QUANTUM COMPUTING: AN OVERVIEW

MIKIO NAKAHARA
Department of Physics and Research Center for Quantum Computing,
Kinki University, Higashi-Osaka 577-8502, Japan
E-mail: nakahara@math.kindai.ac.jp

Elements of quantum computing and quantum infromation processing are introduced for mathematics students. Subjects includle quantum physics, qubits, quantum gates, quantum algorithms, decoherece, quantum error correcting codes and physical realizations. My lectures should serve as introduction to other lectures.

I. INTRODUCTION

Quantum computing and quantum information processing are emerging disciplines in which the principles of quantum physics are employed to store and process information. We use the classical digital technology at almost every moment in our lives: computers, mobile phones, mp3 players, just to name a few. Even though quantum mechanics is used in the design of devices such as LSI, the logic is purely classical. This means that an AND circuit, for example, produces *definitely* 1 when the inputs are 1 and 1. One of the most remarkable aspects of the principles of quantum physics is the *superposition principle* by which a quantum system can take several different states *simultaneously*. The input for a quantum computing device may be a superposition of many possible inputs, and accordingly the output is also a superposition of the corresponding output states. Another aspect of quantum physics, which is far beyond the classical description, is *entanglement*. Given several objects in a classical world, they can be described by specifying each object separately. In a quantum world, however, only a very tiny fraction of all possible states can be described by such separate specifications. In other words, most quantum states cannot be described by such individual specifications, thereby being called "entangled". Why and how these two features give rise to the enormous computational power in quantum computing ane quantum information processing will be explained in this contribution.

A part of this lecture note is based on our book [1]. General references are [2–4].

II. QUANTUM PHYSICS

A. Notation and conventions

We will exclusively work with a finite-dimensional complex vector space $\mathbb{C}^n$ with an inner product $\langle\quad,\quad\rangle$ (Hilbert spaces). A vector in $\mathbb{C}^n$ is called a ket vector or a ket and is denoted as

$$|x\rangle = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad x_i \in \mathbb{C}$$

while a vector in the dual space $\mathbb{C}^{n*}$ is called a bra vector or a bra and denoted $\langle\alpha| = (\alpha_1, \dots, \alpha_n),\ \alpha_i \in \mathbb{C}$. Index i sometimes runs from 0 to $n-1$. The inner product of $|x\rangle$ and $\langle\alpha|$ is

$$\langle\alpha|x\rangle = \sum_{i=1}^{n} \alpha_i x_i.$$

This inner product naturally introduces a correspondence $|x\rangle = (x_1, \dots, x_n)^t \leftrightarrow \langle x| = (x_1^*, \dots, x_n^*)$, by which an inner product of two vectors are defined as $\langle x|y\rangle = \sum_{i=1}^{n} x_i^* y_i$. The inner product naturally defines the norm of a vector $|x\rangle$ as $\||x\rangle\| = \sqrt{\langle x|x\rangle}$.

Pauli matrices are generators of $\mathfrak{su}(2)$ and denoted

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

in the basis in which σ_z is diagonalized. Symbols $X = \sigma_x, Y = -i\sigma_y$ and $Z = \sigma_z$ are also employed.

Let A be an $m \times n$ matrix and B be a $p \times q$ matrix. Then

$$A \otimes B = \begin{pmatrix} a_{11}B, a_{12}B, \ldots, a_{1n}B \\ a_{21}B, a_{22}B, \ldots, a_{2n}B \\ \cdots \\ a_{m1}B, a_{m2}B, \ldots, a_{mn}B \end{pmatrix}$$

is an $(mp) \times (nq)$ matrix called the tensor product of A and B. As a special case, the tensor product of two vectors $|x\rangle = (x_1, x_2, \ldots, x_p)^t$ and $|y\rangle = (y_1, y_2, \ldots, y_q)^t$ is given by

$$|x\rangle \otimes |y\rangle = (x_1y_1, \ldots, x_1y_q, x_2y_1, \ldots, x_2y_q, \ldots, x_py_1, \ldots x_py_q)^t .$$

The tensor product $|x\rangle \otimes |y\rangle$ is often abbreviated as $|x\rangle|y\rangle$ or $|xy\rangle$. Note however that the tensor product of two matrices A and B cannot be written as AB for an obvious reason.

B. Axioms of quantum mechanics

Quantum mechanics was discovered roughly a century ago [5–10]. In spite of its long history, the interpretation of the wave function remains an open question. Here we adopt the most popular one, called the Copenhagen interpretation.

A 1 A pure state in quantum mechanics is represented by a normalized vector $|\psi\rangle$ in a Hilbert space $\mathcal{H}$ associated with the system. If two states $|\psi_1\rangle$ and $|\psi_2\rangle$ are physical states of the system, their linear superposition $c_1|\psi_1\rangle + c_2|\psi_2\rangle$ ($c_k \in \mathbb{C}$), with $\sum_{i=1}^{2} |c_i|^2 = 1$, is also a possible state of the same system (superposition principle).

A 2 For any physical quantity (observable) a, there exists a corresponding Hermitian operator A acting on $\mathcal{H}$. When a measurement of a is made, the outcome is one of the eigenvalues λ_j of A. Let λ_1 and λ_2 be two eigenvalues of A: $A|\lambda_i\rangle = \lambda_i|\lambda_i\rangle$. Consider a superposition state $c_1|\lambda_1\rangle + c_2|\lambda_2\rangle$. If we measure a in this state, the state undergoes an abrupt change (wave function collapse) to one of the eigenstates $|\lambda_i\rangle$ corresponding to the observed eigenvalue λ_i. Suppose we prepare many copies of the state $c_1|\lambda_1\rangle + c_2|\lambda_2\rangle$. The probability of collapsing to the state $|\lambda_i\rangle$ is given by $|c_i|^2$ $(i = 1, 2)$. The complex coefficient c_i is called the probability amplitude in this sense. It should be noted that a measurement produces one outcome λ_i and the probability of obtaining it is experimentally evaluated only after repeating measurements with many copies of the same state. These statements are easily generalized to superposition states of more than two states.

A 3 The time dependence of a state is governed by the Schrödinger equation

$$i\hbar \frac{\partial |\psi\rangle}{\partial t} = H|\psi\rangle, \tag{1}$$

where $\hbar$ is a physical constant known as the Planck constant and H is a Hermitian operator (matrix) corresponding to the energy of the system and is called the Hamiltonian.

Several comments are in order.

- In Axiom A 1, the phase of the vector may be chosen arbitrarily; $|\psi\rangle$ in fact represents the "ray" $\{e^{i\alpha}|\psi\rangle \; | \alpha \in \mathbb{R}\}$. This is called the ray representation. The overall phase is not observable and has no physical meaning.

- Axiom A 2 may be formulated in a different but equivalent way as follows. Suppose we would like to measure an observable a. Let the spectral decomposition of the corresponding operator A be $A = \sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i|$, where $A|\lambda_i\rangle = \lambda_i|\lambda_i\rangle$. Then the expectation value $\langle A\rangle$ of a after measurements with respect to many copies of $|\psi\rangle$ is

$$\langle A\rangle = \langle\psi|A|\psi\rangle. \tag{2}$$

Let us expand $|\psi\rangle$ in terms of $|\lambda_i\rangle$ as $|\psi\rangle = \sum_i c_i|\lambda_i\rangle$. According to A 2, the probability of observing λ_i upon measurement of a is $|c_i|^2$ and therefore the expectation value after many measurements is $\sum_i \lambda_i |c_i|^2$. If, conversely, Eq. (2) is employed, we will obtain the same result since $\langle\psi|A|\psi\rangle = \sum_{i,j} c_j^* c_i \langle\lambda_j|A|\lambda_i\rangle = \sum_{i,j} \lambda_i c_j^* c_i \delta_{ij} = \sum_i \lambda_i |c_i|^2$. This measurement is called the projective measurement. Any particular outcome λ_i will be found with the probability $|c_i|^2 = \langle\psi|P_i|\psi\rangle$, where $P_i = |\lambda_i\rangle\langle\lambda_i|$ is the projection operator and the state immediately after the measurement is $|\lambda_i\rangle$ or equivalently $P_i|\psi\rangle/\sqrt{\langle\psi|P_i|\psi\rangle}$.

- The Schrödinger equation (1) in Axiom A 3 is formally solved to yield

$$|\psi(t)\rangle = e^{-iHt/\hbar}|\psi(0)\rangle, \tag{3}$$

if the Hamiltonian H is time-independent, while

$$|\psi(t)\rangle = \mathcal{T}\exp\left[-\frac{i}{\hbar}\int_0^t H(t)dt\right]|\psi(0)\rangle \tag{4}$$

if H depends on t, where $\mathcal{T}$ is the time-ordering operator. The state at $t > 0$ is $|\psi(t)\rangle = U(t)|\psi(0)\rangle$. The operator $U(t) : |\psi(0)\rangle \mapsto |\psi(t)\rangle$, called the time-evolution operator, is unitary. Unitarity of $U(t)$ guarantees that the norm of $|\psi(t)\rangle$ is conserved: $\langle\psi(0)|U^\dagger(t)U(t)|\psi(0)\rangle = \langle\psi(0)|\psi(0)\rangle = 1 \quad (\forall t > 0)$.

Two mutually commuting operators A and B have simultaneous eigenstates. If, in contrast, they do not commute, the measurement outcomes of these operators on any state $|\psi\rangle$ satisfy the following uncertainty relations. Let $\langle A\rangle = \langle\psi|A|\psi\rangle$ and $\langle B\rangle = \langle\psi|B|\psi\rangle$ be their respective expectation values and $\Delta A = \sqrt{\langle(A - \langle A\rangle)^2\rangle}$ and $\Delta B = \sqrt{\langle(B - \langle B\rangle)^2\rangle}$ be respective standard deviations. Then they satisfy

$$\Delta A \Delta B \geq \frac{1}{2}|\langle\psi|[A, B]|\psi\rangle|. \tag{5}$$

C. Simple example

Examples to clarify the axioms introduced in the previous subsection are given. They are used to controll quantum states in physical realizations of a quantum computer. A spin-1/2 particle has two states, which we call spin-up state $|\uparrow\rangle$ and spin-down state $|\downarrow\rangle$. It is common to assign components $|\uparrow\rangle = (1, 0)^t$ and $|\downarrow\rangle = (0, 1)^t$. They form a basis of a vector space $\mathbb{C}^2$.

Let us consider a time-independent Hamiltonian

$$H = -\frac{\hbar}{2}\omega\sigma_x \tag{6}$$

acting on the spin Hilbert space $\mathbb{C}^2$. Suppose the system is in the eigenstate of σ_z with the eigenvalue $+1$ at time $t = 0$; $|\psi(0)\rangle = |\uparrow\rangle$. The wave function $|\psi(t)\rangle$ $(t > 0)$ is then found from Eq. (3) as

$$|\psi(t)\rangle = \exp\left(i\frac{\omega}{2}\sigma_x t\right)|\psi(0)\rangle = \begin{pmatrix} \cos\omega t/2 & i\sin\omega t/2 \\ i\sin\omega t/2 & \cos\omega t/2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\omega t/2 \\ i\sin\omega t/2 \end{pmatrix} = \cos\frac{\omega}{2}t|\uparrow\rangle + i\sin\frac{\omega}{2}t|\downarrow\rangle. \tag{7}$$

Suppose we measure σ_z in $|\psi(t)\rangle$. The spin is found spin-up with probability $P_\uparrow(t) = \cos^2(\omega t/2)$ and spin-down with probability $P_\downarrow(t) = \sin^2(\omega t/2)$.

Consider a more general Hamiltonian

$$H = -\frac{\hbar}{2}\omega\hat{\boldsymbol{n}}\cdot\boldsymbol{\sigma}, \tag{8}$$

where $\hat{\boldsymbol{n}}$ is a unit vector in $\mathbb{R}^3$. The time-evolution operator is readily obtained, by making use of a well known formula

$$e^{i\alpha(\hat{\boldsymbol{n}}\cdot\boldsymbol{\sigma})} = \cos\alpha I + i(\hat{\boldsymbol{n}}\cdot\boldsymbol{\sigma})\sin\alpha \tag{9}$$

as

$$U(t) = \exp(-iHt/\hbar) = \cos\omega t/2\ I + i(\hat{\boldsymbol{n}}\cdot\boldsymbol{\sigma})\sin\omega t/2. \tag{10}$$

Suppose the initial state is $|\psi(0)\rangle = (1, 0)^t$ for example. Then we find, at a later time $t > 0$,

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle = \begin{pmatrix} \cos(\omega t/2) + in_z\sin(\omega t/2) \\ i(n_x + in_y)\sin(\omega t/2) \end{pmatrix}. \tag{11}$$

D. Multipartite system, tensor product and entangled state

So far, we have implictly assumed that the system is made of a single component. Suppose a system is made of two components, one lives in a Hilbert space $\mathcal{H}_1$ and the other in $\mathcal{H}_2$. A system composed of two separate components is called bipartite. The system as a whole lives in a Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, whose general vector is written as

$$|\psi\rangle = \sum_{i,j} c_{ij} |e_{1,i}\rangle \otimes |e_{2,j}\rangle, \tag{12}$$

where $\{|e_{a,i}\rangle\}$ $(a = 1, 2)$ is an orthonormal basis in $\mathcal{H}_a$ and $\sum_{i,j} |c_{ij}|^2 = 1$.

A state $|\psi\rangle \in \mathcal{H}$ written as a tensor product of two vectors as $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle$, $(|\psi_a\rangle \in \mathcal{H}_a)$ is called a separable state or a tensor product state. A separable state admits a classical interpretation "The first system is in the state $|\psi_1\rangle$ while the second system is in $|\psi_2\rangle$". It is clear that the set of separable state has dimension $\dim \mathcal{H}_1 + \dim \mathcal{H}_2$. Note, however, that the total space $\mathcal{H}$ has different dimension than this: $\dim \mathcal{H} = \dim \mathcal{H}_1 \dim \mathcal{H}_2$. This number is considerably larger than the dimension of the sparable states when $\dim \mathcal{H}_a$ $(a = 1, 2)$ are large. What are the missing states then? Let us consider a spin state

$$|\psi\rangle = \frac{1}{\sqrt{2}} \left(|\uparrow\rangle \otimes |\uparrow\rangle + |\downarrow\rangle \otimes |\downarrow\rangle \right) \tag{13}$$

of two electrons. Suppose $|\psi\rangle$ may be decomposed as

$$|\psi\rangle \;=\; (c_1|\uparrow\rangle + c_2|\downarrow\rangle) \otimes (d_1|\uparrow\rangle + d_2|\downarrow\rangle) = c_1 d_1 |\uparrow\rangle \otimes |\uparrow\rangle + c_1 d_2 |\uparrow\rangle \otimes |\downarrow\rangle + c_2 d_1 |\downarrow\rangle \otimes |\uparrow\rangle + c_2 d_2 |\downarrow\rangle \otimes |\downarrow\rangle.$$

However this decomposition is not possible since we must have $c_1 d_2 = c_2 d_1 = 0$, $c_1 d_1 = c_2 d_2 = 1/\sqrt{2}$ simultaneously and it is clear that the above equations have no common solution, showing $|\psi\rangle$ is not separable.

Such non-separable states are called entangled. Entangled states refuse classical descriptions. Entanglement is used extensively as a powerful computational resource in the following.

Suppose a bipartite state (12) is given. We are interested in when the state is separable and when entangled. The criterion is given by the Schmidt decomposition of $|\psi\rangle$.

Theorem II.1 *Let $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ be the Hilbert space of a bipartite system. Then a vector $|\psi\rangle \in \mathcal{H}$ admits the Schmidt decomposition*

$$|\psi\rangle = \sum_{i=1}^{r} \sqrt{s_i} |f_{1,i}\rangle \otimes |f_{2,i}\rangle, \tag{14}$$

where $s_i > 0$ are called the Schmidt coefficients satisfying $\sum_i s_i = 1$ and $\{|f_{a,i}\rangle\}$ is an orthonormal set of $\mathcal{H}_a$. The number $r \in \mathbb{N}$ is called the Schmidt number of $|\psi\rangle$.

It follows from the above theorem that a bipartite state $|\psi\rangle$ is separable if and only if its Schmidt number r is 1. See [1] for the proof.

E. Mixed states and density matrices

It might happen in some cases that a quantum system under considertation is in the state $|\psi_i\rangle$ with a probability p_i. In other words, we cannot say definitely which state the system is in. Therefore some random nature comes into the description of the system. Such a system is said to be in a mixed state while a system whose vector is uniquely specified is in a pure state. A pure state is a special case of a mixed state in which $p_i = 1$ for some i and $p_j = 0$ $(j \neq i)$.

A particular state $|\psi_i\rangle \in \mathcal{H}$ appears with probability p_i in an ensemble of a mixed state, in which case the expectation value of an observable a is $\langle\psi_i|A|\psi_i\rangle$. The mean value of a averaged over the ensemble is then given by

$$\langle A \rangle = \sum_{i=1}^{N} p_i \langle\psi_i|A|\psi_i\rangle, \tag{15}$$

where N is the number of available states. Let us introduce the density matrix by

$$\rho = \sum_{i=1}^{N} p_i |\psi_i\rangle\langle\psi_i|. \tag{16}$$

Then Eq. (15) is rewritten in a compact form as $\langle A\rangle = \mathrm{Tr}(\rho A)$.

Let A be a Hermitian matrix. A is called positive-semidefinite if $\langle\psi|A|\psi\rangle \geq 0$ for any $|\psi\rangle \in \mathcal{H}$. It is easy to show all the eigenvalues of a positive-semidefinite Hermitian matrix are non-negative. Conversely, a Hermitian matrix A whose every eigenvalue is non-negative is positive-semidefinite.

Properties which a density matrix ρ satisfies are very much like axioms for pure states.

A 1' A physical state of a system, whose Hilbert space is $\mathcal{H}$, is completely specified by its associated density matrix $\rho : \mathcal{H} \to \mathcal{H}$. A density matrix is a positive-semidefinite Hermitian operator with $\mathrm{tr}\,\rho = 1$, see remarks below.

A 2' The mean value of an observable a is given by

$$\langle A\rangle = \mathrm{tr}\,(\rho A). \tag{17}$$

A 3' The temporal evolution of the density matrix follows the Liouville-von Neumann equation

$$i\hbar\frac{d}{dt}\rho = [H, \rho] \tag{18}$$

where H is the system Hamiltonian, see remarks below.

Several remarks are in order.

- The density matrix (16) is Hermitian since $p_i \in \mathbb{R}$. It is positive-semidefinite since $\langle\psi|\rho|\psi\rangle = \sum_i p_i|\langle\psi_i|\psi\rangle|^2 \geq 0$.
- Each $|\psi_i\rangle$ follows the Schrödinger equation $i\hbar\frac{d}{dt}|\psi_i\rangle = H|\psi_i\rangle$ in a closed quantum system. Its Hermitian conjugate is $-i\hbar\frac{d}{dt}\langle\psi_i| = \langle\psi_i|H$. We prove the Liouville-von Neumann equation from these equations as

$$i\hbar\frac{d}{dt}\rho = i\hbar\frac{d}{dt}\sum_i p_i|\psi_i\rangle\langle\psi_i| = \sum_i p_i H|\psi_i\rangle\langle\psi_i| - \sum_i p_i|\psi_i\rangle\langle\psi_i|H = [H, \rho].$$

We denote the set of all possible density matrices as $\mathcal{S}(\mathcal{H})$.

Example II.2 *A pure state $|\psi\rangle$ is a special case in which the corresponding density matrix is $\rho = |\psi\rangle\langle\psi|$. Therefore ρ is nothing but the projection operator onto the state. Observe that $\langle A\rangle = tr\rho A = \sum_i\langle e_i|\psi\rangle\langle\psi|A|e_i\rangle = \langle\psi|A\sum_i|e_i\rangle\langle e_i|\psi\rangle = \langle\psi|A|\psi\rangle$, where $\{|e_i\rangle\}$ is an orthonormal set.*

Let us consider a beam of photons. We take a horizontally polarized state $|e_1\rangle = |\leftrightarrow\rangle$ and a vertically polarized state $|e_2\rangle = |\updownarrow\rangle$ as orthonormal basis vectors. If the photons are a totally uniform mixture of two polarized states, the density matrix is given by

$$\rho = \frac{1}{2}|e_1\rangle\langle e_1| + \frac{1}{2}|e_2\rangle\langle e_2| = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{1}{2}I.$$

This state is called a maximally mixed state.

If photons are in a pure state $|\psi\rangle = (|e_1\rangle + |e_2\rangle)/\sqrt{2}$, the density matrix, with $\{|e_i\rangle\}$ as basis, is

$$\rho = |\psi\rangle\langle\psi| = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

We are interested in when ρ represents a pure state or a mixed state.

Theorem II.3 *A state ρ is pure if and only if $tr\rho^2 = 1$.*

Proof: Since ρ is Hermitian, all its eigenvalues λ_i ($1 \leq i \leq \dim\mathcal{H}$) are real and the corresponding eigenvectors $\{|\lambda_i\rangle\}$ are made orthonormal. Then $\rho^2 = \sum_{i,j}\lambda_i\lambda_j|\lambda_i\rangle\langle\lambda_i|\lambda_j\rangle\langle\lambda_j| = \sum_i\lambda_i^2|\lambda_i\rangle\langle\lambda_i|$. Therefore $\mathrm{tr}\,\rho^2 = \sum_i\lambda_i^2 \leq \lambda_{\max}\sum_i\lambda_i = \lambda_{\max} \leq 1$, where $\lambda_{\max}$ is the largest eigenvalue of ρ. Therefore $\mathrm{tr}\,\rho^2 = 1$ implies $\lambda_{\max} = 1$ and all the other eigenvalues are zero. The converse is trivial. ∎

We classify mixed states into two classes, similarly to the classification of pure states into separable states and entangled states. We use a bipartite system in the definition but generalization to multipartitle systems should be obvious.

Definition II.4 *A state ρ is called separable if it is written in the form*

$$\rho = \sum_i p_i\rho_{1,i}\otimes\rho_{2,i}, \tag{19}$$

where $0 \leq p_i \leq 1$ and $\sum_i p_i = 1$. It is called inseparable, if ρ does not admit the decompostion (19).

In the next subsection, we discuss how to find whether a given bipartite density matrix is separable or inseparable.

F. Negativity

Let ρ be a bipartite state and define the partial transpose $\rho^{\rm pt}$ of ρ with respect to the second Hilbert space as

$$\rho_{ij,kl} \to \rho_{il,kj}, \tag{20}$$

where $\rho_{ij,kl} = (\langle e_{1,i}| \otimes \langle e_{2,j}|)\ \rho\ (|e_{1,k}\rangle \otimes |e_{2,l}\rangle)$. Here $\{|e_{1,k}\rangle\}$ is the orthonormal basis of the first system while $\{|e_{2,k}\rangle\}$ of the second system. Suppose ρ takes a separable form (19). Then the partial transpose yields

$$\rho^{\rm pt} = \sum_i p_i \rho_{1,i} \otimes \rho_{2,i}^t. \tag{21}$$

Note here that ρ^t for any density matrix ρ is again a density matrix since it is still positive semi-definite Hermitian with unit trace. Therefore the partial transposed density matrix (21) is another density matrix. It was conjectured by Peres [12] and subsequently proved by the Horodecki family [13] that positivity of the partially transposed density matrix is necessary and sufficient condition for ρ to be separable in the cases of $\mathbb{C}^2 \otimes \mathbb{C}^2$ systems and $\mathbb{C}^2 \otimes \mathbb{C}^3$ systems. Conversely, if the partial transpose of ρ of these systems is not a density matrix, then ρ is inseparable. Instead of giving the proof, we look at the following example.

Example II.5 *Let us consider the Werner state*

$$\rho = \begin{pmatrix} \frac{1-p}{4} & 0 & 0 & 0 \\ 0 & \frac{1+p}{4} & -\frac{p}{2} & 0 \\ 0 & -\frac{p}{2} & \frac{1+p}{4} & 0 \\ 0 & 0 & 0 & \frac{1-p}{4} \end{pmatrix}, \tag{22}$$

where $0 \le p \le 1$. *Here the basis vectors are arranged in the order*

$$|e_{1,1}\rangle|e_{2,1}\rangle, |e_{1,1}\rangle|e_{2,2}\rangle, |e_{1,2}\rangle|e_{2,1}\rangle, |e_{1,2}\rangle|e_{2,2}\rangle.$$

Partial transpose of ρ *yields*

$$\rho^{\rm pt} = \begin{pmatrix} \frac{1-p}{4} & 0 & 0 & -\frac{p}{2} \\ 0 & \frac{1+p}{4} & 0 & 0 \\ 0 & 0 & \frac{1+p}{4} & 0 \\ -\frac{p}{2} & 0 & 0 & \frac{1-p}{4} \end{pmatrix}.$$

$\rho^{\rm pt}$ *must have non-negative eigenvalues to be a physically acceptable state. The characteristic equation of* $\rho^{\rm pt}$ *is*

$$D(\lambda) = \det(\rho^{\rm pt} - \lambda I) = \left(\lambda - \frac{p+1}{4}\right)^3 \left(\lambda - \frac{1-3p}{4}\right) = 0.$$

There are threefold degenerate eigenvalue $\lambda = (1+p)/4$ *and nondegenerate eigenvalue* $\lambda = (1-3p)/4$. *This shows that* $\rho^{\rm pt}$ *is an unphysical state for* $1/3 < p \le 1$. *If this is the case,* ρ *is inseparable.*

From the above observation, entangled states are characterized by nonvanishing negativity defined as

$$N(\rho) \equiv \frac{1}{2}(\sum_i |\lambda_i| - 1). \tag{23}$$

Note that negativity vanishes if and only if all the eigenvalues of $\rho^{\rm pt}$ are nonnegative.

Exercise II.6 *(1) Show that*

$$\rho = \begin{pmatrix} \frac{p}{2} & 0 & 0 & \frac{p}{2} \\ 0 & \frac{1-p}{2} & \frac{1-p}{2} & 0 \\ 0 & \frac{1-p}{2} & \frac{1-p}{2} & 0 \\ \frac{p}{2} & 0 & 0 & \frac{p}{2} \end{pmatrix} \qquad (0 \le p \le 1) \tag{24}$$

is a density matrix. Show also that the negativity of ρ *vanishes only for* $p = 1/2$.
(2) Show that

$$\rho_1 = \begin{pmatrix} \frac{1+p}{4} & 0 & 0 & \frac{p}{2} \\ 0 & \frac{1-p}{4} & 0 & 0 \\ 0 & 0 & \frac{1-p}{4} & 0 \\ \frac{p}{2} & 0 & 0 & \frac{1+p}{4} \end{pmatrix} \qquad (0 \le p \le 1) \tag{25}$$

is a density matrix. Show also that the negativity does not vanish for $p > 1/3$.

G. Partial trace and purification

Let $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$ be a Hilbert space of a bipartite system made of components 1 and 2 and let A be an arbitrary operator acting on $\mathcal{H}$. The partial trace of A over $\mathcal{H}_2$ generates an operator acting on $\mathcal{H}_1$ defined as

$$A_1 = \mathrm{tr}_2 A \equiv \sum_k (I \otimes \langle k|) A (I \otimes |k\rangle). \tag{26}$$

We will be concerned with the partial trace of a density matrix in practical applications. Let $\rho = |\psi\rangle\langle\psi| \in \mathcal{S}(\mathcal{H})$ be a density matrix of a pure state $|\psi\rangle$. Suppose we are interested only in the first system and have no access to the second system. Then the partial trace allows us to "forget" about the second system.

To be concrete, consider a pure state $|\psi\rangle = \frac{1}{\sqrt{2}}(|e_1\rangle|e_1\rangle + |e_2\rangle|e_2\rangle)$, where $\{|e_i\rangle\}$ is an orthonormal basis of $\mathbb{C}^2$. The corresponding density matrix is

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

where the basis vectors are ordered as $\{|e_1\rangle|e_1\rangle, |e_1\rangle|e_2\rangle, |e_2\rangle|e_1\rangle, |e_2\rangle|e_2\rangle\}$. The partial trace of ρ is

$$\rho_1 = \mathrm{tr}_2 \rho = \sum_{i=1,2} (I \otimes \langle e_i|) \rho (I \otimes |e_i\rangle) = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{27}$$

Note that a pure state $|\psi\rangle$ is mapped to a maximally mixed state ρ_1.

We have seen above that the partial trace of a pure-state density matrix of a bipartite system over one of the constituent Hilbert spaces yields a mixed state. How about the converse? Given a mixed state density matrix, is it always possible to find a pure state density matrix whose partial trace over the extra Hilbert space yields the given density matrix? The answer is yes and the process to find the pure state is called the purification. Let $\rho_1 = \sum_k p_k |\psi_k\rangle\langle\psi_k|$ be a general density matrix of a system 1 with the Hilbert space $\mathcal{H}_1$. Now let us introduce the second Hilbert space $\mathcal{H}_2$ whose dimension is the same as that of $\mathcal{H}_1$. Then formally introduce a normalized vector

$$|\Psi\rangle = \sum_k \sqrt{p_k} |\psi_k\rangle \otimes |\phi_k\rangle, \tag{28}$$

where $\{|\phi_k\rangle\}$ is an orthonormal basis of $\mathcal{H}_2$. We find

$$\mathrm{tr}_2 |\Psi\rangle\langle\Psi| = \sum_{i,j,k} (I \otimes \langle\phi_i|) \left[\sqrt{p_j p_k} |\psi_j\rangle|\phi_j\rangle\langle\psi_k|\langle\phi_k|\right] (I \otimes |\phi_i\rangle) = \sum_k p_k |\psi_k\rangle\langle\psi_k| = \rho_1. \tag{29}$$

It is always possible to purify a mixed state by tensoring an extra Hilbert space of the same dimension as that of the original Hilbert space. Purification is far from unique.

Exercise II.7 *(1) Let*

$$\rho_1 = \frac{1}{4}\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$$

be a density matrix with a basis $\{|\psi_i\rangle\}$*. Find a purification of* ρ_1*.*
(2) Let

$$|\Psi\rangle = \sum_k \sqrt{p_k} |\psi_k\rangle \otimes |\phi_k\rangle$$

be a purification of $\rho_1 = \sum_k p_k |\psi_k\rangle\langle\psi_k| \in \mathcal{S}(\mathcal{H})$*. Show that*

$$|\Psi'\rangle = \sum_k \sqrt{p_k} |\psi_k\rangle \otimes U|\phi_k\rangle$$

is another purification of ρ_1*, where* U *is an arbitrary unitary matrix in* $\mathrm{U}(\dim \mathcal{H})$*.*

H. von Neumann Entropy

1. Shannon Entropy

Entropy is a measure of randomness of a probability distribution. It also quantifies information gained when measurement is made on the random variable. Let us start with classical entropy also know as the Shannon entropy. Let $p(x)$ be a probability distribution for some random variable X. Then the entropy of this distribution is defined as

$$S = -\sum_x p(x)\log_2 p(x).$$

We drop the base 2 in $\log_2$ and simply write log hereafter. There must be a minus sign to make S non-negative. Let us consider two special cases. (i) $p(x) = 1$ for $x = x_0$ and $p(x) = 0$ for $x \neq x_0$. Then

$$S = -1\log 1 = 0.$$

(ii) There are N possibilities for x and $p(x) = 1/N$ independently of x. The distribution is maximally uniform in this case. Then

$$S = -N\frac{1}{N}\log\frac{1}{N} = \log N.$$

This is the maximal possible value for S. In fact, let us maximize S with respect to $p(x)$. We introduce the Lagrange multiplier λ and optimize $-\sum_x p(x)\log p(x) - \lambda(\sum_x p(x) - 1)$ to obtain

$$\delta S = -(\log p(x) + 1 + \lambda)\delta p(x) = 0,$$

from which we find $p(x) = 2^{-1-\lambda}$ is independent of x. The normalization condition fixes λ to $\log N - 1$ so that $p(x) = 1/N$. The entropy S takes an intermediate value between 0 and $\log N$ for a general distribution function $p(x)$.

Example II.8 *(1) Let us consider a coin toss. The outcome is either H (head) or T (tail) with probability* $1/2$*. The entropy for this process is* $S = -2 \times (1/2)\log(1/2) = \log 2 = 1$*. It also implies that the number of bits required to store this information is one.*
(2) Let us consider throwing a die. The outcome is one of the numbers $1, 2, \ldots, 6$ *each with probability* $1/6$*. The entropy for this process is* $S = -6 \times (1/6)\log(1/6) = \log 6 = 2.58\ldots$*. he number of bits required to store this information is three.*

log function in the definition of entropy makes S additive. Let X, Y be two indepenedent random variables and let $p(x)$ and $q(y)$ be their respetive probability distributions. The measurement of X and Y produces outomes x and y with probability $p(x)q(y)$ by definition. Then the entropy of this process is

$$S = -\sum_{x,y} p(x)q(y)\log p(x)q(y) = -\sum_{x,y} p(x)q(y)[\log p(x) + \log q(y)] = -\sum_x p(x)\log p(x) - \sum_y q(y)\log q(y),$$

which is a sum of two entropies associated with X and Y.

Entropy is also regarded as the average number of bits to record the outcome. The following example is taken from [2]. Suppose some source produces one of four numbers $1, 2, 3$ and 4. If they appear with equal probability $1/4$, the entropy of this process is $S = -4 \times (1/4)\log 4 = 2$. Clearly we need two bits to record the outcome. Let us consider another case in which $p(1) = 1/2, p(2) = 1/4$ and $p(3) = p(4) = 1/8$. The entropy is $S = -(1/2)\log(1/2) - (1/4)\log(1/4) + 2 \times (1/8)\log(1/8) = 7/4$. This shows that there is a scheme under which the outcome can be stored with a number of bits less than 2. This can be realized if a small number of bits is assigned for a frequent outcome, 1 in our case. In fact, let the outcome 1 be stored as a bit string 0, 2 as 10, 3 as 110 and 4 as 111. Then the average number of bits required to store N such outcomes is $N/2 + 2N/4 + 2 \times 3N/8 = N(4/7)$.

In view of the additivity mentioned above, the statement of the previous paragraph claims that two pages of a newspaper contains twice as much information as a page of the same newspaper in average.

Exercise II.9 *Let* X *be a two-valued random variable. Let us call the outcomes 0 and 1, which appear with probabilities* p *and* $1 - p$*, respectively. Then the "binary entropy" of* X *is* $S(p) = -p\log p - (1 - p)\log(1 - p)$.
(1) Show that $S(p)$ *takes its maximum value 1 at* $p = 1/2$.
(2) Show that $S(p)$ *is a concave function, that is,*

$$S(px + (1 - p)y) \geq pS(x) + (1 - p)S(y) \quad (0 \leq x, y, p \leq 1).$$

Show also that the equality is true only when $x = y$ *or* $p = 0$ *or* $p = 1$.

Let a random variable X have two probability distributions $p(x)$ and $q(x)$. The relative entropy of $p(x)$ to $q(x)$ is defined as

$$H(p(x)\|q(x)) = \sum_x p(x)\log\frac{p(x)}{q(x)} = -S(p) - \sum_x p(x)\log q(x).$$

The relative entropy vanishes whe $p(x) = q(x)$ and is positive if $p(x) \neq q(x)$. In this sense, it measures the distance between two distributions $p(x)$ and $q(x)$ corresponding to the same random variable X.

Exercise II.10 *Let us prove the positivity mentioned above.*
(1) Show that $-\log x \geq (1-x)/\ln 2$ *for* $x > 0$*, where the quality is satisfied if and only if* $x = 1$*. (Hint: Prove* $\log x \ln 2 = lnx \leq x - 1$ *for* $x > 0$*.)*
(2) Use this fact to prove

$$H(p(x)\|q(x)) \geq 0$$

where the equality is satisfied if and only if $p(x) = q(x)$*. (Hint: Fix* $p(x)$ *and find the variation of* H *with respect to* $\delta q(x)$*. Do not forget to take the constraint* $\sum_x q(x) = 1$ *into account.)*

Let X be a random variable with n outcomes. Then the entropy $S(p)$ is maximized when $p(x)$ is a uniform distribution $q(x) = 1/d$ as was proved before. As an application of the positivity of the relative entropy, we give an another proof of this fact. We find

$$H(p(x)\|q(x) = 1/d) = -S(p) - \sum_x p(x)\log(1/d) = \log d - S(p) \geq 0,$$

which shows verifies $S(p) \leq \log d$ for any $p(x)$.

2. von Neumann Entropy

A natural generalization of the Shannon entropy to a quantum system is the von Neumann entropy. We drop the base 2 in log hereafter unless otherwise stated explicitly. Let us consider a single quantum system with the Hilbert space $\mathbb{C}^n$. Suppose the state is described by a density matrix ρ. The von Neumann entropy is defined as

$$S(\rho) = -\mathrm{tr}\,(\rho\log\rho).$$

Again two extremal cases deserve special study. (i) A pure state $\rho = |\psi\rangle\langle\psi|$. If $|\psi\rangle$ is taken to be one of the basis vectors of the Hilbert space $\mathbb{C}^n$, the state ρ take the form $\rho = \mathrm{diag}\,(1, 0, \ldots, 0)$ and S is evaluated as $S(\rho) = -\mathrm{tr}\,(\rho\log\rho) = 0$. (ii) The maximally mixed state is expressed as $\rho = I_n/n$, where I_n is the unit matrix of dimension n and $S(\rho)$ is evaluated as $S(\rho) = \log n$. This is the maximal possible value S may take. In fact, let us extermize $\tilde{S}(\rho) = -\mathrm{tr}\,(\rho\log\rho) - \lambda(\mathrm{tr}\,\rho - 1)$. We obtain $\delta\tilde{S}(\rho) = -\mathrm{tr}\,[(\log\rho + 1 + \lambda)\delta\rho] = 0$, from which we obtain $\rho = 2^{-1-\lambda}I_n$. The Lagrange multiplier is fixed as $\lambda = \log n - 1$ from the normalization condition $\mathrm{tr}\,\rho = 1$, for which $\rho = I_n/n$. For a general state ρ in $\mathbb{C}^n$, the entropy takes an intermediate value between 0 and $\log n$.

Let $\rho = \sum_i \lambda_i|\lambda_i\rangle\langle\lambda_i|$ be a spectral decomposition of ρ. Then the von Neumann entropy is expressed as

$$S(\rho) = -\sum_i \lambda_i\log\lambda_i.$$

Note that $\lambda_i \geq 0$ due to non-negativity of ρ.

Exercise II.11 *Calculate the entropy of the following states.*

$$\rho_1 = \frac{1}{2}\begin{pmatrix}1&1\\1&1\end{pmatrix},\ \rho_2 = \frac{1}{5}\begin{pmatrix}1&2\\2&4\end{pmatrix},\ \rho_3 = \frac{1}{2}\begin{pmatrix}1&0&0&1\\0&0&0&0\\0&0&0&0\\1&0&0&1\end{pmatrix}.$$

Quantum relative entropy is defined similarly to the classical case. We define the relative entropy by

$$S(\rho\|\sigma) = \mathrm{tr}\,(\rho\log\rho) - \mathrm{tr}\,(\rho\log\sigma).$$

$S(\rho\|\sigma)$ is also non-negative and it vanishes if and only if $\rho = \sigma$ as we now prove.

Theorem II.12 *$S(\rho\|\sigma)$ satisfies*

$$S(\rho\|\sigma) \geq 0,$$

where the equality is satisfied if and only if $\rho = \sigma$.

Proof: Let us consider the variation of $S(\rho\|\sigma) - \lambda(\mathrm{tr}\,\rho - 1)$ under $\delta\rho$ for a fixed σ,

$$\delta S(\rho\|\sigma) - \lambda \mathrm{tr}\,\delta\rho = \mathrm{tr}\,(\delta\rho \log\rho + \delta\rho - \delta\rho \log\sigma - \lambda\delta\rho) = \mathrm{tr}\,\delta\rho(\log\rho + 1 - \log\sigma - \lambda) = 0,$$

from which we find $\log\rho - \log\sigma = (\lambda - 1)I_n$. By exponentiating both sides, we obtain $\rho = e^{\lambda-1}\sigma$. Then we find $\lambda = 1$ since $\mathrm{tr}\,\rho = \mathrm{tr}\,\sigma = 1$. Now we find the relative entropy takes its extremum value 0 if and only if $\rho = \sigma$. This is a minumum since $S(\rho\|\sigma = I_n/n) = \log n - S(\rho) \geq 0$, where the equality is satified iff ρ is maximally mixed. ∎

I. Nonclassical Correlation other than Entanglement

It is important to realize that only inseparable states have quantum correlations analogous to entangled pure states. It does not necessarily imply all separable states have no non-classical correlation though. It is pointed out that useful non-classical correlation exists in a subset of separable states.

Let us consider a bipartite system with two subsystems A and B of dimensions m and n, respectively. A state ρ^{AB} is called (properly) classically correlated if it has a biproduct eigenvectors. If this is the case, the spectral decomposition of ρ^{AB} is

$$\rho^{AB} = \sum_{1\leq i\leq m, 1\leq j\leq n} c_{ij} |i\rangle_A\langle i| \otimes |j\rangle_B\langle j|.$$

If ρ^{AB} has no such eigenvectors, it is called nonclassically correlated. Obviously, entangled state or inseparable state is nonclassically correlated but the converse is not true. There are nonclassically correlated separable states.

III. QUBITS

A (Boolean) bit assumes two distinct values, 0 and 1, and it constitutes the building block of the classical information theory. Quantum information theory, on the other hand, is based on qubits.

A. One qubit

A qubit is a (unit) vector in the vector space $\mathbb{C}^2$, whose basis vectors are denoted as

$$|0\rangle = (1,0)^t \text{ and } |1\rangle = (0,1)^t. \tag{30}$$

What these vectors physically mean depends on the physical realization employed for quantum information processing.

They might represent spin states of an electron, $|0\rangle = |\uparrow\rangle$ and $|1\rangle = |\downarrow\rangle$. Electrons are replaced by nuclei with spin 1/2 in NMR (Nuclear Magnetic Resonance).

In some cases, $|0\rangle$ stands for a vertically polarized photon $|\updownarrow\rangle$ while $|1\rangle$ represents a horizontally polarized photon $|\leftrightarrow\rangle$. Alternatively they might correspond to photons polarized in different directions. For example, $|0\rangle$ may represent a polarization state $|\nearrow\rangle = \frac{1}{\sqrt{2}}(|\updownarrow\rangle + |\leftrightarrow\rangle)$ while $|1\rangle$ represents a state $|\nwarrow\rangle = \frac{1}{\sqrt{2}}(|\updownarrow\rangle - |\leftrightarrow\rangle)$.

Truncated two states from many levels may be employed as a qubit. We may assign $|0\rangle$ to the ground state and $|1\rangle$ to the first excited state of an atom or an ion.

In any case, we have to fix a set of basis vectors when we carry out quantum information processing. In the following, the basis is written in an abstract form as $\{|0\rangle, |1\rangle\}$, unless otherwise stated.

It is convenient to assume the vector $|0\rangle$ corresponds to the classical bit 0, while $|1\rangle$ to 1. Moreover a qubit may be in a superposition state: $|\psi\rangle = a|0\rangle + b|1\rangle$ with $|a|^2 + |b|^2 = 1$. If we measure $|\psi\rangle$ to see whether it is in $|0\rangle$ or $|1\rangle$, the outcome will be 0 (1) with the probability $|a|^2$ ($|b|^2$) and the state immediately after the measurement is $|0\rangle$ ($|1\rangle$).

Although a qubit may take infinitely many different states, it should be kept in mind that we can extract from it as the same amount of information as that of a classical bit. Information can be extracted only through measurements. When we measure a qubit, the state vector 'collapses' to the eigenvector that corresponds to the eigenvalue observed. Suppose a spin is in the state $a|0\rangle + b|1\rangle$. If we observe that the z-component of the spin is $+1/2$, the system immediately after the measurement is in $|0\rangle$. This happens with probability $\langle\psi|0\rangle\langle 0|\psi\rangle = |a|^2$. The measurement outcome of a qubit is always one of the eigenvalues, which we call abstractly 0 and 1.

B. Bloch sphere

It is useful, for many purposes, to express a state of a single qubit graphically. Let us parameterize a one-qubit pure state $|\psi\rangle$ with θ and ϕ as

$$|\psi(\theta,\phi)\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle. \tag{31}$$

The phase of $|\psi\rangle$ is fixed in such a way that the coefficient of $|0\rangle$ is real. It is easy to verify that $(\hat{\boldsymbol{n}}(\theta,\phi)\cdot\boldsymbol{\sigma})|\psi(\theta,\phi)\rangle = |\psi(\theta,\phi)\rangle$, where $\boldsymbol{\sigma} = (\sigma_x,\sigma_y,\sigma_z)$ and $\hat{\boldsymbol{n}}(\theta,\phi)$ is a real unit vector called the Bloch vector with components $\hat{\boldsymbol{n}}(\theta,\phi) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)^t$. It is therefore natural to assign $\hat{\boldsymbol{n}}(\theta,\phi)$ to a state vector $|\psi(\theta,\phi)\rangle$ so that $|\psi(\theta,\phi)\rangle$ is expressed as a unit vector $\hat{\boldsymbol{n}}(\theta,\phi)$ on the surface of the unit sphere, called the Bloch sphere. This correspondence is one-to-one if the ranges of θ and ϕ are restricted to $0 \le \theta \le \pi$ and $0 \le \phi < 2\pi$.

It is verified that state (31) satisfies

$$\langle\psi(\theta,\phi)|\boldsymbol{\sigma}|\psi(\theta,\phi)\rangle = \hat{\boldsymbol{n}}(\theta,\phi). \tag{32}$$

A density matrix ρ of a qubit can be represented as a point on a unit ball. Since ρ is a positive semi-definite Hermitian matrix with unit trace, its most general form is

$$\rho = \frac{1}{2}\left(I + \sum_{i=x,y,z} u_i\sigma_i\right), \tag{33}$$

where $\vec{u} \in \mathbb{R}^3$ satisfies $|\boldsymbol{u}| \le 1$. The reality follows from the Hermiticity requirement and $\mathrm{tr}\,\rho = 1$ is obvious. The eigenvalues of ρ are $\lambda_\pm = \frac{1}{2}\left(1 \pm \sqrt{|\boldsymbol{u}|}\right)/2$ and therefore non-negative. The eigenvalue λ_- vanishes in case $|\boldsymbol{u}| = 1$, for which rank $\rho = 1$. Therefore the surface of the unit sphere corresponds to pure states. The converse is also shown easily. In contrast, all the points $\boldsymbol{u}$ inside a unit ball correspond to mixed states. The ball is called the Bloch ball and the vector $\boldsymbol{u}$ is also called the Bloch vector.

It is easily verified that ρ given by Eq. (33) satisfies

$$\langle\boldsymbol{\sigma}\rangle = \mathrm{tr}\,(\rho\boldsymbol{\sigma}) = \boldsymbol{u}. \tag{34}$$

Exercise III.1 *Prove Eqs. (32) and (34).*

C. Multi-qubit systems and entangled states

Let us consider a group of many (n) qubits next. Such a system behaves quite differently from a classical one and this difference gives a distinguishing aspect to quantum information theory. An n-qubit system is often called a (quantum) register in the context of quantum computing.

As an example, let us consider an n-qubit register. Suppose we specify the state of each qubit separately like a classical case. Each of the qubit is then described by a 2-d complex vector of the form $a_i|0\rangle + b_i|1\rangle$ and we need $2n$ complex numbers $\{a_i, b_i\}_{1\le i\le n}$ to specify the state. This corresponds the a tensor product state $(a_1|0\rangle + b_1|1\rangle)\otimes\ldots\otimes(a_n|0\rangle + b_n|1\rangle) \in \mathbb{C}^{2n}$. If the system is treated in a fully quantum-mechanical way, however, a general state vector of the register is represented as

$$|\psi\rangle = \sum_{i_k=0,1} a_{i_1 i_2 \ldots i_n}|i_1\rangle\otimes|i_2\rangle\otimes\ldots\otimes|i_n\rangle \in \mathbb{C}^{2^n}.$$

Note that $2^n \gg 2n$ for a large number n. The ratio $2^n/2n$ is $\sim 10^{298}$ for $n = 1000$. Most quantum states in a Hilbert space with large n are entangled having no classical analogues. Entanglement is an extremely powerful resource for quantum computation and quantum communication.

Let us consider a 2-qubit system for definiteness. The system has a binary basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. More generally, a basis for a system of n qubits may be $\{|b_{n-1}b_{n-2}\ldots b_0\rangle\}$, where $b_{n-1}, b_{n-2}, \ldots, b_0 \in \{0,1\}$. It is also possible to express the basis in terms of the decimal system. We write $|x\rangle$, instead of $|b_{n-1}b_{n-2}\ldots b_0\rangle$, where $x = b_{n-1}2^{n-1} + b_{n-2}2^{n-2} + \ldots + b_0$. The basis for a 2-qubit system may be written also as $\{|0\rangle, |1\rangle, |2\rangle, |3\rangle\}$ with this decimal notation.

The set

$$\begin{aligned}\{|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \quad |\Phi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle),\\ |\Psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \quad |\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)\}\end{aligned} \tag{35}$$

is an orthonormal basis of a two-qubit system and is called the Bell basis. Each vector is called the Bell state or the Bell vector. Note that all the Bell states are entangled.

IV. QUANTUM GATES, QUANTUM CIRCUIT AND QUANTUM COMPUTATION

A. Introduction

Now that we have introduced qubits to store information, it is time to consider operations acting on them. If they are simple, these operations are called gates, or quantum gates, in analogy with those in classical logic circuits. More complicated quantum circuits are composed of these simple gates. A collection of quantum circuits for executing a complicated algorithm, a quantum algorithm, is a part of a quantum computation.

Definition IV.1 *(Quantum Computation) A quantum computation is a collection of the following three elements:*

(1) A register or a set of registers,

(2) A unitary matrix u, which is taylored to execute a given quantum algorithm and

(3) Measurements to extract information we need.

More formally, a quantum computation is the set $\{\mathcal{H}, U, \{M_m\}\}$*, where* $\mathcal{H} = \mathbb{C}^{2^n}$ *is the Hilbert space of an n-qubit register,* $U \in U(2^n)$ *represents a quantum algorithm and* $\{M_m\}$ *is the set of measurement operators. The hardware* (1) *is called a quantum computer.*

Suppose the register is set to a fiducial initial state, $|\psi_{\rm in}\rangle = |00\ldots0\rangle$ for example. A unitary matrix $U_{\rm alg}$ is generated by an algorithm which we want to execute. Operation of $U_{\rm alg}$ on $|\psi_{\rm in}\rangle$ yields the output state $|\psi_{\rm out}\rangle = U_{\rm alg}|\psi_{\rm in}\rangle$. Information is extracted from $|\psi_{\rm out}\rangle$ by appropriate measurements.

B. Quantum gates

We have so far studied the change of a state upon measurements. When measurements are not made, the time evolution of a state is described by the Schrödinger equation. The time evolution operator U is unitary: $UU^\dagger = U^\dagger U = I$. We will be free from the Schrödinger equation in the following and assume there always exist unitary matrices which we need.

One of the important conclusions derived from the unitarity of gates is that the computational process is reversible.

1. Simple quantum gates

Examples of quantum gates which transform a one-qubit state are given below. We call them one-qubit gates in the following. Linearity guarantees that the action of a gate is completely specified if its action on the basis $\{|0\rangle,\ |1\rangle\}$ is given. Consider the gate I whose action on the basis vectors is $I:\ |0\rangle \to |0\rangle,\ |1\rangle \to |1\rangle$. The matrix expression of this gate is

$$I = |0\rangle\langle 0| + |1\rangle\langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{36}$$

Similarly we introduce $X:\ |0\rangle \to |1\rangle,\ |1\rangle \to |0\rangle$, $Y:\ |0\rangle \to -|1\rangle,\ |1\rangle \to |0\rangle$ and $Z:\ |0\rangle \to |0\rangle,\ |1\rangle \to -|1\rangle$ by

$$X = |1\rangle\langle 0| + |0\rangle\langle 1| = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \sigma_x, \tag{37}$$

$$Y = |0\rangle\langle 1| - |1\rangle\langle 0| = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = -i\sigma_y, \tag{38}$$

$$Z = |0\rangle\langle 0| - |1\rangle\langle 1| = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \sigma_z. \tag{39}$$

The transformation I is the identity transformation, while X is the negation (NOT), Z the phase shift and $Y = XZ$ the combination thereof.

CNOT (controlled-NOT) gate is a 2-qubit gate, which plays an important role. The gate flips the second qubit (the target qubit) when the first qubit (the control qubit) is $|1\rangle$, while leaving the second bit unchanged when the first bit is $|0\rangle$. Let $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ be a basis for the 2-qubit system. We use the standard basis vectors with components

$$|00\rangle = (1,0,0,0)^t,\ |01\rangle = (0,1,0,0)^t,\ |10\rangle = (0,0,1,0)^t,\ |11\rangle = (0,0,0,1)^t.$$

The action of CNOT gate, whose matrix expression will be written as $U_{\rm CNOT}$, is $U_{\rm CNOT}:\ |00\rangle \mapsto |00\rangle,\ |01\rangle \mapsto |01\rangle,\ |10\rangle \mapsto |11\rangle,\ |11\rangle \mapsto |10\rangle$. It has two equivalent expressions

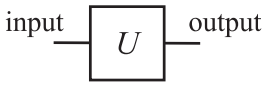
$$U_{\rm CNOT} = |00\rangle\langle 00| + |01\rangle\langle 01| + |11\rangle\langle 10| + |10\rangle\langle 11| = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes X, \tag{40}$$

having a matrix form

$$U_{\rm CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{41}$$

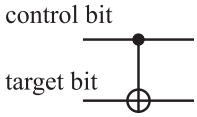
Let $\{|i\rangle\}$ be the basis vectors, where $i \in \{0,1\}$. The action of CNOT on the input state $|i,j\rangle$ is written as $|i, i \oplus j\rangle$, where $i \oplus j$ is an addition mod2.

A 1-qubit gate whose unitary matrix is U is graphically depicted as

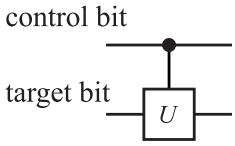


The left horizontal line is the input qubit while the right horizontal line is the output qubit: time flows from the left to the right.

A CNOT gate is expressed as



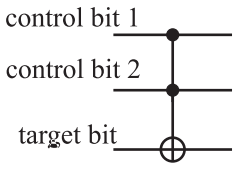
where $\bullet$ denotes the control bit, while $\bigoplus$ denotes the conditional negation. There may be many control bits (see CCNOT gate below). More generally, we consider a controlled-U gate, $V = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes U$, in which the target bit is acted on by a unitary transformation U only when the control bit is $|1\rangle$. This gate is denoted graphically as



CCNOT (Controlled-Controlled-NOT) gate has three inputs and the third qubit flips only when the first two qubits are both in the state $|1\rangle$. The explicit form of the CCNOT gate is

$$U_{\rm CCNOT} = (|00\rangle\langle 00| + |01\rangle\langle 01| + |10\rangle\langle 10|) \otimes I + |11\rangle\langle 11| \otimes X. \tag{42}$$

This gate is graphically expressed as



2. Walsh-Hadamard transformation

The Hadamard gate or the Hadamard transformation H is an important unitary transformation defined by

$$\begin{aligned} U_{\rm H}: \quad & |0\rangle \to \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\ & : |1\rangle \to \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \end{aligned} \tag{43}$$

The matrix representation of H is

$$U_{\rm H} = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)\langle 0| + \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\langle 1| = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{44}$$

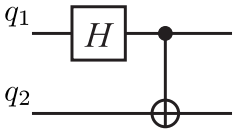
A Hadamard gate is depicted as

There are numerous important applications of the Hadamard transformation. All possible 2^n states are generated when $U_{\rm H}$ is applied on each qubit of the state $|00\ldots0\rangle$:

$$\begin{aligned} & (U_{\rm H} \otimes U_{\rm H} \otimes \ldots \otimes U_{\rm H})|00\ldots0\rangle \\ & = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \ldots \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle. \end{aligned} \tag{45}$$

Therefore, we produce a superposition of all the states $|x\rangle$ with $0 \le x \le 2^n - 1$ simultaneously. The transformation $U_{\rm H}^{\otimes n}$ is called the Walsh transformation, or Walsh-Hadamard transformation and denoted as W_n.

Exercise IV.2 *Show that the quantum circuit*



generates Bell states from inputs $|q_1 q_2\rangle = |00\rangle, |01\rangle, |10\rangle$ *and* $|11\rangle$.

3. SWAP gate and Fredkin gate

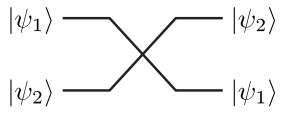
The SWAP gate acts on a tensor product state as

$$U_{\rm SWAP}|\psi_1, \psi_2\rangle = |\psi_2, \psi_1\rangle. \tag{46}$$

The explict form of $U_{\rm SWAP}$ is given by

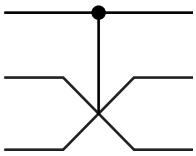
$$U_{\rm SWAP} = |00\rangle\langle 00| + |01\rangle\langle 10| + |10\rangle\langle 01| + |11\rangle\langle 11| = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{47}$$

The SWAP gate is expressed as



Note that the SWAP gate is a special gate which maps an arbitrary tensor product state to a tensor product state. In contrast, most 2-qubit gates map a tensor product state to an entangled state.

The controlled-SWAP gate



is also called the Fredkin gate. It flips the second (middle) and the third (bottom) qubits only when the first (top) qubit is in the state $|1\rangle$. Its explicit form is $U_{\text{Fredkin}} = |0\rangle\langle 0| \otimes I_4 + |1\rangle\langle 1| \otimes U_{\text{SWAP}}$.

C. No-cloning theorem

Theorem IV.3 *(Wootters and Zurek [14]) An unknown quantum system cannot be cloned by unitary transformations.*

Proof: Suppose there would exist a unitary transformation U that makes a clone of a quantum system. Namely, suppose U acts, for any state $|\varphi\rangle$, as $U : |\varphi 0\rangle \to |\varphi\varphi\rangle$. Let $|\varphi\rangle$ and $|\phi\rangle$ be two states that are linearly independent. Then we should have $U|\varphi 0\rangle = |\varphi\varphi\rangle$ and $U|\phi 0\rangle = |\phi\phi\rangle$ by definition. Then the action of U on $|\psi\rangle = \frac{1}{\sqrt{2}}(|\varphi\rangle + |\phi\rangle)$ yields

$$U|\psi 0\rangle = \frac{1}{\sqrt{2}}(U|\varphi 0\rangle + U|\phi 0\rangle) = \frac{1}{\sqrt{2}}(|\varphi\varphi\rangle + |\phi\phi\rangle).$$

If U were a cloning transformation, we must also have

$$U|\psi 0\rangle = |\psi\psi\rangle = \frac{1}{2}(|\varphi\varphi\rangle + |\varphi\phi\rangle + |\phi\varphi\rangle + |\phi\phi\rangle),$$

which contradicts the previous result. Therefore, there does not exist a unitary cloning transformation. ∎

Note however that the theorem does not apply if the states to be cloned are limited to $|0\rangle$ and $|1\rangle$. For these cases, the copying operator U should work as $U : |00\rangle \mapsto |00\rangle, \quad : |10\rangle \mapsto |11\rangle$. We can assign arbitrary action of U on a state whose second input is $|1\rangle$ since this case will never happen. What we have to keep in our mind is only that U be unitary. An example of such U is $U = (|00\rangle\langle 00| + |11\rangle\langle 10|) + (|01\rangle\langle 01| + |10\rangle\langle 11|)$, where the first set of operators renders U the cloning operator and the second set is added just to make U unitary. We immediately notice that U is nothing but the CNOT gate.

Therefore, if the data under consideration is limited within $|0\rangle$ and $|1\rangle$, we can copy the qubit states even in a quantum computer. This fact is used to construct quantum error correcting codes.

D. Quantum teleportation

The purpose of quantum teleportation is to transmit an unknown quantum *state* of a qubit using two classical bits in such a way that the recipient reproduces the same state as the original qubit state. Note that the qubit itself is not transported but the information required to reproduce the quantum state is transmitted. The original state is destroyed such that quantum teleportation is not in contradition with the no-cloning theorem.

Alice: Alice has a qubit, whose state she does *not* know. She wishes to send Bob the quantum state of this qubit through a classical communication channel. Let $|\phi\rangle = a|0\rangle + b|1\rangle$ be the state of the qubit. Both of them have been given one of the qubits of the entangled pair $|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ in advance. They start with the state

$$|\phi\rangle \otimes |\Phi^+\rangle \;=\; \frac{1}{\sqrt{2}}\left(a|000\rangle + a|011\rangle + b|100\rangle + b|111\rangle\right), \tag{48}$$

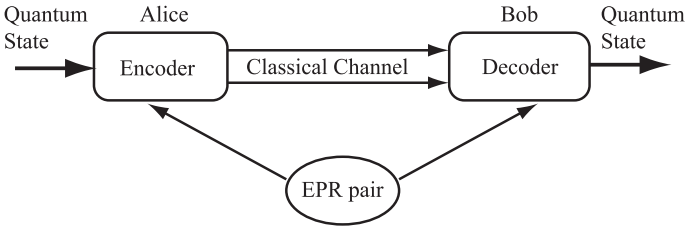


FIG. 1: In quantum teleportation, Alice sends Bob two classical bits so that Bob reproduces a qubit state Alice initially had.

where Alice possesses the first two qubits while Bob has the third. Alice applies $U_{\rm CNOT} \otimes I$ followed by $U_{\rm H} \otimes I \otimes I$ to this state, which results in

$$\begin{aligned}&(U_{\rm H} \otimes I \otimes I)(U_{\rm CNOT} \otimes I)(|\phi\rangle \otimes |\Phi^+\rangle) \\ &= \frac{1}{2}[|00\rangle(a|0\rangle + b|1\rangle) + |01\rangle(a|1\rangle + b|0\rangle) + |10\rangle(a|0\rangle - b|1\rangle) + |11\rangle(a|1\rangle - b|0\rangle)].\end{aligned} \tag{49}$$

If Alice measures the 2 qubits in her hand, she will obtain one of the states $|00\rangle, |01\rangle, |10\rangle$ or $|11\rangle$ with equal probability 1/4. Bob's qubit (one of the EPR pair previously) collapses to $a|0\rangle + b|1\rangle, a|1\rangle + b|0\rangle, a|0\rangle - b|1\rangle$ or $a|1\rangle - b|0\rangle$, respectively, depending on the result of Alice's measurement. Alice then sends Bob her result of the measurement using two classical bits.

Bob: After receiving two classical bits, Bob knows the state of the qubit in his hand;

received bits	Bob's state	decoding
00	$a\|0\rangle + b\|1\rangle$	I
01	$a\|1\rangle + b\|0\rangle$	X
10	$a\|0\rangle - b\|1\rangle$	Z
11	$a\|1\rangle - b\|0\rangle$	Y

(50)

Bob reconstructs the intial state $|\phi\rangle$ by applying the decoding process shown above. Suppose Alice sends Bob classical bits 10, for example. Then Bob applies Z on his qubit to reconstruct $|\phi\rangle$ as $Z: \ (a|0\rangle - b|1\rangle) \mapsto (a|0\rangle + b|1\rangle) = |\phi\rangle$.

Figure 2 shows the actual quantum circuit for quantum teleportation.

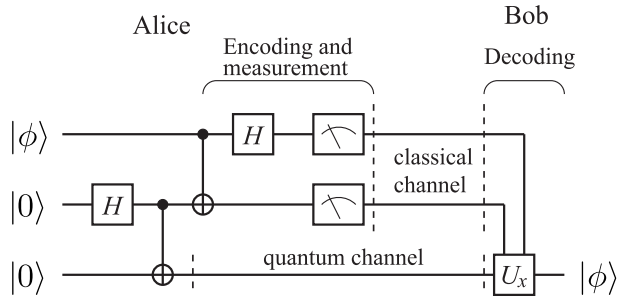


FIG. 2: Quantum circuit implementation of quantum teleportation.

E. LOCC

LOCC is the set of manipulations in quantum information processing. It can be defined for any multipartite systems but we concentrate on a bipartite system for definiteness. Suppose Alice has a subsystem A and Bob has a subsystem B of a given bipartite system AB. Each of them is allowed to make unitary operations on his/her own subsystem and make measurements on own subsystem (**L**ocal **O**perations) and they are allowed to communicate classically (**C**lassical **C**ommunication) using telephone or internet, hence the name LOCC. Let Alice's Hilber space be $\mathbb{C}^m$ and Bob's be $\mathbb{C}^n$. Then local operations are elements of $\mathrm{U}(m) \otimes \mathrm{U}(n)$. By using the second manipulation, Alice can reflect data Bob supplies on her subsystem, and vice versa, but possible operations are still restricted within $\mathrm{U}(m) \otimes \mathrm{U}(n)$. An important operation ruled out from LOCC is to entangle two remote qubits in a tensor product state.

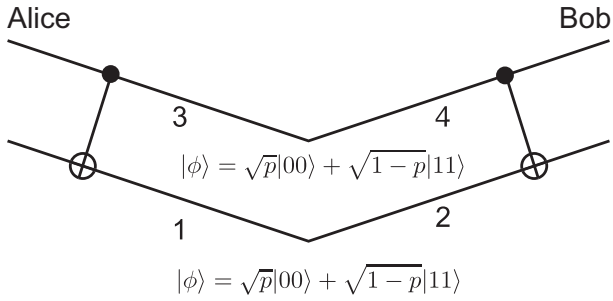


FIG. 3: Entanglement distillation of the first kind. Alice and Bob share a Bell state $|\Phi_+\rangle$ when their measurement outcomes of the third and the fourth qubits are 11.

Example IV.4 *Suppose Alice and Bob share one of two Bell states*

$$|\Phi_+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle),\ |\Psi_+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle).$$

Alice has the first qubit while Bob has the second. They have picked up one of the state. Then they can tell which state they have chosen by LOCC. Alice measures her qubit and send the result to Bob by classical communication. Then Bob measures his own qubit. Suppose Alice measured 1 on her qubit and told Bob that she got 1. If Bob's readout is 1, the state they have shared was $|\Phi_+\rangle$, while Bob's readout is 0, their state was $\|Psi_+\rangle$.

Let us consider $SU(4)$, the set of two-qubit operations. It will be shown later that any element $U \in SU(4)$ can be decomposed as $U = K_1 H K_2$, where $K_i \in SU(2) \otimes SU(2)$ and $H = \exp[-i(c_x\sigma_x \otimes \sigma_x + c_y\sigma_y \otimes \sigma_y + c_z\sigma_z \otimes \sigma_z)$. Since

$$(U_1 \otimes U_2)|\psi_1\rangle|\psi_2\rangle = (U_1|\psi_1\rangle) \otimes (U_2|\psi_2\rangle),$$

any element of $SU(2) \otimes SU(2)$ fails to entangle tensor product states. The element H of SU(4) is in charge of entanglement. Due to the same reason, LOCC cannot increase or decrease entanglement. It is interesting to note that $SU(2) \otimes SU(2)$ is isomorhic to $O(4)$.

We look at an interesting example of LOCC protocol in the next subsection.

F. Entenglement Distillation I

Entanglement distillation is an LOCC protocol for distributing an EPR pair between two parties, Alice and Bob. There are two types of entanglement distillation. The second one requires knowledge of a quantum channel and will be explained after a quantum channel is introduced.

The first one creates a maximally entangled state $|\Phi_+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ from a pair of less entangled state $|\phi\rangle = \sqrt{p}|00\rangle + \sqrt{1-p}|11\rangle$. Suppose Alice and Bob share two $|\phi\rangle$ states. Alice keeps the first qubit while Bob has the second qubit of each pair. Now the initial state is

$$|\Psi_1\rangle = p|00\rangle|00\rangle + \sqrt{p(1-p)}(|00\rangle|11\rangle + |11\rangle|00\rangle) + (1-p)|11\rangle|11\rangle.$$

Alice has the first and the third qubits while Bob has the second and the fourth. After the two pairs are distributed, both Alice and Bob applies CNOT gate on their qubits. The first (second) qubit is the control bit while the third (fourth) qubit is the target bit, see Fig. 3. The resulting state is

$$|\Psi_2\rangle = p|00\rangle|00\rangle + \sqrt{p(1-p)}(|00\rangle + |11\rangle)|11\rangle + (1-p)|11\rangle|00\rangle.$$

Subsequently, Alice and Bob measure the third and the fourth qubits and exchange their measurement outcomes using classical communication, such as telephone or email. When the measurement outcomes are 00, they discard the qubits and start again from distribution of a pair of $|\psi\rangle$. This happens with a probability $p^2 + (1-p)^2 = 1 - 2p + 2p^2$. When the measurment outcomes are 11, the state of the first and the second qubits is

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

as promised. This happens with a probability $2p(1-p) = 2p - 2p^2$. We have learned that a maximally entangled state $|\Phi_+\rangle$ is obtained from a pair of less entangled states under entanglement distillation.

G. Universal quantum gates

It can be shown that any classical logic gate can be constructed by using a small set of gates, AND, NOT and XOR for example. Such a set of gates is called the *universal* set of gates. It can be shown that the CCNOT gate simulates these classical gates, and hence quantum circuits simulate any classical circuits. The set of quantum gates is, however, much larger than those classical gates. Thus we want to find a universal set of *quantum* gates from which any quantum circuits can be constructed.

It can be shown that
(1) the set of single qubit gates and
(2) CNOT gate
form a universal set of quantum circuits (universality theorem). The proof is highly technical and is not given here [1, 2, 16]. We, instead, sketch the proof in several lines.

It can be shown that any $U \in \mathrm{U}(n)$ is written as a product of N two-level unitary matrices, where $N \leq n(n-1)/2$ and a two-level unitary matrix is a unit matrix I_n in which only four components V_{aa}, V_{ab}, V_{ba} and V_{bb} are different from I_n. Moreover $V = (V_{ij})$ is an element of U(2). An example of a two-level unitary matrix is

$$V = \begin{pmatrix} \alpha^* & 0 & 0 & \beta^* \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta & 0 & 0 & \alpha \end{pmatrix}, \quad (|\alpha|^2 + |\beta|^2 = 1)$$

where $a = 1$ and $b = 4$.

Now we need to prove the universality theorem for two-level unitary matrices, which is certainly simpler than the general proof. By employing CNOT gates and their generalizations, it is possible to move the elements V_{aa}, V_{ab}, V_{ba} and V_{bb} so that they acts on a single qubit in the register. We need to implement the controlled-V gate whose target qubit is the one on which V acts. Implementation of the controlled-V gate requires generalized CNOT gates and several U(2) gates [1, 2, 16].

H. Quantum parallelism and entanglement

Given an input x, a typical quantum computer "computes" $f(x)$ as

$$U_f : |x\rangle|0\rangle \mapsto |x\rangle|f(x)\rangle, \tag{51}$$

where U_f is a unitary matrix which implements the function f.

Suppose U_f acts on an input which is a superposition of many $|x\rangle$. Since U_f is a linear operator, it acts on all the constituent vectors of the superposition simultaneously. The output is also a superposition of all the results;

$$U_f : \sum_x |x\rangle|0\rangle \mapsto \sum_x |x\rangle|f(x)\rangle. \tag{52}$$

This feature, called the *quantum parallelism*, gives quantum computer an enormous power. A quantum computer is advantageous over a classical counterpart in that it makes use of this quantum parallelism and also entanglement.

A unitary transformation acts on a superposition of all possible states in most quantum algorithms. This superposition is prepared by the action of the Walsh-Hadamard transformation on an n-qubit register in the initial state $|00\ldots0\rangle = |0\rangle \otimes |0\rangle \otimes \ldots \otimes |0\rangle$ resulting in $\sum_{x=0}^{2^n-1} |x\rangle/\sqrt{2^n}$. This state is a superposition of vectors encoding all the integers between 0 and $2^n - 1$. Then the linearlity of U_f leads to

$$U_f \left(\frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle|0\rangle \right) = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} U_f|x\rangle|0\rangle = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle|f(x)\rangle. \tag{53}$$

Note that the superposition is made of $2^n = e^{n \ln 2}$ states, which makes quantum computation exponentially faster than classical counterpart in a certain kind of computation.

What about the limitation of a quantum computer[3]? Let us consider the CCNOT gate for example. This gate flips the third qubit if and only if the first and the second qubits are both in the state $|1\rangle$ while it leaves the third qubit unchanged otherwise. Let us fix the third input qubit to $|0\rangle$. The third output qubit state is $|x \wedge y\rangle$, where $|x\rangle$ and $|y\rangle$ are the first and the second input qubits respectively. Suppose the input state of the first and the second qubits is a

superposition of all possible states while the third qubit is fixed to $|0\rangle$. This can be achieved by the Walsh-Hadamard transformation as

$$\begin{aligned} U_{\rm H}|0\rangle \otimes U_{\rm H}|0\rangle \otimes |0\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |0\rangle \\ &= \frac{1}{2}(|000\rangle + |010\rangle + |100\rangle + |110\rangle). \end{aligned} \tag{54}$$

By operating CCNOT on this state, we obtain

$$U_{\rm CCNOT}(U_{\rm H}|0\rangle \otimes U_{\rm H}|0\rangle \otimes |0\rangle) = \frac{1}{2}(|000\rangle + |010\rangle + |100\rangle + |111\rangle). \tag{55}$$

This output may be thought of as the truth table of AND: $|x, y, x \wedge y\rangle$. It is extremely important to note that the output is an entangled state and the measurement projects the state to *one line* of the truth table, i.e., a single term in the RHS of Eq. (55).

There is no advantage of quantum computation over classical one at this stage. This is because only *one* result may be obtained by a single set of measurements. What is worse, we cannot choose a specific vector $|x, y, x \wedge y\rangle$ at our will! Thus any quantum algorithm should be programmed so that the particular vector we want to observe should have larger probability to be measured compared to other vectors. The programming strategies to deal with this feature are

1. to amplify the amplitude, and hence the probability, of the vector that we want to observe. This strategy is employed in the Grover's database search algorithm.

2. to find a common property of all the $f(x)$. This idea was employed in the quantum Fourier transform to find the order[81] of f in the Shor's factoring algorithm.

Now we consider the power of entanglement. Suppose we have an n-qubit register, whose Hilbert space is 2^n-dimensional. Since each qubit has two basis states $\{|0\rangle, |1\rangle\}$, there are $2n$ basis states, i.e., n $|0\rangle$'s and n $|1\rangle$'s, involved to span this Hilbert space. Imagine that we have a single quantum system, instead, which has the same Hilbert space. One might think that the system may do the same quantum computation as the n-qubit register does. One possible problem is that one cannot "measure the kth digit" leaving other digits unaffected. Even worse, consider how many different basis vectors are required for this system. This single system must have an enormous number, 2^n, of basis vectors! Multipartite implementation of a quantum algorithm requires exponentially smaller number of basis vectors than monopartite implementation since the former makes use of entanglement as a computational resource.

V. SIMPLE QUANTUM ALGORITHMS

Let us introduce a few simple quantum algorithms which will be of help to understand how quantum algorithms are different from and superior to classical algorithms.

A. Deutsch algorithm

The Deutsch algorithm is one of the first quantum algorithms which showed quantum algorithms may be more efficient than their classical counterparts. In spite of its simplicty, full usage of superposition principle and entanglement has been made here.

Let $f : \{0, 1\} \to \{0, 1\}$ be a binary function. Note that there are only four possible f, namely

$$\begin{aligned} f_1 &: 0 \mapsto 0,\ 1 \mapsto 0, \quad f_2 : 0 \mapsto 1,\ 1 \mapsto 1, \\ f_3 &: 0 \mapsto 0,\ 1 \mapsto 1, \quad f_4 : 0 \mapsto 1,\ 1 \mapsto 0. \end{aligned}$$

First two cases, f_1 and f_2, are called *constant*, while the rest, f_3 and f_4, are *balanced*. If we only have classical resources, we need to evaluate f twice to tell if f is constant or balanced. There is a quantum algorithm, in contrast, with which it is possible to tell if f is constant or balanced with a single evaluation of f, as was shown by Deutsch [18].

Let $|0\rangle$ and $|1\rangle$ correspond to classical bits 0 and 1, respectively, and consider the state $|\psi_0\rangle = \frac{1}{2}(|00\rangle - |01\rangle + |10\rangle - |11\rangle)$. We apply f on this state in terms of the unitary operator $U_f : |x, y\rangle \mapsto |x, y \oplus f(x)\rangle$, where $\oplus$ is an addition mod 2. To be explicit, we obtain

$$|\psi_1\rangle = U_f|\psi_0\rangle = \frac{1}{2}(|0, f(0)\rangle - |0, \neg f(0)\rangle + |1, f(1)\rangle - |1, \neg f(1)\rangle),$$

where $\neg$ stands for negation. Therefore this operation is nothing but the CNOT gate with the control bit $f(x)$; the target bit y is flipped if and only if $f(x) = 1$ and left unchanged otherwise. Subsequently we apply the Hadamard gate on the first qubit to obtain

$$\begin{aligned} |\psi_2\rangle &= U_{\rm H}|\psi_1\rangle \\ &= \frac{1}{2\sqrt{2}}\left[(|0\rangle + |1\rangle)(|f(0)\rangle - |\neg f(0)\rangle) + (|0\rangle - |1\rangle)(|f(1)\rangle - |\neg f(1)\rangle)\right] \end{aligned}$$

The wave function reduces to

$$|\psi_2\rangle = \frac{1}{\sqrt{2}}|0\rangle(|f(0)\rangle - |\neg f(0)\rangle) \tag{56}$$

in case f is constant, for which $|f(0)\rangle = |f(1)\rangle$, and

$$|\psi_2\rangle = \frac{1}{\sqrt{2}}|1\rangle(|f(0)\rangle - |f(1)\rangle) \tag{57}$$

if f is balanced, for which $|\neg f(0)\rangle = |f(1)\rangle$. Therefore the measurement of the first qubit tells us whether f is constant or balanced.

Let us consider a quantum circuit which implements the Deutsch algorithm. We first apply the Walsh-Hadamard transformation $W_2 = U_{\rm H} \otimes U_{\rm H}$ on $|01\rangle$ to obtain $|\psi_0\rangle$. We need to introduce a conditional gate U_f, i.e., the controlled-NOT gate with the control bit $f(x)$, whose action is $U_f : |x, y\rangle \to |x, y \oplus f(x)\rangle$. Then the Hadamard gate is applied on the first qubit before it is measured. Figure 4 depicts this implementation.

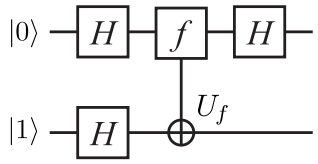


FIG. 4: Implementation of the Deutsch algorithm.

In the quauntum circuit, we assume the gate U_f is a black box for which we do not ask the explicit implementation. We might think it is a kind of subroutine. Such a black box is often called an oracle. The gate U_f is called the Deutsch oracle. Its implementation is given only after f is specified.

Then what is the merit of the Deutsch algorithm? Suppose your friend gives you a unitary matrix U_f and asks you to tell if f is constant or balanced. Instead of applying $|0\rangle$ and $|1\rangle$ separately, you may contruct the circuit in Fig. 4 with the given matrix U_f and apply the circuit on the input state $|01\rangle$. Then you can tell your friend whether f is constant or balanced with a single use of U_f.

B. Deutsch-Jozsa algorithm

The Deutsch algorithm introduced in the previous section may be generalized to the Deutsch-Jozsa algorithm [19]. Let us first define the Deutsch-Jozsa problem. Suppose there is a binary function

$$f : S_n \equiv \{0, 1, \ldots, 2^n - 1\} \to \{0, 1\}. \tag{58}$$

We require f be either *constant* or *balanced* as before. When f is constant, it takes a constant value 0 or 1 irrespetive of the input value x. When it is balanaced the value $f(x)$ for a half of $x \in S_n$ is 0 while it is 1 for the rest of x. Although there are functions which are neither constant nor balanced, we will not consider such cases here. Our task is to find an algorithm which tells if f is constant or balanced with the least possible number of evaluations of f.

It is clear that we need at least $2^{n-1}+1$ steps, in the worst case with classical manipulations, to make sure if $f(x)$ is constant or balanced with 100 % confidence. It will be shown below that the number of steps reduces to a single step if we are allowed to use a quantum algorithm.

The algorithm is divided into the following steps:

1. Prepare an $(n+1)$-qubit register in the state $|\psi_0\rangle = |0\rangle^{\otimes n} \otimes |1\rangle$. First n qubits work as input qubits while the $(n+1)$st qubit serves as a "scratch pad". Such qubits, which are neither input qubits nor output qubits, but work as a scratch pad to store temporary information are called ancillas or ancillary qubits.

2. Apply the Walsh-Hadamard transforamtion to the register. Then we have the state

$$\begin{aligned} |\psi_1\rangle &= U_{\rm H}^{\otimes n+1}|\psi_0\rangle = \frac{1}{\sqrt{2^n}}(|0\rangle+|1\rangle)^{\otimes n} \otimes \frac{1}{\sqrt{2}}(|0\rangle-|1\rangle) \\ &= \frac{1}{\sqrt{2^n}}\sum_{x=0}^{2^n-1}|x\rangle \otimes \frac{1}{\sqrt{2}}(|0\rangle-|1\rangle). \end{aligned} \tag{59}$$

3. Apply the $f(x)$-controlled-NOT gate on the register, which flips the $(n+1)$st qubit if and only if $f(x)=1$ for the input x. Therefore we need a U_f gate which evaluates $f(x)$ and acts on the register as $U_f|x\rangle|c\rangle = |x\rangle|c\oplus f(x)\rangle$, where $|c\rangle$ is the one-qubit state of the $(n+1)$st qubit. Observe that $|c\rangle$ is flipped if and only if $f(x)=1$ and left unchanged otherwise. We then obtain a state

$$\begin{aligned} |\psi_2\rangle &= U_f|\psi_1\rangle = \frac{1}{\sqrt{2^n}}\sum_{x=0}^{2^n-1}|x\rangle\frac{1}{\sqrt{2}}(|f(x)\rangle-|\neg f(x)\rangle) \\ &= \frac{1}{\sqrt{2^n}}\sum_x(-1)^{f(x)}|x\rangle\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle). \end{aligned} \tag{60}$$

   Although the gate U_f is applied once for all, it is applied to *all* the n-qubit states $|x\rangle$ simultaneously.

4. The Walsh-Hadamard transformation (45) is applied on the first n qubits next. We obtain

$$|\psi_3\rangle = (W_n \otimes I)|\psi_2\rangle = \frac{1}{\sqrt{2^n}}\sum_{x=0}^{2^n-1}(-1)^{f(x)}U_{\rm H}^{\otimes n}|x\rangle\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle). \tag{61}$$

   It is instructive to write the action of the one-qubit Hadamard gate as

$$U_{\rm H}|x\rangle = \frac{1}{\sqrt{2}}(|0\rangle+(-1)^x|1\rangle) = \frac{1}{\sqrt{2}}\sum_{y\in\{0,1\}}(-1)^{xy}|y\rangle,$$

   where $x \in \{0,1\}$, to find the resulting state. The action of the Walsh-Hadamard transformation on $|x\rangle = |x_{n-1}\dots x_1x_0\rangle$ yields

$$\begin{aligned} W_n|x\rangle &= (U_{\rm H}|x_{n-1}\rangle)(U_{\rm H}|x_{n-2}\rangle)\dots(U_{\rm H}|x_0\rangle) \\ &= \frac{1}{\sqrt{2^n}}\sum_{y_{n-1},y_{n-2},\dots,y_0\in\{0,1\}}(-1)^{x_{n-1}y_{n-1}+x_{n-2}y_{n-2}+\dots+x_0y_0} \\ &\quad \times|y_{n-1}y_{n-2}\dots y_0\rangle = \frac{1}{\sqrt{2^n}}\sum_{y=0}^{2^n-1}(-1)^{x\cdot y}|y\rangle, \end{aligned} \tag{62}$$

   where $x\cdot y = x_{n-1}y_{n-1}\oplus x_{n-2}y_{n-2}\oplus\dots\oplus x_0y_0$. Substituting this result into Eq. (61), we obtain

$$|\psi_3\rangle = \frac{1}{2^n}\left(\sum_{x,y=0}^{2^n-1}(-1)^{f(x)}(-1)^{x\cdot y}|y\rangle\right)\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle). \tag{63}$$

5. The first n qubits are measured. Suppose $f(x)$ is constant. Then $|\psi_3\rangle$ is put in the form

$$|\psi_3\rangle = \frac{1}{2^n}\sum_{x,y}(-1)^{x\cdot y}|y\rangle\frac{1}{\sqrt{2}}(|0\rangle-|1\rangle)$$

up to an overall phase. Let us consider the summation $\frac{1}{2^n}\sum_{x=0}^{2^n-1}(-1)^{x\cdot y}$ for a fixed $y \in S_n$. Clearly it vanishes since $x \cdot y$ is 0 for half of x and 1 for the other half of x unless $y = 0$. Therefore the summation yields δ_{y0}. Now the state reduces to $|\psi_3\rangle = |0\rangle^{\otimes n}\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ and the measurement outcome of the first n qubits is always $00\ldots0$. Suppose $f(x)$ is balanced next. The probability amplitude of $|y = 0\rangle$ in $|\psi_3\rangle$ is proportional to $\sum_{x=0}^{2^n-1}(-1)^{f(x)}(-1)^{x\cdot y} = \sum_{x=0}^{2^n-1}(-1)^{f(x)} = 0$. Therefore the probability of obtaining measurement outcome $00\ldots0$ for the first n qubits vanishes. In conclusion, the function f is constant if we obtain $00\ldots0$ upon the meaurement of the first n qubits in the state $|\psi_3\rangle$ and it is balanced otherwise.

VI. DECOHERENCE

A quantum system is always in interaction with its environment. This interaction inevitably alter the state of the quantum system, which causes loss of information encoded in this system. The system under consideration is not a *closed* system when interaction with outside world is in action. We formulate the theory of *open* quantum system in this section by regarding the combined system of the quantum system and its environment as a closed system and subsequently trace out the environment degrees of freedom. Let ρ_S and ρ_E be the initial density matrices of the system and the environment, respectively. Even when the initial state is an uncorrelated state $\rho_S \otimes \rho_E$, the system-environment interaction entangles the total system so that the total state develops to an inseparable entangled state in general. Decoherence is a process in which environment causes various changes in the quantum system, which manifests itself as undesirable noise.

A. Open quantum system

Let us start our exposition with some mathematical background materials [1, 2, 24].

We deal with general quantum states described by density matrices. We are interested in a general evolution of a quantum system, which is described by a powerful tool called a quantum operation. One of the simplest quantum operations is a unitary time evolution of a closed system. Let ρ_S be a density matrix of a closed system at $t = 0$ and let $U(t)$ be the time evolution operator. Then the corresponding quantum map $\mathcal{E}$ is defined as

$$\mathcal{E}(\rho_S) = U(t)\rho_S U(t)^\dagger. \tag{64}$$

One of our primary aims in this section is to generalize this map to cases of open quantum systems.

1. Quantum operations and Kraus operators

Suppose a system of interest is coupled with its environment. We must specify the details of the environment and the coupling between the system and the environment to study the effect of the environment on the behavior of the system. Let H_S, H_E and H_{SE} be the system Hamiltonian, the environment Hamiltonian and their interaction Hamiltonian, respectively. We assume the system-environment interaction is weak enough so that this separation into the system and its environment makes sense. To avoid confusion, we often call the system of interest the principal system. The total Hamiltonian H_T is then

$$H_T = H_S + H_E + H_{SE}. \tag{65}$$

Correspondingly, we denote the system Hilbert space and the environment Hilbert space as $\mathcal{H}_S$ and $\mathcal{H}_E$, respectively, and the total Hilbert space as $\mathcal{H}_T = \mathcal{H}_S \otimes \mathcal{H}_E$. The condition of weak system-environment interaction may be lifted in some cases. Let us consider a qubit propagating through a noisy quantum channel, for example. "Propagating" does not necessarily mean propagating in space. The qubit may be spatially fixed and subject to time-dependent noise. When the noise is localized in space and time, the input and the output qubit states belong to a well defined Hilbert space $\mathcal{H}_S$ and the above separation of the Hamiltonian is perfectly acceptable even for strongly interacting cases. We consider, in the following, how the principal system state ρ_S at $t = 0$ evolves in time in the presence of its environment. A map which describes a general change of the state from ρ_S to $\mathcal{E}(\rho_S)$ is called a quantum operation. We have already noted that the unitary time evolution is an example of a quantum operation. Other quantum operations include state change associated with measurement and state change due to noise. The latter quantum map is our primary interest in this section.

The state of the total system is described by a density matrix ρ. Suppose ρ is uncorrelated initally at time $t = 0$,

$$\rho(0) = \rho_S \otimes \rho_E, \tag{66}$$

where ρ_S (ρ_E) is the initial density matrix of the principal system (environment). The total system is assumed to be closed and to evolve with a unitary matrix $U(t)$ as

$$\rho(t) = U(t)(\rho_S \otimes \rho_E)U(t)^\dagger. \tag{67}$$

Note that the resulting state is not a tensor product state in general. We are interested in extracting information on the state of the principal system at some later time $t > 0$.

Even under these circumstances, however, we may still define the system density matrix $\rho_S(t)$ by taking partial trace of $\rho(t)$ over the environment Hilbert space as

$$\rho_S(t) = \mathrm{tr}_E[U(t)(\rho_S \otimes \rho_E)U(t)^\dagger]. \tag{68}$$

We may forget about the environment by taking a trace over $\mathcal{H}_E$. This is an example of a quantum operation, $\mathcal{E}(\rho_S) = \rho_S(t)$. Let $\{|e_j\rangle\}$ be a basis of the system Hilbert space while $\{|\varepsilon_a\rangle\}$ be that of the environment Hilbert space. We may take the basis of $\mathcal{H}_T$ to be $\{|e_j\rangle \otimes |\varepsilon_a\rangle\}$. The initial density matrices may be written as $\rho_S = \sum_j p_j |e_j\rangle\langle e_j|$, $\rho_E = \sum_a r_a |\varepsilon_a\rangle\langle\varepsilon_a|$.

Action of the time evolution operator on a basis vector of $\mathcal{H}_T$ is explicitly written as

$$U(t)|e_j, \varepsilon_a\rangle = \sum_{k,b} U_{kb;ja}|e_k, \varepsilon_b\rangle, \tag{69}$$

where $|e_j, \varepsilon_a\rangle = |e_j\rangle \otimes |\varepsilon_a\rangle$ for example. Using this expression, the density matrix $\rho(t)$ is written as

$$\begin{aligned} U(t)(\rho_S \otimes \rho_E)U(t)^\dagger &= \sum_{j,a} p_j r_a U(t)|e_j, \varepsilon_a\rangle\langle e_j, \varepsilon_a|U(t)^\dagger \\ &= \sum_{j,a,k,b,l,c} p_j r_a U_{kb;ja}|e_k, \varepsilon_b\rangle\langle e_l, \varepsilon_c|U^*_{lc;ja}. \end{aligned} \tag{70}$$

The partial trace over $\mathcal{H}_E$ is carried out to yield

$$\begin{aligned} \rho_S(t) &= \mathrm{tr}_E[U(t)(\rho_S \otimes \rho_E)U(t)^\dagger] = \sum_{j,a,k,b,l} p_j r_a U_{kb;ja}|e_k\rangle\langle e_l|U^*_{lb;ja} \\ &= \sum_{j,a,b} p_j \left(\sum_k \sqrt{r_a} U_{kb;ja}|e_k\rangle\right)\left(\sum_l \sqrt{r_a}\langle e_l|U^*_{lb;ja}\right). \end{aligned} \tag{71}$$

To write down the quantum operation in a closed form, we assume the initial environment state is a pure state, which we take, without loss of generality, $\rho_E = |\varepsilon_0\rangle\langle\varepsilon_0|$. Even when ρ_E is a mixed state, we may always complement $\mathcal{H}_E$ with a fictitious Hilbert space to "purify" ρ_E, see § II G. With this assumption, $\rho_S(t)$ is written as

$$\begin{aligned} \rho_S(t) &= \mathrm{tr}_E[U(t)(\rho_S \otimes |\varepsilon_0\rangle\langle\varepsilon_0|)U(t)^\dagger] \\ &= \sum_a (I \otimes \langle\varepsilon_a|)U(t)(\rho_S \otimes |\varepsilon_0\rangle\langle\varepsilon_0|)|U(t)^\dagger(I \otimes |\varepsilon_a\rangle) \\ &= \sum_a (I \otimes \langle\varepsilon_a|)U(t)(I \otimes |\varepsilon_0\rangle)\rho_S(I \otimes \langle\varepsilon_0|)U(t)^\dagger(I \otimes |\varepsilon_a\rangle). \end{aligned}$$

We will drop $I\otimes$ from $I \otimes \langle\varepsilon_a|$ hereafter, whenever it does not cause confusion. Let us define the Kraus operator $E_a(t) : \mathcal{H}_S \to \mathcal{H}_S$ by

$$E_a(t) = \langle\varepsilon_a|U(t)|\varepsilon_0\rangle. \tag{72}$$

Then we may write

$$\mathcal{E}(\rho_S) = \rho_S(t) = \sum_a E_a(t)\rho_S E_a(t)^\dagger. \tag{73}$$

This is called the operator-sum representation (OSR) of a quantum operation $\mathcal{E}$. Note that $\{E_a\}$ satisfies the completeness relation

$$\left[\sum_a E_a(t)^\dagger E_a(t)\right]_{kl} = \left[\sum_a \langle\varepsilon_0|U(t)^\dagger|\varepsilon_a\rangle\langle\varepsilon_a|U(t)|\varepsilon_0\rangle\right]_{kl} = \delta_{kl}, \tag{74}$$

where I is the unit matrix in $\mathcal{H}_S$. This is equivalent with the trace-preserving property of $\mathcal{E}$ as $1 = \mathrm{tr}_S\rho_S(t) = \mathrm{tr}_S(\mathcal{E}(\rho_S)) = \mathrm{tr}_S\left(\sum_a E_a^\dagger E_a\rho_S\right)$ for any $\rho_S \in \mathcal{S}(\mathcal{H}_S)$. Completeness relation and trace-preserving property are satisfied since our total system is a closed system. A general quantum map does not necessarily satisfy these properties [25].

At this stage, it turns out to be useful to relax the condition that $U(t)$ be a time evolution operator. Instead, we assume U be any operator including an arbitrary unitary gate. Let us consider a two-qubit system on which the CNOT gate acts. Suppose the principal system is the control qubit while the environment is the target qubit. Then we find

$$E_0 = (I \otimes \langle 0|)U_{\mathrm{CNOT}}(I \otimes |0\rangle) = P_0,\ E_1 = (I \otimes \langle 1|)U_{\mathrm{CNOT}}(I \otimes |0\rangle) = P_1,$$

where $P_i = |i\rangle\langle i|$, and consequently

$$\mathcal{E}(\rho_S) = P_0\rho_S P_0 + P_1\rho_S P_1 = \rho_{00}P_0 + \rho_{11}P_1 = \begin{pmatrix} \rho_{00} & 0 \\ 0 & \rho_{11} \end{pmatrix}, \tag{75}$$

where $\rho_S = \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix}$. Unitarity condition may be relaxed when measurements are included as quantum operations, for example.

Tracing out the extra degrees of freedom makes it impossible to invert a quantum operation. Given an initial principal system state ρ_S, there are infinitely many U that yield the same $\mathcal{E}(\rho_S)$. Therefore even though it is possible to compose two quantum operations, the set of quantum operations is not a group but merely a semigroup. [82]

2. Operator-sum representation and noisy quantum channel

Operator-sum representation (OSR) introduced in the previous subsection seems to be rather abstract. Here we give an interpretation of OSR as a noisy quantum channel. Suppose we have a set of unitary matrices $\{U_a\}$ and a set of non-negative real numbers $\{p_a\}$ such that $\sum_a p_a = 1$. By choosing U_a randomly with probability p_a and applying it to ρ_S, we define the expectation value of the resulting density matrix as

$$\mathcal{M}(\rho_S) = \sum_a p_a U_a \rho_S U_a^\dagger, \tag{76}$$

which we call a mixing process [26]. This occurs when a flying qubit is sent through a noisy quantum channel which transforms the density matrix by U_a with probability p_a, for example. Note that no enviroment has been introduced in the above definition, and hence no partial trace is involved.

Now the correspondence between $\mathcal{E}(\rho_S)$ and $\mathcal{M}(\rho_S)$ should be clear. Let us define $E_a \equiv \sqrt{p_a}U_a$. Then Eq. (76) is rewritten as

$$\mathcal{M}(\rho_S) = \sum_a E_a \rho_S E_a^\dagger \tag{77}$$

and the equivalence has been shown. Operators E_a are identified with the Kraus operators. The system transforms, under the action of U_a, as

$$\rho_S \to E_a\rho_S E_a^\dagger/\mathrm{tr}\left(E_a\rho_S E_a^\dagger\right). \tag{78}$$

Conversely, given a noisy quantum channel $\{U_a, p_a\}$ we may introduce an "environment" with the Hilbert space $\mathcal{H}_E$ as follows. Let $\mathcal{H}_E = \mathrm{Span}(|\varepsilon_a\rangle)$ be a Hilbert space with the dimension equal to the number of the unitary matrices $\{U_a\}$, where $\{|\varepsilon_a\rangle\}$ is an orthonormal basis. Define formally the environment density matrix $\rho_E = \sum_a p_a|\varepsilon_a\rangle\langle\varepsilon_a|$ and

$$U \equiv \sum_a U_a \otimes |\varepsilon_a\rangle\langle\varepsilon_a| \tag{79}$$

which acts on $\mathcal{H}_S \otimes \mathcal{H}_E$. It is easily verified from the orthonormality of $\{|\varepsilon_a\rangle\}$ that U is indeed a unitary matrix. Partial trace over $\mathcal{H}_E$ then yields

$$\begin{aligned}\mathcal{E}(\rho_S) &= \mathrm{tr}_E[U(\rho_S \otimes \rho_E)U^\dagger] \\ &= \sum_a (I \otimes \langle\varepsilon_a|)\left(\sum_b U_b \otimes |\varepsilon_b\rangle\langle\varepsilon_b|\right)\left(\rho_S \otimes \sum_c p_c|\varepsilon_c\rangle\langle\varepsilon_c|\right) \\ &\quad \times \left(\sum_d U_d \otimes |\varepsilon_d\rangle\langle\varepsilon_d|\right)(I \otimes |\varepsilon_a\rangle) \\ &= \sum_a p_a U_a \rho_S U_a^\dagger = \mathcal{M}(\rho_S) \end{aligned} \tag{80}$$

showing that the mixing process is also decribed by a quantum operation with a fictitious environment.

3. Completely positive maps

All linear operators we have encountered so far map vectors to vectors. A quantum operation maps a density matrix to another density matrix linearly.[83] A linear operator of this kind is called a superoperator. Let Λ be a superoperator acting on the system density matrices, $\Lambda : \mathcal{S}(\mathcal{H}_S) \to \mathcal{S}(\mathcal{H}_S)$. The operator Λ is easily extended to an operator acting on $\mathcal{H}_T$ by $\Lambda_T = \Lambda \otimes I_E$, which acts on $\mathcal{S}(\mathcal{H}_S \otimes \mathcal{H}_E)$. Note, however, that Λ_T is not necessarily a map $\mathcal{S}(\mathcal{H}_T) \to \mathcal{S}(\mathcal{H}_T)$. It may happen that $\Lambda_T(\rho)$ is not a density matrix any more. We have already encountered this situation when we have introduced partial transpose operation in § II G. Let $\mathcal{H}_T = \mathcal{H}_1 \otimes \mathcal{H}_2$ be a two-qubit Hilbert space, where $\mathcal{H}_k$ is the kth qubit Hilbert space. It is clear that the transpose operation $\Lambda_t : \rho_1 \to \rho_1^t$ on a single-qubit state ρ_1 preserves the density matrix properties. For a two-qubit density matrix ρ_{12}, however, this is not always the case. In fact, we have seen that $\Lambda_t \otimes I : \rho_{12} \to \rho_{12}^{\mathrm{pt}}$ defined by Eq. (20) maps a density matrix to a matrix which is not a density matrix when ρ_{12} is inseparable.

A map Λ which maps a positive operator acting on $\mathcal{H}_S$ to another positive operator on $\mathcal{H}_S$ is said to be positive. Moreover, it is called a completely positive map (CP map), if its extension $\Lambda_T = \Lambda \otimes I_n$ remains a positive operator for an arbitrary $n \in \mathbb{N}$.

Theorem VI.1 *A linear map Λ is CP if and only if there exists a set of operators $\{E_a\}$ such that $\Lambda(\rho_S)$ can be written as*

$$\Lambda(\rho_S) = \sum_a E_a \rho_S E_a^\dagger. \tag{81}$$

We require not only that Λ be CP but also $\Lambda(\rho)$ be a density matrix:

$$\mathrm{tr}\,\Lambda(\rho_S) = \mathrm{tr}\left(\sum_a E_a \rho E_a^\dagger\right) = \mathrm{tr}\left(\sum_a E_a^\dagger E_a \rho\right) = 1. \tag{82}$$

This condition is satisfied for any ρ if and only if

$$\sum_a E_a^\dagger E_a = I_S. \tag{83}$$

Therefore, any quantum operation obtained by tracing out the environment degrees of freedom is CP and preserves trace.

B. Measurements as quantum operations

We have already seen that a unitary evoluation $\rho_S \to U\rho_S U^\dagger$ and a mixing process $\rho_S \to \sum_i p_i U_i \rho_S U_i^\dagger$ are quantum operations. We will see further examples of quantum operations in this section and the next. This section deals with measurements as quantum operations.

1. Projective measurements

Suppose we measure an observable $A = \sum_i \lambda_i P_i$, where $P_i = |\lambda_i\rangle\langle\lambda_i|$ is the projection operator corresponding to the eigenvector $|\lambda_i\rangle$. We have seen in Chapter 2 that the probability of observing λ_i upon a measurement of A in a state ρ is

$$p(i) = \langle\lambda_i|\rho|\lambda_i\rangle = \mathrm{tr}\,(P_i\rho) \tag{84}$$

and the state changes as $\rho \to P_i\rho P_i/p(i)$. This process happens with a probability $p(i)$. Thus we may regard the measurement process as a quantum operation

$$\rho_S \to \sum_i p(i)\frac{P_i\rho_S P_i}{p(i)} = \sum_i P_i\rho_S P_i, \tag{85}$$

where the set $\{P_i\}$ satisifes the completeness relation $\sum_i P_i P_i^\dagger = I$.

The projective measurement is a special case of a quantum operation in which the Kraus operators are $E_i = P_i$.

2. POVM

We have been concerned with projective measurements so far. However, it should be noted that they are not unique type of measurements. Here we will deal with the most general framework of measurement and show that it is a quantum operation.

Suppose a system and an environment, prepared initially in a product state $|\psi\rangle|e_0\rangle$, are acted by a unitary operator U, which applies an operator M_i on the system and, at the same time, put the environment to $|e_i\rangle$ for various i. It is written explicitly as

$$|\Psi\rangle = U|\psi\rangle|e_0\rangle = \sum_i M_i|\psi\rangle|e_i\rangle. \tag{86}$$

The system and its environment are correlated in this way. This state must satisfy the normalization condition since U is unitary; $\langle\psi|\langle e_0|U^\dagger U|\psi\rangle|e_0\rangle = \sum_{i,j}\langle\psi|\langle e_i|M_i^\dagger M_j\otimes I|\psi\rangle|e_j\rangle = \langle\psi|\sum_i M_i^\dagger M_i|\psi\rangle = 1$. Since $|\psi\rangle$ is arbitrary, we must have

$$\sum_i M_i^\dagger M_i = I_S, \tag{87}$$

where I_S is the unit matrix acting on the system Hilbert space $\mathcal{H}_S$. Operators $\{M_i^\dagger M_i\}$ are said to form a POVM (positive operator-valued measure).

Suppose we measure the environment with a measurement operator

$$O = I_S\otimes\sum_i \lambda_i|e_i\rangle\langle e_i| = \sum_i \lambda_i\left(I_S\otimes|e_i\rangle\langle e_i|\right).$$

We obtain a measurement outcome λ_k with a probability

$$\begin{aligned} p(k) &= \langle\Psi|(I_S\otimes|e_k\rangle\langle e_k|)|\Psi\rangle \\ &= \sum_{i,j}\langle\psi|\langle e_i|M_i^\dagger(I_S\otimes|e_k\rangle\langle e_k|)M_j|\psi\rangle|e_j\rangle = \langle\psi|M_k^\dagger M_k|\psi\rangle, \end{aligned} \tag{88}$$

where $|\Psi\rangle = U|\psi\rangle|e_0\rangle$. The combined system immediately after the measurement is

$$\begin{aligned} \frac{1}{\sqrt{p(k)}}(I_S\otimes|e_k\rangle\langle e_k|)U|\psi\rangle|e_0\rangle &= \frac{1}{\sqrt{p(k)}}(I_S\otimes|e_k\rangle\langle e_k|)\sum_i M_i|\psi\rangle|e_i\rangle \\ &= \frac{1}{\sqrt{p(k)}}M_k|\psi\rangle|e_k\rangle. \end{aligned} \tag{89}$$

Let $\rho_S = \sum_i p_i|\psi_i\rangle\langle\psi_i|$ be an arbitrary density matrix of the principal system. It follows from the above observation for a pure state $|\psi\rangle\langle\psi|$ that the reduced density matrix immediately after the measurement is

$$\sum_k p(k)\frac{M_k\rho_S M_k^\dagger}{p(k)} = \sum_k M_k\rho_S M_k^\dagger. \tag{90}$$

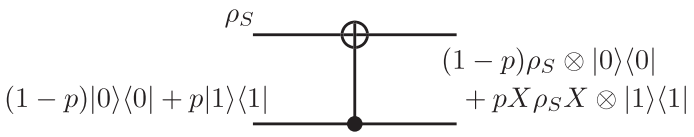


FIG. 5: Quantum circuit modelling a bit-flip channel. The gate is the inverted CNOT gate $I \otimes |0\rangle\langle 0| + \sigma_x \otimes |1\rangle\langle 1|$.

This shows that POVM measurement is a quantum operation in which the Kraus operators are given by the generalized measurement operators $\{M_i\}$. The projective measurement is a special class of POVM, in which $\{M_i\}$ are the projective operators.

C. Examples

Now we examine several important examples which have relevance in quantum information theory. Decoherence appears as an error in quantum information processing. The next chapter is devoted to strategies to fight against errors introduced in this section.

1. Bit-flip channel

Consider a closed two-qubit system with a Hilbert space $\mathbb{C}^2 \otimes \mathbb{C}^2$. We call the first qubit the "(principal) system" while the second qubit the "environment". A bit-flip channel is defined by a quantum operation

$$\mathcal{E}(\rho_S) = (1-p)\rho_S + p\sigma_x\rho_S\sigma_x,\ 0 \le p \le 1. \tag{91}$$

The input ρ_S is bit-flipped with a probability p while it remains in its input state with a probability $1-p$. The Kraus operators are read off as

$$E_0 = \sqrt{1-p}I,\ E_1 = \sqrt{p}\sigma_x. \tag{92}$$

The circuit depicted in Fig. 5 models the bit-flip channel provided that the second qubit is in a mixed state $(1-p)|0\rangle\langle 0| + p|1\rangle\langle 1|$. The circuit is nothing but the inverted CNOT gate $V = I \otimes |0\rangle\langle 0| + \sigma_x \otimes |1\rangle\langle 1|$. The output of this circuit is

$$\begin{aligned} & V\left(\rho_S \otimes [(1-p)|0\rangle\langle 0| + p|1\rangle\langle 1|]\right)V^\dagger \\ &= (1-p)\rho_S \otimes |0\rangle\langle 0| + p\sigma_x\rho_S\sigma_x|1\rangle\langle 1|, \end{aligned} \tag{93}$$

from which we obtain

$$\mathcal{E}(\rho_S) = (1-p)\rho_S + p\sigma_x\rho_S\sigma_x \tag{94}$$

after tracing over the environment Hilbert space.

The choice of the second qubit input state is far from unique and so is the choice of the circuit. Suppose the initial state of the environment is a pure state $|\psi_E\rangle = \sqrt{1-p}|0\rangle + \sqrt{p}|1\rangle$, for example. Then the output of the circuit in Fig. 5 is

$$\mathcal{E}(\rho_S) = \mathrm{tr}_E[V\rho_S \otimes |\psi_E\rangle\langle\psi_E|V^\dagger] = (1-p)\rho_S + p\sigma_x\rho_S\sigma_x, \tag{95}$$

producing the same result as before.

Let us see what transformation this quantum operation brings about in ρ_S. We parametrize ρ_S using the Bloch vector as

$$\rho_S = \frac{1}{2}\left(I + \sum_{k=x,y,z} c_k\sigma_k\right), \quad (c_k \in \mathbb{R}) \tag{96}$$

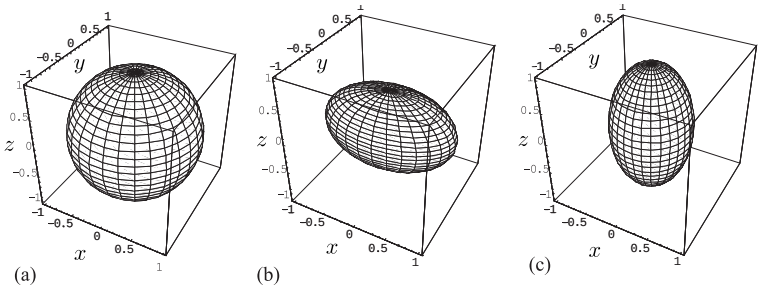


FIG. 6: Bloch sphere of the input state ρ_S (a) and output states of (b) bit-flip channel and (c) phase-flip channel. The probability $p = 0.2$ is common to both channels.

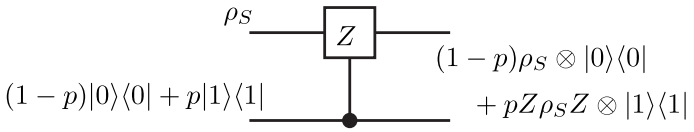


FIG. 7: Quantum circuit modelling a phase-flip channel. The gate is the inverted controlled-σ_z gate.

where $\sum_k c_k^2 \leq 1$. We obtain

$$\begin{aligned}\mathcal{E}(\rho_S) &= (1-p)\rho_S + p\sigma_x\rho_S\sigma_x \\ &= \frac{1-p}{2}(I + c_x\sigma_x + c_y\sigma_y + c_z\sigma_z) + \frac{p}{2}(I + c_x\sigma_x - c_y\sigma_y - c_z\sigma_z) \\ &= \frac{1}{2}\begin{pmatrix} 1+(1-2p)c_z & c_x - i(1-2p)c_y \\ c_x + i(1-2p)c_y & 1-(1-2p)c_z \end{pmatrix}. \end{aligned} \tag{97}$$

Observe that the radius of the Bloch sphere is reduced along the y- and the z-axes so that the radius in these directions is $|1-2p|$. Equation (97) shows that the quantum operation has produced a mixture of the Bloch vector states (c_x, c_y, c_z) and $(c_x, -c_y, -c_z)$ with weights $1-p$ and p respectively. Figure 6 (a) shows the Bloch sphere which represents the input qubit states. The Bloch sphere shrinks along the y- and z-axes, which results in the ellipsoid shown in Fig. 6 (b).

2. Phase-flip channel

Consider again a closed two-qubit system with the "(principal) system" and its "environment".
The phase-flip channel is defined by a quantum operation

$$\mathcal{E}(\rho_S) = (1-p)\rho_S + p\sigma_z\rho_S\sigma_z,\ 0 \leq p \leq 1. \tag{98}$$

The input ρ_S is phase-flipped ($|0\rangle \mapsto |0\rangle$ and $|1\rangle \mapsto -|1\rangle$) with a probability p while it remains in its input state with a probability $1-p$. The corresponding Kraus operators are

$$E_0 = \sqrt{1-p}I,\ E_1 = \sqrt{p}\sigma_z. \tag{99}$$

A quantum circuit which models the phase-flip channel is shown in Fig. 7. Let ρ_S be the first qubit input state while $(1-p)|0\rangle\langle 0| + p|1\rangle\langle 1|$ be the second qubit input state. The circuit is the inverted controlled-σ_z gate

$$V = I \otimes |0\rangle\langle 0| + \sigma_z \otimes |1\rangle\langle 1|.$$

The output of this circuit is

$$\begin{aligned} &V\left(\rho_S \otimes [(1-p)|0\rangle\langle 0| + p|1\rangle\langle 1|]\right)V^\dagger \\ &\quad = (1-p)\rho_S \otimes |0\rangle\langle 0| + p\sigma_z\rho_S\sigma_z \otimes |1\rangle\langle 1|, \end{aligned} \tag{100}$$

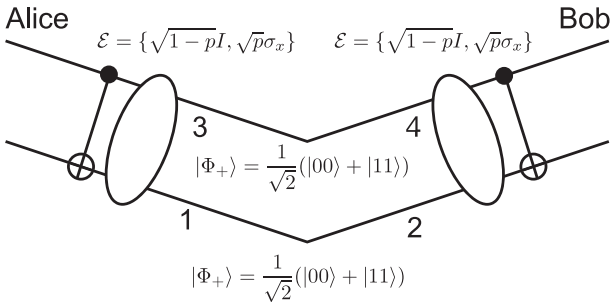


FIG. 8: Entanglement distillation of the second kind. Alice and Bob share a Bell state $|\Phi_+\rangle$ with a good precision if $p \ll 1$ when their measurement outcomes of the third and the fourth qubits are 00 or 11.

from which we obtain

$$\mathcal{E}(\rho_S) = (1-p)\rho_S + p\sigma_z\rho_S\sigma_z. \tag{101}$$

The second qubit input state may be a pure state

$$|\psi_E\rangle = \sqrt{1-p}|0\rangle + \sqrt{p}|1\rangle, \tag{102}$$

for example. Then we find

$$\mathcal{E}(\rho_S) = \mathrm{tr}_E[V\rho_S \otimes |\psi_E\rangle\langle\psi_E|V^\dagger] = E_0\rho_S E_0^\dagger + E_1\rho_S E_1^\dagger, \tag{103}$$

where the Kraus operators are

$$E_0 = \langle 0|V|\psi_E\rangle = \sqrt{1-p}I,\ E_1 = \langle 1|V|\psi_E\rangle = \sqrt{p}\sigma_z. \tag{104}$$

Let us work out the transformation this quantum operation brings about to ρ_S. We parametrize ρ_S using the Bloch vector as before. We obtain

$$\begin{aligned}
\mathcal{E}(\rho_S) &= (1-p)\rho_S + p\sigma_z\rho_S\sigma_z \\
&= \frac{1-p}{2}(I + c_x\sigma_x + c_y\sigma_y + c_z\sigma_z) + \frac{p}{2}(I - c_x\sigma_x - c_y\sigma_y + c_z\sigma_z) \\
&= \frac{1}{2}\begin{pmatrix} 1+c_z & (1-2p)(-c_x - ic_y) \\ (1-2p)(c_x + ic_y) & 1-c_z \end{pmatrix}.
\end{aligned} \tag{105}$$

Observe that the off-diagonal components decay while the diagonal components remain the same. Equation (105) shows that the quantum operation has produced a mixture of the Bloch vector states (c_x, c_y, c_z) and $(-c_x, -c_y, c_z)$ with weights $1-p$ and p respectively. The initial state has a definite phase $\phi = \tan^{-1}(c_y/c_x)$ in the off-diagonal components. The phase after the quantum operaition is applied is a mixture of states with ϕ and $\phi + \pi$. This process is called the phase relaxation process, or the T_2 process in the context of NMR. The radius of the Bloch sphere is reduced along the x- and the y-axes as $1 \to |1-2p|$. Figure 6 (c) shows the effect of the phase-flip channel on the Bloch sphere for $p = 0.2$.

Other examples will be found in [1, 2].

D. Entenglement Distillation II

The second entanglement distillation protocol recovers the EPR state $|\Phi_+\rangle$ from a pair of the EPR states on which noisy channels are applied, see Fig. 8. Let $\mathcal{E} = \{\sqrt{1-p}I, \sqrt{p}\sigma_x\}$ be the set of error operators describing to the channel. It is assumed that p is a small positive number. The initial state is

$$|\Psi_1\rangle = \frac{1}{2}(|00\rangle|00\rangle + |00\rangle|11\rangle + |11\rangle|00\rangle + |11\rangle|11\rangle).$$

The state after the error operators are applied is mixed as

$$\begin{aligned}\rho_2 &= (p^2+(1-p)^2)^2|\psi_{00}\rangle\langle\psi_{00}| + (p^2+(1-p)^2)(2p(1-p))|\psi_{0x}\rangle\langle\psi_{0x}| \\ &\quad +(2p(1-p))(p^2+(1-p)^2)|\psi_{x0}\rangle\langle\psi_{x0}| + (2p(1-p))^2|\psi_{xx}\rangle\langle\psi_{xx}|,\end{aligned}$$

where

$$\begin{aligned}|\psi_{00}\rangle &= \frac{1}{2}(|00\rangle|00\rangle + |00\rangle|11\rangle + |11\rangle|00\rangle + |11\rangle|11\rangle), \\ |\psi_{0x}\rangle &= \frac{1}{2}(|00\rangle|01\rangle + |00\rangle|10\rangle + |11\rangle|01\rangle + |11\rangle|10\rangle), \\ |\psi_{x0}\rangle &= \frac{1}{2}(|01\rangle|00\rangle + |01\rangle|11\rangle + |10\rangle|00\rangle + |10\rangle|11\rangle), \\ |\psi_{xx}\rangle &= \frac{1}{2}(|01\rangle|01\rangle + |01\rangle|10\rangle + |10\rangle|01\rangle + |10\rangle|10\rangle).\end{aligned}$$

Then CNOT gates are applied as before as shown in Fig. 8. The resulting state is

$$\begin{aligned}\rho_3 &= (p^2+(1-p)^2)^2|\tilde{\psi}_{00}\rangle\langle\tilde{\psi}_{00}| + (p^2+(1-p)^2)(2p(1-p))|\tilde{\psi}_{0x}\rangle\langle\tilde{\psi}_{0x}| \\ &\quad +(2p(1-p))(p^2+(1-p)^2)|\tilde{\psi}_{x0}\rangle\langle\tilde{\psi}_{x0}| + (2p(1-p))^2|\tilde{\psi}_{xx}\rangle\langle\tilde{\psi}_{xx}|,\end{aligned}$$

where

$$\begin{aligned}|\tilde{\psi}_{00}\rangle &= \frac{1}{2}(|00\rangle|00\rangle + |00\rangle|11\rangle + |11\rangle|11\rangle + |11\rangle|00\rangle), \\ |\tilde{\psi}_{0x}\rangle &= \frac{1}{2}(|00\rangle|01\rangle + |00\rangle|10\rangle + |11\rangle|10\rangle + |11\rangle|01\rangle), \\ |\tilde{\psi}_{x0}\rangle &= \frac{1}{2}(|01\rangle|01\rangle + |01\rangle|10\rangle + |10\rangle|10\rangle + |10\rangle|01\rangle), \\ |\tilde{\psi}_{xx}\rangle &= \frac{1}{2}(|01\rangle|00\rangle + |01\rangle|11\rangle + |10\rangle|11\rangle + |10\rangle|00\rangle).\end{aligned}$$

Now they measure the third and the fourth qubits and exchange there outcomes using a classical communication. Suppose their readings are 00 or 11. Then, with probability $(p^2+(1-p)^2)^2$, the first and the second qubits are in the state $|\Phi_+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. This probability $\sim 1-4p$ is close to 1 for $p \ll 1$. There is a small probability $(2p(1-p))^2 \sim 4p^2$ with which the resulting state of the first and the second qubits is $|\Psi_+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle)$ even though the readouts of the third and the fourth qubits are 00 or 11.

VII. QUANTUM ERROR CORRECTING CODES

A. Introduction

It has been shown in the previous chapter that interactions between a quantum system with environment cause undesirable changes in the state of the quantum system. In the case of qubits, they appear as bit-flip and phase-flip errors, for example. To reduce such errors, we must implement some sort of error correcting mechanism in the algorithm.

Before we introduce quantum error correcting codes, we have a brief look at the simplest version of error correcting code in classical bits. Suppose we transmit a serise of 0's and 1's through a noisy classical channel. Each bit is assumed to flip independently with a probability p. Thus a bit 0 sent through the channel will be received as 0 with probability $1-p$ and as 1 with probability p. To reduce channel errors, we may invoke to majority vote. Namely, we encode logical 0 by 000 and 1 by 111, for example. When 000 is sent through this channel, it will be received as 000 with probability $(1-p)^3$, as 100, 010 or 001 with probability $3p(1-p)^2$, as 011, 101 or 110 with probability $3p^2(1-p)$ and finally as 111 with probability p^3. By taking the majority vote, we correctly reproduce the desired result 0 with probability $p_0 = (1-p)^3 + 3p(1-p)^2 = (1-p)^2(1+2p)$ while fails with probability $p_1 = 3p^2(1-p) + p^3 = (3-2p)p^2$. We obtain $p_0 \gg p_1$ for sufficiently small $p \geq 0$. In fact, we find $p_0 = 0.972$ and $p_1 = 0.028$ for $p = 0.1$. The success probability p_0 increases as p approaches to 0, or alternatively, if we use more bits to encode 0 or 1.

This method cannot be applicable to qubits, however, due to no-cloning theorem. We have to somehow think out the way to overcome this theorem.

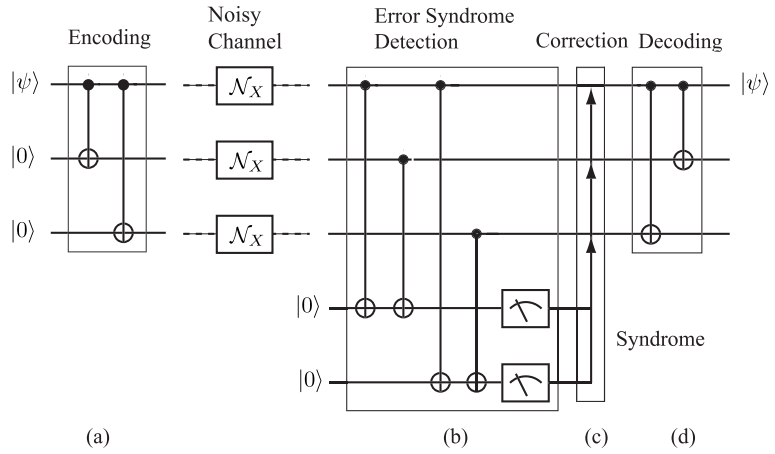


FIG. 9: Quantum circuits to (a) encode, (b) detect bit-flip error syndrome, (c) make correction to a relevant qubit and (d) decode. The gate $\mathcal{N}_X$ stands for the bit-flip noise.

B. Three-qubit bit-flip code: the simplest example

It is instructive to introduce a simple example of quantum error correcting codes (QECC). We closely follow Steane [30] here.

1. *Bit-flip QECC*

Suppose Alice wants to send a qubit or a serise of qubits to Bob through a noisy quantum channel. Let $|\psi\rangle = a|0\rangle + b|1\rangle$ be the state she wants to send. If she is to transmit a serise of qubits, she sends them one by one and the following argument applies to each of the qubits. Let p be the probability with which a qubit is flipped and we assume there are no other types of errors in the channel. In other words, the operator X is applied to the qubit with probability p and consequently the state is mapped to

$$|\psi\rangle \to |\psi'\rangle = X|\psi\rangle = a|1\rangle + b|0\rangle. \tag{106}$$

We have already seen in the previous section that this channel is described by a quantum operation (91).

2. *Encoding*

To reduce the error probability, we want to mimic somehow the classical counterpart without using a clone machine. Let us recall that the action of a CNOT gate is CNOT : $|j0\rangle \to |jj\rangle$, $j \in \{0, 1\}$ and therefore it duplicates the control bit $j \in \{0, 1\}$ when the target bit is initially set to $|0\rangle$. We use this fact to *triplicate* the basis vectors as

$$|\psi\rangle|00\rangle = (a|0\rangle + b|1\rangle)|00\rangle \to |\psi\rangle_E = a|000\rangle + b|111\rangle, \tag{107}$$

where $|\psi\rangle_E$ denotes the encoded state. The state $|\psi\rangle_E$ is called the logical qubit while each constituent qubit is called the physical qubit. We borrow terminologies from classical error correcting code (ECC) and call the set

$$C = \{a|000\rangle + b|111\rangle | a, b \in \mathbb{C}, |a|^2 + |b|^2 = 1\} \tag{108}$$

the code and each member of C a codeword. It is important to note that the state $|\psi\rangle$ is not triplicated but only the basis vectors are triplicated. This redundancy makes it possible to detect errors in $|\psi\rangle_E$ and correct them as we see below.

A quantum circuit which implements the encoding (107) is easily found from our experience in CNOT gate. Let us consider the circuit shown in Fig. 9 (a) whose input state is $|\psi\rangle|00\rangle$. It is immediately found that the output of this circuit is $|\psi\rangle_E = a|000\rangle + b|111\rangle$ as promised.

TABLE I: State Bob receives and the probability which this may happen.

State Bob receives	Probability
$a\|000\rangle + b\|111\rangle$	$(1-p)^3$
$a\|100\rangle + b\|011\rangle$	$p(1-p)^2$
$a\|010\rangle + b\|101\rangle$	$p(1-p)^2$
$a\|001\rangle + b\|110\rangle$	$p(1-p)^2$
$a\|110\rangle + b\|001\rangle$	$p^2(1-p)$
$a\|101\rangle + b\|010\rangle$	$p^2(1-p)$
$a\|011\rangle + b\|100\rangle$	$p^2(1-p)$
$a\|111\rangle + b\|000\rangle$	p^3

TABLE II: States after error extraction is made and the probabilities with which these states are produced.

State after error syndrome extraction	Probability
$(a\|000\rangle + b\|111\rangle)\|00\rangle$	$(1-p)^3$
$(a\|100\rangle + b\|011\rangle)\|11\rangle$	$p(1-p)^2$
$(a\|010\rangle + b\|101\rangle)\|10\rangle$	$p(1-p)^2$
$(a\|001\rangle + b\|110\rangle)\|01\rangle$	$p(1-p)^2$
$(a\|110\rangle + b\|001\rangle)\|01\rangle$	$p^2(1-p)$
$(a\|101\rangle + b\|010\rangle)\|10\rangle$	$p^2(1-p)$
$(a\|011\rangle + b\|100\rangle)\|11\rangle$	$p^2(1-p)$
$(a\|111\rangle + b\|000\rangle)\|00\rangle$	p^3

3. Transmission

Now the state $|\psi\rangle_E$ is sent through a quantum channel which introduces bit-flip error with a rate p for each qubit independently. We assume p is sufficiently small so that not many errors occur during qubit transmission. The received state depends on in which physical qubit(s) the bit-flip error occurred. Table I lists possible received states and the probabilities with which these states are received.

4. Error syndrome dectection and correction

Now Bob has to extract from the received state which error occurred during qubits transmission. For this purpose, Bob prepares two ancillary qubits in the state $|00\rangle$ as depicted in Fig. 9 (b) and apply four CNOT operations whose control bits are the encoded qubits while the target qubits are Bob's two ancillary qubits. Let $|x_1x_2x_3\rangle$ be a basis vectors Bob has received and let A (B) be the output state of the first (second) ancilla qubit. It is seen from Fig. 9 (b) that $A = x_1 \oplus x_2$ and $B = x_1 \oplus x_3$. Let $a|100\rangle + b|011\rangle$ be the received logical qubit for example. Note that the first qubit state in both of the basis vectors is different from the second and the third qubit states. These difference are detected by the pairs of CNOT gates in Fig. 9 (b). The error extracting sequence transforms the anicillary qubits as

$$(a|100\rangle + b|011\rangle)|00\rangle \to a|10011\rangle + b|01111\rangle = (a|100\rangle + b|011\rangle)|11\rangle.$$

Both of the ancillary qubits are flipped since $x_1 \oplus x_2 = x_1 \oplus x_3 = 1$ for both $|100\rangle$ and $|011\rangle$. It is important to realize that (i) the syndrome is independent of a and b and (ii) the received state $a|100\rangle + b|011\rangle$ remains intact; we have detected an error without measuring the received state! These features are common to all QECC.

We list the result of other cases in Table II. Note that among eight possible states, there are exactly two states with the same ancilla state. Does it mean this error extraction scheme does not work? Now let us compare the probabilities associated with the same ancillary state. When the ancillary state is $|10\rangle$, for example, there are two possible reseived states $a|010\rangle + b|101\rangle$ and $a|101\rangle + b|010\rangle$. Note that the former is received with probability $p(1-p)^2$ while that latter with $p^2(1-p)$. Therefore the latter probability is negligible compared to the former for sufficiently small p.

It is instructive to visualize what errors do to the encoded basis vectors. Consider a cube with the unit length. The vertices of the cube have coordinates (i, j, k) where $i, j, k \in \{0, 1\}$. We assign a vector $|ijk\rangle$ to the vertex (i, j, k), under which the vectors $|000\rangle$ and $|111\rangle$ correspond to diagonally separated vertices. An action of X_i, the operator $X = \sigma_x$ acting on the ith qubit, sends these basis vectors to the nearest neighbor vetices, which differ from the correct

basis vectors in the ith position. The intersection of the sets of vectors obtained by a single action of X_i on $|000\rangle$ and $|111\rangle$ is an empty set. Therefore an action of a single error operator X can be corrected with no ambiguity.

Now Bob measures his ancillary qubits and obtains two bits of classical information. The set of two bits is called the (error) syndrome and it tells Bob in which physical qubit the error occurred during transmission. Bob applies correcting procedure to the received state according to the error sydrome he has obtained. Ignoring extra error states with small probabilities, we immediately find that the following action must be taken:

error syndrome	correction to be made
00	identity operation (nothing is required)
01	apply σ_x to the third qubit
10	apply σ_x to the second qubit
11	apply σ_x to the first qubit

Suppose the syndrome is 01, for example. The state Bob received is likely to be $a|001\rangle + b|110\rangle$. Bob recovers the initial state Alice has sent by applying $I \otimes I \otimes \sigma_x$ on the received state:

$$(I \otimes I \otimes \sigma_x)(a|001\rangle + b|110\rangle) = a|000\rangle + b|111\rangle.$$

If Bob receives the state $a|110\rangle + b|001\rangle$, unfortunately, he will obtain

$$(I \otimes I \otimes \sigma_x)(a|110\rangle + b|001\rangle) = a|111\rangle + b|000\rangle.$$

In fact, for any error sydrome, Bob obtaines either $a|000\rangle + b|111\rangle$ or $a|111\rangle + b|000\rangle$. The latter case occurs if and only if more than one qubit are flipped, and hence it is less likely to happen for sufficiently small error rate p. The probability with which multiple error occurs is found from Table I as

$$P(\text{error}) = 3p^2(1-p) + p^3 = 3p^2 - 2p^3. \tag{109}$$

This error rate is less than p if $p < 1/2$. In contrast, success probability has been enhanced from $1-p$ to $1-P(\text{error}) = 1-3p^2+2p^3$. Let $p = 0.1$, for example. Then the error rate is lowered to $P(\text{error}) = 0.028$, while the success probability is enhanced from 0.9 to 0.972.

5. Decoding

Now that Bob has corrected an error, what is left for him is to decode the encoded state. This is nothing but the inverse transformation of the encoding (107). It can be seen from Fig. 9 (d) that

$$\text{CNOT}_{12}\text{CNOT}_{13}(a|000\rangle + b|111\rangle) = a|000\rangle + b|100\rangle = (a|0\rangle + b|1\rangle)|00\rangle. \tag{110}$$

6. Miracle of entanglement

This example, albeit simple, contains almost all fundamental ingredients of QECC. We prepare some redundant qubits which somehow "triplicate" the original qubit state to be sent without violating no-cloning theorem. Then the encoded qubits are sent through a noisy channel, which causes a bit-flip in at most one of the qubits. The received state, which may be subject to an error, is then entangled with ancillary qubits, whose state reflects the error which occurred during the state transmission. This results in an entangled state

$$\sum_k |\text{A bit-flip error in the } k\text{th qubit}\rangle \otimes |\text{corresponding error syndrome}\rangle. \tag{111}$$

The wave function, upon the measurement of the ancillary qubits, collapses to a state with a bit-flip error corresponding to the observed error syndrome. In a sense, syndrome measurement singles out a particular error state which produces the observed syndrome.

Once syndrome is found, it is an easy task to transform the received state back to the original state. Note that everything is done without knowing what the origial state is.

7. Continuous rotations

We have considered noise X so far. Suppose noise in the channel is characterized by a contiuous paramter α as

$$U_\alpha = e^{i\alpha X} = \cos\alpha I + iX\sin\alpha, \tag{112}$$

which maps a state $|\psi\rangle$ to

$$U_\alpha|\psi\rangle = \cos\alpha|\psi\rangle + i\sin\alpha X|\psi\rangle. \tag{113}$$

Suppose U_α acts on the first qubit, for example. Bob then receives

$$\begin{aligned}&(U_\alpha\otimes I\otimes I)(a|000\rangle + b|111\rangle)\\ &= \cos\alpha(a|000\rangle + b|111\rangle) + i\sin\alpha(a|100\rangle + b|011\rangle).\end{aligned}$$

The output of the error syndrome detection circuit, before the syndrome measurement is made, is an entangled state

$$\cos\alpha(a|000\rangle + b|111\rangle)|00\rangle + i\sin\alpha(a|100\rangle + b|111\rangle)|11\rangle, \tag{114}$$

see Table II. Measurement of the error syndrome yields either 00 or 11. In the former case the state collapses to $|\psi\rangle = a|000\rangle + b|111\rangle$ and this happens with a probability $\cos^2\alpha$. In the latter case, on the other hand, the received state collapses to $X|\psi\rangle = a|100\rangle + b|011\rangle$ and this happens with a probability $\sin^2\alpha$. Bob applies I (X) to the first qubit to correct the error when the syndrome readout is 00 (11).

It is clear that error U_α may act on the second or the third qubit. Continuous rotation U_α for any α may be corrected in this way. In general, linearity of a quantum circuit guarantees that any QECC, which corrects the bit-flip error X, corrects continuous error U_α.

VIII. DIVINCENZO CRITERIA

We have learned so far that information may be encoded and processed in a quantum-mechanical way. This new discipline called quantum information processing (QIP) is expected to solve a certain class of problems that current digital computers cannot solve in a practical time scale. Although a small scale quantum information processor is already available commercially, physical realization of large scale quantum information processors is still beyond the scope of our currently available technology.

A quantum computer should have at least $10^2 \sim 10^3$ qubits to be able to execute algorithms that are more efficient than their classical counterparts. DiVincenzo proposed necessary conditions, so-called the *DiVincenzo criteria* that any physical system has to fulfill to be a candidate for a viable quantum computer [38]. In the next section, we outline these conditions as well as two additional criteria for networkability.

A. DiVincenzo criteria

In his influential article [38], DiVincenzo proposed five criteria that any physical system must satisfy to be a viable quantum computer. We summarize the relevant parts of these criteria in this section.

1. *A scalable physical system with well characterized qubits.*

   To begin with, we need a quantum register made of many qubits to store information. Recall that a classical computer also requires memory to store information. The simplest way to realize a qubit physically is to use a two-level quantum system. For example, an electron, a spin 1/2 nucleus or two mutually orthogonal polarization states (horizontal and vertical, for example) of a single photon can be a qubit. We may also employ a two-dimensional subspace, such as the ground state and the first excited state, of a multi-dimensional Hilbert space, such as atomic energy levels. In any case, the two states are identified as the basis vectors, $|0\rangle$ and $|1\rangle$, of the Hilbert space so that a general single qubit state takes the form $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, $|\alpha|^2 + |\beta|^2 = 1$. A multi-qubit state is expanded in terms of the tensor products of these basis vectors. Each qubit must be separately addressable. Moreover it should be scalable up to a large number of qubits. The two-dimensional vector space of a qubit may be extended to three-dimensional (qutrit) or, more generally, d-dimensional (qudit).

   A system may be made of several different kinds of qubits. Qubits in an ion trap quantum computer, for instance, may be defined as: (1) hyperfine/Zeeman sublevels in the electronic ground state of ions (2) a ground

state and an excited state of a weakly allowed optical transition and (3) normal mode of ion oscillation. A similar scenario is also proposed for Josephson junction qubits, in which two flux qubits are coupled through a quantized LC circuit. Simultaneous usage of several types of qubits may be the most promising way to achieving a viable quantum computer.

2. *The ability to initialize the state of the qubits to a simple fiducial state, such as* $|00\ldots0\rangle$.

   Suppose you are not able to reset your (classical) computer. Then you will never trust the output of some computation even though processing is done correctly. Therefore initialization is an important part of both quantum and classical information processors.

   In many realizations, initialization may be done simply by cooling to bring the system into its ground state. Let ΔE be the difference between energies of the first excited state and the ground state. The system is in the ground state with a good precision at low temperatures satisfying $k_BT \ll \Delta E$. Alternatively, we may use projective measurement to project the system onto a desired state. In some cases, we observe the system to be in an undesired state upon such measurement. Then we may transform the system to the desired fiducial state by applying appropriate gates.

   For some realizations, such as liquid state NMR, however, it is impossible to cool the system down to extremely low temperatures. In those cases, we are forced to use a thermally populated state as an initial state. This seemingly difficult problem may be amended by several methods if some computational resources are sacrificed. We then obtain an "effective" pure state, so-called the pseudopure state, which works as an initial state for most purposes.

   Continuous fresh supply of qubits in a specified state, such as $|0\rangle$, is also an important requirement for successful quantum error correction. as we have seen in Section VII.

3. *Long decoherence times, much longer than the gate operation time.*

   Hardware of a classical computer lasts long, for on the order of 10 years. Things are totally different for a quantum computer, which is fragile against external disturbance called decoherence, see Section VI.

   Decoherence is probably the hardest obstacle to building a viable quantum computer. Decoherence means many aspects of quantum state degradation due to interactions of the system with the environment and sets the maximum time available for quantum computation. Decoherence time itself is not very important. What matters is the ratio "decoherence time/gate operation time". For some realizations, decoherence time may be as short as $\sim \mu$s. This is not necessarily a big problem provided that the gate operation time, determined by the Rabi oscillation period and the qubit-coupling strength, for example, is much shorter than the decoherence time. If the typical gate operation time is $\sim$ ps, say, the system may execute $10^{12-6} = 10^6$ gate operations before the quantum state decays. We quote the number $\sim 10^5$ of gates required to factor 21 into 3 and 7 by using Shor's algorithm [40].

   There are several ways to effectively prolong decoherence time. A closed-loop control method incorporates QECC, while an open-loop control method incorporates noiseless subsystem [41] and decoherence free subspace (DFS) [42].

4. *A "universal" set of quantum gates.*

   Suppose you have a classical computer with a big memory. Now you have to manipulate the data encoded in the memory by applying various logic gates. You must be able to apply arbitrary logic operations on the memory bits to carry out useful information processing. It is known that the NAND gate is universal, i.e., any logic gates may be implemented with NAND gates.

   Let $H(\gamma(t))$ be the Hamiltonian of an n-qubit system under consideration, where $\gamma(t)$ collectively denotes the control parameters in the Hamiltonian. The time-development operator of the system is $U[\gamma(t)] = \mathcal{T}\exp\left[-\frac{i}{\hbar}\int^T H(\gamma(t))dt\right] \in \mathrm{U}(2^n)$, where $\mathcal{T}$ is the time-ordering operator. Our task is to find the set of control parameters $\gamma(t)$, which implements the desired gate U_{gate} as $U[\gamma(t)] = U_{\mathrm{gate}}$. Although this "inverse problem" seems to be difficult to solve, a theorem by Barenco *et al.* guarantees that any $\mathrm{U}(2^n)$ gate may be decomposed into single-qubit gates $\in \mathrm{U}(2)$ and CNOT gates [16]. Therefore it suffices to find the control sequences to implement U(2) gates and a CNOT gate to construct an arbitrary gate. Naturally, implementation of a CNOT gate in any realization is considered to be a milestone in this respect. Note, however, that any two-qubit gates, which are neither a tensor product of two one-qubit gates nor a SWAP gate, work as a component of a universal set of gates [43].

5. *A qubit-specific measurement capability.*

   The result of classical computation must be displayed on a screen or printed on a sheet of paper to readout the result. Although the readout process in a classical computer is regarded as too trivial a part of computation, it is a vital part in quantum computing.

   The state at the end of an execution of quantum algorithm must be measured to extract the result of the computation. The measurement process depends heavily on the physical system under consideration. For most realizations, projective measurements are the primary method to extract the outcome of a computation. In liquid state NMR, in contrast, a projective measurement is impossible, and we have to resort to ensemble averaged measurements.

   Measurement in general has no 100% efficiency due to decoherence, gate operation error and many more reasons. If this is the case, we have to repeat the same computation many times to achieve reasonably high reliability.

Moreover, we should be able to send and store quantum information to construct a quantum data processing network. This "networkability" requires following two additional criteria to be satisfied.

(6) *The ability to interconvert stationary and flying qubits.*

   Some realizations are excellent in storing quantum information while long distant transmission of quantum information might require different physical resources. It may happen that some system has a Hamiltonian which is easily controllable and is advantageous in executing quantum algorithms. Compare this with a current digital computer, in which the CPU and the system memory are made of semiconductors while a hard disk drive is used as a mass storage device. Therefore a working quantum computer may involve several kinds of qubits and we are forced to introduce distributed quantum computing. Interconverting ability is also important in long distant quantum teleportation using quantum repeaters.

(7) *The ability to faithfully transmit flying qubits between specified locations.*

   Needless to say, this is an indispensable requirement for quantum communication such as quantum key distribution. This condition is also important in distributed quantum computing mentioned above.

B. Physical realizations

There are numerous physical systems proposed as possible candidates for a viable quantum computer to date [44]. Here is the list of the candidates;

1. Liquid-state/Solid-state NMR and ENDOR
2. Trapped ions
3. Neutral atoms in optical lattice
4. Cavity QED with atoms
5. Linear optics
6. Quantum dots (spin-based, charge-based)
7. Josephson junctions (charge, flux, phase qubits)
8. Electrons on liquid helium surface

and other unique realizations. ARDA QIST roadmap [44] evaluates each of these realizations.

IX. NMR QUANTUM COMPUTER

In the following, a short introduction to an NMR quantum computer is given. Mathematical, rather than physical, aspects will be emphasized.

NMR quantum computer is one of the most established systems among many physical realizations of a quantum computer. In spite of its peculiar character associated with initialization and lack of scalability, it still works as a prototypical quantum computer, with at most 10 qubits, on which small-scale quantum algorithms can be executed. Qubits in this realization are spin-1/2 nuclei. Molecules with a certain number of such nuclei are employed as a quantum register. The system is made of a macroscopic number ($\sim 10^{20}$) of moleclues in thermal equilibrium, and we have to take care of these aspects in initilization and measurements. Our exposition follows mostly [1]. Other useful review is [47]. Interested readers may concult with there references. The symbol $I_k = \sigma_k/2 \quad (k = x, y, z)$ is employed throughout this section.

Molecules with a certain number of spin 1/2 nuclei are required to construct an NMR quantum computer. Figure 10 lists typical molecules employed in NMR QC to date.

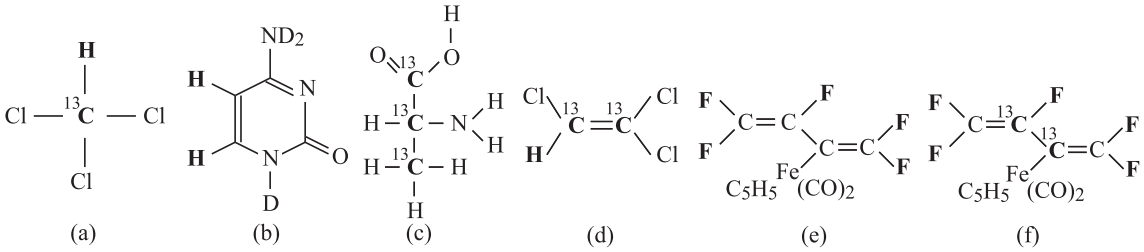


FIG. 10: Structure of molecules listed employed in NMR QC. Nuclei working as qubits are indicated in boldface. (a) Chloroform. (b) Partially deuterated cytosine. (c) ^{13}C labelled carbons in alanine. (d) Trichloroethylene. (e) Pentafluorobutadienl cyclopentadienyldicarbonyliron complex. (f) Perfluorobutadienyl iron complex with the inner two carbons ^{13}C-labelled.

A. NMR Spectrometer

Figure 11 shows the schematic diagram of the NMR spectrometer setup. A test tube containing molecules is placed in an NMR spectrometer. It is under a strong magnetic field $\boldsymbol{B}_0$ on the order of 10 T, which introduces well-defined spin-up and spin-down eigenstates of each nucleus. The energy difference between two spin states is $\hbar\gamma B_0 \equiv \hbar\omega_0$, where γ is the gyromagentic ratio of the nucleus, where ω_0 is called the **Larmor frequency**. The direction of $\boldsymbol{B}_0$ is taken as the z-axis. A radio frequency (rf) magnetic field $\boldsymbol{B}_1(t)$ along the x-axis is applied through a coil to implement one-qubit gates as will be shown later. It selectively accesses each spin by tuning its rf frequency $\omega_{\rm rf}$ with the Larmor frequency of the target nucleus. The amplitude B_1, the frequency $\omega_{\rm rf}$, the phase ϕ and the pulse shape (square-well, Gaussian and so on) are controllable parameters. The same coil is also used to pick up signals from rotating spins through magnetic induction when measurement is done. Several coils are introduced to control several nuclear species simultaneously. Each coil produces rf pulses for a particular nuclear species and receives induction signals from them.

B. Hamiltonian

1. *Single-Spin Hamiltonian*

We are exclusively concerned with room-temperature liquid state NMR here. Due to rapid random motion of molecules in a liquid at room temperature, both rotational and translational intermolecular interactions are averaged

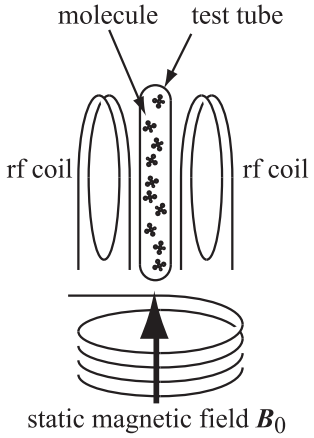


FIG. 11: Test tube with a macroscopic number of molecules is placed in a strong static field $\boldsymbol{B}_0$ and an rf magnetic field $\boldsymbol{B}_1(t)$ generated by a pair of rf coils.

to vanish, and each molecule may be regarded as being isolated from other molecules.

Let us consider a nucleus with spin 1/2 in a strong static magnetic field $\boldsymbol{B}_0$ on the order of 10 T along the z-axis. The Hamiltonian of this nucleus is

$$H_0 = -\hbar\gamma \boldsymbol{B}_0 \cdot \boldsymbol{I} = -\hbar\omega_0 I_z, \tag{115}$$

where γ is the nuclear gyromagnetic ratio and $\omega_0 = \gamma B_0$ is called the **Larmor frequency**. The eigenvalues of the Hamiltonian are $E_0 = -\hbar\omega_0/2$ and $E_1 = \hbar\omega_0/2$, and the corresponding eigenstates are

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{116}$$

respectively. Note that the state $|0\rangle$ ($|1\rangle$) denotes the spin up (down) state in physicists' terminology. Table III shows the Larmor frequencies of several nuclei which are often employed in NMR QC.

TABLE III: Larmor frequecies of typical nuclei at $B_0 = 11.74$ T.

Nucleus	^{1}H	^{13}C	^{19}F	^{31}P
ω_0[MHz]	500	126	470	202

The spin state of a nucleus can be controlled by applying a radio frequency (rf) magnetic field in the xy-plane. Here we take its direction along the $-x$-axis as

$$\boldsymbol{B}_1(t) = -B_1(t)\cos(\omega_{\rm rf}t - \phi)\hat{\boldsymbol{x}}, \tag{117}$$

where $\hat{\boldsymbol{x}}$ is the unit vector along the x-axis and $\omega_{\rm rf}$ and $-\phi$ are the angular frequency and the initial phase of the rf field, respectively. This field induces an extra term in the Hamiltonian of the form

$$H_{\rm rf} = 2\hbar\omega_1 \cos(\omega_{\rm rf}t - \phi) I_x, \tag{118}$$

where $2\hbar\omega_1 = \hbar\gamma B_1$. The factor 2 has been multiplied to make the corresponding Hamiltonian simpler. The total Hamiltonian in the laboratory frame (i.e., fixed coordinate axes) is therefore

$$H = H_0 + H_{\rm rf} = -\hbar\omega_0 I_z + 2\hbar\omega_1 \cos(\omega_{\rm rf}t - \phi) I_x. \tag{119}$$

The parameters ω_1, $\omega_{\rm rf}$ and ϕ are controllable as functions of time, while ω_0 (i.e., B_0) is fixed. It is always assumed in the following that the condition $\omega_0 \gg \omega_1$ is satisfied. Therefore a nuclear spin has two well-defined eigenstates $|0\rangle = |\uparrow\rangle$, and $|1\rangle = |\downarrow\rangle$ and the rf field acts as a perturbation to control the spin states.

The ratio $k_BT/\hbar\omega_0$ is on the order of 8×10^4, at room temperature of $T \sim 300$ K, for $\omega_0 \sim 500$ MHz, which is the hydrogen Larmor frequency at $B_0 = 11$ T. Therefore, the liquid is in a thermal mixed state. For this reason, we use density matrices rather than wave functions to describe NMR quantum states. The one-spin density matrix of a thermal equilibrium state is

$$\rho(T) = \frac{e^{-H/k_BT}}{Z(T)}, \tag{120}$$

where T is the temperature and $Z(T) = \mathrm{tr}\, e^{-H/k_BT}$ is the partition function. The density matrix in the absence of an rf field is

$$\rho(T) = \frac{e^{\hbar\omega_0 I_z/k_BT}}{\mathrm{tr}\, e^{\hbar\omega_0 I_z/k_BT}} = \frac{1}{2}\left[I + \frac{\hbar\omega_0}{k_BT} I_z + O\left(\frac{\hbar\omega_0}{k_BT}\right)^2\right]. \tag{121}$$

The dynamics of a density matrix is given by the Liouville-von Neumann equation

$$i\hbar\frac{d\rho}{dt} = [H, \rho]. \tag{122}$$

The Hamiltonian (119) has explicit time-dependence through coupling with an rf field. This is inconvenient in integrating the Liouville-von Neumann equation. This problem is solved if we change the frame of reference from the laboratory frame to a frame rotating with the Larmor frequency around the z-axis. Let

$$U_R = e^{-i\hbar\omega I_z t} \tag{123}$$

be a unitary transformation to a rotating frame with the angular velocity ω in general (we put $\omega = \omega_0$ later). Here we regard I_z as the generator of rotations around the z-axis. The density matrix is now transformed into

$$\rho_R = U_R \rho U_R^\dagger. \tag{124}$$

The Hamiltonian is also transformed to $\tilde{H}$, whose form is derived below. The Liouville-von Neumann equation in the rotating frame takes the same form as Eq. (122) and is given by

$$i\hbar \frac{d\rho_R}{dt} = [\tilde{H}, \rho_R]. \tag{125}$$

We substitute Eq. (123) into the above equation to obtain

$$\begin{aligned}
\tilde{H} &= U_R H U_R^\dagger - i\hbar U_R \frac{dU_R^\dagger}{dt} \\
&= \hbar \begin{pmatrix} -(\omega_0 - \omega)/2 & \omega_1 e^{-i\omega t}\cos(\phi - \omega_{\rm rf} t) \\ \omega_1 e^{i\omega t}\cos(\phi - \omega_{\rm rf} t) & (\omega_0 - \omega)/2 \end{pmatrix} \\
&= \frac{\hbar}{2} \begin{pmatrix} -\omega_0 + \omega & \omega_1 \left[e^{-i[(\omega - \omega_{\rm rf})t + \phi]} + e^{-i[(\omega + \omega_{\rm rf})t - \phi]} \right] \\ \omega_1 \left[e^{i[(\omega - \omega_{\rm rf})t + \phi]} + e^{i[(\omega + \omega_{\rm rf})t - \phi]} \right] & \omega_0 - \omega \end{pmatrix}.
\end{aligned} \tag{126}$$

Note that the main contribution $-\hbar\omega_0 I_z$ in the laboratory frame disappears under this transformation if we set $\omega = \omega_0$, which we will assume hereafter. Now we further simplify this Hamiltonian (126) by taking the "resonance condition" $\omega_{\rm rf} = \omega_0$. Namely, we take $\omega_{\rm rf}$ in resonance with the Larmor frequency ω_0 of the spin. Moreover, we note that the terms oscillating rapidly with the frequency $2\omega_0$ are averaged to vanish if we are interested in the time scale much longer than $1/\omega_0$. This approximation is known as the **rotating wave approximation**. We will see later that flipping a spin by angle π by making use of the Rabi oscillation takes time $\sim 1/\omega_1$, and this is much longer than $1/\omega_0$ due to the assumition $\omega_1 \ll \omega_0$. Therefore it is legitimate to replace the Hamiltonian (126) with a simpler time-independent Hamiltonian

$$\tilde{H} = \hbar\omega_1(\cos\phi I_x + \sin\phi I_y) = \hbar\omega_1 \begin{pmatrix} 0 & e^{-i\phi} \\ e^{i\phi} & 0 \end{pmatrix}. \tag{127}$$

Note that this Hamiltonian is traceless: $\mathrm{tr}\,\tilde{H} = 0$. A traceless Hamiltonian generates only elements of SU(2). Therefore, the one-qubit NMR Hamiltonian generates SU(2) gates only. Note, however, that this is by no means a restriction. Any $U \in \mathrm{U}(2)$ may be mapped to $e^{i\alpha}U \in \mathrm{SU}(2)$ by multiplying a proper phase factor. Since this extra overall phase is not observable, we may replace a U(2) gate U with an equivalent SU(2) gate $\tilde{U}$. This remains true for a multi-qubit unitary gate; see the next subsection.

2. Multi-Spin Hamiltonian

Molecules with n spins are required to execute n-qubit quantum algorithms. Let us consider a linear molecule in which each spin is coupled only to its nearest neighbor spins to simplify our argument. Although a more complicated spin network will be advantageous in saving the number of gates and the execution time, actual implementation requires more elabolated techniques in this case. We will take the natural unit in which $\hbar = 1$ hereafter to simplify mathematical expressions. It will be recovered whenever necessary.

Let us consider a molecule with two spins to begin with. We denote the Larmor frequency of the ith spin by $\omega_{0,i}$ $(i = 1, 2)$. We assume there is a Heiseberg type interaction of the form

$$H_{\rm int} = J \sum_{k=x,y,z} I_k \otimes I_k \tag{128}$$

between spins, where J is the coupling strength. In fact there are other types of interaction including inter-molecular interaction. These interactions are averaged out, thanks to rapid translational and rotational motions of molecules at room temeparatures, and give no contribution to the Hamiltonian (128).

Suppose there are two oscillating magnetic fields along the $-x$-axis with frequency $\omega_{\rm rf,i}$ and amplitude $B_{1,i}$ $(i = 1, 2)$. The Hamiltonian in the laboratory frame is

$$H = H_0 + H_{\rm rf,1} + H_{\rm rf,2}, \tag{129}$$

where

$$H_0 = -\omega_{0,1} I_z \otimes I - \omega_{0,2} I \otimes I_z + J \sum_{k=x,y,z} I_k \otimes I_k, \tag{130}$$

while

$$H_{\rm rf,1} = 2\omega_{1,1} \cos(\omega_{\rm rf,1} t - \phi_1)(I_x \otimes I + g I \otimes I_x) \tag{131}$$

and

$$H_{\rm rf,2} = 2\omega_{1,2} \cos(\omega_{\rm rf,2} t - \phi_2)(g^{-1} I_x \otimes I + I \otimes I_x), \tag{132}$$

where $2\omega_{1,i} = \gamma_i B_{1,i}$ and $g = \gamma_2/\gamma_1$ is the ratio of the gyromagnetic ratios of two nuclei. Here I is the unit matrix of dimension 2. The first (second) term in the parentheses in Eqs. (131) and (132) is the interaction Hamiltonian describing the coupling between the first (second) spin and the oscillating fields.

The transformation to a rotating frame of respective spin proceeds similarly to the single-spin case. Let us introduce the transformation

$$U_R = e^{-i\omega_{0,1} I_z t} \otimes e^{-i\omega_{0,2} I_z t}. \tag{133}$$

The Hamitonian $\tilde{H}$ of the spins in respective rotating frames is defined as before as

$$\tilde{H} = \tilde{H}_0 + \tilde{H}_{\rm rf,1} + \tilde{H}_{\rm rf,2}, \tag{134}$$

where

$$\begin{aligned}
\tilde{H}_0 &= U_R H_0 U_R^\dagger - i U_R \frac{d}{dt} U_R^\dagger \\
&= J \left(e^{-i\omega_{0,1} I_z t} \otimes e^{-i\omega_{0,2} I_z t} \right) \sum_{k=x,y,z} I_k \otimes I_k \left(e^{i\omega_{0,1} I_z t} \otimes e^{i\omega_{0,2} I_z t} \right) \\
&= \pi J \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & e^{i\Delta\omega_0 t} & 0 \\ 0 & e^{-i\Delta\omega_0 t} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} + J I_z \otimes I_z.
\end{aligned} \tag{135}$$

Here $\Delta\omega_0 \equiv \omega_{0,2} - \omega_{0,1}$ is the difference in the Larmor frequencies of the spins. The matrix elements $e^{\pm i\Delta\omega_0 t}$ are averaged to vanish for the time scale τ satisfying $\Delta\omega_0 \tau \gg 2\pi$. Table IV shows relevant parameters for typical two-

TABLE IV: Physical parameters of two-spin molecules, ^{13}C labelled chloroform and cytosine. The magnetic field is set to $B_0 = 11.74$[T].

	$\omega_{0,1}$	$\omega_{0,2}$	$\Delta\omega_0$	J
Chloroform	500 MHz	100 MHz	400 MHz	200 Hz
Cytosine	500 MHz	500 MHz	765 Hz	7.1 Hz

qubit molecules, ^{13}C labelled chloroform and cytosine. ^{13}C labelled chloroform is a **heteronucleus molecule** whose qubits are hydrogen and ^{13}C nuclei. Cytosine is a **homonucleus molecule**, both qubits of which are hydrogen nuclei. It seems impossible at first glance to address a particular spin in the presence of other spins of the same species since they have the same resonance frequency. However, selective addressing is made possible through the so-called **chemical shift**. The Larmor frequency of a nucleus in a molecule depends not only on the nuclear species but also on its position in the molecule. The electron density at each nucleus varies according to the bonds around it, and therefore the effective magnetic field depends on where a particular nucleus sits in the molecule. This shift in the Larmor frequency is called the chemical shift and allows us to selectively address each nucleus of a properly designed

molecule. We cannot employ methane (CH_4) as a four-qubit molecule since all the hydrogen nuclei sit in equivalent positions and therefore have the same chemical shift. Symmetry of the molecule must be broken to produce different chemical shifts.

The pulse width for one-qubit control is typically $\tau \sim 10\ \mu$s for ^{13}C labelled chloroform for which $\Delta\omega_0\tau \sim 4000 \gg 1$. For cytosine, the one-qubit control pulse width τ cannot be too short. Let τ be the pulse width. Then its Fourier transform has a width $\sim 1/\tau$ in the frequency domain. Therefore selective addressing to each spin is impossible unless τ satisfies the condition $1/\tau \ll \Delta\omega_0$, In actual implementation, the pulse width τ is taken such that the condition

$$\Delta\omega_0\tau \gg 1 \gg J\tau \tag{136}$$

is satisfied. The second inequality must be satisfied for the effect of the J-coupling to be negligible during the one-qubit operation. Due to a large ratio $\Delta\omega_0/J \sim 10^2$ for cytosine, there always exists such τ which satisfies the condition (136). We have to resort to numerical optimization if one of the inequalities is not satified.

Now the interaction Hamiltonian takes a simple Ising form

$$\tilde{H}_0 = JI_z \otimes I_z \tag{137}$$

for both heteronucleus and homonucleus molecules, where a time scale $\tau \gg 1/\Delta\omega_0$ is assumed in the latter case. Disappearance of $I_x \otimes I_x$ and $I_y \otimes I_y$ is understood intuitively as follows. Suppose the rf fields are turned off. Then the i-th spin executes free precession with frequency $\omega_{0,i}$ around the z-axis. Since $\omega_{0,1}$ and $\omega_{0,2}$ differ by $\Delta\omega_0$, their x- and y-axes in the rotating frames rotate with relative angular frequencey $\Delta\omega_0$. Therefore, for a time scale τ such that $\Delta\omega_0\tau \gg 1$, the contribution from $I_x \otimes I_x$ and $I_y \otimes I_y$ is averaged out to vanish. The term $I_z \otimes I_z$ does not vanish since the z-axes in the rotating frame remain the same as the laboratory frame for both spins. Application of rf fields merely introduces slow motions of spins in the rotating framesand it does not alter this conclusion.

As for $H_{\mathrm{rf},1}$, we obtain

$$\begin{aligned}
\tilde{H}_{\mathrm{rf},1} &= U_R H_{\mathrm{rf},1} U_R^\dagger \\
&= \omega_{1,1}\Big[\left(e^{i(\omega_{\mathrm{rf},1}t-\phi_1)} + e^{-i(\omega_{\mathrm{rf},1}t-\phi_1)}\right) \left\{ \left(e^{-i\omega_{0,1}I_z t} I_x e^{i\omega_{0,1}I_z t}\right) \otimes I \right\} \\
&\quad + g\left(e^{i(\omega_{\mathrm{rf},1}t-\phi_1)} + e^{-i(\omega_{\mathrm{rf},1}t-\phi_1)}\right) \left\{ I \otimes \left(e^{-i\omega_{0,2}I_z t} I_x e^{i\omega_{0,2}I_z t}\right) \right\} \Big].
\end{aligned}$$

Now we take the resonance condition $\omega_{\mathrm{rf,i}} = \omega_{0,i}$ $(i = 1, 2)$. Then $\tilde{H}_{\mathrm{rf},1}$ is simplified as

$$\begin{aligned}
\tilde{H}_{\mathrm{rf},1} &= \frac{\omega_{1,1}}{2}\Big[\begin{pmatrix} 0 & e^{-i\phi_1} \\ e^{i\phi_1} & 0 \end{pmatrix} \otimes I \\
&\quad + gI \otimes \begin{pmatrix} 0 & e^{-i(\Delta\omega_0 t+\phi_1)} + e^{-i(\Omega_0 t-\phi_1)} \\ e^{i(\Delta\omega_0 t+\phi_1)} + e^{i(\Omega_0 t-\phi_1)} & 0 \end{pmatrix} \Big],
\end{aligned}$$

where $\Omega_0 \equiv \omega_{0,1} + \omega_{0,2}$. The second matrix vanishes for τ such that $\Omega\tau, \Delta\omega_0\tau \gg 1$, and finally we obtain

$$\tilde{H}_{\mathrm{rf},1} = \omega_{1,1}\left[\cos\phi_1 I_x \otimes I + \sin\phi_1 I_y \otimes I\right]. \tag{138}$$

Similarly we prove that

$$\tilde{H}_{\mathrm{rf},2} = \omega_{1,2}\left[\cos\phi_2 I \otimes I_x + \sin\phi_2 I \otimes I_y\right]. \tag{139}$$

In summary, the Hamiltonian for a two-qubit molecule in the rotating frames with respective Larmor frequency is

$$\begin{aligned}
\tilde{H} &= JI_z \otimes I_z + \omega_{1,1}\left[\cos\phi_1 I_x \otimes I + \sin\phi_1 I_y \otimes I\right] \\
&\quad + \omega_{1,2}\left[\cos\phi_2 I \otimes I_x + \sin\phi_2 I \otimes I_y\right].
\end{aligned} \tag{140}$$

From a control theoretical point of view, the first term is out of our control and is called the **drift term**, while the second and the third terms, altogether, are called the **control terms** since $\omega_{1,i}$ and ϕ_i are controllable.

Generalization of the above two-qubit Hamiltonian to an n-qubit Hamiltonian is straightforward. For a molecule with n spins coupled linearly, the Hamiltonian in the rotating frame of each spin with angluar frequency $\omega_{0,i}$ takes the form

$$\tilde{H} = \sum_{i=1}^{n-1} J_{i,i+1} I_{z,i} \otimes I_{z,i+1} + \sum_{i=1}^{n} \omega_{1i}(\cos\phi_i I_{x,i} + \sin\phi_i I_{y,i}), \tag{141}$$

where $J_{i,i+1}$ stands for the coupling strength between spins i and $i+1$ and $I_{k,i} = I \otimes \ldots \otimes I_k \otimes \ldots \otimes I$ with I_k in the ith position. The resonance condition $\omega_{\mathrm{rf,i}} = \omega_{0,i}$ and linear configuration of n spins are understood in deriving Eq. (141).

We will work exclusively with Hamiltonians in the rotating frame of each spin in the rest of this section.

C. Implementation of Gates and Algorithms

The Hamiltonians introduced in the previous section are employed to implement quantum gates. Here we consider one-, two-, and multi-qubit gates separately.

1. *One-Qubit Gates in One-Qubit Molecule*

The Hamiltonian

$$\tilde{H} = \omega_1(\cos\phi I_x + \sin\phi I_y)$$

contains only I_x and I_y as SU(2) generators. This is not a problem though since rotations generated by I_z can be implemented with $I_{x,y}$ generators as we see below. Let us define SU(2) gates which are often employed as building blocks of quantum circuits. Let $X, Y, Z, \bar{X}, \bar{Y}$ and $\bar{Z}$ be rotations by $\pi/2$ around $\tilde{x}$-, $\tilde{y}$-, $\tilde{z}$-, $-\tilde{x}$-, $-\tilde{y}$- and $-\tilde{z}$-axes respectively. Their explicit forms as SU(2) matrices are

$$\begin{aligned} X &= e^{-i(\pi/2)I_x} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix},\ Y = e^{-i(\pi/2)I_y} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \\ Z &= e^{-i(\pi/2)I_z} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1-i & 0 \\ 0 & 1+i \end{pmatrix},\ \bar{X} = e^{i(\pi/2)I_x} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i \\ i & 1 \end{pmatrix}, \\ \bar{Y} &= e^{i(\pi/2)I_y} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \bar{Z} = e^{i(\pi/2)I_z} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1+i & 0 \\ 0 & 1-i \end{pmatrix}. \end{aligned} \tag{142}$$

It is useful for later purposes to write down the explicit form of a gate $R(\theta, \phi)$, whose rotation angle is θ and phase angle is ϕ in the xy-plane,

$$R(\theta,\phi) \ = \ e^{-i\theta(\cos\phi I_x + \sin\phi I_y)} = \cos\frac{\theta}{2} I - 2i\sin\frac{\theta}{2}(\cos\phi I_x + \sin\phi I_y) = \begin{pmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2}e^{-i\phi} \\ -i\sin\frac{\theta}{2}e^{i\phi} & \cos\frac{\theta}{2} \end{pmatrix}. \tag{143}$$

Let us consider implementing X, for example. We need to find parameters ϕ, ω_1 and τ such that

$$e^{-i\int_0^\tau \tilde{H}dt} = e^{-i\omega_1\tau(\cos\phi I_x + \sin\phi I_y)} = e^{-i\pi I_x/2}.$$

It is easily found that a pulse with phase $\phi = 0$, amplitude ω_1 and duration τ satisfying $\omega_1\tau = \pi/2$ does the job. We assume here the pulse shape is square and express it graphically as in Fig. 12. The parameter τ is called the **pulse width**. More sophisticated pulses are available, but we restrict ourselves within square pulses to simplify our calculation. Similarly Y, $\bar{X}$, $\bar{Y}$ are obtaind by applying pulses with $\phi = \pi/2, \pi$ and $-\pi/2$ and pulse duration $\tau = \pi/2\omega_1$, respectively. A typical value for ω_1 is ~ 100 kHz for heteronucleus molecules, and the above operation is implemented with the pulse width $\tau \sim 1/\omega_1 \sim 10\ \mu$s.

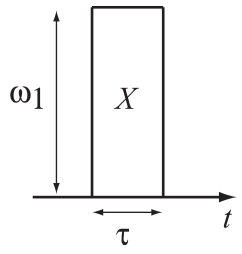


FIG. 12: Square pulse. Amplitude of a continuous wave with frequency $\omega_{\rm rf}$ is modulated by this pulse. The amplitude corresponds to ω_1 and the pulse width to τ. They satisfy $\omega_1\tau = \pi/2$.

Exercise IX.1 *Let $\phi = \pi/4$ in $\tilde{H}$ and write down the unitary matrix which $\tilde{H}$ generates when $\omega_1\tau = \pi/2$. Apply this unitary matrix to $|0\rangle$ and $|1\rangle$ and find the states obtained.*

The Hamiltonian lacks the generator I_z. This does not imply Z and $\bar{Z}$ cannot be implemented with the Hamiltonian $\tilde{H}$. There are three ways to implement $U = e^{-i\alpha I_z}$ with $\tilde{H}$. The simplest one is to shift the clock of the NMR by a certain amount of time τ_z. The frame is rotating with the angular velocity ω_0 around the z-axis, and if we shift the clock of the NMR sequencer by τ_z, we will obtain a rotation equivalent with $U = e^{-i\omega_0\tau_z I_z}$. Thus shifting the clock by $\tau_z = \alpha/\omega_0$ implements the gate $e^{-i\alpha I_z}$. The second one is to literally generate $U = e^{-i\alpha I_z}$ with $I_{x,y}$. By noting the identity $I_z = e^{-i(\pi/2)I_x} I_y e^{i(\pi/2)I_x}$ we immediately obtain

$$e^{-i\alpha I_z} = e^{-i(\pi/2)I_x} e^{-i\alpha I_y} e^{i(\pi/2)I_x}. \tag{144}$$

The third one is applicable only in the end of the computation and when the spin is in one of the eigenstates of I_z. It is clear that $e^{-i\alpha I_z}|j\rangle \sim |j\rangle$ if the phase is ignored since $\sigma_z|j\rangle = \pm|j\rangle$. Therefore if we can shift some of the I_z rotation matrices toward the very end of the algorithm, we may ignore them at all. Now we have shown that $\tilde{H}$ generates all SU(2) rotations.

The following relations are useful in designing pulse sequences for NMR quantum computing:

$$\begin{array}{cccc} XY\bar{X} = Z, & \bar{Y}XY = Z, & \bar{X}\bar{Y}X = Z, & Y\bar{X}\bar{Y} = Z \\ \bar{X}YX = \bar{Z}, & YX\bar{Y} = \bar{Z}, & X\bar{Y}\bar{X} = \bar{Z}, & \bar{Y}\bar{X}Y = \bar{Z} \\ XY = ZX, & XY = YZ, & \bar{Y}X = XZ, & Y\bar{X} = ZY \\ XZ = Z\bar{Y}, & \bar{Y}Z = Z\bar{X}, & \bar{X}Z = ZY, & YZ = ZX \\ XZZ = ZZ\bar{X}, & & YZZ = ZZ\bar{Y}. & \end{array} \tag{145}$$

By making use of these relations, it becomes possible to replace Z and $\bar{Z}$ with other rotations. It also becomes possible to eliminate some of Z and $\bar{Z}$ by sending them to the both ends of a pulse sequence.

Exercise IX.2 *Verify the above relations.*

It is clear that the Hamiltonian $\tilde{H}$ is independent of t so far as ω_1 and ϕ are time-independent. In general, ω_1 and ϕ may change as functions of time. In actual experiments, they are often taken to be piecewise constant, for which case the time-evolution operator is given by

$$U = \mathcal{T} e^{-i\int_0^T \tilde{H}(t)dt} \equiv e^{-i\tilde{H}(t_n)\Delta t_n} e^{-i\tilde{H}(t_{n-1})\Delta t_{n-1}} \cdots e^{-i\tilde{H}(t_1)\Delta t_1}, \tag{146}$$

where $\mathcal{T}$ stands for the time-ordered product and

$$\tilde{H}(t_k) = \omega_1(t_k)\left[\cos\phi(t_k) I_x + \sin\phi(t_k) I_y\right]$$

is the Hamiltonian at the kth step whose temporal duration is Δt_k.

Example IX.3 *Let us consider implementing the Hadamard gate*

$$U_{\rm H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

with our Hamiltonian $\tilde{H}$. Since $\det U_{\rm H} = -1$, we have to multiply i to $U_{\rm H}$ to make it an element of SU(2). (The factor $-i$ also does the job.) We are tempted to use Eq. (9) to find parameters ω_1, ϕ and τ such that

$$\tilde{H}\tau = -\frac{\pi}{\sqrt{2}}(I_x + I_z),$$

which certainly satisfies $e^{-i\tilde{H}\tau} = U_{\rm H}$. However, this does not work since we do not have an I_z term in $\tilde{H}$. Therefore, we have to implement $U_{\rm H}$ using the formula

$$\begin{aligned} & e^{-i\alpha I_x} e^{-i\beta I_y} e^{-i\gamma I_x} \\ &= \begin{pmatrix} \cos\left(\frac{\beta}{2}\right)\cos\left(\frac{\alpha+\gamma}{2}\right) - i\sin\left(\frac{\beta}{2}\right)\sin\left(\frac{\alpha-\gamma}{2}\right) & -\cos\left(\frac{\alpha-\gamma}{2}\right)\sin\left(\frac{\beta}{2}\right) - i\cos\left(\frac{\beta}{2}\right)\sin\left(\frac{\alpha+\gamma}{2}\right) \\ \cos\left(\frac{\alpha-\gamma}{2}\right)\sin\left(\frac{\beta}{2}\right) - i\cos\left(\frac{\beta}{2}\right)\sin\left(\frac{\alpha+\gamma}{2}\right) & \cos\left(\frac{\beta}{2}\right)\cos\left(\frac{\alpha+\gamma}{2}\right) + i\sin\left(\frac{\beta}{2}\right)\sin\left(\frac{\alpha-\gamma}{2}\right) \end{pmatrix}. \end{aligned} \tag{147}$$

Comparison between $U_{\rm H}$ and the above expression immediately leads to the following solution:

$$\alpha = -\pi, \beta = \frac{\pi}{2}, \gamma = 0, \tag{148}$$

for example. Therefore $U_{\rm H}$ is implemented by two square pulses as

$$U_{\rm H} = e^{i\omega_1\tau_2 I_x} e^{-i\omega_1\tau_1 I_y} = \bar{X}^2 Y, \tag{149}$$

where $\omega_1\tau_1 = \pi/2$ and $\omega_1\tau_2 = \pi$. The amplitude ω_1 need not be the same for the two pulses, but there is no reason to employ different amplitude either. The amplitude should be large to implement a gate with a shorter pulse width. However, a large amplitude pulse leads to overcurrent in the rf coil and eventually damages the coil. A typical pulse width for a π-pulse is on the order of 10 μs as mentioned before.

Using the symbols introduced above, this pulse sequence is conveniently expressed as

$$U_{\rm H}: \ -Y - \bar{X}^2 - . \tag{150}$$

The time flows from left to right as before. We also describe the pulse sequence graphically as in Fig. 13.

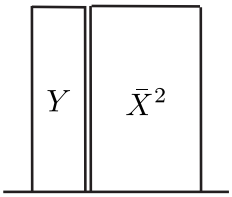


FIG. 13: Control pulse sequence to implement the Hadamard gate H. $\bar{X}^2$ is a π-pulse around $-x$-axis. The time flows from left to right.

Exercise IX.4 *Implement the phase shift gate*

$$U(\theta) = \begin{pmatrix} 1 & 0 \\ 0 & e^{-i\theta} \end{pmatrix} \tag{151}$$

using the Hamiltonian $\tilde{H}$. Here θ is a real constant. Note that $U \notin \mathrm{SU}(2)$ and a phase must be multiplied to make it an element of SU(2).

2. One-Qubit Operation in Two-Qubit Molecule: Bloch-Siegert Effect

Let us consider the effect of an off-resonance pulse on a qubit. We have to consider this effect when we have a multi-qubit molecule with several nuclei of the same species; addressing to one qubit may affect the other qubits of the same species since they have close resonance frequencies.

We first consider the effect of an off-resonance pulse on a one-qubit molecule. Let $\omega_{\rm rf} = \omega_0 + \delta$, δ being the detuning parameter. Then we find from Eqs. (126) and (127) that

$$\begin{aligned} \tilde{H} &= \delta I_z + \omega_1 (\cos\phi I_x + \sin\phi I_y) = \delta (\epsilon \cos\phi I_x + \epsilon \sin\phi I_y + I_z) \\ &= \delta\sqrt{1+\epsilon^2}\hat{\boldsymbol{n}} \cdot \boldsymbol{I}, \end{aligned} \tag{152}$$

where $\epsilon = \omega_1/\delta$ and

$$\hat{\boldsymbol{n}} = \frac{1}{\sqrt{1+\epsilon^2}} (\epsilon\cos\phi, \epsilon\sin\phi, 1)^t \tag{153}$$

is a unit vector. The time-development operator is

$$U(t) = e^{-i\tilde{H}t} = e^{-i\delta\sqrt{1+\epsilon^2}\hat{\boldsymbol{n}}\cdot\boldsymbol{I}t}. \tag{154}$$

Suppose the detuning is large enough compared to ω_1 so that $|\epsilon| \ll 1$. Then it follows that $\hat{\boldsymbol{n}} \simeq (0,0,1)^t$, and we have an approximation

$$U(t) \simeq e^{-i\delta\sqrt{1+\epsilon^2}I_z t}. \tag{155}$$

In fact, the rotation axis $\hat{\boldsymbol{n}}$ is slightly tilted from the z-axis and the spin precesses around this axis, which remains near the z-axis. This observation justifies the negligence of components $I_{x,y}$ in Eq. (155). However, the effect of ϵ in the square root is not negligible if we are concerned with a long-term behavior of the spin, in which $t\delta\epsilon^2$ is sizeable. This effect is called the **Bloch-Siegert effect** [59], and this shift in the reference phase must be taken into account when designing pulse sequences which involve detuned rf fields. Suppose, for example, that $\epsilon = 10^{-1}$ and $\delta t = 20\pi$. Then we obtain $\delta\sqrt{1+\epsilon^2}t - \delta t \simeq 0.31$ rad $\simeq 18°$, which is not negligible at all.

Exercise IX.5 *Suppose a spin is in the state $|\uparrow\rangle$ at $t = 0$ and its time-development is driven by the operator (154). Find the spin wave function at later time $t > 0$. Find when the spin comes back to the initial state up to an overall phase.*

Next we consider manipulating a single qubit in a two-qubit molecule. In case of a heteronucleus molecule, an rf field in resonance with one of the qubits has no effect on the other qubit. In this case, $\epsilon = \omega_1/\delta$ is typically on the order of 10^{-3}. For $\delta t = 20\pi$ as before, we obtain $\delta\sqrt{1+\epsilon^2}t - \delta t \simeq 3 \times 10^{-5}$ rad $\simeq 1.8 \times 10^{-3}$ deg. If, in contrast, a homonucleus molecule is considered, we have to take a small amplitude pulse with $\omega_1 \ll \Delta\omega_0$, $\Delta\omega_0$ being the difference in the Larmor frequencies of two nuclei of the same species, for selective addressing to a particular qubit. This makes the pulse width τ longer, since $\omega_1\tau$ specifies the rotation angle. The effect of the J-coupling may not be negligible if $\tau \gtrsim 1/J$.

Let us consider the opposite limit in which $\delta \ll \omega_1$. This takes place when we apply a hard pulse (i.e., very short pulse) in resonance with one of the qubits, qubit 2, say, in a homonucleus molecule. Equation (152) with $\delta \ll \omega_1$ leads to a Hamiltonian $\tilde{H} \simeq \omega_1(\cos\phi I_x + \sin\phi I_y)$ acting on qubit 1. Therefore qubit 1 also gets rotated by the same amount as qubit 2. In other words, by applying a hard pulse in resonance with one of the qubits, both qubits are rotated similtaneously by the same angle. Therefore a gate $I \otimes U$, which is meant to act on the second qubit, works as $U \otimes U$ if it is implemented with a hard pulse.

3. Two-Qubit Gates

Any n-qubit gate may be implemented with single-qubit gates and the CNOT gates according to the universality theorem by Barenco *et al.* [60]. We have shown in the previous subsection how single-qubit gates are implemented. Let us consider the CNOT gate here. We recall that

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

Note again that $\det U_{\rm CNOT} = -1$ and we have to multiply $U_{\rm CNOT}$ by $e^{\pm i\pi/4}$, for example, to make it an element of SU(4). We must employ the J-coupling term to implement the CNOT gate since it cannot be decomposed into a tensor product of two SU(2) gates. A standard implementation of the CNOT gate is [46]

$$U_{\rm CNOT} = Z_1\bar{Z}_2X_2U_J(\pi/J)Y_2, \tag{156}$$

where X_j is a $\pi/2$-rotation around the x-axis of the jth qubit while $\bar{X}_j$ is a $\pi/2$-rotation around the $-x$ axis of the jth qubit, for example. Explicitly,

$$X_1 = e^{-i\pi I_x/2} \otimes I, \quad \bar{X}_2 = I \otimes e^{-i\pi I_x/2},$$

for example. The matrix $U_J(\tau)$ is generated solely by the J-coupling term, without any rf pulses applied during period of time τ, as

$$U_J(\tau) = e^{-iJI_z\otimes I_z\tau} = \begin{pmatrix} e^{-iJ\tau/4} & 0 & 0 & 0 \\ 0 & e^{iJ\tau/4} & 0 & 0 \\ 0 & 0 & e^{iJ\tau/4} & 0 \\ 0 & 0 & 0 & e^{-iJ\tau/4} \end{pmatrix}. \tag{157}$$

Therefore

$$U_J(\pi/J) = e^{-i\pi I_z \otimes I_z} = \begin{pmatrix} e^{-i\pi/4} & 0 & 0 & 0 \\ 0 & e^{i\pi/4} & 0 & 0 \\ 0 & 0 & e^{i\pi/4} & 0 \\ 0 & 0 & 0 & e^{-i\pi/4} \end{pmatrix}. \tag{158}$$

Then it is easy to find that the LHS of Eq. (156) takes the form

$$Z_1 \bar{Z}_2 X_2 U_J(\pi/J) Y_2 = e^{-i\pi/4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

as promised.

Exercise IX.6 *Implement the "inverted" CNOT gate*

$$U_{\mathrm{CNOT}'} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

with the two-qubit NMR Hamiltonian.

It is clear from the construction that the expansion such as Eq. (156) requires certain degrees of expertise in NMR pulse programming and/or trial and error to adjust all the matrix elements. We introduce in §IX D a remarkable technique fully utilizing the theory of Lie algebras and Lie groups to obtain implementations of any two-qubit unitary gates. Although this technique is model independent, it is best suited for NMR quantum computing due to the reasons to be clarified below.

In principle, therefore, an NMR quantum computer is universal, and any U(2^n) gate may be implemented by properly choosing the control parameters. One might wonder how a one- or two-qubit gate is embedded in a multi-qubit molecule. This is the subject of the next subsection.

4. Multi-Qubit Gates

Suppose we have a molecule with many qubits coupled linearly. Clearly we cannot turn off inter-qubit couplings even when we do not need them. This is rather inconvenient if we want to employ one-qubit gates and CNOT gates as building blocks of quantum algorithms. One-qubit operations are executed faster compared to $1/J$, and the effect of J-couplings is safely negligible. In contrast, a two-qubit operation involves a particular J-coupling, and we have to get rid of the time-evolution of the state due to other J-couplings. This "interaction on demand" is possible if a technique called **refocusing** is employed. Refocusing cancels unwanted inter-qubit couplings.

5. Three-Qubit Case

This is best understood from the following example of a three-qubit molecule. Suppose there are three spins 1, 2 and 3 in the molecule, each with the Larmor frequency $\omega_{0,i}$, $(i = 1, 2, 3)$. The coupling strength between 1 and 2 is J_{12} and that between 2 and 3 is J_{23}. It is assumed that the spins are linear so that $J_{31} = 0$. The Hamiltonian in the rotating frame of each qubit is

$$\begin{aligned} \tilde{H} &= J_{12} I_z \otimes I_z \otimes I + J_{23} I \otimes I_z \otimes I_z + \omega_{11}(\cos\phi_1 I_{1x} + \sin\phi_1 I_{1y}) \\ &\quad + \omega_{12}(\cos\phi_2 I_{2x} + \sin\phi_2 I_{2y}) + \omega_{13}(\cos\phi_3 I_{3x} + \sin\phi_3 I_{3y}), \end{aligned} \tag{159}$$

where $I_{1x} = I_x \otimes I \otimes I$, for example.

Suppose we want to implement a gate

$$U_\alpha = \exp(-i\alpha I_z \otimes I_z \otimes I). \tag{160}$$

If it were not for the J_{23} coupling, we just need to turn off all the rf pulses and wait for a duration $\tau = \alpha/J_{12}$. In our case, however, the coupling between qubits 2 and 3 is also active, producing unwanted contribution

$$\exp\left(-iJ_{23}(\alpha/J_{12})I \otimes I_z \otimes I_z\right),$$

which must be somehow nullified. The trick is to use the identity

$$e^{i\pi I_x} I_z e^{-i\pi I_x} = \bar{X}^2 I_z X^2 = -I_z.$$

To get rid of undesirable time evolution due to J_{23}, we apply the first π-pulse $e^{-i\pi I_{x3}}$ on the third qubit at $\tau/2$ and then allow the molecule to evolve freely for another duration $\tau/2$. The extra contribution cancels out by this flipping of the third qubit. Finally apply the second π-pulse $e^{i\pi I_{x3}}$ on the third qubit so that it comes back to its correct history.

More explicitly we verify that

$$\begin{aligned} U &= \mathcal{T}\exp\left[-i\int_0^\tau \tilde{H}(t)dt\right] \\ &= (I \otimes I \otimes \bar{X})\exp\left(-i\frac{\tau}{2}\tilde{H}_0\right)(I \otimes I \otimes X)\exp\left(-i\frac{\tau}{2}\tilde{H}_0\right) \\ &= \exp\left[-i\frac{\tau}{2}\left(J_{12}I_z \otimes I_z \otimes I - J_{23}I \otimes I_z \otimes I_z\right)\right] \\ &\quad \times \exp\left[-i\frac{\tau}{2}\left(J_{12}I_z \otimes I_z \otimes I + J_{23}I \otimes I_z \otimes I_z\right)\right] \\ &= \exp\left(-i\alpha I_z \otimes I_z \otimes I\right), \end{aligned} \tag{161}$$

where $\tilde{H}_0$ is the Hamiltonian (159) without rf pulses and use has been made of the identity

$$[I_z \otimes I_z \otimes I, I \otimes I_z \otimes I_z] = 0.$$

This result shows that we can eliminate the effect of J_{23} coupling by applying a pair of π-pulses to qubit 3. This technique is called **refocusing** or **decoupling**. Refocusing is also used to cancel field inhomogeneity and reduce transverse relaxation.

D. Time-Optimal Control of NMR Quantum Computer

We have implemented the CNOT gate in the end of §IX C 3. Although CNOT plays a particularly important role in the universality theorem, almost any two-qubit gate not in $\mathrm{SU}(2) \otimes \mathrm{SU}(2)$ does the job, an important exception being the SWAP gate.

In the present section, we consider a general strategy to implement any two-qubit gate. Our implementation is also optimal in terms of gate execution time. Let us start with some mathematical background materials.

1. A Brief Introduction to Lie Algebras and Lie Groups

It is assumed that the reader has some familiarity with the elementary theory of Lie algebras and Lie groups, such as SO(3) and SU(2). See [65] and [66], for example.

A Lie group G is a group equipped with a structure of an analytic manifold, where the group operations $G \times G \to G, G \to G$ defined by

$$xy \mapsto xy, \;\; x \mapsto x^{-1}, \tag{162}$$

respectively, are analytic with respect to local coordinates [66, 67].

Given a Lie group G, consider the tangent space $\mathfrak{g}$ of G at the unit element $I \in G$. In other words, $\mathfrak{g}$ is nothing but the vector space $T_I(G)$, which is constructed as follows [67]. Consider a curve $c : (a, b) \to G$ such that $c(0) = I$ and $c'(0) = X_c$, where it is assumed that $a < 0 < b$ and the curve belongs to C^1 class. For each choice of c, there exists a tangent vector X_c.[84] Suppose we take all the curves that pass I at $t = 0$ and consider the set of tangent vectors $\mathfrak{g} = \{X_c | c(0) = I, c \in C^1\}$. Then the set $\mathfrak{g}$ has a structure of a vector space, in which an addition and a scalar multiplication are well defined:

$$\forall X, Y \in \mathfrak{g}, \forall c_k \in \mathbb{R} \Rightarrow c_1 X + c_2 Y \in \mathfrak{g}. \tag{163}$$

Moreover, being a tangent vector space $\mathfrak{g}$ a Lie group G, the Lie bracket is well defined too:

$$X, Y \in \mathfrak{g} \to [X, Y] \in \mathfrak{g}, \tag{164}$$

where $[X, Y] = XY - YX$ is the Lie bracket of X and Y. The vector space $\mathfrak{g}$ is called the Lie algebra associated with a Lie group G. It is common to denote the Lie algebra of a Lie group G by the corresponding lower case German letter: the Lie algebra of $\mathrm{SU}(n)$ is denoted as $\mathfrak{su}(n)$, for example. Alternatively, the exponential map $\exp : \mathfrak{g} \to G$, $X \mapsto \exp X$ maps $\mathfrak{g}$ to a component G_0 of G, which contains the unit element I. By definition, this means that $G_0 = G$ for a simply connected Lie group G.

Let us work out an example $G = \mathrm{SU}(n)$. Consider a curve $c(t) : (a, b) \to \mathrm{SU}(n)$. It satisfies $c(t)^\dagger c(t) = I$ and $\det c(t) = 1$ for any $t \in (a, b)$. There exists a vector $X \in \mathfrak{g}$ such that $c(t) = \exp(Xt)$ in the vicinity of $t \sim 0$. The vector X satisfies the corresponding conditions

$$\det e^{Xt} = \exp(\mathrm{tr}\, X)t = 1, e^{Xt} e^{X^\dagger t} = e^{(X+X^\dagger)t} = I.$$

It is found from these conditions that

$$\mathrm{tr}\, X = 0 \quad \text{and} \quad X + X^\dagger = 0, \tag{165}$$

that is, X is traceless and skew-Hermitian. Conversely, any traceless skew-Hermitian matrix X defines $U = e^{Xt}$, which satisfies $\det U = 1$ and $U^\dagger U = I$. In summary

$$\mathfrak{su}(n) = \{X \in M(n, \mathbb{C}) | \mathrm{tr}\, X = 0, X + X^\dagger = 0\}. \tag{166}$$

The set $M(n, \mathbb{C})$ of $n \times n$ complex matrices has $2n^2$ real free parameters. The conditions $X = -X^\dagger$ reduces this down to n^2. In particular, the diagonal elements d_i of X must be pure imaginary. The condition $\mathrm{tr}\, X = 0$ introduces an additional condition $\sum_i d_i = 0$, which reduces the degrees of freedom to $n^2 - 1$ and hence $\dim\ \mathfrak{su}(n) = n^2 - 1$. Let X_k $(1 \le k \le n^2 - 1)$ be the generators of $\mathfrak{su}(n)$. Any element $U \in \mathrm{SU}(n)$ is then expressed as

$$U = \exp\left(\sum_{k=1}^{n^2-1} \alpha_k X_k \right). \tag{167}$$

For $\mathrm{SU}(2)$, for example, the vector space $\mathfrak{su}(2)$ is spanned by three traceless anti-Hermitian matrices, which we often take $i\sigma_k$ $(k = x, y, z)$.

We note that the condition $\det U = 1$ does not apply for $U \in \mathrm{U}(n)$, and accordingly the corresponding Lie algebra is

$$\mathfrak{u}(n) = \{X \in M(n, \mathbb{C}) | X + X^\dagger = 0\}, \tag{168}$$

for which $\dim\ \mathfrak{u}(n) = n^2$.

It is convenient to take the set of generators of $\mathrm{U}(2^n)$ as

$$I_{k_1} \otimes I_{k_2} \otimes \ldots \otimes I_{k_n}, \tag{169}$$

where $I_k \in \{I, I_x, I_y, I_z\}$. The generator $I \otimes I \otimes \ldots \otimes I$ must be excluded as a generator of $\mathfrak{su}(2^n)$ since it does not satisfy the traceless condition. In this way, we find there are $4^n - 1$ generators for $\mathfrak{su}(2^n)$.

Example IX.7 *Generators of* $\mathfrak{su}(2^2)$ *are*

$$I_k \otimes I,\ I \otimes I_k,\ I_j \otimes I_k \quad (j, k = x, y, z).$$

Observe that there are $3 + 3 + 9 = 4^2 - 1$ *generators.*

2. Cartan Decomposition and Optimal Implementation of Two-Qubit Gates

We have seen in the preceeding sections that one-qubit operation takes a short time on the order $10\,\mu$s for a heteronucleus molecule, while a two-qubit entangling operation takes time typically $\sim 1/J \sim 10\,$ms. Therefore one-qubit operation time may be neglected in estimating the total execution time of a quantum algorithm [68]. Let us consider a molecule with two heteronucleus spins for definiteness, whose Hamiltonian, in the rotating frame with

respective Larmor frequency, is given in Eq. (140), in which $\omega_{1,i}$ and ϕ_i are control parameters. Typically we have $\omega_{1,i} \gg J$, which justifies the above assumption of negligible one-qubit operation time compared to two-qubit operation time. This Hamiltonian generates a unitary matrix $U_{\rm alg} \in {\rm SU}(4)$ via the time-evolution equation

$$U_{\rm alg} = \mathcal{T} e^{-i \int_0^\tau \tilde{H}(t) dt}. \tag{170}$$

One may naively think that the path providing the shortest execution time corresponds to the shortest path connecting the unit matrix I (at $t = 0$) and $U_{\rm alg}$ at $t = T$. Note however that the one-qubit operation time is negligible and we may use one-qubit gates as many times as necessary. Thus we may identify $U_1, U_2 \in {\rm SU}(4)$ which differ by an element of $K \equiv {\rm SU}(2) \otimes {\rm SU}(2)$. This means that the relevant space for evaluating the time-optimal path is the coset space ${\rm SU}(4)/{\rm SU}(2) \otimes {\rm SU}(2)$ in which U_1 and $U_2 = KU_1$ are identified. To find the time-optimal path connecting the unit matrix I and the matrix $U_{\rm alg}$, therefore, amounts to finding the time-optimal path connecting cosets $[I]$ and $[U_{\rm alg}]$, where $[U] \equiv \{kU | k \in K\}$. The Lie algebra $\mathfrak{su}(4)$ is decomposed as $\mathfrak{su}(4) = \mathfrak{k} \oplus \mathfrak{p}$ [68–70], where

$$\mathfrak{k} = {\rm Span}(\{iI \otimes I_k,\, iI_k \otimes I\}), \quad (k = x, y, z), \tag{171}$$
$$\mathfrak{p} = \mathfrak{k}^\perp = {\rm Span}(\{iI_j \otimes I_k\}), \quad (j, k = x, y, z). \tag{172}$$

They satisfy the commutation relations

$$[\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k}, \qquad [\mathfrak{p}, \mathfrak{k}] \subset \mathfrak{p}, \qquad [\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k}. \tag{173}$$

Decomposition of a Lie algebra $\mathfrak{g}$ into $\mathfrak{k}$ and $\mathfrak{p}$, satisfying the above commutation relations, is called a **Cartan decomposition**. The **Cartan subalgebra** $\mathfrak{h} = {\rm Span}(\{iI_j \otimes I_j\}) \subset \mathfrak{p}$ plays an important role in our construction. A general theorem of Lie algebras proves that any element $U_{\rm alg} \in {\rm SU}(4)$ has a KP decomposition $U_{\rm alg} = kp$ with $k \in K \equiv \exp \mathfrak{k}$ and $p \in P \equiv \exp \mathfrak{p}$. Moreover, any matrix $p \in P$ is rewritten in a conjugate form $p = k_1^\dagger h k_1$, where $k_1 \in K$ and h is an element of the **Cartan subgroup** H of ${\rm SU}(4)$ defined as

$$H \equiv \exp \mathfrak{h} = \Big\{ \exp\Big(i \sum_{j=x,y,z} \alpha_j I_j \otimes I_j\Big) \Big| \, \alpha_j \in \mathbb{R} \Big\}. \tag{174}$$

Therefore we have a corresponding Cartan decomposition of a group element as $U_{\rm alg} = kp = kk_1^\dagger h k_1 = k_2 h k_1$, where $k_i \in K$, $h \in H$ and $k_2 = kk_1^\dagger$. The quantum algorithm $U_{\rm alg}$ is now decomposed into one-qubit operations k_1, k_2 and a two-qubit entangling operation h. This decomposition determines an optimized pulse sequence of the NMR quantum computer as discussed in [68–70].

Cartan decomposition of an arbitrary $U \in {\rm SU}(4)$ proceeds explicitly as follows. We take the magic basis [71] defined as

$$\begin{aligned}
|\Psi_0\rangle &= \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \\
|\Psi_1\rangle &= \frac{i}{\sqrt{2}}(|01\rangle + |10\rangle), \\
|\Psi_2\rangle &= \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle), \\
|\Psi_3\rangle &= \frac{i}{\sqrt{2}}(|00\rangle - |11\rangle),
\end{aligned} \tag{175}$$

which is different from an ordinary Bell basis by phase. The transformation rule of a matrix U with respect to the standard binary basis $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ into that with the magic basis $\{|\Psi_i\rangle\}$ is $U \to U_B \equiv Q^\dagger U Q$, where

$$Q = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 & i \\ 0 & i & 1 & 0 \\ 0 & i & -1 & 0 \\ 1 & 0 & 0 & -i \end{pmatrix}. \tag{176}$$

The matrix Q defines an isomorphism (1:1 linear map preserving the group product) between $K = {\rm SU}(2) \otimes {\rm SU}(2)$ and ${\rm SO}(4)$ and is used to classify two-qubit gates [69, 71]. In fact, it is easy to verify that $Q^\dagger k Q$ is an element of

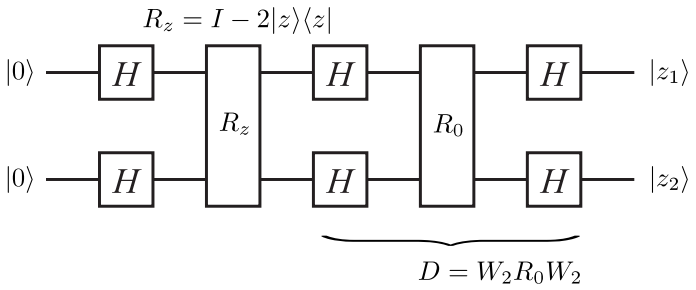


FIG. 14: Implementation of the Grover database search algorithm for $n = 2$ qubits case. H is the Hadamard gate, $W_2 = U_{\rm H}^{\otimes 2}$, while $R_z = I - 2|z\rangle\langle z|$.

SO(4) for $k \in K$. Moreover, Q diagonalizes elements of the Cartan subgroup, viz $Q^\dagger h Q = \mathrm{diag}(e^{i\theta_0}, e^{i\theta_1}, e^{i\theta_2}, e^{i\theta_3})$ for $h \in H$. We find for $U = k_2 h k_1$ that

$$U_B = Q^\dagger U Q = Q^\dagger k_2 Q \cdot Q^\dagger h Q \cdot Q^\dagger k_1 Q = O_2 h_D O_1,$$

where $O_i \equiv Q^\dagger k_i Q$ is an element of SO(4) and $h_D \equiv Q^\dagger h Q$ is a diagonal matrix. From $U_B^t U_B = O_1^t h_D^2 O_1$, we notice that $U_B^t U_B$ is diagonalized by O_1 and its eigenvalues form the diagonal elements of h_D^2. Finally O_2 is found as $O_2 = U_B (h_D O_1)^{-1}$.

Example IX.8 *Let us consider implementing two-qubit Grover's database search algorithm U_z as a concrete example. The data are encoded in one of the basis vectors $|00\rangle, |01\rangle, |10\rangle, |11\rangle$, and the gate U_z picks out a particular binary basis vector $|z\rangle = |ij\rangle$ as a "target file" upon acting on $|00\rangle$ [77, 78]. Figure 14 shows the actual quantum cirucit implementation of the Grover algorithm. Here H is the Hadamard gate and $R_z = I - 2|z\rangle\langle z|$, $R_0 = I - 2|0\rangle\langle 0|$, cf Fig.* **??**. *Here we do not explicitly give oracle cirucits R_z and R_0, but they are treated as black boxes.*

Here we consider U_{10} which picks out the file $|10\rangle$ with a single step. The unitary matrix representing this algorithm takes the form

$$U_{10} = W_2 R_0 W_2 R_{10} W_2 = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix}. \tag{177}$$

We apply the above strategy and find the Cartan decomposition of $U_{10} = k_2 h k_1$ as

$$\begin{aligned} k_1 &= I_2 \otimes I_2, \\ h &= e^{i\pi(I_x \otimes I_x - I_y \otimes I_y)}, \\ k_2 &= e^{-i(\pi/2) I_z} \otimes e^{i(\pi/\sqrt{2})(I_x + I_y)}. \end{aligned} \tag{178}$$

Actually, the decomposition is not unique, and we choose a solution that minimizes the execution time of an NMR quantum computer. To implement this decomposition with NMR, such terms as $e^{i\pi(I_x \otimes I_x)}$ must be rewritten in favor of the subset of generators of SU(4) contained in the Hamiltonian (140). We verify, for example, that

$$\begin{aligned} & e^{i\pi(I_x \otimes I_x)} \\ &= [e^{-i(\pi/2) I_y} \otimes e^{i(\pi/2) I_y}] \cdot e^{-i\pi(I_z \otimes I_z)} \cdot [e^{i(\pi/2) I_y} \otimes e^{-i(\pi/2) I_y}]. \end{aligned} \tag{179}$$

Table V shows the pulse sequence to implement U_{10} with an NMR quantum computer. We call the hydrogen nucleus and the carbon nucleus qubit 1 and qubit 2, respectively. The time-optimal path requires the execution time of $1/J$, which happens to be the same as that for the conventional pulse sequence [79].

Exercise IX.9 *Find a Cartan decomposition of the controlled-Z gate*

$$U_Z = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes \sigma_z$$

and implement the solution with an NMR pulse sequence which implements the decomposition.

Exercise IX.10 *Find a Cartan decomposition of $U_{\rm CNOT}$. Find an NMR pulse sequence which implements the decomposition.*

TABLE V: Time-optimal pulse sequences for Grover's algorithm U_{10}. The number 1 (2) denotes the first (second) qubit. Here X ($\bar{X}$) and Y ($\bar{Y}$) denote $\pi/2$-pulse around x ($-x$) and y ($-y$) axis, respectively. The symbol Pi(θ) denotes a π-pulse around a direction $(\cos\theta, \sin\theta, 0)$ in the Bloch sphere. The symbol $(1/2J)$ indicates the length of the idle time, during which no external pulses are applied.

	Pulse sequence	Execution time
1:	X $(1/2J)$ Xm Y $(1/2J)$ X Ym	$1/J$
2:	X $(1/2J)$ Xm Ym $(1/2J)$ Y Pi($\pi/4$)	

E. DiVincenzo Criteria

DiVincenzo criteria for an NMR quantum computer are evaluated as follows.

1. A scalable physical system with well-characterized qubits:

   Spin 1/2 nuclei in a molecule are used as qubits. They cannot be cooled down to ultralow temperature since a molecule must be solved in a liquid to simplify nucleus-nucleus interaction. Selective addressing to each spin is possible by taking advantage of Larmor frequency differences. Chemical shifts among the same nuclear spices make it possible to access spins selectively even in a homonucleus molecule. However, selective addressing becomes harder and harder as the number of the same nuclei grows. The initialization outlined in §**??** is also difficult for a large number of spins, and the estimated upper bound in the number of qubits in an NMR quantum computer is ~ 10.

2. The ability to initialize the state of the qubits to a simple fiducial state, such as $|00\ldots0\rangle$:

   Molecules in a liquid solvent at room temperature are in a thermal equilibrium state, which is quite close to the uniform mixture of all possible spin states. Since any unitary transformation cannot map a mixed state to a pure state, we need to employ nonunitary operations, such as temporal averaging, spatial averaging or logical labelling, to prepare a pseudopure state $|00\ldots0\rangle$, for example. The number of steps (the number of pulses, say) required to prepare the pseudopure state diverges exponentially as a function of the number of qubits n. The number of steps cannot be too large for an NMR quantum computer because of a finite decoherence time. The maximum number of qubits is estimated to be limited to ~ 10 from this viewpoint as well.

3. Long decoherence times, much longer than the gate operation time:

   Decoherence time depends on the molecule employed as a quantum computer. It may be as large as $10^2 \sim 10^3$ s. Single-qubit gate operation time can be as short as $\sim 10^{-5}$ s, while two-qubit gate operation time, making use of the J-couplings, takes $\sim 10^{-2} \sim 10^{-1}$ s. It has been shown in [80] that a faithful implementation of Shor's algorithm providing the factorization $21 = 3 \times 7$ requires approximately 10^5 gate operations, among which $\sim 10^4$ are two-qubit gate operations. Therefore we need at least $\sim 10^{-2} \times 10^4 = 10^2$ s decoherence time to execute this modest factorization.

4. A "univeral" set of quantum gates:

   One-qubit operations are implemented with rf pulses, by making use of the Rabi oscillations. Two-qubit operations are realized by using the J-coupling between nuclei. Some important two-qubit gates, such as the CNOT gate and the SWAP gate, are realized. In fact, a simplified version of Shor's factorization algorithm has already been demonstrated [57].

5. A qubit-specific measurement capability:

   Measurement of qubit states with the free induction decay (FID) is a well-established measurement technique in NMR, having several decades of history. It is also possible to measure the density matrix itself (quantum state tomography) and the unitary gate (quantum process tomography) within the current technology. However, the signal to noise ratio scales as ne^{-an}, $a \sim 1$ being a constant, and the readout becomes more and more difficult as the number of qubits grows.

In addition to the difficulties listed above, thermal density matrix at room temperature is not entangled, and it is often criticized that NMR is not a true quantum computer. However, it works as a simulator to a real quantum computer on which we can execute quantum algorithms. Several important techniques have been developed from these standpoints in the past. It should also be addressed that NMR is the only quantum computer which is commercially

available. An NMR quantum computer is expected to remain an important tool to develop various techniques necessary to materialize a real working quantum computer to come.

Acknowledgements

I would like to thank Professors Jinchuan Hou and Chi-Kwong Li for giving me the opportunity to lecture at Taiyuan Summer School 2011. I would like to thank Akira SaiToh for discussions. My research is partly supported by "Open Research Center" Project for Private Universities: matching fund subsidy from MEXT (Ministry of Education, Culture, Sports, Science and Technology) and Grants-in-Aid for Scientific Research from the JSPS (Grant No. 23540470).

---

[1] M. Nakahara and T. Ohmi, *Quantum Computing: From Linear Algebra to Physical Realizations*, (Taylor and Francis, 2008).
[2] M. A. Neilsen and I. L. Chuang, *Quantum Computation and Quantum Information*, (Cambridge University Press, 2000).
[3] E. Rieffel and W. Polak, *ACM Computing Surveys (CSUR)* **32** (2000) 300.
[4] Y. Uesaka, *Mathematical Principle of Quantum Computation*, (Corona Publishing, Tokyo, in Japanese, 2000).
[5] P. A. M. Dirac, *Principles of Quantum Mechanics* (4th ed.), (Clarendon Press, 1981).
[6] L. I. Shiff, *Quantum Mechanics* (3rd ed.), (McGraw-Hill, 1968).
[7] A. Messiah, *Quantum Mechanics*, (Dover, 2000).
[8] J. J. Sakurai, *: Modern Quantum Mechanics* (2nd Edition), (Addison Wesley, Boston, 1994).
[9] L. E. Ballentine, *Quantum Mechanics*, (World Scientific, Singapore, 1998).
[10] A. Peres, *Quantum Theory: Concepts and Methods*, (Springer, 2006).
[11] A. SaiToh, R. Rahimi and M. Nakahara, Phys. Rev. A **77**, 052101 (2008)
[12] A. Peres, *Phys. Rev. Lett.* **77** (1996) 1413.
[13] M. Horodecki *et al.*, *Phys. Lett. A* **223** (1996) 1.
[14] W. K. Wootters, and W. H. Zurek, *Nature* **299** (1982) 802.
[15] M. A. Nielsen *et al.*, *Nature* **396** (1998) 52.
[16] A. Barenco *et al.*, *Phys. Rev. A* **52** (1995) 3457.
[17] Z. Meglicki, `http://beige.ucs.indiana.edu/M743/index.html`
[18] D. Deutsch, *Proc. Roy. Soc. Lond. A*, **400** (1985) 97.
[19] D. Deutsch and R. Jozsa, *Proc. Roy. Soc. Lond. A*, **439** (1992) 553.
[20] E. Bernstein and U. Vazirani, *SIAM J. Comput.*, **26** (1997) 1411.
[21] D. R. Simon, Proc. 35th Annual Sympo. Found. Comput. Science, (IEEE Comput. Soc. Press, Los Alamitos, 1994) 116.
[22] T. Mihara and S. C. Sung, *Comput. Complex.* **12** (2003) 162.
[23] M. A. Neilsen and I. L. Chuang, *Quantum Computation and Quantum Information*, (Cambridge University Press, 2000).
[24] K. Hornberger, quant-ph/0612118.
[25] H. Barnum, M. A. Nielsen and B. Schumacher, *Phys. Rev. A* **57** (1998) 4153.
[26] Y. Kondo, *et al.*, *J. Phys. Soc. Jpn.* **76** (2007) 074002.
[27] G. Lindblad, *Commun. Math. Phys.* **48** (1976) 119.
[28] V. Gorini, A. Kossakowski and E. C. G. Sudarshan, *J. Math. Phys.*, **17** (1976) 821.
[29] A. J. Fisher, Lecture note available at `http://www.cmmp.ucl.ac.uk/~ajf/course_notes.pdf`
[30] A. M. Steane, quant-ph/0304016.
[31] P. W. Shor, *Phys. Rev. A* **52** (1995) 2493.
[32] A. Hosoya, *Lectures on Quantum Computation* (Science Sha, in Japanese, 1999).
[33] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, (North-Holland, Amsterdam, 1977).
[34] A. R. Calderbank and P. W. Shor, *Phys. Rev. A* **54** (1996) 1098.
[35] A. M. Steane, *Phys. Rev. Lett.* **77** (1996) 793.
[36] J Niwa, K. Matsumoto and H. Imai, quant-ph/0211071.
[37] D. P. DiVincenzo and P. W. Shor, *Phys. Rev. Lett.* **77** (1996) 3260.
[38] D. P. DiVincenzo, *Fortschr. Phys.* **48** (2000) 771.
[39] M. Nakahara, S. Kanemitsu, M. M. Salomaa and S. Takagi (eds.) "Physical Realization of Quantum Computing: Are the DiVincenzo Criteria Fulfilled in 2004?" (World Scientific, Singapore, 2006).
[40] J. Vartiainen *et al.*, *Phys. Rev. A* **70** (2004) 012319.
[41] E. Knill, R. Laflamme, and L. Viola, Phys. Rev. Lett. **84**, 2525 (2000); P. Zanardi, *Phys. Rev. A* **63**, 012301 (2001); W. G. Ritter, *Phys. Rev. A* **72** (2005) 012305.
[42] G. M. Palma, K. A. Suominen and A. K. Ekert, *Proc. R. Soc. London A* **452** (1996) 567; L. M. Duan and G. C. Guo, *Phys. Rev. Lett.* **79** (1997) 1953; P. Zanardi and M. Rasetti, *Phys. Rev. Lett.* **79** (1997) 3306; D. A. Lidar, I. L. Chuang,

and K. B. Whaley, *Phys. Rev. Lett.* **81** (1998) 2594; P. Zanardi, *Phys. Rev. A* **60** (1999) 729(R); D. Bacon, D. A. Lidar, and K. B. Whaley, *Phys. Rev. A* **60** (1999) 1944.
[43] D. P. DiVincenzo, *Phys. Rev. A* **51** (1995) 1015.
[44] http://qist.lanl.gov/
[45] Y. Kondo, M. Nakahara and S. Tanimura, in *Physical Realizations of Quantum Computing: Are the DiVincenzo Criteria Fulfilled in 2004?*, ed. M. Nakahara *et al.*, World Scientific, Singapore (2006).
[46] L. M. K. Vandersypen and I. L. Chuang, Rev. Mod. Phys. **76**, 1037 (2004).
[47] J. Jones, in *Quantum Entanglement and Information Processing: Lecture Notes of the les Houches Summer School 2003*, eds. J.-M. Raimond, J. Dalibard and D. Esteve, Elsevier Science and Technology (2004).
[48] E. Knill, I. L. Chuang and R. Laflamme, Phys. Rev. A **57**, 3348 (1998).
[49] I. L. Chuang *et al.*, Nature **393**, 143 (1998).
[50] J. A. Jones, M. Mosca and R. H. Hansen, Nature **393**, 344 (1998).
[51] J. A. Jones and M. Mosca, Phys. Rev. Lett. **83**, 1050 (1999).
[52] D. G. Cory *et al.*, Phys. Rev. Lett. **81**, 2152 (1998).
[53] M. D. Price *et al.*, Phys. Rev. A **60**, 2777 (1999).
[54] B. R. Laflamme *et al.*, Phil. Trans. R. Soc. Lond. A **356**, 1941 (1998).
[55] L. M. K. Vandersypen *et al.*, Phys. Rev. Lett. **83**, 3085 (1999).
[56] L. M. K. Vandersypen *et al.*, Phys. Rev. Lett. **85**, 5452 (2000).
[57] L. M. K. Vandersypen *et al.*, Nature **414**, 883 (2001).
[58] R. Brüschweiler, Phys. Rev. Lett. **85**, 4815 (2000).
[59] F. Bloch and A. Siegert, Phys. Rev. **57**, 522 (1940).
[60] A. Barenco *et al.*, Phys. Rev. A **52**, 3457 (1995).
[61] J. J. Sylvester, Phil. Mag. **34**, 461 (1867).
[62] J. Hadamard, Bull. Sci. Math. **17**, 240 (1893).
[63] J. A. Jones and E. Knill, J. Magn. Reson., Ser. A **141**, 322 (1999).
[64] D. W. Leung *et al.*, Phys. Rev. A **61**, 042310 (2000).
[65] W. Rossmann, *Lie Groups*, Oxford Univ. Press, New York (2002).
[66] S. Helgason, *Differential Geometry, Lie Groups and Symmetric Spaces*, Academic Press, New York (1978).
[67] M. Nakahara, *Geometry, Topology and Physics* (2nd ed.), Taylor and Francis, Boca Raton (2003).
[68] N. Khaneja, R. Brockett and S. J. Glaser, Phys. Rev. A **63**, 032308 (2001).
[69] J. Zhang *et al.*, Phys. Rev. A **67**, 042313 (2003).
[70] M. Nakahara *et al.*, Phys. Lett. A **350**, 27 (2006).
[71] Y. Makhlin, Quant. Info. Proc. **1**, 243 (2002).
[72] E. Knill, I. L. Chuang and R. Laflamme, Phys. Rev. A **81**, 5672 (1998).
[73] U. Sakaguchi, H. Ozawa and T. Fukumi, Phys. Rev. A **61**, 042313 (2000).
[74] D. G. Cory, A. F. Fahmy and T. F. Havel, Proc. Natl. Acad. Sci. USA **94**, 1634 (1997).
[75] N. A. Gershenfeld and I. L. Chuang, Science **275**, 350 (1997).
[76] L. M. K. Vandersypen *et al.*, Phys. Rev. Lett. **83**, 3085 (1999).
[77] L. K. Grover, in *Proceedings of the 28th Annual ACM Symposium on the Theory of Computation*, 212, ACM Press, New York (1996).
[78] L. K. Grover, Phys. Rev. Lett. **79**, 325 (1997).
[79] I. L. Chuang, N. Gershenfeld and M. Kubince, Phys. Rev. Lett. **80**, 3408 (1998).
[80] J. J. Vartiainen *et al.*, Phys. Rev. A **70**, 012319 (2004).
[81] Let $m, N \in \mathbb{N}$ $(m < N)$ be numbers coprime to each other. Then there exists $P \in \mathbb{N}$ such that $m^P \equiv 1 \pmod{N}$. The smallest such number P is called the period or the order. It is easily seen that $m^{x+P} \equiv m^x \pmod{N}$, $\forall x \in \mathbb{N}$.
[82] A set S is called a semigroup if S is closed under a product satisfying associativity $(ab)c = a(bc)$. If S has a unit element e, such that $ea = ae = a, \forall a \in S$, it is called a monoid.
[83] Of course, the space of density matrices is not a linear vector space. What is meant hear is a linear operator, acting on the vector space of Hermitian matrices, also acts on the space of density matrices and it maps a density matrix to another density matrix.
[84] More formally, a tangent vector is defined as an equivalence class of curves that satisfies the conditions (162); see [67] for example.