \sigma_{yB}\rho] \\ &= \mathrm{Tr}\left[\sigma_{0A} \otimes \sigma_{zB}\left(R_{xB}(\pi/2)\rho R_{xB}^\dagger(\pi/2)\right)\right] \\ &= P_{00} - P_{01} + P_{10} - P_{11}, \end{aligned}$$

where the quantities in the last line refer to the results of measuring the corresponding diagonal elements *after* applying a $R_x(\pi/2)$ pulse on qubit B. The analysis pulses and

coefficient	operation on A	operation on B	expression for coefficient
λ_1	I	I	$\frac{1}{2}(P_{00}+P_{01}+P_{10}+P_{11})$
λ_2	I	$R_y(-\pi/2)$	$\frac{1}{2}(P_{00}-P_{01}+P_{10}-P_{11})$
λ_3	I	$R_x(\pi/2)$	$\frac{1}{2}(P_{00}-P_{01}+P_{10}-P_{11})$
λ_4	I	I	$\frac{1}{2}(P_{00}-P_{01}+P_{10}-P_{11})$
λ_5	$R_y(-\pi/2)$	I	$\frac{1}{2}(P_{00}+P_{01}-P_{10}-P_{11})$
λ_6	$R_y(-\pi/2)$	$R_y(-\pi/2)$	$\frac{1}{2}(P_{00}-P_{01}-P_{10}+P_{11})$
λ_7	$R_y(-\pi/2)$	$R_x(\pi/2)$	$\frac{1}{2}(P_{00}-P_{01}-P_{10}+P_{11})$
λ_8	$R_y(-\pi/2)$	I	$\frac{1}{2}(P_{00}-P_{01}-P_{10}+P_{11})$
λ_9	$R_x(\pi/2)$	I	$\frac{1}{2}(P_{00}+P_{01}-P_{10}-P_{11})$
λ_{10}	$R_x(\pi/2)$	$R_y(-\pi/2)$	$\frac{1}{2}(P_{00}-P_{01}-P_{10}+P_{11})$
λ_{11}	$R_x(\pi/2)$	$R_x(\pi/2)$	$\frac{1}{2}(P_{00}-P_{01}-P_{10}+P_{11})$
λ_{12}	$R_x(\pi/2)$	I	$\frac{1}{2}(P_{00}-P_{01}-P_{10}+P_{11})$
λ_{13}	I	I	$\frac{1}{2}(P_{00}+P_{01}-P_{10}-P_{11})$
λ_{14}	I	$R_y(-\pi/2)$	$\frac{1}{2}(P_{00}-P_{01}-P_{10}+P_{11})$
λ_{15}	I	$R_x(\pi/2)$	$\frac{1}{2}(P_{00}-P_{01}-P_{10}+P_{11})$
λ_{16}	I	I	$\frac{1}{2}(P_{00}-P_{01}-P_{10}+P_{11})$

Table 10.2: Tomography measurements needed to determine a bipartite density matrix.

measurements needed to define all 16 coefficients are given in Table 10.2. We see that we need only 9 distinct pairs of analysis pulse settings (not 16) and four measurements required for each pair, giving a total of 36 measurements. This corresponds to 9/4 more measurements than for determining the probability values of a CNOT matrix. For three qubits we would need 27 distinct analysis pulse settings (not 64).

The required 36 measurements can be compared with the number of independent real parameters which is $2^{2n}-1=15$. As in the case of tomography of a single qubit the redundancy is due to the fact that we cannot *apriori* assume the measured values give a physically possible density matrix.

10.3 Maximum likelihood reconstruction

Following the procedure of the previous section we can determine the coefficients λ_i and reconstruct the density matrix from

$$\rho=\sum_{i=1}^{2^{2n}}\lambda_i O_i.$$

This procedure is straightforward, but unfortunately it is not useful in practice. Due to measurement errors the reconstructed density matrix is likely to violate the conditions (4.11). In other words the tomographically reconstructed operator may turn out to be nonphysical.

To remedy this problem a maximum likelihood reconstruction which respects conditions (4.11) can be used[117]. Time permitting we will return to this topic.

Chapter 11

Quantum process tomography

Quantum process tomography (QPT) provides a description of a quantum process that maps density matrices onto density matrices

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

A quantum process can be written in an operator sum representation as

$$\mathcal{E}(\rho) = \sum_i A_i \rho A_i^\dagger.$$

Take an orthonormal basis for the density matrix such as Eq. (10.5), $\{O_1..O_{16}\}$ for two qubits and express the Kraus operators in this basis as $A_i = \sum_j a_{i,j} O_j$. Then

$$\mathcal{E}(\rho) = \sum_i A_i \rho A_i^\dagger = \sum_i \left(\sum_j a_{ij} O_j \right) \rho \left(\sum_k a_{ik}^* O_k^\dagger \right) = \sum_{jk} \chi_{jk} O_j \rho O_k^\dagger$$

with the process matrix

$$\chi_{jk} = \sum_i a_{ij} a_{ik}^*.$$

The process matrix has trace

$$\mathrm{Tr}[\chi] = \sum_m \chi_{mm} = \sum_m \sum_i a_{im} a_{im}^* = 2^n.$$

We will instead work with a normalized χ matrix defined by $\chi_{jk} = \frac{1}{2^n} \sum_i a_{ij} a_{ik}^*$.

Let's consider a simple example of a π rotation about x acting on a single qubit. In the basis $\{0, 1\}$ the matrix representation is

$$R_x(\pi) = \begin{pmatrix} 0 & -i \\ -i & 0 \end{pmatrix}.$$

There is one normalized Kraus operator $A_1 = R_x(\pi)$. We find $a_{1,2} = -i\sqrt{2}$ and all other

operation	Kraus operators	non-zero expansion coefficients	χ
$R_x(\pi)$	$E_1 = R_x(\pi)$	$a_{12} = -\sqrt{2}i$	$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$
$R_x(\pi/2)$	$E_1 = R_x(\pi/2)$	$a_{11} = 1,\ a_{12} = -i$	$\begin{pmatrix} 1/2 & i/2 & 0 & 0 \\ -i/2 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$
$R_y(\pi)$	$E_1 = R_y(\pi)$	$a_{13} = -\sqrt{2}i$	$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$
$R_y(\pi/2)$	$E_1 = R_y(\pi/2)$	$a_{11} = 1,\ a_{13} = -i$	$\begin{pmatrix} 1/2 & 0 & i/2 & 0 \\ 0 & 0 & 0 & 0 \\ -i/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$
$R_z(\pi/4)$	$E_1 = R_z(\pi/4)$	$a_{11} = \sqrt{2}\cos(\pi/8),\ a_{14} = -i\sqrt{2}\sin(\pi/8)$	$\begin{pmatrix} \cos^2(\pi/8) & 0 & 0 & i\cos(\pi/8)\sin(\pi/8) \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -i\cos(\pi/8)\sin(\pi/8) & 0 & 0 & \sin^2(\pi/8) \end{pmatrix}$

Table 11.1: Ideal process matrices for several one-qubit gates. The Kraus operators are expanded in the basis $O_i = \sigma_i/\sqrt{2}$, $i = 0, 1, 2, 3$.

coefficients are zero. The process matrix is thus

$$\begin{aligned} \chi &= \frac{1}{2}\sum_i a_{ij}a_{ik}^* \\ &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned} \tag{11.1}$$

For future reference the ideal χ matrices for several one-qubit gates are given in Table 11.1.

Another simple example is a controlled phase gate C_Z acting on two qubits. In the basis $\{00, 01, 10, 11\}$ the matrix representation is

$$C_Z = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Thus there is one Kraus operator $A_1 = C_Z$ which immediately satisfies the normalization condition $A_1^\dagger A_1 = I$.

We find $a_{1,1} = 1, a_{1,4} = 1, a_{1,13} = 1, a_{1,16} = -1$ and all other coefficients are zero. The

process matrix is thus

$$\chi = \frac{1}{4}\sum_i a_{ij}a_{ik}^* = \begin{pmatrix} \frac{1}{4} & 0 & 0 & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & 0 & 0 & -\frac{1}{4} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & 0 & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & 0 & 0 & -\frac{1}{4} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & 0 & \frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{4} & 0 & 0 & -\frac{1}{4} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1}{4} & 0 & 0 & -\frac{1}{4} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{4} & 0 & 0 & \frac{1}{4} \end{pmatrix}. \quad (11.2)$$

Now consider an imperfect C_Z operation

$$C_{Z\phi} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -e^{\imath\phi} \end{pmatrix}.$$

The phase ϕ is a static error. The Kraus operator is $A_1 = C_{Z\phi}$ and the basis expansion coefficients are $a_{1,1} = \frac{3-e^{\imath\phi}}{2}, a_{1,4} = \frac{1+e^{\imath\phi}}{2}, a_{1,13} = \frac{1+e^{\imath\phi}}{2}, a_{1,16} = \frac{-1-e^{\imath\phi}}{2}$. The process matrix is

$$\chi_\phi = \begin{pmatrix} \frac{1}{8}(5-3\cos(\phi)) & 0 & 0 & \frac{1}{8}(\cos(\phi)-2i\sin(\phi)+1) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{8}(\cos(\phi)-2i\sin(\phi)+1) & 0 & 0 & \frac{1}{8}(-\cos(\phi)+2i\sin(\phi)-1) \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{8}(\cos(\phi)+2i\sin(\phi)+1) & 0 & 0 & \frac{1}{8}(\cos(\phi)+1) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{8}(\cos(\phi)+1) & 0 & 0 & \frac{1}{8}(-\cos(\phi)-1) \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{8}(\cos(\phi)+2i\sin(\phi)+1) & 0 & 0 & \frac{1}{8}(\cos(\phi)+1) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{8}(\cos(\phi)+1) & 0 & 0 & \frac{1}{8}(-\cos(\phi)-1) \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{8}(-\cos(\phi)-2i\sin(\phi)-1) & 0 & 0 & \frac{1}{8}(-\cos(\phi)-1) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{8}(-\cos(\phi)-1) & 0 & 0 & \frac{1}{8}(\cos(\phi)+1) \end{pmatrix}$$

The trace overlap between the ideal and nonideal process matrices is

$$\mathrm{Tr}[\chi\chi_\phi^\dagger] = \frac{5+3\cos\phi}{8}$$

while the trace distance is

$$\frac{1}{2}\mathrm{Tr}[|\chi-\chi_\phi|] = \frac{\sqrt{3}}{2}|\sin(\phi/2)|.$$

Note that the fidelity $\sim 1-\mathcal{O}(\phi^2)$ while the trace distance process measure $\sim 1-\mathcal{O}(\phi)$.

# bits n	Hilbert space dimension 2^n	# basis states of ρ 4^n	# measurements state tomography $4^n \times 2^n = 8^n$	# process parameters $(4^n)^2 \rightarrow 16^n - 4^n$	# measurements PT $4^n \times 8^n = 32^n$
1	2	4	8	12	32
2	4	16	64	240	1024
3	8	64	256	4032	32768

Table 11.2: Measurement requirements for process tomography. The number of independent process parameters is less than the number of input - output state mappings (16^n) due to symmetries of the density matrix.

# bits n	# basis states of ρ 4^n	# proj. meas. state tomography 6^n	# proj. meas. PT $4^n \times 6^n = 24^n$	time at 1 s^{-1}	time at 10 s^{-1}	time at 100 s^{-1}	time at 1000 s^{-1}
1	4	6	24	40 min.	4 min.	0.4 min.	2.4 s
2	16	36	576	16 h	1.6 h	9.6 min.	0.96 min.
3	64	216	13824	384 h	38.4 h	3.8 h	22.8 min

Table 11.3: Measurement time for tomography on a $n-$bit process with data rates of $1, 10, 100$ s^{-1} and 100 repetitions for each physical measurement. Note that the number of distinct projective measurements actually needed for state tomography can be reduced to 6^n since all 8^n measurements are not linearly independent.

The corresponding process errors are

$$E_{\text{fidelity}} = 1 - \text{Tr}\left[\sqrt{\sqrt{\chi_{\text{sim}}}\chi_{\text{id}}\sqrt{\chi_{\text{sim}}}}\right], \tag{11.3a}$$

$$E_{\text{distance}} = \frac{1}{2}\text{Tr}\left[\sqrt{(\chi_{\text{id}} - \chi_{\text{sim}})^\dagger (\chi_{\text{id}} - \chi_{\text{sim}})}\right]. \tag{11.3b}$$

11.1 Resource Scaling of Tomography

A problem with using tomography to diagnose the operation of a quantum computer is that the measurement resources needed for full tomography scale exponentially with the number of qubits. The required number of measurements for reconstruction of a quantum black box process appears to be 32^n with n the number of qubits, as shown in the last column in Table 11.2. The number of distinct experimental settings needed for projective measurements is actually only 24^n as shown in Table 11.3. The corresponding measurement time for different data rates, and assuming 100 repetitions for each experimental measurement (10% statistical uncertainty) is also given in the table.

Clearly, even with very fast measurement times, it will be impractical to fully characterize the quantum process of many qubits since the resource requirements scale exponentially. A great deal of current research is directed towards developing useful process characterization techniques that do not scale exponentially with the system size. One line of approach takes

advantage of prior knowledge about the error channels to avoid performing full tomography. There are also compressed sensing methods which seek to reconstruct the density matrix with less than a full set of measurements. Again this requires making some prior assumptions about the structure of the system under study.

11.2 Randomized benchmarking

Another approach which avoids the resource requirements of tomography is randomized benchmarking. This only works for Clifford group gates which are not universal. On the other hand, as we have seen, the Clifford group is needed for error correction so verifying that at least Clifford group gates work well is a useful diagnostic.

The idea was clearly presented in a paper by E. Knill, et al.[118], although there is related earlier work from Emerson, et al.[119]. Consider a single qubit prepared in state $|0\rangle$. We apply a random sequence of gates C_j, $j = 1, 2, ...n$ where each of the C_j is a member of the Clifford group, as shown in Fig. 11.1. We will end up in some state $|\psi\rangle$. We can calculate the inverse operation U^{-1} such that

$$U^{-1}|\psi\rangle = U^{-1}C_n..C_2C_1|0\rangle = |0\rangle.$$

Since the C_j are chosen from the Clifford group the operator U^{-1} is also an element of the Clifford group, or can be composed out of elements of the Clifford group. Recalling the Gottesman-Knill theorem we can efficiently calculate U^{-1}. If state preparation, measurement, and gate operations were all perfect we would simply observe the qubit in state $|0\rangle$ every time. In the presence of gate errors the probability of ending up in $|0\rangle$ will decrease as the number of gates n is increased. Measuring the success probability as a function of n we can extract the average gate fidelity.

There are two important features of this approach. First, it can be extended to multiple qubits, in which case the Clifford group includes CNOT gates, with a resource scaling that is linear in the number of qubits. This is a great advantage compared to the exponential scaling of tomography. Second, we can extract the gate fidelity even if state preparation and state measurement are imperfect. These so-called SPAM errors are always present and when performing tomography it is difficult to separate the gate error from SPAM errors. Using randomized benchmarking we can isolate the gate fidelity which is important for understanding whether or not the gates are good enough for reaching the fault tolerance

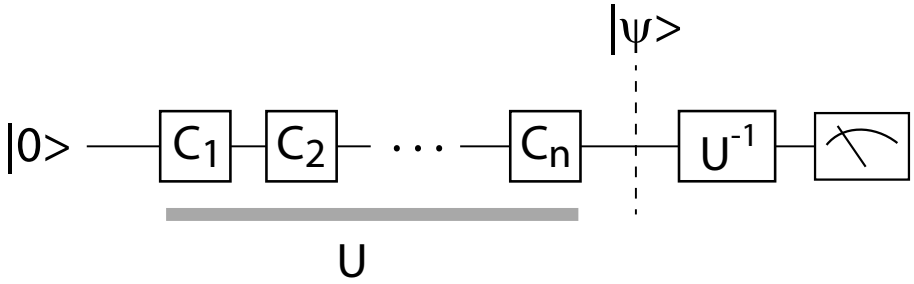


Figure 11.1: Randomized benchmarking sequence.

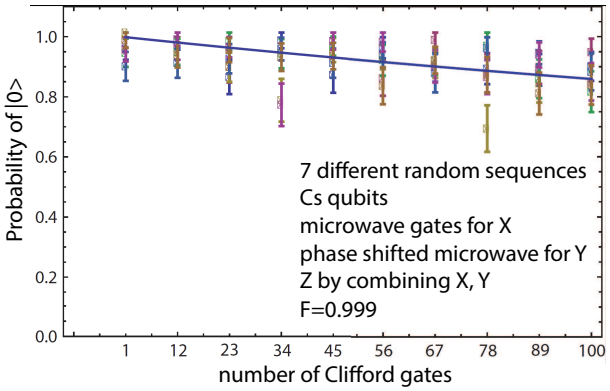


Figure 11.2: Randomized benchmarking data from a Cs atom experiment.

threshold. From a practical point of view experimentalists like randomized benchmarking because it is straightforward to implement, and because it generally gives better fidelity numbers than tomography which is subject to SPAM errors.

The basic equation of randomized benchmarking is[118]

$$P(n) = \frac{1}{2} + \frac{(1-d_{\rm im})(1-d)^n}{2}. \tag{11.4}$$

Here $P(n)$ is the probability of measuring $|0\rangle$ at the end of the sequence of length n, $d_{\rm im}$ is the depolarization probability associated with state preparation, the final inverting gate, and the measurement, and d is the depolarization probability associated with an average Clifford gate. A typical measurement curve is shown in Fig. 11.2.

By fitting the measured $P(n)$ to Eq. (11.4) we can extract d. Since the value of d will depend on the specific sequence used the experiments also average other different randomly chosen Clifford sequences. Given the depolarization probability d the gate fidelity is

$$F = \left(1 - \frac{d}{2}\right)^{1/2}. \tag{11.5}$$

We can derive Eq. (11.5) as follows. The depolarization probability d describes an error model associated with the depolarizing channel. The depolarizing channel is the quantum process (see Sec. 9.3.4)

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

This says that with probability d the qubit state ρ is depolarized and is replaced by the completely mixed state $I/2$. This channel takes a pure state and contracts it into the

interior of the Bloch sphere. The fidelity is defined by

$$\begin{aligned}
F(\rho,\rho') &= \mathrm{Tr}\left[\sqrt{\sqrt{\rho}\rho'\sqrt{\rho}}\right] \\
&= \mathrm{Tr}\left[\sqrt{\sqrt{\rho}\left((1-d)\rho + d\frac{I}{2}\right)\sqrt{\rho}}\right] \\
&= \mathrm{Tr}\left[\sqrt{(1-d)\rho^2 + d\frac{I}{2}\rho}\right] \\
&= \mathrm{Tr}\left[\sqrt{(1-d)\rho + d\frac{\rho}{2}}\right] \quad \text{(for pure states)} \\
&= \mathrm{Tr}\left[\sqrt{\left(1-\frac{d}{2}\right)\rho}\right] \quad \text{(for pure states)} \\
&= \sqrt{1-\frac{d}{2}}\,\mathrm{Tr}\left[\sqrt{\rho}\right] \\
&= \sqrt{1-\frac{d}{2}}
\end{aligned}$$

as desired. Note that the probability of observing the target state after application of a gate is $F^2 = 1 - d/2$, not F[114].

While randomized benchmarking has much more favorable resource requirements than process tomography it does not directly provide information about the causes of the measured errors. Such information can be acquired from an extension of randomized benchmarking known as gate set tomography[120, 121].

Chapter 12

Error Correction

Error correction is needed for reliable computation. This is true for both classical and quantum computers. Despite the fact that the underlying physical components are unreliable it is possible to synthesize reliable computing machines by incorporating redundancy into the design. This insight dates back to von Neumann[122]. Unfortunately quantum error correction is much more demanding of physical resources than correction of errors in classical binary computers. There are at least two reasons for this. One is that the no-cloning theorem prevents copying unknown states so majority voting cannot be used to provide an error syndrome. The second issue is that simply measuring one or more qubits collapses superposition states and removes entanglement with other degrees of freedom, thereby interrupting the computation. The solution is based on encoding logical information in entangled states of multiple qubits and using measurements that reveal just enough information to diagnose and correct errors, but not so much that the quantum computation is stopped.

An example of this is shown in Fig. 12.1. Consider the state of three qubits $a_0|0\rangle + a_1|1\rangle, b_0|0\rangle + b_1|1\rangle, c_0|0\rangle + c_1|1\rangle$, which can be expressed as a superposition of eight possible kets

$$\begin{aligned} |\psi\rangle &= a_0b_0c_0|000\rangle + a_0b_1c_1|011\rangle + a_1b_0c_1|101\rangle + a_1b_1c_0|110\rangle \\ &+ a_0b_0c_1|001\rangle + a_0b_1c_0|010\rangle + a_1b_0c_0|100\rangle + a_1b_1c_1|111\rangle. \end{aligned}$$

This has been written with the even parity kets in the first row and the odd parity kets in the second row. If we make a von Neumann measurement of all three qubits we can calculate the parity from the measurement results but the state will collapse to a single ket of even or odd parity. Alternatively, by using an ancilla qubit, and mapping the joint parity of the input state onto the ancilla with three CNOT gates a measurement of the ancilla collapses the qubit state to an even or odd parity state of the form

$$|\psi_{\rm even}\rangle = \frac{a_0b_0c_0|000\rangle + a_0b_1c_1|011\rangle + a_1b_0c_1|101\rangle + a_1b_1c_0|110\rangle}{(|a_0b_0c_0|^2 + |a_0b_1c_1|^2 + |a_1b_0c_1|^2 + |a_1b_1c_0|^2)^{1/2}}$$

or

$$|\psi_{\rm odd}\rangle = \frac{a_0b_0c_1|001\rangle + a_0b_1c_0|010\rangle + a_1b_0c_0|100\rangle + a_1b_1c_1|111\rangle}{(|a_0b_0c_1|^2 + |a_0b_1c_0|^2 + |a_1b_0c_0|^2 + |a_1b_1c_1|^2)^{1/2}}.$$

A quantum superposition state with definite parity is preserved. This type of ancilla aided parity measurement which preserves quantum superpositions is at the heart of methods for quantum error correction[123] as we will see in this chapter.

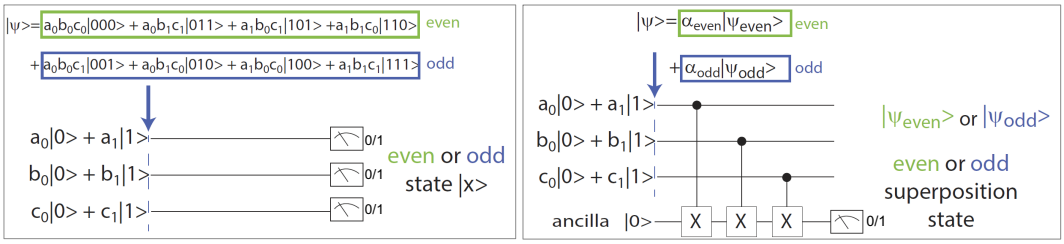


Figure 12.1: von Neumann measurement (left) and ancilla aided parity measurement (right) of a three qubit state.

12.1 Classical error correction

Classical error correction is largely based on Hamming codes[124].

See handwritten notes.

12.2 Quantum error correction

Errors on physical qubits can be corrected by encoding information in logical qubits comprised of several physical qubits. There are many different possible ways of doing this. The first codes that were invented were Shor's 9 qubit code[123] and Steane's 7 qubit code[125]. Many codes currently being studied use large numbers of physical qubits to reduce sensitivity to errors. A prime example of this is the surface code[126]. There is a large literature on quantum error correction. Some valuable early references from the 1990s include [127, 128, 129, 130, 131, 132, 133, 134, 135, 136]. More recent reviews include[137, 138, 139].

Codes are labelled as

$$[[n, k, d]]$$

where n is the number of physical qubits in a codeword, k is the number of encoded logical qubits, and d is the code distance. The classical code distance is just the Hamming distance between codewords which is the number of places where two vectors of binary values differ. Figure 12.2 shows that codewords c_1, c_2 subject to t errors will end up in regions of radius t. For the errors to be distinguishable

$$t = \lfloor (d-1)/2 \rfloor.$$

Thus a distance d code can detect up to $d-1$ bit errors and correct up to $(d-1)/2$ errors.

Codes with larger d can detect, and potentially correct, more errors but at the cost of requiring more physical qubits. The optimal size n of a quantum error correcting code depends on numerous architectural considerations and there is no simple answer to this question. What is known is that the minimum possible codeword size that can correct an arbitrary logical error is a $[[5, 1, 3]]$ code that requires 5 physical qubits per logical qubit[140].

To see that this is the minimum possible, consider a codeword that is n qubits long and encodes a single logical qubit. Due to at most one physical error the number of states that can appear are $3n + 1$ for the logical state $|0\rangle_{\text{enc}}$ and $3n + 1$ for the logical state $|1\rangle_{\text{enc}}$. The number of states is $3n+1$ for each logical value since a Pauli X, Y, or Z error could occur on any of the n bits plus there is the possibility of no error. The n qubits span a 2^n dimensional

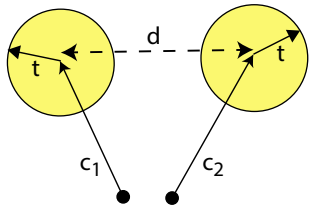


Figure 12.2: Geometrical picture showing that a distance d nondegenerate code can correct $t = \lfloor (d-1)/2 \rfloor$. errors.

space which must not be smaller than the number of possible states after an error. Thus we require

$$2^n \geq 2(3n+1)$$

which implies $n \geq 5$. Strictly speaking this is the lower limit for a nondegenerate code (a code for which all the codewords are distinct), but there are no known degenerate codes with $n < 5$.

If we attempt to directly translate classical correction codes to qubits we encounter several problems. The no-cloning theorem prevents us from making copies of the state. Furthermore measurements will collapse quantum superpositions and halt a computation. There are both bit flip and phase flip errors and more generally continuous analog errors in the quantum state. Furthermore any hardware used to detect and correct errors will also be subject to failure. All in all quantum error correction appears daunting.

To clarify why quantum error correction is possible consider a qubit state $|\psi\rangle = a_0|0\rangle + a_1|1\rangle$ interacting with an environment in state $|E\rangle$. The initial state of qubit and environment is assumed to be a product state $|\psi\rangle|E\rangle$. The following error processes can occur

$$\begin{aligned} |0\rangle|E\rangle &\rightarrow \beta_1|0\rangle|E_1\rangle + \beta_2|1\rangle|E_2\rangle \\ |1\rangle|E\rangle &\rightarrow \beta_3|0\rangle|E_3\rangle + \beta_4|1\rangle|E_4\rangle. \end{aligned}$$

Here $|E_1\rangle, |E_2\rangle, |E_3\rangle, |E_4\rangle$ are states of the environment, that need not be orthogonal, and $\beta_1, \beta_2, \beta_3, \beta_4$ are complex amplitudes. The initial state transforms as

$$\begin{aligned} |\psi\rangle|E\rangle &\rightarrow a_0\beta_1|0\rangle|E_1\rangle + a_0\beta_2|1\rangle|E_2\rangle + a_1\beta_3|0\rangle|E_3\rangle + a_1\beta_4|1\rangle|E_4\rangle \\ &= \frac{1}{2}|\psi\rangle\left(\beta_1|E_1\rangle + \beta_3|E_3\rangle\right) + \frac{1}{2}Z|\psi\rangle\left(\beta_1|E_1\rangle - \beta_3|E_3\rangle\right) \\ &\quad + \frac{1}{2}X|\psi\rangle\left(\beta_2|E_2\rangle + \beta_4|E_4\rangle\right) + \frac{1}{2}XZ|\psi\rangle\left(\beta_2|E_2\rangle - \beta_4|E_4\rangle\right). \end{aligned}$$

If we define new environment states as

$$\begin{aligned} |\tilde{E}_1\rangle &= \beta_1|E_1\rangle + \beta_3|E_3\rangle, & |\tilde{E}_2\rangle &= \beta_1|E_1\rangle - \beta_3|E_3\rangle, \\ |\tilde{E}_3\rangle &= \beta_2|E_2\rangle + \beta_4|E_4\rangle, & |\tilde{E}_4\rangle &= \beta_2|E_2\rangle - \beta_4|E_4\rangle \end{aligned}$$

we can write

$$|\psi\rangle|E\rangle \rightarrow \frac{1}{2}|\psi\rangle|\tilde{E}_1\rangle + \frac{1}{2}Z|\psi\rangle|\tilde{E}_2\rangle + \frac{1}{2}X|\psi\rangle|\tilde{E}_3\rangle + \frac{1}{2}XZ|\psi\rangle|\tilde{E}_4\rangle. \tag{12.1}$$

We see that the qubit - environment interaction can be decomposed into one of four discrete operations acting on the qubit, together with a change to the state of the environment. If we can diagnose which error occurred, without measuring $|\psi\rangle$, then the error can be corrected with single qubit operations. The structure of this result is analogous to how state teleportation works as described in Sec. 8.3.2. It is worth noting that the environment can be a thermal bath that the qubit is coupled to, but it could also be the control system that implements logic gates. Errors due to imperfect control are therefore correctable in the same way as errors due to environmental caused decoherence.

12.3 Bit-flip code

Let's look at a specific example of Eq. (12.1) for the bit-flip channel. The circuit is shown in Fig. 12.3. We encode $|\psi\rangle = a_0|0\rangle + a_1|1\rangle$ in three qubits as a logical state $|\psi\rangle_L = a_0|000\rangle + a_1|111\rangle$. Encoding is done with the first two CNOT operations. A bit-flip error on any of the three qubits can occur. No error is just the identity operation and an error on qubit j is X_j. The lower two qubits in the figure act as ancillas.

The error is decoded with the next two CNOT operations. As shown in the figure the error can then be coherently corrected for with a Toffoli gate where the ancillas are the control qubits and the data qubit is the target. After the Toffoli gate we have the corrected state shown in the last column of Table 12.1. The ancillas must then be reset to $|0\rangle$ so the data can be encoded again.

Alternatively the ancilla states can be measured in the z basis and classical logic used to implement the effect of the Toffoli gate. The classical logic is simply if the syndrome measurement is 00, 01, or 10 do nothing and if the syndrome is 11 perform an X operation

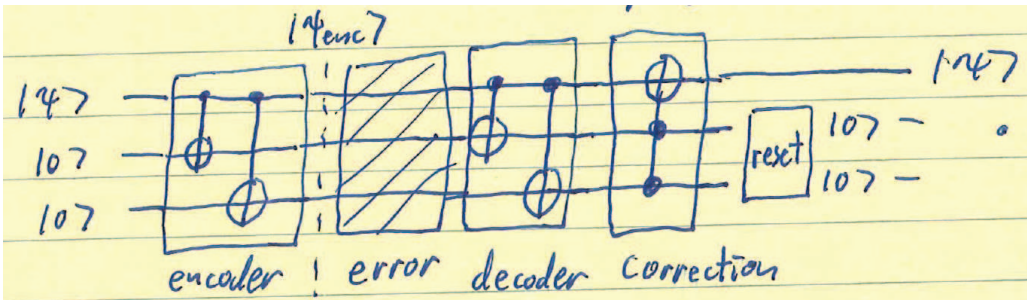


Figure 12.3: Circuit for correcting bit-flip errors.

error operation	error state	decoded state	corrected state
I	$a_0\|000\rangle + a_1\|111\rangle$	$a_0\|000\rangle + a_1\|100\rangle$	$\|\psi\rangle\|00\rangle$
X_1	$a_0\|100\rangle + a_1\|011\rangle$	$a_0\|111\rangle + a_1\|011\rangle$	$\|\psi\rangle\|11\rangle$
X_2	$a_0\|010\rangle + a_1\|101\rangle$	$a_0\|010\rangle + a_1\|110\rangle$	$\|\psi\rangle\|10\rangle$
X_3	$a_0\|001\rangle + a_1\|110\rangle$	$a_0\|001\rangle + a_1\|101\rangle$	$\|\psi\rangle\|01\rangle$

Table 12.1: Operation of the circuit for protecting the state $|\psi\rangle = a_0|0\rangle + a_1|1\rangle$ against bit-flip errors.

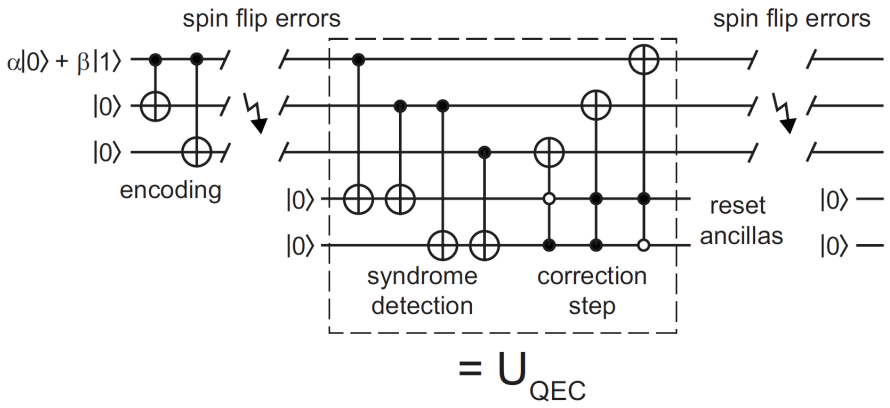


Figure 12.4: Circuit for correcting bit-flip errors from [144] with additional ancilla qubits. The logical qubit is encoded at all times. A CNOT control qubit with an open circle indicates that the qubit should be flipped before and after the CNOT operation.

on the data qubit. The ancillas are then reset to zero so the data can be encoded again.

The first approach is known as coherent error correction. Instead of measuring the ancilla qubits to extract the error syndrome they are used coherently in a circuit to correct the error. Entropy must still be removed, and this is done by resetting the ancillas. This approach has been used to implement repeated cycles of error correction[141].

The second approach is measurement based error correction which has been developed in great detail, and demonstrated experimentally[142, 143]. In this approach the qubits have to be measured, and reset so the data can be encoded again. Although coherent error correction has the advantage that qubit measurements are not required, only qubit resets, it has the disadvantage that the recovery operation which requires a Toffoli gate is more difficult to implement.

In both approaches only a single error can be detected and corrected. If more than one qubit suffers a bit-flip the code fails. A slight modification to the circuit can be used to correct phase-flip errors due to the action of a Z gate. to see how to do this note that in the z basis.

$$\begin{aligned} X|0\rangle &= |1\rangle \qquad X|1\rangle = |0\rangle \\ Z|0\rangle &= |0\rangle \qquad Z|1\rangle = -|1\rangle \end{aligned}$$

If we work in the x basis defined by $|0\rangle_x = (|0\rangle + |1\rangle)/\sqrt{2}$, $|1\rangle_x = (|0\rangle - |1\rangle)/\sqrt{2}$ we find

$$\begin{aligned} X|0\rangle_x &= |0\rangle_x \qquad X|1\rangle_x = -|1\rangle_x \\ Z|0\rangle_x &= |1\rangle_x \qquad Z|1\rangle_x = -|0\rangle_x. \end{aligned}$$

In other words Z acts like X and X acts like Z when we use the x basis. Therefore the bit-flip correction circuit will correct phase-flip errors if we change basis. The operator that changes basis is the Hadamard gate for which

$$\begin{aligned} H|0\rangle &= |0\rangle_x \qquad H|1\rangle = |1\rangle_x \\ H|0\rangle_x &= |0\rangle \qquad H|1\rangle_x = |1\rangle. \end{aligned}$$

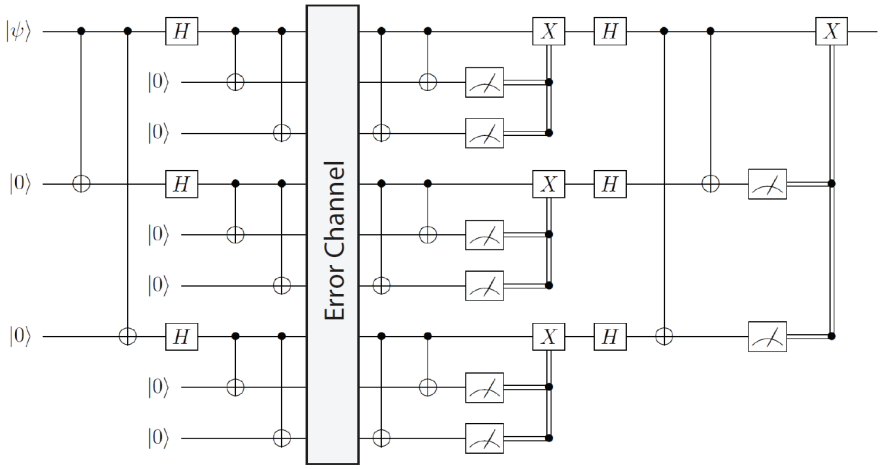


Figure 12.5: Shor's 9 qubit code. Double lines represent classical signals.

The circuit for correcting phase-flips simply involves adding Hadamard gates after the encoding and before the decoding steps.

While both the coherent and measurement based approaches have been demonstrated they are not entirely satisfactory since the qubit is not encoded or protected during the error detection and correction cycle. This deficiency can be corrected for by using more ancilla qubits as shown in Fig. 12.4.

12.4 Shor code

In order to simultaneously correct for all possible qubit errors we need a more complicated circuit. This was first invented by Shor in 1995 [123]. The circuit is shown in Fig. 12.5 and it can correct X, Z, and $Y = iXZ$ errors. To do so the encoding that corrects bit-flip (X errors) is combined with the encoding that corrects phase-flip (Z errors).

We start by encoding

$$|0\rangle \to |000\rangle_x, \quad |1\rangle \to |111\rangle_x.$$

This is done with the first two CNOT gates and three H gates. For each of the three qubits we then perform a bit-flip encoding

$$|0\rangle_x \to \frac{|000\rangle + |111\rangle}{\sqrt{2}}, \quad |1\rangle_x \to \frac{|000\rangle - |111\rangle}{\sqrt{2}}.$$

The resulting logical code words are

$$\begin{aligned} |0\rangle_L &= \frac{1}{2^{3/2}}(|000\rangle + |111\rangle)(|000\rangle + |111\rangle)(|000\rangle + |111\rangle) \\ |1\rangle_L &= \frac{1}{2^{3/2}}(|000\rangle - |111\rangle)(|000\rangle - |111\rangle)(|000\rangle - |111\rangle). \end{aligned}$$

A X or Z error on any of the 9 qubits transforms the code word to an orthogonal state from which the error syndrome can be diagnosed and ψ recovered. A detailed description

can be found in [136]. However many errors give the same state, for example Z_2 and Z_3, so the code is degenerate. This would be a problem for a classical code since the error syndrome is not unique for different errors. In the quantum case it doesn't matter since we only need to know how the code word changed for it to be corrected.

12.5 Steane or color code

Another possible logical encoding is the 7 qubit code due to Steane[125]. This is also the building block used in the color code due to Bombin and Martin-Delgado[145]. This code has several nice features. The fault tolerant error threshold may be as high as $\sim 10\%$ for larger color codes[146, 147], although other studies[148] suggest a threshold error rate of $\sim 8 \times 10^{-4}$. All Clifford group operators can be implemented transversally which makes fault tolerant operation possible. For universal quantum computation we still need to distill magic states which can then be teleported into the computational substrate to perform T gates. Logical encoding and computation on encoded qubits was demonstrated in 2014 using trapped ions by the Blatt group[149].

We will describe the code in terms of colored plaquettes. The seven qubits are divided into three plaquettes of four qubits each, with color labels red, green, blue. The logical code space is defined as the set of states $|\psi\rangle_L$ for which $S_i|\psi\rangle_L = +|\psi\rangle_L$ where the S_i are the stabilizer operators. Each plaquette has an X and a Z stabilizer giving in all six stabilizers

$$\begin{aligned} S_x^{(1)} &= X_1X_2X_3X_4, & S_z^{(1)} &= Z_1Z_2Z_3Z_4, \\ S_x^{(2)} &= X_2X_3X_5X_6, & S_z^{(2)} &= Z_2Z_3Z_5Z_6, \\ S_x^{(3)} &= X_3X_4X_6X_7, & S_z^{(3)} &= Z_3Z_4Z_6Z_7. \end{aligned}$$

The stabilizers impose six constraints on the seven qubits leaving a two-dimensional space which is used to encode a logical qubit.

The logical states $|0\rangle_L, |1\rangle_L$ are defined as eigenstates of $Z_L = Z_1Z_2Z_3Z_4Z_5Z_6Z_7$ where $Z_i|0\rangle_i = |0\rangle_i$, $Z_i|1\rangle_i = -|1\rangle_i$. The eigenvalue conditions are $Z_L|0\rangle_L = |0\rangle_L$ and $Z_L|1\rangle_L = -|1\rangle_L$ and the states can be written explicitly as

$$\begin{aligned} |0\rangle_L &= |1010101\rangle + |0101101\rangle + |1100011\rangle + |0011011\rangle \\ &+ |1001110\rangle + |0110110\rangle + |1111000\rangle + |0000000\rangle, \\ |1\rangle_L &= |0101010\rangle + |1010010\rangle + |0011100\rangle + |1100100\rangle \\ &+ |0110001\rangle + |1001001\rangle + |0000111\rangle + |1111111\rangle, \end{aligned}$$

where we have suppressed obvious normalization factors of $1/\sqrt{8}$. The logical states are orthonormal ${}_L\langle a|b\rangle_L = \delta_{ab}$ and $X_L|0\rangle_L = |1\rangle_L$, $X_L|1\rangle_L = |0\rangle_L$. Each logical state satisfies the six stabilizer conditions as well as the Z_L eigenvalue condition.

It is easily checked that all Clifford operators are transversal with this code. The T gate is not transversal.

12.5.1 Clifford Group

According to the Gottesman-Knill theorem[30] the action of combinations of Clifford gates can be efficiently simulated (in polynomial time) on a classical computer. The power of

index	x axis	y axis	z axis	U	index	x axis	y axis	z axis	U
1	I	I	I	$\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$	2	I	I	$\pi/2$	$e^{-\imath\pi/4}\begin{pmatrix}1 & 0\\ 0 & i\end{pmatrix}$
3	I	I	π	$-i\begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$	4	I	I	$-\pi/2$	$e^{\imath\pi/4}\begin{pmatrix}1 & 0\\ 0 & -i\end{pmatrix}$
5	I	π	I	$-1\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}$	6	I	π	$\pi/2$	$-e^{\imath\pi/4}\begin{pmatrix}0 & 1\\ i & 0\end{pmatrix}$
7	π	I	I	$-i\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}$	8	π	I	$\pi/2$	$e^{-\imath\pi/4}\begin{pmatrix}0 & 1\\ -i & 0\end{pmatrix}$
9	π	$\pi/2$	I	$\frac{-i}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}$	10	I	$-\pi/2$	I	$\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}$
11	$\pi/2$	I	$\pi/2$	$\frac{e^{-\imath\pi/4}}{\sqrt{2}}\begin{pmatrix}1 & 1\\ -i & i\end{pmatrix}$	12	$\pi/2$	π	$\pi/2$	$-\frac{e^{\imath\pi/4}}{\sqrt{2}}\begin{pmatrix}1 & 1\\ i & -i\end{pmatrix}$
13	π	$-\pi/2$	I	$\frac{i}{\sqrt{2}}\begin{pmatrix}1 & -1\\ -1 & -1\end{pmatrix}$	14	$-\pi/2$	I	$\pi/2$	$\frac{e^{-\imath\pi/4}}{\sqrt{2}}\begin{pmatrix}1 & -1\\ i & i\end{pmatrix}$
15	I	$\pi/2$	I	$\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -1\\ 1 & 1\end{pmatrix}$	16	$-\pi/2$	π	$\pi/2$	$\frac{e^{\imath\pi/4}}{\sqrt{2}}\begin{pmatrix}1 & -1\\ -i & -i\end{pmatrix}$
17	$-\pi/2$	$-\pi/2$	I	$\frac{e^{-\imath\pi/4}}{\sqrt{2}}\begin{pmatrix}1 & i\\ -1 & i\end{pmatrix}$	18	$-\pi/2$	$\pi/2$	I	$\frac{e^{\imath\pi/4}}{\sqrt{2}}\begin{pmatrix}1 & i\\ 1 & -i\end{pmatrix}$
19	$-\pi/2$	π	I	$\frac{i}{\sqrt{2}}\begin{pmatrix}1 & i\\ -i & -1\end{pmatrix}$	20	$-\pi/2$	I	I	$\frac{1}{\sqrt{2}}\begin{pmatrix}1 & i\\ i & 1\end{pmatrix}$
21	$\pi/2$	$-\pi/2$	I	$\frac{e^{\imath\pi/4}}{\sqrt{2}}\begin{pmatrix}1 & -i\\ -1 & -i\end{pmatrix}$	22	$\pi/2$	I	I	$\frac{1}{\sqrt{2}}\begin{pmatrix}1 & -i\\ -i & 1\end{pmatrix}$
23	$\pi/2$	π	I	$\frac{-i}{\sqrt{2}}\begin{pmatrix}1 & -i\\ i & -1\end{pmatrix}$	24	$\pi/2$	$\pi/2$	I	$\frac{e^{-\imath\pi/4}}{\sqrt{2}}\begin{pmatrix}1 & -i\\ 1 & i\end{pmatrix}$

Table 12.2: Elements of the Clifford group $\mathcal{C}_1$ for a single qubit. The notation π_j is shorthand for $R_j(\pi) = -i\sigma_j$ and $\pm\pi_j/2$ is shorthand for $R_j(\pm\pi/2)$.The gates are applied along z, y, then x axes, i.e. reading right to left along each row. The operators have been grouped according to the top row of the matrix, and have been factored so that the first nonzero element in the top row is unity.

quantum computers requires non Clifford gates.

The Clifford group consists of those operators which map elements of the Pauli group onto elements of the Pauli group. The Pauli group $\mathcal{P}_1$ for one qubit is the set of Pauli operators $\{I, X, Y, Z\}$ multiplied by ± 1 and $\pm i$, i.e.

$$\mathcal{P}_1 = \{\pm I, \pm iI, \pm X, \pm iX, \pm Y, \pm iY, \pm Z, \pm iZ\}.$$

The number of elements in the Pauli group acting on n qubits is $|\mathcal{P}_n| = 4^{n+1}$.

For one qubit the Clifford group $\mathcal{C}_1$ is generated by the operations $I, R_j(\pm\pi/2), R_j(\pi)$ about axes $j = x, y, z$. We include $R_j(\pm\pi/2)$ since these are distinct operators but do not need $R_j(\pm\pi)$ since $R_j(\pi) = -R_j(-\pi)$. Forming all permutations of the four operators $I, R_j(\pi/2), R_j(-\pi/2), R_j(\pi)$ acting on axes $j = x, y, z$ gives $4^3 = 64$ possible operators. We eliminate operators that are the same up to a global phase leaving the 24 distinct Clifford group elements shown in Table 12.2.

description	$[[n,k,d]]$	threshold (units of 10^{-3})	transversal gate set
smallest code[140]	$[[5,1,3]]$	0.03 [150]	PH, M_3[27]
Steane/color	$[[7,1,3]]$	$0.001 - 0.1$	Clifford group
Shor	$[[9,1,3]]$	0.2 [150, 151]	CNOT[27]
4.8.8 color	$[[17,1,5]]$	0.8 [148]	Clifford group
small surface code	$[[17,1,3]]$	0.8 [152]	Pauli group
Golay	$[[23,1,7]]$	1.3 [153]	Clifford group

Table 12.3: Characteristics of some small to medium size codes. Each $[[n,k,d]]$ code has k logical qubits encoded in n physical qubits and has distance d.

This method can be extended to generate the 11520 elements of $\mathcal{C}_2$ acting on two qubits. For two qubits the Clifford group is generated by the set $\{H, S, CNOT\}$. The operators are $I_a, R_{aj}(\pm\pi/2), R_{aj}(\pi)$ on each of two qubits. In addition to the single qubit operators we allow for three additional two-qubit operators $CNOT_{12}, CNOT_{21}, CNOT_{12}CNOT_{21}$ where $CNOT_{ab}$ is a CNOT gate with qubit a the control and qubit b the target. This gives $4(4^3)^2 = 16384$ possible operators. A large number of these are the same up to a global phase. The number of distinct elements in the Clifford group acting on n qubits is [M. Ozols 2008]

$$|\mathcal{C}_n| = 2^{n^2+2n} \prod_{j=1}^{n} \left(4^j - 1\right),$$

which gives $|\mathcal{C}_2| = 11520$, $|\mathcal{C}_3| = 92\,897\,280$, and $|\mathcal{C}_4| = 12\,128\,668\,876\,800$.

12.5.2 Gottesman-Knill theorem

The Gottesman-Knill theorem[30] says:

Suppose a quantum computation is performed which involves only the following elements: state preparations in the computational basis, Clifford group gates, and measurements of observables in the Pauli group (which includes measurement in the computational basis as a special case). Such a computation may be efficiently simulated on a classical computer.

The proof of this theorem is most easily presented using the stabilizer formalism - but we have not discussed stabilizer operators yet. Aaronson and Gottesman have provided a computer program which can be used for such simulations[10].

The implication of this theorem is that the Clifford group is not sufficient for universal quantum computation that can efficiently solve classically hard problems. Nevertheless Clifford group operators are important for error correction, can be used to create entangled states, and are important for randomized benchmarking tests of gate fidelity.

12.5.3 Towards fault tolerance

There are many different possible error correcting codes. A comparison of some small codes is given in Table 12.3. The indicated threshold is the "pseudo threshold" for which the logical error rate is equal to the underlying physical error rate. To further reduce the logical error

rate the physical rate must be smaller than the threshold in which case the logical error rate is

$$p_{\text{logical}} = p_{\text{th}} \left(\frac{p}{p_{\text{th}}} \right)^2 = p \frac{p}{p_{\text{th}}}.$$

Suppose the code threshold is $p_{\text{th}} = 10^{-3}$ and the physical error rate is an impressive $p = 10^{-4}$, the logical error rate will be $p_{\text{logical}} = 10^{-5}$, an improvement by an order of magnitude.

Nevertheless if the quantum computation requires more than 10^5 gate operations this is not good enough. The logical error rate can be further reduced by code concatenation in which each physical qubit that comprises the logical encoding is again protected by encoding in multiple physical qubits. At q layers of concatenation we need at least n^q physical qubits for each protected logical qubit that consists of n physical qubits. The logical error rate of the concatenated code is

$$p_{\text{logical}} = p_{\text{th}} \left(\frac{p}{p_{\text{th}}} \right)^{2^q}.$$

To get a feel for what this formula implies suppose we use a code with $p_{\text{th}} = 0.001$ and our hardware has $p = p_{\text{th}}/10 = 0.0001$. At $q = 1, 2, 3, 4$ the logical error rate will be $10^{-5}, 10^{-7}, 10^{-11}, 10^{-19}$. While the reduction in the logical error rate is dramatic due to the factor of 2^q in the exponent we should recall that the resource requirements are daunting. Using, for example, the 4.8.8 color code with $n = 17$ we need $17, 289, 4913, 83521$ qubits per logical qubit for $q = 1, 2, 3, 4$. To achieve low logical error rates and reduce the resource requirements we should either build better hardware with smaller p, design a code with larger p_{th}, invent a more efficient approach to error correction that does not rely on code concatenation.

Such a more efficient approach is provided by the surface code[154, 126]. The surface code does not work by concatenation but reduces the logical error rate by encoding in larger and larger tilings of qubits with error detection and correction provided by repeated measurement of X and Z stabilizers on 4-qubit plaquettes.

Even if we have a code design and hardware that provide a low logical error rate there is still one more outstanding challenge. It can be shown[26, 27] that it is not possible to implement a universal gate set transversally. In a transversal implementation a gate acting on a logical qubit is performed pairwise on the physical qubits. The advantage of this is that errors in one physical qubit will at most affect the corresponding physical qubit in the other logical qubit, and will not spread to additional physical qubits.

In order to nevertheless achieve fault tolerance there are several possible strategies.

a) Use a code that is transversal for Clifford gates and inject the T gates needed for universality using what is called "magic state" distillation which effectively teleports T gates into the circuit[155]. The distillation and teleportation procedure can be made fault tolerant but requires large resources. This type of approach can be implemented in the framework of concatenated codes or the surface code.

b) Combine different codes that are separately transversal for some gates and together are transversal for a universal gate set[156, 157]. This is known as code switching. The smallest known code size for which this is possible is $n = 49$[158].

c) Give up the requirement of a transversal code and nevertheless achieve fault tolerance by adding additional error checking to the circuit[159].

At the present time the most efficient route towards a fault tolerant encoding that provides universal computation appears to be the surface code with magic state distillation for injection of T gates[160]. This is a very active area of research and it is likely we will see new ideas and results in the coming years[139].

Chapter 13

Implementations

13.1 Physical resource requirements

13.2 DiVincenzo criteria

Chapter 14

Trapped ion qubits

To be added....

Chapter 15

Neutral atom qubits

To be added....

Chapter 16

Superconducting qubits

We can encode qubits in single particles such as ions, neutral atoms, quantum dots, or photons. Superconducting qubits are different. The qubit is encoded in the state of a circuit element composed of a large number of atoms. This is an example of "macroscopic" quantum coherence. There has been remarkable progress in developing supeconducting circuits and circuit elements for quantum computing in the last two decades. Small circuits with less than 10 qubits and gate fidelities above 0.999, as well as larger circuits containing more than 50 qubits, but with as yet unreported gate fidelities, have been fabricated. Many believe that superconductors represent the leadingpath towards a practical quantum computer and the bulk of commercial investmant is currently directed at superconducting devices.

16.1 Circuit quantization

We are familiar with basic linear circuit elements: the resistor R, the capacitor C and the inductor L. The relations between charge Q, magnetic flux Φ, current I, and voltage V for these elements are

$$R = \frac{V}{I}, \quad C = \frac{Q}{V}, \quad L = \frac{\Phi}{I}. \tag{16.1}$$

Charge and current on a capacitor are related by

$$I = \frac{dQ}{dt} = C\frac{dV}{dt}$$

while magnetic flux and voltage in an inductor are related by

$$V = -\int d\ell \cdot \mathbf{E} = \frac{d\Phi}{dt} = L\frac{dI}{dt}.$$

Note that correct definition of signs in the above relations can be confusing and is best done with reference to a circuit diagram. The power dissipated in a resistor is $P = IV = \frac{V^2}{R} = RI^2$ and the stored energy in a capacitor is $U_C = \frac{1}{2}CV^2$ and in an inductor $U_L = \frac{1}{2}LI^2$.

The resistor is dissipative and heat is generated by a current. In contrast ideal capacitors and inductors are lossless. In reality circuit elements are predominantly of one type by contain small effective contributions of other types of elements.

mechanical		electronic	
position	x	flux	Φ
momentum	p	charge	Q
mass	m	capacitance	C
spring constant	κ	inverse inductance	$1/L$
frequency	$\omega = \sqrt{\kappa/m}$		$\omega = 1/\sqrt{LC}$
quantization			
Hamiltonian	$H = \frac{p^2}{2m} + \frac{\kappa x^2}{2}$		$H = \frac{1}{2}\frac{\Phi^2}{L} + \frac{1}{2}\frac{Q^2}{C}$
coordinate	$x \to \hat{x}$		$\Phi \to \hat{\Phi}$
	$= \sqrt{\frac{\hbar}{2m\omega}}(\hat{a} + \hat{a}^\dagger)$		$= \sqrt{\frac{\hbar\sqrt{L}}{2\sqrt{C}}}(\hat{a} + \hat{a}^\dagger)$
momentum	$p \to \hat{p} = -i\hbar\frac{\partial}{\partial x}$		$Q \to \hat{Q} = -i\hbar\frac{\partial}{\partial \Phi}$
	$= i\sqrt{\frac{m\hbar\omega}{2}}(\hat{a}^\dagger - \hat{a})$		$= \sqrt{\frac{\hbar\sqrt{C}}{2\sqrt{L}}}(\hat{a}^\dagger - \hat{a})$
commutator	$[\hat{x}, \hat{p}] = i\hbar$		$[\hat{\Phi}, \hat{Q}] = i\hbar$

Table 16.1: Mechanical and electronic oscillators.

An LC circuit containing an ideal inductor and capacitor connected in series will oscillate at frequency ω. The circuit relations are

$$V = L\frac{dI}{dt}$$

and

$$-\frac{dV}{dt} = \frac{I}{C} = -L\frac{d^2I}{dt^2}$$

so

$$\frac{d^2I}{dt^2} = -\frac{1}{LC}I.$$

The solution is $I = I_0\cos(\omega t)$ with $\omega = \frac{1}{\sqrt{LC}}$.

Some typical values of small elements are $C = 1$ pF, $L = 10$ nH giving $\omega = 10^{10}$ rad/s, or in cycles per second $f = \frac{\omega}{2\pi} \simeq 1.6$ GHz. The dimensions of C and L with these values are less than 1 mm. On the other hand the wavelength of the oscillation is

$$\lambda = \frac{c}{f} \simeq 19 \text{ cm}.$$

We see that $\lambda \gg$ the size of the circuit elements. This situation is referred to as a lumped element circuit.

The magnetic flux Φ is a collective degree of freedom. It corresponds to the position of a mass in a mechanical oscillator. The charge Q corresponds to the momentum and Φ and Q are conjugate coordinates. This analogy follows from the Hamiltonian for the circuit. For a mechanical oscillator the energy or Hamiltonian is

$$H = \frac{p^2}{2m} + \frac{\kappa x^2}{2}$$

with m the mass, p the momentum, and κ the spring constant. For the LC circuit the stored energy is

$$H = \frac{1}{2}CV^2 + \frac{1}{2}LI^2 = \frac{1}{2}\frac{Q^2}{C} + \frac{1}{2}\frac{\Phi^2}{L}.$$

Both Hamiltonians are quadratic in a pair of conjugate variables. We can therefore quantize the electronic oscillator by analogy with quantization of a mechanical oscillator. The analogs between the mechanical and LC oscillators are listed in Table 16.1.

The energy levels of the LC oscillator are quantized, but they are equidistant so this is not directly useful as a qubit. If we were to say encode a qubit in the lowest two levels, a resonant field would also couple the second level to the third level, and so on, leading to leakage out of the computational basis. While it is interesting to note that there is a way to encode a qubit in a harmonic oscillator[161], other challenges arise such as the need to prepare highly nonclassical states to perform gate operations.

The approach that is instead followed with superconducting qubits is to introduce a nonlinear element to so that the energy levels are no longer equally spaced. In this way the two lowest levels can be independently addressed and used as a qubit. The nonlinear element that is used is the Josephson junction.

16.2 Superconductivity

In order to understand how Josephson junctions operate we should first learn something about superconductivity. In any normal metal electrons, which carry current, scatter off impurities and are repelled by other electrons all of which have the same electronic charge. The result is a finite resistivity and dissipation of energy in the current carrying material. Dissipation is undesirable for qubit devices and can be avoided at low temperatures where many materials become superconducting and carry current while exhibiting zero resistivity. A microscopic theory of superconductivity was first developed by Bardeen, Cooper and Schreiffer[162]. Here we give a simplified argument for how electrons, even though they electrically repel each other, can exhibit a weak attraction inside a material.

The physical picture is that a conduction electron will exert an attractive force on a positive ion in the valence band, giving a small charge displacement. A second conduction electron will be attracted by the displaced ion leading to an effective attraction between electrons. The effect is weak so in order for the attractive force to dominate over thermal fluctuations the material must be sufficiently cold. The attractive force leads to electron pairing, into so-called Cooper pairs[163]. The paired electrons form an integer spin boson, with many Cooper pairs contributing to a supercurrent.

A derivation of the effective attractive potential proceeds as follows. Consider two electrons with wave functions expressed as sums over plane wave states with momenta $\mathbf{k}$

$$\begin{aligned}\psi_1(\mathbf{r}_1) &= \sum_{\mathbf{k}} a'_{\mathbf{k}} e^{\imath \mathbf{k}\cdot\mathbf{r}_1} \\ \psi_2(\mathbf{r}_2) &= \sum_{\mathbf{k}} a'_{\mathbf{k}} e^{\imath(-\mathbf{k})\cdot\mathbf{r}_2}.\end{aligned}$$

The a_k are plane wave expansion coefficients and the electrons are assumed to have opposite

momenta. The joint wavefunction is

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2) = \sum_{\mathbf{k}} a_{\mathbf{k}} e^{\imath \mathbf{k}\cdot(\mathbf{r}_1 - \mathbf{r}_2)} \tag{16.2}$$

with $a_{\mathbf{k}} = (a'_{\mathbf{k}})^2$.

The time independent Schrödinger equation is

$$-\frac{\hbar^2}{2m}(\nabla_1^2 + \nabla_2^2)\psi + V(\mathbf{r}_1 - \mathbf{r}_2)\psi = E\psi.$$

Here $V(\mathbf{r}_1 - \mathbf{r}_2)$ is the interaction potential which we assume only depends on the separation of the electrons, m is the mass of an electron, and E is the energy. Using (16.2) we find

$$\frac{\hbar^2}{m}\sum_{\mathbf{k}} a_{\mathbf{k}} k^2 e^{\imath \mathbf{k}\cdot(\mathbf{r}_1 - \mathbf{r}_2)} + [V(\mathbf{r}_1 - \mathbf{r}_2) - E]\sum_{\mathbf{k}} a_{\mathbf{k}} e^{\imath \mathbf{k}\cdot(\mathbf{r}_1 - \mathbf{r}_2)} = 0.$$

Then multiply by $e^{\imath \mathbf{k}'\cdot(\mathbf{r}_1 - \mathbf{r}_2)}$ which gives

$$\frac{\hbar^2}{m}\sum_{\mathbf{k}} a_{\mathbf{k}} k^2 e^{\imath (\mathbf{k}+\mathbf{k}')\cdot(\mathbf{r}_1 - \mathbf{r}_2)} + [V(\mathbf{r}_1 - \mathbf{r}_2) - E]\sum_{\mathbf{k}} a_{\mathbf{k}} e^{\imath (\mathbf{k}+\mathbf{k}')\cdot(\mathbf{r}_1 - \mathbf{r}_2)} = 0.$$

Integrate over $\mathbf{r}_1 - \mathbf{r}_2$ using

$$\int d^3r\, e^{\imath(\mathbf{k}+\mathbf{k}')\cdot\mathbf{r}} = (2\pi)^3\delta(\mathbf{k}+\mathbf{k}')$$

to get

$$\sum_{\mathbf{k}}\left(E - \frac{\hbar^2 k^2}{m}\right) a_{\mathbf{k}}\delta(\mathbf{k}+\mathbf{k}') = \frac{1}{(2\pi)^3}\sum_{\mathbf{k}}\int d^3r\, V(\mathbf{r}) e^{\imath(\mathbf{k}+\mathbf{k}')\cdot\mathbf{r}} a_{\mathbf{k}}.$$

The left hand side is zero unless $\mathbf{k} = -\mathbf{k}'$ in which case

$$\left(E - \frac{\hbar^2 k'^2}{m}\right) a_{-\mathbf{k}'} = \frac{1}{(2\pi)^3}\sum_{\mathbf{k}''}\int d^3r\, V(\mathbf{r}) e^{\imath(\mathbf{k}''+\mathbf{k}')\cdot\mathbf{r}} a_{\mathbf{k}''}.$$

Relabeling dummy indices this becomes

$$(E - 2\epsilon_k)\, a_{\mathbf{k}} = \sum_{\mathbf{k}'} V_{\mathbf{k},\mathbf{k}'} a_{\mathbf{k}'}$$

with $\epsilon_k = \frac{\hbar^2 k^2}{2m}$ and $V_{\mathbf{k},\mathbf{k}'} = \frac{1}{(2\pi)^3}\int d^3r\, V(\mathbf{r}) e^{\imath(\mathbf{k}'-\mathbf{k})\cdot\mathbf{r}}$ is the interaction energy between momentum eigenstates.

We then assume that $V_{\mathbf{k},\mathbf{k}'} = 0$ for $k' > k_c$ with k_c a high momentum critical wavenumber and $V_{\mathbf{k},\mathbf{k}'} = 0$ for $k' < k_F$, the Fermi wavenumber. We also approximate $V_{\mathbf{k},\mathbf{k}'}$ by a constant $-V$, with $V > 0$ for $k_F < k, k' < k_c$. With these approximations and replacing the sum

by an integral over the density of states $N(\epsilon)\sum_{\mathbf{k}} \to \int d\epsilon\, N(\epsilon) \simeq N_F \int d\epsilon$, where N_F is the density of states at the Fermi energy, we find

$$e^{2/N_F V} = \frac{2E_F + E_c - E}{2E_F - E}.$$

Assuming a weak interaction so that $N_F V \ll 1$ we arrive at

$$E \simeq 2E_F - 2E_c e^{-2/N_F V}.$$

The second term on the right hand side is negative so $E < 2E_F$ which implies an attractive potential and results in Cooper pairs.

16.3 Flux Quantization

We introduce a collective wavefunction for the supercurrent of Cooper pairs

$$\psi(\mathbf{r}, t) = \sqrt{n(\mathbf{r}, t)} e^{\imath\phi(\mathbf{r},t)}$$

with $n = |\psi|^2$ the charge density and ϕ the phase.

The electric current is

$$\mathbf{J} = (2e) \int d^3r\, \psi^* \mathbf{v} \psi$$

with the velocity $\mathbf{v} = -i\frac{\hbar}{m}\nabla - \frac{q}{mc}\mathbf{A}$ where $q = 2e$ is the charge of a Cooper pair and $|bfA$ is the vector potential. Assuming uniform charge density and potential and linear phase gradient we find

$$\mathbf{J} = \frac{2en}{m}\left(\hbar\nabla\phi - \frac{2e}{c}\mathbf{A}\right). \tag{16.3}$$

It then follows from the Maxwell equations that the magnetic field satisfies

$$\nabla^2 \mathbf{B} = \frac{1}{\lambda^2}\mathbf{B}$$

with $\lambda^2 = \frac{c}{4\pi}\frac{mc}{4ne^2}$. This is known as the London equation. Solving in a one-dimensional greometry shows that the magnetic field inside a superconductor vanishes exponentially with a penetration length λ. Taking $n \sim 10^{28}$ m^{-3} for the charge density gives $\lambda \sim 25$ nm. Vanishing of the magnetic field inside a superconductor is known as the Meissner effect.

Consider a superconducting ring as in Fig. 16.1 with zero current and magnetic field inside the material of the ring. From Eq. (16.3) zero current implies

$$\nabla\phi = \frac{2e}{\hbar c}\mathbf{A}$$

where Φ is the magnetic flux inside the ring. Integrating the phase around the dashed

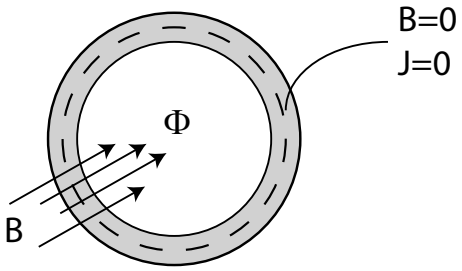


Figure 16.1: Superconducting ring in an applied magnetic field.

contour we find

$$\begin{aligned}\oint d\mathbf{l}\cdot\nabla\phi &= \frac{2e}{\hbar c}\oint d\mathbf{l}\cdot\mathbf{A}\\ &= \frac{2e}{\hbar c}\oint d\mathbf{a}\cdot\nabla\times\mathbf{A}\\ &= \frac{2e}{\hbar c}\oint d\mathbf{a}\cdot\mathbf{B}\\ &= \frac{2e}{\hbar c}\Phi.\end{aligned}$$

Requiring that the wavefunction be single valued implies that the change in phase after one loop around the ring is a multiple of 2π or

$$\Phi = \Phi_0 p$$

with p an integer and $\Phi_0^{(\mathrm{cgs})} = \frac{\hbar c}{2e}$ $\left(\Phi_0^{(\mathrm{SI})} = \frac{h}{2e}\right)$ the quantum of magnetic flux. Note the charge is not e but $2e$ which corresponds to a Cooper pair.

16.4 Josephson Junction

The Josephson junction can be modeled as a thin piece of insulating material separating two superconductors. See handwritten notes.....

Chapter 17

Quantum dot qubits

To be added....

Chapter 18

Photonic qubits

To be added....

Appendix A

Probability distributions

A few of the most important probability distributions are defined here.

A.1 Binomial distribution

Consider n independent trials of a random event, each of which gives "heads" with probability p and tails with probability $1-p$. The probability of observing x heads is

$$\begin{aligned} P_{\rm B}(x) &= \binom{n}{x} p^x (1-p)^{n-x} \\ &= \frac{n!}{(n-x)!x!} p^x (1-p)^{n-x}. \end{aligned}$$

If $p = 1/2$ then

$$P_{\rm B}(n/2) = \frac{n!}{(n/2)!(n/2)!}(1/2)^n.$$

A.2 Poisson distribution

The binomial distribution is exact, but inconvenient for mathematical analysis. For many trials (large n) we can approximate the binomial distribution with a continuous function. The Poisson distribution is

$$P_{\rm P}(x) = \frac{e^{-np}(np)^x}{x!}.$$

This can be derived from the binomial distribution using Sterling's formula $n! \simeq \sqrt{2\pi} n^{n+1/2} e^{-n}$. The mean and standard deviation of the Poisson distribution are

$$\langle x \rangle = np = N, \quad \sigma = \sqrt{np} = \sqrt{N}$$

so the fractional error in estimating the mean is

$$\frac{\sigma}{\langle x \rangle} = \frac{1}{\sqrt{n}p} = \frac{1}{\sqrt{N}}.$$

A.3 Gaussian distribution

In the limit of large n the Poisson distribution can be approximated by a Gaussian distribution

$$P_{\rm G}(x) = \left(\frac{1}{2\pi\sigma^2}\right)^{1/2} e^{-(x-\langle x\rangle)^2/(2\sigma^2)}.$$

The standard deviation is $\sigma = \sqrt{\langle x\rangle}$ which is the same as for the Poisson distribution.

A.4 Central Limit Theorem

This theorem says that the average value of N independent random variables as $N \to \infty$ is Gaussian. Let

$$y = \frac{x_1 + x_2 + ...x_N}{N}$$

with the x_i independent and identically distributed random variables. Then $\langle x_i\rangle = \langle x_j\rangle = \langle x\rangle$ is the mean and σ is the standard deviation of each random variable. The probability distribution of y is

$$P(y - \langle x\rangle) = \left(\frac{1}{2\pi\sigma_y^2}\right)^{1/2} e^{-(y-\langle x\rangle)^2/(2\sigma_y^2)}$$

which is a Gaussian distribution with standard deviation $\sigma_y = \sigma/\sqrt{N}$.

Bibliography

[1] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, *A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem*, Science **292**, 472 (2001).

[2] R. Raussendorf and H. J. Briegel, *A one-way quantum computer*, Phys. Rev. Lett. **86**, 5188 (2001).

[3] P. G. Kwiat, A. J. Berglund, J. B. Altepeter, and A. G. White, *Experimental Verification of Decoherence-Free Subspaces*, Science **290**, 498–501 (2000).

[4] R. J. C. Spreeuw, *Classical wave-optics analogy of quantum-information processing*, Phys. Rev. A **63**, 062302 (2001).

[5] N. Bhattacharya, H. B. van Linden van den Heuvell, and R. J. C. Spreeuw, *Implementation of quantum search algorithm using classical Fourier optics*, Phys. Rev. Lett. **88**, 137901 (2002).

[6] R. Jozsa and N. Linden, *On the role of entanglement in quantum-computational speed-up*, Proc. Roy. Soc. Lond. A **459**, 2011 (2003).

[7] E. Schrödinger, *Die gegenwärtige situation in der quantenmechanik*, Naturwissenschaften **23**, 807–812; 823–828; 844–849 (1935).

[8] E. Schrödinger, *Discussion of probability relations between separated systems*, Math. Proc. Cambr. Philos. Soc. **31**, 555 (1935).

[9] E. Schrödinger, *Probability relations between separated systems*, Math. Proc. Cambr. Philos. Soc. **32**, 446 (1936).

[10] S. Aaronson and D. Gottesman, *Improved simulation of stabilizer circuits*, Phys. Rev. A **70**, 052328 (2004).

[11] M. Van den Nest, *Universal quantum computation with little entanglement*, Phys. Rev. Lett. **110**, 060504 (2013).

[12] M. Howard, J. Wallman, V. Veitch, and J. Emerson, *Contextuality supplies the 'magic' for quantum computation*, Nature **510**, 351 (2014).

[13] R. P. Feynman, *Simulating physics with computers*, Int. J. Theor. Phys. **21**, 467 (1982).

[14] S. Lloyd, *Univeral quantum simulators*, Science **273**, 1073 (1996).

[15] K. D. Raedt, K. Michielsen, H. D. Raedt, B. Trieuc, G. Arnold, M. Richter, T. Lippert, H. Watanabe, and N. Itoe, *Massively parallel quantum computer simulator*, Comp. Phys. Commun. **176**, 121 (2007).

[16] T. Häner and D. S. Steiger, *0.5 petabyte simulation of a 45-qubit quantum circuit*, arXiv:1704.01127 (2017).

[17] E. Pednault, J. A. Gunnels, G. Nannicini, L. Horesh, T. Magerlein, E. Solomonik, and R. Wisnieff, *Breaking the 49-qubit barrier in the simulation of quantum circuits*, arXiv:1710.05867 (2017).

[18] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, and H. Neven, *Simulation of low-depth quantum circuits as complex undirected graphical models*, arXiv:1712.05384 (2017).

[19] J. W. Cooley and J. Tukey, *An algorithm for the machine calculation of complex Fourier series*, Math. Comp. **19**, 297 (1965).

[20] A. G. Fowler and L. C. L. Hollenberg, *Scalability of Shor's algorithm with a limited set of rotation gates*, Phys. Rev. A **70**, 032329 (2004).

[21] R. B. Griffiths and C.-S. Niu, *Semiclassical Fourier transform for quantum computation*, Phys. Rev. Lett. **76**, 3228–3231 (1996).

[22] J. Chiaverini, J. Britton, D. Leibfried, E. Knill, M. D. Barrett, R. B. Blakestad, W. M. Itano, J. D. Jost, C. Langer, R. Ozeri, T. Schaetz, and D. J. Wineland, *Implementation of the semiclassical quantum Fourier transform in a scalable system*, Science **308**, 997 (2005).

[23] D. P. DiVincenzo, *Two-bit gates are universal for quantum computation*, Phys. Rev. A **51**, 1015–1022 (1995).

[24] T. Sleator and H. Weinfurter, *Realizable universal quantum logic gates*, Phys. Rev. Lett. **74**, 4087 (1995).

[25] A. Barenco, C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. A. Smolin, and H. Weinfurter, *Elementary gates for quantum computation*, Phys. Rev. A **52**, 3457–3467 (1995).

[26] B. Eastin and E. Knill, *Restrictions on transversal encoded quantum gate sets*, Phys. Rev. Lett. **102**, 110502 (2009).

[27] B. Zeng, A. Cross, and I. L. Chuang, *Transversality versus universality for additive quantum codes*, IEEE Trans. Inf. Th. **57**, 6272 (2011).

[28] D. Beckman, A. N. Chari, S. Devabhaktuni, and J. Preskill, *Efficient networks for quantum factoring*, Phys. Rev. A **54**, 1034 (1996).

[29] A. Y. Kitaev, *Quantum computations: algorithms and error correction*, Russ. Math. Surv. **52**, 1191 (1997).

[30] D. Gottesman, *The Heisenberg representation of quantum computers*, quant-ph/9807006 (1998).

[31] W. K. Wootters and W. H. Zurek, *A single quantum cannot be cloned*, Nature **299**, 802 (1982).

[32] D. Dieks, *Communication by EPR devices*, Phys. Lett. **92A**, 271 (1982).

[33] S. Weinberg, *Precision tests of quantum mechanics*, Phys. Rev. Lett. **62**, 485–488 (1989).

[34] J. J. Bollinger, D. J. Heinzen, W. M. Itano, S. L. Gilbert, and D. J. Wineland, *Test of the linearity of quantum mechanics by rf spectroscopy of the* $^9Be^+$ *ground state*, Phys. Rev. Lett. **63**, 1031–1034 (1989).

[35] D. Gottesman, *Fault-tolerant quantum computation with higher-dimensional systems*, in Quantum Computing and Quantum Communications, edited by C. P. Williams, Lecture Notes in Computer Science, (Springer, Berlin/Heidelberg), **1509**, 302 (1999).

[36] M. Howard and J. Vala, *Qudit versions of the qubit* $\pi/8$ *gate*, Phys. Rev. A **86**, 022316 (2012).

[37] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press) (2000).

[38] D. Deutsch, *Quantum-theory, the Church-Turing principle and the universal quantum computer*, Proc. Roy. Soc. **400**, 97 (1985).

[39] D. Deutsch and R. Jozsa, *Rapid solution of problems by quantum computation*, Proc. Roy. Soc. **439**, 553 (1992).

[40] S. Gulde, M. Riebe, G. P. T. Lancaster, C. Becher, J. Eschner, H. Häffner, F. Schmidt-Kaler, I. Chuang, and R. Blatt, *Implementation of the Deutsch-Jozsa algorithm on an ion-trap quantum computer*, Nature **421**, 48 (2003).

[41] L. DiCarlo, J. M. Chow, J. M. Gambetta, L. S. Bishop, B. R. Johnson, D. I. Schuster, J. Majer, A. Blais, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, *Demonstration of two-qubit algorithms with a superconducting quantum processor*, Nature (London) **460**, 240 (2009).

[42] Z. Wu, J. Li, W. Zheng, J. Luo, M. Feng, and X. Peng, *Experimental demonstration of the Deutsch-Jozsa algorithm in homonuclear multispin systems*, Phys. Rev. A **84**, 042312 (2011).

[43] E. Brainis, L.-P. Lamoureux, N. J. Cerf, P. Emplit, M. Haelterman, and S. Massar, *Fiber-optics implementation of the Deutsch-Jozsa and Bernstein-Vazirani quantum algorithms with three qubits*, Phys. Rev. Lett. **90**, 157902 (2003).

[44] G. Vallone, G. Donati, N. Bruno, A. Chiuri, and P. Mataloni, *Experimental realization of the Deutsch-Jozsa algorithm with a six-qubit cluster state*, Phys. Rev. A **81**, 050302 (2010).

[45] L. K. Grover, *Quantum mechanics helps in searching for a needle in a haystack*, Phys. Rev. Lett. **79**, 325 (1997).

[46] C. Zalka, *Grovers quantum searching algorithm is optimal*, Phys. Rev. A **60**, 2746 (1999).

[47] G. F. Viamontes, I. L. Markov, and J. P. Hayes, *Improving gate-level simulation of quantum circuits*, Quant. Inf. Proc. **2**, 347 (2003).

[48] G. L. Long, *Grover algorithm with zero theoretical failure rate*, Phys. Rev. A **64**, 022307 (2001).

[49] T. Xia, M. Lichtman, K. Maller, A. W. Carr, M. J. Piotrowicz, L. Isenhower, and M. Saffman, *Randomized benchmarking of single-qubit gates in a 2D array of neutral-atom qubits*, Phys. Rev. Lett. **114**, 100503 (2015).

[50] L. Isenhower, M. Saffman, and K. Mølmer, *Multibit C_kNOT quantum gates via Rydberg blockade*, Quant. Inf. Proc. **10**, 755 (2011).

[51] K. Mølmer, L. Isenhower, and M. Saffman, *Efficient Grover search with Rydberg blockade*, J. Phys. B: At. Mol. Opt. Phys. **44**, 184016 (2011).

[52] Z. Diao, M. S. Zubairy, and G. Chen, *A quantum circuit design for Grover's algorithm*, Z. Naturforsch. **57a**, 701 (2002).

[53] P. W. Shor, *Algorithms for quantum computation: Discrete logarithms and factoring*, in Proc. 35th Annual Symposium on Foundations of Computer Science, IEEE Computer Society Press pp. 124–134 (1994).

[54] P. W. Shor, *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM J. Comput. **26**, 1484 (1997).

[55] C. Pomerance, *A tale of two sieves*, Notices AMS **43**, 1473 (1996).

[56] A. Pavlidis and D. Gizopoulos, *Fast quantum modular exponentiation architecture for Shor's factoring algorithm*, Qu. Inf. Comput. **14**, 649 (2014).

[57] T. Monz, D. Nigg, E. A. Martinez, M. F. Brandl, P. Schindler, R. Rines, S. X. Wang, I. L. Chuang, and R. Blatt, *Realization of a scalable Shor algorithm*, Science **351**, 1068 (2016).

[58] E. Martín-López, A. Laing, T. Lawson, R. Alvarez, X.-Q. Zhou, and J. L. O'Brien, *Experimental realisation of Shor's quantum factoring algorithm using qubit recycling*, Nat. Photon. **6**, 773 (2012).

[59] N. Xu, J. Zhu, D. Lu, X. Zhou, X. Peng, and J. Du, *Quantum factorization of 143 on a dipolar-coupling nuclear magnetic resonance system*, Phys. Rev. Lett. **108**, 130501 (2012).

[60] R. Dridi and H. Alghassi, *Prime factorization using quantum annealing and computational algebraic geometry*, Sci. Rep. **7**, 43048 (2017).

[61] R. L. Rivest, A. Shamir, and L. Adleman, *A method for obtaining digital signatures and public key cryptosystems*, Commun. ACM **21**, 120 (1978).

[62] M. Agrawal, N. Kayal, and N. Saxena, *Primes is in P*, Ann. Math. **160**, 781 (2004).

[63] P. S. Bourdon and H. T. Williams, *Sharp probability estimates for Shor's order-finding algorithm*, Quant. Inf. Comput. **7**, 522 (2007).

[64] G. Schaller and R. Schützhold, *The role of symmetries in adiabatic quantum algorithms*, Quant. Inf. Comput. **10**, 0109 (2010).

[65] D. Coppersmith, *An approximate Fourier transform useful in quantum factoring*, IBM Research Report p. RC 19642 (1994).

[66] Y. S. Nam and R. Blümel, *Scaling laws for Shor's algorithm with a banded quantum Fourier transform*, Phys. Rev. A **87**, 032333 (2013).

[67] I. L. Markov and M. Saeedi, *Faster quantum number factoring via circuit synthesis*, Phys. Rev. A **87**, 012310 (2013).

[68] R. F. Werner, *Quantum states with einstein-podolsky-rosen correlations admitting a hidden-variable model*, Phys. Rev. A **40**, 4277 (1989).

[69] O. Gühne and G. Tóth, *Entanglement detection*, Phys. Rep. **474**, 1 (2009).

[70] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, *Quantum entanglement*, Rev. Mod. Phys. **81**, 865–942 (2009).

[71] A. Peres, *Seperability criterion for density matrices*, Phys. Rev. Lett. **77**, 1413 (1996).

[72] M. Horodecki, P. Horodecki, and R. Horodecki, *Separability of mixed states: Necessary and sufficient conditions,*, Phys. Lett. A **223**, 1 (1996).

[73] W. K. Wootters, *Entanglement of formation of an arbitrary state of two qubits*, Phys. Rev. Lett. **80**, 2245 (1998).

[74] C. A. Sackett, D. Kielpinski, B. E. King, C. Langer, V. Meyer, C. J. Myatt, M. Rowe, Q. A. Turchette, W. M. Itano, D. J. Wineland, and C. Monroe, *Experimental entanglement of four particles*, Nature (London) **404**, 256 (2000).

[75] Q. A. Turchette, C. S. Wood, B. E. King, C. J. Myatt, D. Leibfried, W. M. Itano, C. Monroe, and D. J. Wineland, *Deterministic entanglement of two trapped ions*, Phys. Rev. Lett. **81**, 3631 (1998).

[76] L.-M. Duan, G. Giedke, J. I. Cirac, and P. Zoller, *Inseparability criterion for continuous variable systems*, Phys. Rev. Lett. **84**, 2722–2725 (2000).

[77] R. Simon, *Peres-Horodecki separability criterion for continuous variable systems*, Phys. Rev. Lett. **84**, 2726–2729 (2000).

[78] A. Peres, *Quantum Theory: Concepts and Methods* (Kluwer, Dordrecht, The Netherlands) (1995).

[79] M. Arndt, O. Nairz, J. Vos-Andreae, C. Keller, G. van der Zouw, and A. Zeilinger, *Wave-particle duality of c_{60} molecules*, Nature **401**, 680 (1999).

[80] A. Einstein, B. Podolsky, and N. Rosen, *Can quantum-mechanical description of physical reality be considered complete?*, Phys. Rev. **47**, 777–780 (1935).

[81] A. Peres, *Unperformed experiments have no results*, Am. J. Phys. **46**, 745 (1978).

[82] D. Bohm, *Quantum theory* (Prentice-Hall, Inc., Englewood Cliffs, USA) (1951).

[83] J. S. Bell, *On the einstein podolsky rosen paradox*, Physics **1**, 195 (1964).

[84] G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger, *Violation of Bell's inequality under strict Einstein locality conditions*, Phys. Rev. Lett. **81**, 5039–5043 (1998).

[85] M. A. Rowe, D. Kielpinski, V. Meyer, C. A. Sackett, W. M. Itano, C. Monroe, and D. J. Wineland, *Experimental violation of a Bell's inequality with efficient detection*, Nature **409**, 791 (2001).

[86] B. Hensen, H. Bernien, A. E. Dréau, A. Reiserer, N. Kalb, M. S. Blok, J. Ruitenberg, R. F. L. Vermeulen, R. N. Schouten, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, M. Markham, D. J. Twitchen, D. Elkouss, S. Wehner, and T. H. T. amd R. Hanson, *Loophole-free Bell inequality violation using electron spins separated by 1.3 kilometres*, Nature **526**, 682 (2015).

[87] M. Giustina, M. A. M. Versteegh, S. Wengerowsky, J. Handsteiner, A. Hochrainer, K. Phelan, F. Steinlechner, J. Kofler, J.-A. Larsson, C. Abellán, W. Amaya, V. Pruneri, M. W. Mitchell, J. Beyer, T. Gerrits, A. E. Lita, L. K. Shalm, S. W. Nam, T. Scheidl, R. Ursin, B. Wittmann, and A. Zeilinger, *Significant-loophole-free test of Bell's theorem with entangled photons*, Phys. Rev. Lett. **115**, 250401 (2015).

[88] L. K. Shalm, E. Meyer-Scott, B. G. Christensen, P. Bierhorst, M. A. Wayne, M. J. Stevens, T. Gerrits, S. Glancy, D. R. Hamel, M. S. Allman, K. J. Coakley, S. D. Dyer, C. Hodge, A. E. Lita, V. B. Verma, C. Lambrocco, E. Tortorici, A. L. Migdall, Y. Zhang, D. R. Kumor, W. H. Farr, F. Marsili, M. D. Shaw, J. A. Stern, C. Abellán, W. Amaya, V. Pruneri, T. Jennewein, M. W. Mitchell, P. G. Kwiat, J. C. Bienfang, R. P. Mirin, E. Knill, and S. W. Nam, *Strong loophole-free test of local realism*, Phys. Rev. Lett. **115**, 250402 (2015).

[89] C. Abellán, W. Amaya, D. Mitrani, V. Pruneri, and M. W. Mitchell, *Generation of fresh and pure random numbers for loophole-free Bell tests*, Phys. Rev. Lett. **115**, 250403 (2015).

[90] T. Scheidl, R. Ursin, J. Kofler, S. Ramelow, X.-S. Ma, T. Herbst, L. Ratschbacher, A. Fedrizzi, N. K. Langford, T. Jennewein, and A. Zeilinger, *Violation of local realism with freedom of choice*, PNAS **107**, 19708 (2010).

[91] J. Gallicchio, A. S. Friedman, and D. I. Kaiser, *Testing Bell's inequality with cosmic photons: Closing the setting-independence loophole*, Phys. Rev. Lett. **112**, 110405 (2014).

[92] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, *Proposed experiment to test local hidden-variable theories*, Phys. Rev. Lett. **23**, 880–884 (1969).

[93] P. G. Kwiat and L. Hardy, *The mystery of the quantum cakes*, Am. J. Phys. **68**, 33 (2000).

[94] B. S. Cirel'son, *Quantum generalizations of bell's inequality*, Lett. Math. Phys. **4**, 93 (1980).

[95] A. K. Ekert, *Quantum cryptography based on Bell's theorem*, Phys. Rev. Lett. **67**, 661–663 (1991).

[96] W. Dür, G. Vidal, and J. I. Cirac, *Three qubits can be entangled in two inequivalent ways*, Phys. Rev. A **62**, 062314 (2000).

[97] N. Gisin, *Hidden quantum nonlocality revealed by local filters*, Phys. Lett. A **210**, 151 (1996).

[98] F. Haas, J. Volz, R. Gehr, J. Reichel, and J. Estève, *Entangled states of more than 40 atoms in an optical fiber cavity*, Science **344**, 180–183 (2014).

[99] S.-K. Liao, W.-Q. Cai, J. Handsteiner, B. Liu, J. Yin, L. Zhang, D. Rauch, M. Fink, J.-G. Ren, W.-Y. Liu, Y. Li, Q. Shen, Y. Cao, F.-Z. Li, J.-F. Wang, Y.-M. Huang, L. Deng, T. Xi, L. Ma, T. Hu, L. Li, N.-L. Liu, F. Koidl, P. Wang, Y.-A. Chen, X.-B. Wang, M. Steindorfer, G. Kirchner, C.-Y. Lu, R. Shu, R. Ursin, T. Scheidl, C.-Z. Peng, J.-Y. Wang, A. Zeilinger, and J.-W. Pan, *Satellite-relayed intercontinental quantum network*, Phys. Rev. Lett. **120**, 030501 (2018).

[100] A. S. Holevo, *Bounds for the quantity of information transmitted by a quantum communication channel*, Probl. Pered. Infor. **9**, 3 (1973), translation: Information theoretical aspects of quantum measurements, in Probl. Inf. Trans. **9**, 177 (1973).

[101] B. C. H. and S. J. Wiesner, *Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states*, Phys. Rev. Lett. **69**, 2881 (1992).

[102] K. Mattle, H. Weinfurter, P. G. Kwiat, and A. Zeilinger, *Dense coding in experimental quantum communication*, Phys. Rev. Lett. **76**, 4656 (1996), translation: Information theoretical aspects of quantum measurements, in Probl. Inf. Trans. **9**, 177 (1973).

[103] B. C. H., G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, *Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels*, Phys. Rev. Lett. **70**, 1895 (1993).

[104] D. Gottesman and I. L. Chuang, *Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations*, Nature **402**, 390 (1999).

[105] E. Knill, R. Laflamme, and G. J. Milburn, *A scheme for efficient quantum computation with linear optics*, Nature **409**, 46 (2001).

[106] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, *Concentrating partial entanglement by local operations*, Phys. Rev. A **53**, 2046 (1996).

[107] N. F. Ramsey, *A new molecular beam resonance method*, Phys. Rev. **76**, 996 (1949).

[108] N. F. Ramsey, *A molecular beam resonance method with separated oscillating fields*, Phys. Rev. **78**, 695 (1950).

[109] M. A. Lombardi, T. P. Heavner, and S. R. Jefferts, *NIST primary frequency standards and the realization of the SI second*, Measure **2**, 74 (2007).

[110] R. P. Feynman, F. L. Vernon, and R. W. Hellwarth, *Geometrical representation of the Schrödinger equation for solving maser problems*, J. Appl. Phys. **28**, 49 (1957).

[111] E. Paladino, Y. M. Galperin, G. Falci, and B. L. Altshuler, $1/f$ *noise: Implications for solid-state quantum information*, Rev. Mod. Phys. **86**, 361 (2014).

[112] G. Lindblad, *On the generators of quantum dynamical semigroups*, Commun. Math. Phys. **48**, 119 (1976).

[113] M. R. Geller and Z. Zhou, *Efficient error models for fault-tolerant architectures and the Pauli twirling approximation*, Phys. Rev. A **88**, 012314 (2013).

[114] A. Gilchrist, N. K. Langford, and M. A. Nielsen, *Distance measures to compare real and ideal quantum processes*, Phys. Rev. A **71**, 062310 (2005).

[115] R. Jozsa, *Fidelity for mixed quantum states*, J. Mod. Opt. **41**, 2315 (1994).

[116] X. L. Zhang, A. T. Gill, L. Isenhower, T. G. Walker, and M. Saffman, *Fidelity of a Rydberg blockade quantum gate from simulated quantum process tomography*, Phys. Rev. A **85**, 042310 (2012).

[117] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, *Measurement of qubits*, Phys. Rev. A **64**, 052312 (2001).

[118] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland, *Randomized benchmarking of quantum gates*, Phys. Rev. A **77**, 012307 (2008).

[119] J. Emerson, R. Alicki, and K. Życzkowski, *Scalable noise estimation with random unitary operators*, J. Opt. B: Quantum Semiclass. Opt. **7**, S347 (2005).

[120] R. Blume-Kohout, J. K. Gamble, E. Nielsen, J. Mizrahi, J. D. Sterk, and P. Maunz, *Robust, self-consistent, closed-form tomography of quantum logic gates on a trapped ion qubit*, arXiv:1310.4492 (2013).

[121] R. Blume-Kohout, J. K. Gamble, E. Nielsen, K. Rudinger, J. Mizrahi, K. Fortier, and P. Maunz, *Demonstration of qubit operations below a rigorous fault tolerance threshold with gate set tomography*, Nat. Commun. **8**, 14485 (2017).

[122] J. von Neumann, *Probabilistic logic and the synthesis of reliable organisms from unreliable components*, in Automata Studies, C. E. Shannon and J. McCarthy, eds., Princeton University Press, Princeton, NJ, (1956) pp. 43–98.

[123] P. W. Shor, *Scheme for reducing decoherence in quantum computer memory*, Phys. Rev. A **52**, R2493–R2496 (1995).

[124] R. Hamming, *Error detecting and error correcting codes*, Bell Sys. Tech. J. **29**, 147 (1950).

[125] A. M. Steane, *Error correcting codes in quantum theory*, Phys. Rev. Lett. **77**, 793 (1996).

[126] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, *Surface codes: Towards practical large-scale quantum computation*, Phys. Rev. A **86**, 032324 (2012).

[127] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, *Mixed-state entanglement and quantum error correction*, Phys. Rev. A **54**, 3824 (1996).

[128] P. W. Shor, *Fault-tolerant quantum computation*, in proceedings 37th annual symposium on foundaitons of computer science, FOCS 96 (IEEE Computer Society, Washington, DC, 1996) p. 56 (1996).

[129] D. Gottesman, *Class of quantum error-correcting codes saturating the quantum Hamming bound*, Phys. Rev. A **54**, 1862–1868 (1996).

[130] D. Gottesman, *Stabilizer codes and quantum error correction*, arXiv:quant-ph/9705052 (1997).

[131] E. Knill and R. Laflamme, *Theory of quantum error-correcting codes*, Phys. Rev. A **55**, 900–911 (1997).

[132] D. Gottesman, *Theory of fault-tolerant quantum computation*, Phys. Rev. A **57**, 127 (1998).

[133] J. Preskill, *Fault-tolerant quantum computation*, in H. Lo, S. Popescu, and T. Spiller (Eds.), *Introduction to Quantum Computation*, pp. 213–269 (World Scientific, Singapore) (1998).

[134] J. Preskill, *Reliable quantum computers*, Proc. R. Soc. Lond. A **454**, 385 (1998).

[135] A. M. Steane, *Introduction to quantum error correction*, Phil. Trans. R. Soc. Lond. A **256**, 1739 (1998).

[136] C. M. Caves, *Quantum error correcetion and reversible operations*, J. Superconductivity **12**, 707 (1999).

[137] S. J. Devitt, W. J. Munro, and K. Nemoto, *Quantum error correction for beginners*, Rep. Prog. Phys. **76**, 076001 (2013).

[138] B. M. Terhal, *Quantum error correction for quantum memories*, Rev. Mod. Phys. **87**, 307–346 (2015).

[139] E. T. Campbell, B. M. Terhal, and C. Vuillot, *Roads towards fault-tolerant universal quantum computation*, Nature **549**, 172 (2017).

[140] R. Laflamme, C. Miquel, J. P. Paz, and W. H. Zurek, *Perfect quantum error correcting code*, Phys. Rev. Lett. **77**, 198 (1996).

[141] P. Schindler, J. T. Barreiro, T. Monz, V. Nebendahl, D. Nigg, M. Chwalla, M. Hennrich, and R. Blatt, *Experimental repetitive quantum error correction*, Science **332**, 1059 (2011).

[142] J. Chiaverini, D. Leibfried, T. Schaetz, M. D. Barrett, R. B. Blakestad, J. Britton, W. M. Itano, J. D. Jost, E. Knill, C. Langer, R. Ozeri, and D. J. Wineland, *Realization of quantum error correction*, Nature **432**, 602–605 (2004).

[143] D. Ristè, S. Poletto, M.-Z. Huang, A. Bruno, V. Vesterinen, O.-P. Saira, and L. DiCarlo, *Detecting bit-flip errors in a logical qubit using stabilizer measurements*, Nat. Commun. **6**, 6983 (2015).

[144] V. Nebendahl, H. Häffner, and C. F. Roos, *Optimal control of entangling operations for trapped-ion quantum computing*, Phys. Rev. A **79**, 012312 (2009).

[145] H. Bombin and M. A. Martin-Delgado, *Topological quantum distillation*, Phys. Rev. Lett. **97**, 180501 (2006).

[146] H. G. Katzgraber, H. Bombin, and M. A. Martin-Delgado, *Error threshold for color codes and random three-body Ising models*, Phys. Rev. Lett. **103**, 090501 (2009).

[147] M. Ohzeki, *Accuracy thresholds of topological color codes on the hexagonal and square-octagonal lattices*, Phys. Rev. E **80**, 011141 (2009).

[148] A. J. Landahl, J. T. Anderson, and P. R. Rice, *Fault-tolerant quantum computing with color codes*, arXiv:1108.5738 (2011).

[149] D. Nigg, M. Müller, E. A. Martinez, P. Schindler, M. Hennrich, T. Monz, M. A. Martin-Delgado, and R. Blatt, *Quantum computations on a topologically encoded qubit*, Science **345**, 302 (2014).

[150] A. W. Cross, D. P. DiVincenzo, and B. M. Terhal, *A comparative code study for quantum fault tolerance*, Qu. Inf. Comp. **9**, 0541 (2009).

[151] P. Brooks and J. Preskill, *Fault-tolerant quantum computation with asymmetric Bacon-Shor codes*, Phys. Rev. A **87**, 032310 (2013).

[152] Y. Tomita and K. M. Svore, *Low-distance surface codes under realistic quantum noise*, Phys. Rev. A **90**, 062320 (2014).

[153] A. Paetznick and B. W. Reichardt, *Fault-tolerant ancilla preparation and noise threshold lower bounds for the 23-qubit Golay code*, Qu. Inf. Comp. **12**, 1034 (2012).

[154] S. Bravyi and A. Y. Kitaev, *Quantum codes on a lattice with boundary*, arXiv:quant-ph/9811052 (1998).

[155] S. Bravyi and A. Kitaev, *Universal quantum computation with ideal Clifford gates and noisy ancillas*, Phys. Rev. A **71**, 022316 (2005).

[156] T. Jochym-O'Connor and R. Laflamme, *Using concatenated quantum codes for universal fault-tolerant quantum gates*, Phys. Rev. Lett. **112**, 010505 (2014).

[157] A. Paetznick and B. W. Reichardt, *Universal fault-tolerant quantum computation with only transversal gates and error correction*, Phys. Rev. Lett. **111**, 090505 (2013).

[158] E. Nikahd, M. Sedighi, and M. Saheb Zamani, *Nonuniform code concatenation for universal fault-tolerant quantum computing*, Phys. Rev. A **96**, 032337 (2017).

[159] T. J. Yoder, R. Takagi, and I. L. Chuang, *Universal fault-tolerant gates on concatenated stabilizer codes*, Phys. Rev. X **6**, 031039 (2016).

[160] C. Chamberland, T. Jochym-O'Connor, and R. Laflamme, *Overhead analysis of universal concatenated quantum codes*, Phys. Rev. A **95**, 022313 (2017).

[161] D. Gottesman, A. Kitaev, and J. Preskill, *Encoding a qubit in an oscillator*, Phys. Rev. A **64**, 012310 (2001).

[162] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, *Theory of superconductivity*, Phys. Rev. **108**, 1175 (1957).

[163] L. N. Cooper, *Bound electron pairs in a degenerate Fermi gas*, Phys. Rev. **104**, 1189 (1956).