3

Density Matrices and Quantum Operations

In this chapter, we introduce the concepts and properties of general quantum states - mixed states, and also the concept of quantum operations on an open system. Using these, we can present the general strategy of quantum computing, and possible applications to other problems such as quantum communication and quantum sensing. While the discussion can apply to infinite dimensional system, we will focus on the finite dimensional cases so that the Hilbert spaces under considerations are identified with the set of complex column vectors.

3.1 Mixed States and Density Matrices

Recall that if a quantum state is represented by a unit vector $|\psi\rangle \in \mathbb{C}^n$, it is convenient to represent it as a density matrix $|\psi\rangle\langle\psi|$ because $|\psi\rangle$ and $|\phi\rangle$ will generate the same density matrix if and only if $|\psi\rangle = e^{i\alpha}|\phi\rangle$ for some $\alpha \in \mathbb{R}$ so that they represent the same system. It might happen in some cases that a quantum system under consideration 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. This random nature should not be confused with a probabilistic behavior of a quantum 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 $p_j = 0$ $(j \neq i)$.

Mixed states arise naturally in physical systems, for example.

- Suppose we observe a beam of totally unpolarized light and measure whether photons are polarized vertically or horizontally. The measurement outcome of a particular photon is *either* horizontal *or* vertical. Therefore when the beam passes through a linear polarizer, the intensity is halved. The beam is a uniform mixture of horizontally polarized photons and vertically polarized photons.

- A particle source emits a particle in a state $|\psi_i\rangle$ with a probability p_i $(1 \leq i \leq N)$.

- Consider an ensemble of identical closed systems in contact with a reservoir with temperature T. If we pick up one of the members in the ensemble, it is in a state $|\psi_i\rangle$ with energy E_i with a probability $p_i = e^{-E_i/k_BT}/Z(T)$, where $Z(T) = \text{Tr}e^{-H/k_BT}$ is the partition function.

In each of these examples, a particular state $|\psi_i\rangle \in \mathcal{H} \equiv \mathbb{C}^n$ appears with probability p_i, in which case the expectation value of the observable a is $\langle\psi_i|A|\psi_i\rangle$, where we assume $|\psi_i\rangle$ is normalized; $\langle\psi_i|\psi_i\rangle = 1$. The mean value of a is then given by

$$\langle A\rangle = \sum_{i=1}^{N} p_i \langle\psi_i|A|\psi_i\rangle, \tag{3.1}$$

where N is the number of available states. Let us introduce the **density matrix** (**operator**) by

$$\rho = \sum_{i=1}^{N} p_i |\psi_i\rangle\langle\psi_i|. \tag{3.2}$$

Note that $Z \mapsto \text{Tr}\,(Z)$ is a linear function on $\mathbf{M}_n$, and for $X = (x_{rs}) \in \mathbf{M}_{m,n}$ and $Y = (y_{uv}) \in \mathbf{M}_{n,m}$, we have

$$\text{Tr}\,(XY) = \sum_{r,s} x_{rs} y_{sr} = \text{Tr}\,(YX).$$

Thus, Eq. (3.1) is rewritten in a compact form as

$$\langle A\rangle = \text{Tr}(\rho A). \tag{3.3}$$

because the left hand side equals

$$\begin{aligned}\sum_{i=1}^{N} p_j \langle\psi_j|A|\psi_j\rangle &= \sum_{j=1}^{N} p_j \text{Tr}\,[\langle\psi_j|A|\psi_j\rangle] = \sum_{j=1}^{N} p_j \text{Tr}\,(|\psi_j\rangle\langle\psi_j|A) \\ &= \text{Tr}\,[(\sum_{j=1}^{N} p_j (|\psi_j\rangle\langle\psi_j|)A] = \text{Tr}\,(\rho A).\end{aligned}$$

Here we use the fact that $\text{Tr}\,[\mu] = \mu \in \mathbb{C}$ for the first equality, and $\text{Tr}\,(XY) = \text{Tr}\,(YX)$ for the second equality.

Suppose we have two mixed states obtained by convex combinations of pure states $|\psi_1\rangle\langle\psi_1|, \ldots, |\psi_r\rangle\langle\psi_r|$ and $|\phi_1\rangle\langle\phi_1|, \ldots, |\phi_s\rangle\langle\phi_s|$ and end up as the same density matrix ρ, i.e., for two probability vectors $(p_1, \ldots, p_r), (q_1, \ldots, q_s)$,

$$\rho = \sum_{j}^{r} p_j |\psi_j\rangle\langle\psi_j| = \sum_{j=1}^{s} q_j |\phi_j\rangle\langle\phi_j|.$$

Then the mean value for any observable will be the same. So, we will not be able to distinguish the two states; practically, they are the same. So, we

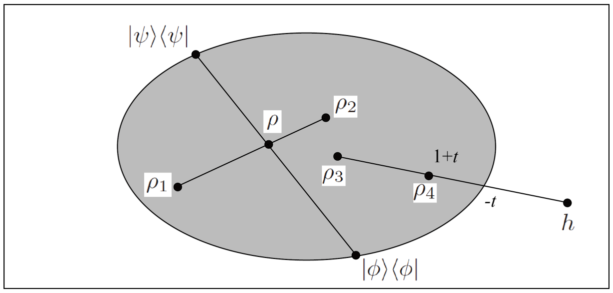

FIGURE 3.1
Space of density matrices $\mathbf{D}_n$ (gray oval) as a subset of unit-trace Hermitian matrices (white square). The boundary of $\mathbf{D}_n$ represents pure states while the interior represents mixed states that are not pure.

will focus only on the density matrices when we consider a system in mixed states. We will denote by $\mathbf{D}_n$ the set of density matrices in $\mathbf{M}_n$.

Figure 3.1 depicts the space of density matrices $\mathbf{D}_n$ (the gray oval) which is a subset of unit-trace Hermitian matrices (the white square). The state ρ is represented by a mixture $\rho = (1-t)|\psi\rangle\langle\psi| + t|\phi\rangle\langle\phi|$, where $t \in [0,1]$, of two pure states as well as a mixture of two mixed states $\rho = (1-s)\rho_1 + s\rho_2$. A general unit-trace Hermitian matrix h may be also represented as $h = (1+t)\rho_3 - t\rho_4$ but the coefficient must be negative in this case.

EXAMPLE 3.1.1. *Let* $p_1 = p_2 = 1/2$ *and* $|\psi_1\rangle = (1,0)^t, |\psi_2\rangle = (1,1)^t/\sqrt{2}$. *The density matrix of this state is*

$$\rho = \frac{1}{4}\begin{pmatrix} 3 & 1 \\ 1 & 1 \end{pmatrix}.$$

This density matrix also represents a mixed state with $q_1 = \left(2+\sqrt{2}\right)/4, q_2 = \left(2-\sqrt{2}\right)/4$ *and* $|\phi_1\rangle = (1+\sqrt{2},1)^t/\sqrt{2\left(2+\sqrt{2}\right)}, |\phi_2\rangle = (1-\sqrt{2},1)^t/\sqrt{2(2-\sqrt{2})}$. *This mixture is obtained by the spectral decomposition of* ρ *and hence* $\langle\phi_1|\phi_2\rangle = 0$.

We may also decompose ρ *into two mixed states as* $\rho = \frac{1}{2}\rho_1 + \frac{3}{4}\rho_2$, *where*

$$\rho_1 = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \rho_2 = \frac{1}{6}\begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix}.$$

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

*The postulates work for trace class operators in infinite dimensional Hilbert space $\mathcal{H}$.

A1′ A physical state of a system, whose Hilbert space is $\mathbb{C}^n$, is completely specified by its associated density matrix $\rho : \mathbb{C}^n \to \mathbb{C}^n$. A density matrix is a positive semi-definite Hermitian operator with $\operatorname{Tr}\rho = 1$ (see remarks below).

A2′ The mean value of an observable a is given by

$$\langle A \rangle = \operatorname{Tr}(\rho A). \tag{3.4}$$

A3′ The temporal evolution of the density matrix of a closed system is given by the **Liouville-von Neumann equation**,

$$i\hbar \frac{d}{dt}\rho = [H, \rho], \tag{3.5}$$

where H is the system Hamiltonian (see remarks below).

Several remarks are in order.

- The set $\{|\psi_1\rangle, \ldots, |\psi_N\rangle\rangle\}$ associated with the density matrix (3.2) may not be orthonormal although $\langle\psi_j|\psi_j\rangle = 1$ for each j. Nevertheless, ρ is Hermitian since $p_i \geq 0$, and it is positive semi-definite

$$\langle\phi|\rho|\phi\rangle = \sum_i p_i \langle\phi|\psi_i\rangle\langle\psi_i|\phi\rangle = \sum_i p_i |\langle\psi_i|\phi\rangle|^2 \geq 0.$$

  We also have

$$\begin{aligned}\operatorname{Tr}\rho &= \sum_k \langle e_k|\rho|e_k\rangle = \sum_{i,k} \langle e_k|p_i|\psi_i\rangle\langle\psi_i|e_k\rangle \\ &= \sum_i p_i \langle\psi_i| \left(\sum_k |e_k\rangle\langle e_k|\right) |\psi_i\rangle = \sum_i p_i \langle\psi_i|\psi_i\rangle = 1,\end{aligned}$$

  where $\{|e_k\rangle\}$ is an orthonormal basis of $\mathcal{H}$.

- 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 equalities as

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

It is easy to verify that $\mathbf{D}_n$ is a convex set, i.e., $r\rho_1 + (1-r)\rho_2$ with $r \in [0,1]$ for $\rho_{1,2} \in \mathbf{D}_n$ is also a density matrix.

Note that one can always do a spectral decomposition $\rho = \sum_{j=1}^n \lambda_j |\lambda_j\rangle\langle\lambda_j|$ so that it is a convex combination of its eigenprojections.

EXAMPLE 3.1.2. *A pure state $|\psi\rangle$ is a special case in which the corresponding density matrix is*

$$\rho = |\psi\rangle\langle\psi|. \tag{3.6}$$

Therefore ρ in this case is nothing but the projection operator onto the state $|\psi\rangle$. Observe that

$$\langle A\rangle = \mathrm{Tr}\,\rho A = \sum_i \langle e_i|\psi\rangle\langle\psi|A|e_i\rangle = \sum_i \langle\psi|A|e_i\rangle\langle e_i|\psi\rangle = \langle\psi|A|\psi\rangle.$$

Let us consider a beam of photons. We take a horizontally polarized state $|0\rangle = |\leftrightarrow\rangle$ and a vertically polarized state $|1\rangle = |\updownarrow\rangle$ as orthonormal basis vectors. If the photons are a totally uniform mixture of two polarized states, the density matrix is given by

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

This state is a uniform mixture of $|\updownarrow\rangle$ and $|\leftrightarrow\rangle$ and called a **maximally or uniformly mixed state**.

If photons are in a pure state $|\psi\rangle = (|0\ \rangle + |1\rangle)/\sqrt{2}$, the density matrix, with $\{|0\rangle, |1\rangle\}$ as basis, is

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

If $|\psi\rangle$ itself is used as a basis vector, the other vector being $|\phi\rangle = (|0\ \rangle - |1\rangle)/\sqrt{2}$, the density matrix with respect to the basis $\{|\psi\rangle, |\phi\rangle\}$ has a component expression

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

Verify that they all satisfy Hermiticity, positive semi-definitness and $\mathrm{Tr}\,\rho = 1$.

More generally, if we let $|x\rangle = (x_1, \ldots, x_n)^t$ and $|y\rangle = (y_1, \ldots, y_n)^t$ be a pair of orthonormal vectors in $\mathbb{C}^n$ so that $|v_j\rangle = (x_j, y_j)^t \neq 0$ for every j. Let $\rho_j = \frac{1}{p_j}|v_j\rangle\langle v_j|$ with $p_j = \||v_j\rangle\|^2$ for $j = 1, \ldots, n$. Then $\rho = \sum_{j=1}^n p_j^2 \rho_j = \frac{1}{2}I_2$ is the maximally mixed state.

Let $A = \sum_a \lambda_a |\lambda_a\rangle\langle\lambda_a|$ be the spectral decomposition of an observable A and let $\rho = \sum_i p_i|\psi_i\rangle\langle\psi_i|$ be an arbitrary state. Then the measurement outcome of A is λ_a with the probability

$$p(a) = \sum_i p_i |\langle\lambda_a|\psi_i\rangle|^2 = \langle\lambda_a|\rho|\lambda_a\rangle = \mathrm{Tr}\,(P_a\rho), \tag{3.7}$$

where $P_a = |\lambda_a\rangle\langle\lambda_a|$ is the projection operator. The state changes to a pure state $|\lambda_a\rangle\langle\lambda_a|$ immediately after the measurement with the outcome λ_a. This change is written as $\rho \mapsto P_a\rho P_a/p(a)$.

The following determines when ρ represents a pure state.

THEOREM 3.1.3. *A state $\rho \in \mathbf{D}_n$ is pure if and only if any one of the following condition holds.*

$$\text{(a)}\quad \rho^2 = \rho. \qquad \text{(b)}\quad Tr\,\rho^2 = 1.$$

Proof. Suppose $\rho = |\psi\rangle\langle\psi|$ is a pure state. Then $\rho^2 = (|\psi\rangle\langle\psi|)(|\psi\rangle\langle\psi|) = |\psi\rangle\langle\psi| = \rho$. Thus, the condition (a) holds. If (a) holds, then $\mathrm{Tr}\,\rho^2 = \mathrm{Tr}\,\rho = 1$. Thus, the condition (b) holds. If (b) holds, and $\rho = \sum_{j=1}^n \lambda_j|\lambda_j\rangle\langle\lambda_j|$, where $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$ and $\sum_{j=1}^n \lambda_j = 1$. Then $\rho^2 = \sum_{j=1}^n \lambda_j^2|\lambda_j\rangle\langle\lambda_j|$ has eigenvalues $\lambda_1^2, \ldots, \lambda_n^2$. So, if $\mathrm{Tr}\,\rho^2 = 1 = \mathrm{Tr}\,\rho$, then $0 = \sum_{j=1}^n(\lambda_j - \lambda_j^2) = \sum_{j=1}^n \lambda_j(1-\lambda_j)$ so that all the nonnegative numbers $\lambda_j(1-\lambda_j)$ is zero. Thus, $\lambda_j \in \{0,1\}$. Since $\sum_{j=1}^n \lambda_j = 1$, we see that $\lambda_1 = 1$ and $\lambda_j = 0$ for $j > 1$. Thus, $\rho = |\lambda_1\rangle\langle\lambda_1|$ is a pure state. ▮

3.2 Uncorrelated, separable and inseparable states

We classify mixed states of a multipartite system into three classes, namely, uncorrelated, separable and inseparable states; see the definition below. We use a bipartite system in the definition, but generalization to multipartitle systems should be obvious. Recall that $|\psi\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \in \mathbb{C}^m \otimes \mathbb{C}^n$ is a tensor product state in the bipartite system $\mathbb{C}^m \otimes \mathbb{C}^n$. We have $|\psi\rangle\langle\psi| = |\psi_1\rangle\langle\psi_1| \otimes |\psi_2\rangle\langle\psi_2| \in \mathbf{M}_m \otimes \mathbf{M}_n$. More generally, if $\rho_1 \in \mathbf{D}_m$ has a spectral decomposition $\rho_1 = \sum_{j=1}^m \lambda_j|\lambda_j\rangle\langle\lambda_j|$ and $\rho_2 \in \mathbf{D}_n$ has a spectral decomposition $\rho_2 = \sum_{j=1}^n \mu_j|\mu_j\rangle\langle\mu_j|$, then

$$\rho_1 \otimes \rho_2 = \sum_{r,s} \lambda_r\mu_s|\lambda_r\mu_s\rangle\langle\lambda_r\mu_s\rangle.$$

Note that $\mathbf{M}_m \otimes \mathbf{M}_n \equiv \mathbf{M}_N$ whenever $mn = N$. For $\rho \in \mathbf{D}_{mn}$, we will write $\rho \in \mathbf{M}_m \otimes \mathbf{M}_n$ to emphasize that ρ is in the bipartite system composed of subsystems with (mixed) states in $\mathbf{M}_m$ and $\mathbf{M}_n$.

DEFINITION 3.2.1. *A state $\rho \in \mathbf{M}_m \otimes \mathbf{M}_n$ is called* **uncorrelated** *if it is written as*

$$\rho = \rho_1 \otimes \rho_2, \tag{3.8}$$

with $\rho_1 \in \mathbf{D}_m, \rho_2 \in \mathbf{D}_n$. It is called **separable** *if it is written in the form*

$$\rho = \sum_j p_j\rho_{1,j} \otimes \rho_{2,j}, \quad \rho_{1,j} \in \mathbf{D}_m, \rho_{2,j} \in \mathbf{D}_n, \tag{3.9}$$

where $0 \leq p_i \leq 1$ *and* $\sum_i p_i = 1$*. It is called* **inseparable** *if* ρ *does not admit the decompostion (3.9),*

It is important to note that while not every density matrix $\rho \in \mathbf{D}_{mn}$ is separable, i.e., a convex combination of uncorrelated quantum states $\rho_{1,j} \otimes \rho_{2,j}$, it is always possible to express ρ as a real linear combination of uncorrelated quantum states with negative coefficients. See Exercise 3.9.

It is also worthwhile to realize that only inseparable states have quantum correlations analogous to that of an entangled pure state. However, it does not necessarily imply separable states have no non-classical correlation. It was pointed out that useful non-classical correlation exists in the subset of separable states [24].

It is easy to determine whether a given mixed state $\rho \in \mathbf{M}_m \otimes \mathbf{M}_n$ is an uncorrelated state. Let $\rho = (P_{rs})_{1 \leq r,s \leq m}$, where $P_{rs} \in M_n$. $\sigma_2 = P_{jj}/\mathrm{Tr}\,(P_{jj})$ for any $P_{jj} \neq 0$. Then ρ is a uncorrelated state if and only if $P_{rs} = a_{rs}\sigma_2$ with $a_{rs} \in \mathbb{C}$ for all $1 \leq r, s \leq m$. If a_{rs} does exist for all $1 \leq r, s \leq m$, then $\rho = \sigma_1 \otimes \sigma_2$ with $\sigma_1 = (a_{rs}) \in \mathbf{M}_m$. See Exercise 3.1

3.3 Partial Trace and Purification

Let $A \in \mathbf{M}_{mn} \equiv \mathbf{M}_m \otimes \mathbf{M}_n$. The **partial trace** of A over $\mathbb{C}^n$ is matrix in $\mathbf{M}_m$ defined as

$$A_1 = \mathrm{Tr}\,_2 A \equiv \sum_k (I \otimes \langle f_k|) A (I \otimes |f_k\rangle) \tag{3.10}$$

where $\{f_1, \ldots, f_n\}$ is an orthonormal basis for $\mathbb{C}^n$. It is easy to see that $\mathrm{Tr}\,_2(A_1 \otimes A_2) = (\mathrm{Tr}\, A_2)A_1$, and $\mathrm{Tr}\,_2$ is the unique linear map from $\mathbf{M}_{mn}$ to $\mathbf{M}_m$ satisfying $A_1 \otimes A_2 \mapsto (\mathrm{Tr}\, A_2)A_1$. As a result, if $A = \sum c_j A_{1,j} \otimes A_{2,j}$ for some $A_{1,j} \in \mathbf{M}_m, A_{2,j} \in \mathbf{M}_n$ and $c_j \in \mathbb{C}$, then $\mathrm{Tr}\,_2(A) = \sum_j c_j (\mathrm{Tr}\, A_{2,j}) A_{1,j}$. Consequently, $\mathrm{Tr}\,_2(A)$ is the same for any choice of orthonormal basis, see Exercise 3.6.

Suppose ρ is a bipartite state. If we are interested only in the first system and have no access to the second system, then the partial trace $\mathrm{Tr}\,_2(\rho)$ allows us to "forget" about the second system. In other words, the partial trace quantifies our ignorance of the second system.

The situation is particularly interesting when $\rho = |\psi\rangle\langle\psi| \in \mathbf{M}_{mn}$ is a density matrix of a pure state $|\psi\rangle$. To be more concrete, let us consider the 2-qubit state

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

The corresponding density matrix is

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

The partial trace of ρ over the second Hilbert space is

$$\rho_1 = \mathrm{Tr}_2\rho = \sum_{i=0,1}(I \otimes \langle i|)\rho(I \otimes |i\rangle) = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{3.11}$$

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

Observe that

$$\mathrm{Tr}\,(\rho_1 A) = \mathrm{Tr}\,(\rho(A \otimes I)) \tag{3.12}$$

for an observable A acting on the first Hilbert space. The expectation value of $A \otimes I$ under that state ρ is equally obtained by using 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 a pure state is called the **purification**. Let $\rho = \sum_{k=1}^{s} p_k|\psi_k\rangle\langle\psi_k| \in \mathbf{D}_m$, where $|\psi_k\rangle\langle\psi_k|$ is a pure state for each k. Then we can let

$$|\Psi\rangle = \sum_{k=1}^{s}\sqrt{p_k}|\psi_k\rangle \otimes |e_k\rangle, \tag{3.13}$$

where $\{|e_1\rangle, \ldots, |e_s\rangle\}$ is an orthonormal basis for $\mathbb{C}^s$. We find

$$\begin{aligned}\mathrm{Tr}_2|\Psi\rangle\langle\Psi| &= (I \otimes \langle e_i|)\left[\sqrt{p_j p_k}|\psi_j\rangle|e_j\rangle\langle\psi_k|\langle e_k|\right](I \otimes |e_i\rangle) \\ &= \sum_k p_k|\psi_k\rangle\langle\psi_k| = \rho. \end{aligned}\tag{3.14}$$

Now, for any mixed state $\rho \in \mathbf{M}_m$, one can always use its spectral decomposition $\rho = \sum_{j=1}^{r}\lambda_j|\lambda_j\rangle\langle\lambda_j|$, where r is the number of positive eigenvalues of ρ. Thus it is always possible to purify a mixed state by tensoring an extra Hilbert space of dimension equal to the rank of ρ. It is easy to see, by construction, that purification is far from unique. In fact, there are an infinite number of purifications of a given mixed state density matrix; again see Exercise 3.6.

3.4 Quantum Operations for open systems

3.4.1 Operator sum representation and Kraus operators

Recall that a quantum operation on a closed quantum system has the form $\Phi : \mathbf{M}_n \to \mathbf{M}_n$ such that

$$\Phi(A) = UAU^\dagger \qquad \text{for all } A \in \mathbf{M}_n.$$

Here, the unitary matrix U may be time-dependent, i.e., $U = U(t)$. In general, one has to consider open quantum systems, i.e., quantum systems interacting with other quantum systems; see [2] and [9] for some general background. This is unavoidable because of the following.

1. The quantum system will always interact with the environment.
2. In quantum computing one often introduce auxiliary system to the process to help the computing process.

As a result, if one considers a quantum state $\rho \in \mathbf{D}_n$ corresponding to the principal system, i.e., the system one is interested in, there is always another quantum state, say, $\sigma \in \mathbf{D}_k$ corresponding to the environment or the auxiliary system, such that $\sigma \otimes \rho$ is regarded as the initial state of the total system, which is a closed system. Then the evolution and transformation of the bipartite system will be described by unitary similarity transforms

$$\sigma \otimes \rho \mapsto U(\sigma \otimes \rho)U^\dagger.$$

However, one cannot or need not have complete control of the environment (auxiliary) system. So, one will apply a partial trace operation to $U(\sigma \otimes \rho)U^* \in \mathbf{M}_m \otimes \mathbf{M}_r$, where $nk = mr$, to obtain $\tau \in \mathbf{M}_m$ for our investigation. This will be the general quantum operations one can apply to a quantum system. Under this mathematical framework, we have the following theorem.

THEOREM 3.4.1. *For every quantum operation $\Phi : \mathbf{M}_n \to \mathbf{M}_m$ there exist $r \in \mathbb{N}$ and $F_1, \ldots, F_r \in \mathbf{M}_{m,n}$ such that $\sum_{j=1}^r F_j^\dagger F_j = I_n$ and*

$$\Phi(A) = \sum_{j=1}^{r} F_j A F_j^\dagger \qquad \text{for all } A \in \mathbf{M}_n. \tag{3.15}$$

Proof. Let $\Phi : \mathbf{M}_n \to \mathbf{M}_m$ be a quantum operation. By the previous discussion, it can be realized as the partial trace of $U(\sigma \otimes \rho)U^\dagger$ for a suitable choice of $\sigma \in \mathbf{D}_k$ and $U \in \mathrm{U}(kn)$, where U may be time dependent, governing the dynamics of the closed system with initial state $\sigma \otimes \rho$. By purification, we may assume that $\sigma \in \mathbf{D}_k$ is a pure state $(1, 0, \ldots, 0)^t(1, 0, \ldots, 0)$ so that

$\sigma \otimes \rho = \rho \oplus 0_{(k-1)n}$. Suppose $\rho \mapsto U(\sigma \otimes \rho)U^\dagger = (B_{pq})$, where $B_{jj} \in \mathbf{M}_m$ for $j = 1, \dots, r$ with $r = nk/m$, and we will apply partial trace to obtain $B_{11} + \cdots + B_{rr}$ as the image of $\Phi(\rho)$. We may let the first n columns of U to from the matrix

$$F = \begin{pmatrix} F_1 \\ \vdots \\ F_r \end{pmatrix} \in \mathbf{M}_{rm,n} \tag{3.16}$$

with $F_j \in \mathbf{M}_{m,n}$ for each j, then by block multiplication we have

$$U(\sigma \otimes \rho)U^\dagger = F\rho F^\dagger$$

so that $B_{jj} = F_j \rho F_j^\dagger$ for $j = 1, \dots, r$, and

$$\Phi(\rho) = B_{11} + \cdots + B_{rr} = \sum_{j=1}^{r} F_j \rho F_j^\dagger.$$

Thus, the action of Φ on density matrices has the asserted form. Since the set of density matrices generate all matrices in $\mathbf{M}_n$, we see that Φ has the asserted form. ∎

One can reverse the above proof to show that every operator Φ of the form (3.15) corresponds to a quantum operation $\Phi : \mathbf{M}_n \to \mathbf{M}_m$ of an open system as follows. For the given $F_1, \dots, F_r$ in (3.15), we can form the matrix F in (3.16). The condition $\sum_{j=1}^{r} F_j^\dagger F_j = I_n$ means that the F has orthonormal columns. Hence, we can extend F to a unitary $U = [F \,|\, \tilde{F}] \in \mathbf{M}_{mr}$ such that $\Phi(\rho)$ is the partial trace of $U(\sigma \otimes \rho)U^\dagger \in \mathbf{M}_{mr}$ for all density matrices ρ. Thus, Φ is a quantum operation.

Kraus [8] obtained the result in the context of quantum mechanics. The quantum operation expressed in the form (3.15) is called the **operator sum representation (OSR)** of the quantum operation, and the matrices $F_1, \dots, F_r$ are called the **Kraus operators** of the quantum operation Φ. In fact, a quantum operation can also be viewed as a **quantum channel**, which describes the change of quantum states ρ as they go through a quantum device. In such a context, the Kraus operators are also known as **error operators**, and we need to find a quantum operation Ψ known as the **recovery channel** such that $\Psi \circ \Phi(\rho) = \rho$ for specific choice of ρ in the code space. This will be the main topic in Chapter 7.

EXAMPLE 3.4.2. *Let $U_1, \dots, U_r \in \mathrm{U}(n)$ and $p_1, \dots, p_r$ be positive numbers summing up to 1. Then $\Phi : \mathbf{M}_n \to \mathbf{M}_n$ defined by*

$$\Phi(A) = \sum_{j=1}^{r} p_j U_j A U_j^\dagger \qquad \text{for all } A \in \mathbf{M}_n$$

is a quantum channel known as the **random unitary channel** *or* **mixed unitary channel**. *It is easy to construct a unitary $V \in M_{nr} \in M_n \otimes M_r$ with the first n columns forming the matrix* $\begin{pmatrix} \sqrt{p_1}U_1 \\ \vdots \\ \sqrt{p_r}U_r \end{pmatrix}$ *with orthonormal columns. Then*

$$\Phi(A) = \operatorname{Tr}_2 \left[V \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} V^\dagger \right] \qquad \textit{for all } A \in M_n.$$

Alternatively, we may let $V = U_1 \oplus \cdots \oplus U_r \in \mathbf{M}_{nr}$ and $\sigma \in \mathbf{D}_r$ be the diagonal matrix with diagonal entries $p_1, \ldots, p_r$. Then $A \mapsto \operatorname{Tr}_2(V(\sigma \otimes A)V^\dagger) = \sum_{j=1}^r p_j U_j A U_j^\dagger$.

3.4.2 Quantum channels and Measurements

930978025

Also, quantum measurements can be viewed as quantum operations on open systems. As mentioned before a Hermitian matrix $A = \sum_{j=1}^n \lambda_j |\lambda_j\rangle\langle\lambda_j|$ is associated with an observable. If a state $\rho \in \mathbf{D}_n$ goes through the measurement process corresponding to A, the state ρ will "collapse" to one of the pure states $|\lambda_j\rangle\langle\lambda_j|$ with a probability $\operatorname{Tr}(A\rho)$.

More generally, if some eigenvalues of A have multiplicities larger than one, we may write $A = \sum_{j=1}^s \lambda_j P_j$, where P_j is the projection operator corresponding to the eigenvalue λ_j for the distinct eigenvalues $\lambda_1, \ldots, \lambda_s$ of A. In such a case, the **projective measurement** of ρ under the measurement associated with A is the quantum operation

$$\rho \to \sum_j P_j \rho P_j, \tag{3.17}$$

where $p_j = \operatorname{Tr}(P_j \rho P_j) = \operatorname{Tr}(\rho P_j)$ and the set $\{P_1, \ldots, P_r\}$ satisfies the completeness relation $\sum_j P_j P_j^\dagger = \sum_j P_j = I$. Clearly, the projective measurement is a special case of a quantum operation in which the Kraus operators are $F_j = P_j$. Upon measurement of ρ, we get the state $\frac{1}{p_j} P_j \rho P_j \in \mathbf{D}_n$ with a probability p_j.

More generally, for any positive semidefinite matrices $Q_1, \ldots, Q_r \in \mathbf{M}_n$ such that $Q_1 + \cdots + Q_r = I_n$, there are $M_1, \ldots, M_r \in \mathbf{M}_n$ such that $M_j^\dagger M_j = Q_j$. The measurement operators are then associated with the quantum operation

$$\rho \mapsto \sum_{j=1}^r M_j \rho M_j^\dagger$$

so that ρ will change to the quantum state $\frac{1}{p_j} M_j \rho M_j^\dagger$ with a probability $p_j = \operatorname{Tr}(M_j \rho M_j^\dagger) = \operatorname{Tr}(\rho Q_j)$. The set $\{Q_1, \ldots, Q_r\} = \{M_1^\dagger M_1, \ldots, M_r^\dagger M_r\}$ is known as the **positive operator-valued measure (POVM)**.

Note that there is flexibility in the choice or construction of $M_1, \ldots, M_r$. In fact, one may always change M_j to $U_j M_j$ for unitary matrices $U_1, \ldots, U_r$ and we still have $(U_j M_j)^\dagger (U_j M_j) = Q_j$. Of course, the resulting measurement of ρ will become $\frac{1}{p_j} U_j M_j \rho M_j^\dagger U_j^\dagger$ with $p_j = \mathrm{Tr}\,(U_j M_j \rho M_j^\dagger U_j^\dagger) = \mathrm{Tr}\,(\rho Q_j)$. Note that the state after the POVM measurement is not determined due to the ambiguity caused by the choice of $U_1, \ldots, U_r$. In fact, the choice (construction) of M_j depends on one's control of the system as mentioned before. For example, if the measurement process $\rho \mapsto \sum M_j \rho M_j^\dagger$ is associated with the effect of the environment, we can have limited control on $M_1, \ldots, M_r$. On the other hand, if we create an auxiliary system in the laboratory, we may have more control on constructing $M_1, \ldots, M_r$. Measurement with $\{M_1, \ldots, M_r\}$ is called the **generalized measurement**.

EXAMPLE 3.4.3. *Suppose Bob will be given a quantum state chosen from the linearly independent set of unit vectors $\{|\psi_1\rangle, \ldots, |\psi_m\rangle\}$, which may not be orthonormal. He can construct the following POVM $\{Q_1, \ldots, Q_{m+1}\}$ such that he will know for sure that $|\psi_j$ is sent to him if the measurement of the received state yields Q_j if $Q_j = |\phi_j\rangle\langle\phi_j|/m$, where $\langle\phi_j|\phi_j\rangle = 1$ and $\langle\phi_j|\psi_i\rangle = 0$ for all $i \neq j$ for $j = 1, \ldots, m$ and $Q_{m+1} = I - \sum_{j=1}^m Q_j$. Evidently, a measurement of $|\psi_j\rangle\langle\psi_j|$ will yield Q_j or Q_{m+1}.*

In fact, one can associate a POVM with a projective measurement on a bipartite system with the principal system as a subsystem by the following proposition.

PROPOSITION 3.4.4. *Let $\{Q_1, \ldots, Q_r\} \subseteq \mathbf{M}_n$ be a POVM. Then there are projective measurements $\{P_1, \ldots, P_r\}$ in $\mathbf{M}_r \otimes \mathbf{M}_n$ and a unitary $V \in \mathbf{M}_r \otimes \mathbf{M}_n$ such that for $j = 1, \ldots, r$, Q_j is the leading $n \times n$ submatrix of $V^\dagger P_j V$ equivalently, $V_1^\dagger P_j V_1 = Q_j$ if V_1 is the first n columns of V. Consequently, we have the following realization of the POVM as a quantum operation Φ on an open system*

$$\Phi(A) = \sum_{j=1}^r Q_j^{1/2} A Q_j^{1/2} = \mathrm{Tr}\,_1(\mathrm{Tr}\, V(E_{11} \otimes A)V^\dagger) \qquad \textit{for all } A \in \mathbf{M}_n.$$

Proof. Let $\{|e_1\rangle, \ldots, |e_r\rangle\}$ be the standard basis for $\mathbb{C}^r$ and $E_{jj} = |e_j\rangle\langle e_j|$ for $j = 1, \ldots, r$. Let $P_j = E_{jj} \otimes I_n \in \mathbf{M}_r \otimes \mathbf{M}_n$, and set $V_1 = \begin{bmatrix} Q_1^{1/2} \\ \vdots \\ Q_r^{1/2} \end{bmatrix}$. Then $V_1^\dagger V_1 = \sum_{j=1}^r Q_j = I_n$. So, V_1 has orthonormal columns and we can extend V_1 to a unitary matrix $V \in \mathbf{M}_r \otimes \mathbf{M}_n$. It is easy to verify that

$$[Q_1^{1/2} \cdots Q_r^{1/2}] P_j \begin{bmatrix} Q_1^{1/2} \\ \vdots \\ Q_r^{1/2} \end{bmatrix} = Q_j$$

so that Q_j is the leading submatrix of $V^\dagger P_j V$, and

$$\mathrm{Tr}\,_1(V(E_{11}\otimes\rho)V^\dagger)=\sum_{j=1}^{r}Q_j^{1/2}\rho Q_j^{1/2}.$$

∎

3.5 Partial transpose, Entanglement witness

Suppose ρ is not an uncorrelated state. It is hard to determine whether ρ is separable or inseparable. In fact, this is an NP-hard problem;[†] see [14]. Nevertheless, we have the following simple test to identify inseparable states.

Define the **partial transpose** ρ^{pt} of $\rho \in \mathbf{D}_{mn}$ with respect to the second Hilbert space $\mathbb{C}^n$ as

$$\rho_{ij,kl} \to \rho_{il,kj}, \tag{3.18}$$

where

$$\rho_{ij,kl} = (\langle e_{1,i}| \otimes \langle e_{2,j}|)\ \rho\ (|e_{1,k}\rangle \otimes |e_{2,l}\rangle).$$

Here $\{|e_{1,k}\rangle : 1 \le k \le m\}$ is the basis for $\mathbb{C}^m$, while $\{|e_{2,k}\rangle : 1 \le k \le n\}$ is the basis for $\mathbb{C}^n$. In particular, if $\rho = \rho_1 \otimes \rho_2$, then $\rho^{\mathrm{pt}} = \rho_1 \otimes \rho_2^t$, and if $\rho = (P_{ij})$ with $P_{ij} \in \mathbf{M}_n$, then $\rho^{\mathrm{pt}} = (P_{ij}^t)$.

Now, suppose ρ is separable and has the form (3.9). Then the partial transpose yields

$$\rho^{\mathrm{pt}} = \sum_i p_j \rho_{1,j} \otimes \rho_{2,j}^t. \tag{3.19}$$

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 (3.19) is another density matrix. It was conjectured by Peres [6] and subsequently proved by the Hordecki family [13] that positivity of the partially transposed density matrix is a 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 a proof of the assertion, we look at the following example.

EXAMPLE 3.5.1. *Let us consider the Werner state*

$$\rho = \begin{pmatrix} \frac{1-p}{4} & 0 & 0 & 0 \\ 0 & \frac{1+p}{4} & -\frac{p}{2} & 0 \\ 0 & -\frac{p}{2} & \frac{1+p}{4} & 0 \\ 0 & 0 & 0 & \frac{1-p}{4} \end{pmatrix}, \tag{3.20}$$

[†] Interested readers can consult the paper [14] for the definition.

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

$$|0\rangle|0\rangle, |1\rangle|1\rangle, |1\rangle|0\rangle, |1\rangle|1\rangle.$$

Partial transpose of ρ *yields*

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

Note that we need to consider off-diagonal matrix elements only when we partically transpose the matrix. We have, for example,

$$\begin{aligned} \rho_{01,10} &= (\langle 0| \otimes \langle 1|)\ \rho\ (|1\rangle \otimes |0\rangle) \\ &\to (\langle 0| \otimes \langle 0|)\ \rho\ (|1\rangle \otimes |1\rangle) = \rho^{\mathrm{pt}}_{00,11}. \end{aligned}$$

For ρ^{pt} *to be a physically acceptable state, it must have non-negative eigenvalues. The characteristic equation of* ρ^{pt} *is*

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

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

Suppose a Hermitian matrix $H \in \mathbf{M}_n$ has eigenvalues $\lambda_1, \ldots, \lambda_n$ summing up to 1. Then $H \in \mathbf{D}_n$ if and only if all the eigenvalues of H are nonnegative. Thus, inseparable states ρ can be determined by non-vanishing **negativity** defined as

$$N(\rho) \equiv \frac{\sum_j |\lambda_i| - 1}{2}, \tag{3.21}$$

where λ_j's are the eigenvalues of ρ^{pt}.

Negativity is one of the so-called entanglement monotones [14], which also include concurrence, entanglement of formation and entropy of entanglement.[‡]

EXAMPLE 3.5.2. *It was mentioned above that vanishing negativity is equivalent with separability only for* $\mathbb{C}^2 \otimes \mathbb{C}^2$ *systems and* $\mathbb{C}^2 \otimes \mathbb{C}^3$ *systems. A counter*

[‡] Interested readers can see the definitions and background of these concepts in [14] and its references.

example in a $\mathbb{C}^2 \otimes \mathbb{C}^4$ system has been given in Horodecki [15]. Let us consider

$$\rho = \frac{1}{7b+1} \begin{pmatrix} b & 0 & 0 & 0 & 0 & b & 0 & 0 \\ 0 & b & 0 & 0 & 0 & 0 & b & 0 \\ 0 & 0 & b & 0 & 0 & 0 & 0 & b \\ 0 & 0 & 0 & b & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1+b}{2} & 0 & 0 & \frac{\sqrt{1-b^2}}{2} \\ b & 0 & 0 & 0 & 0 & b & 0 & 0 \\ 0 & b & 0 & 0 & 0 & 0 & b & 0 \\ 0 & 0 & b & 0 & \frac{\sqrt{1-b^2}}{2} & 0 & 0 & \frac{1+b}{2} \end{pmatrix} \qquad (0 \le b \le 1), \tag{3.22}$$

which is known to be inseparable. The partial transposed matrix with respect to the second system is

$$\rho^{\mathrm{pt}} = \frac{1}{7b+1} \begin{pmatrix} b & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & b & 0 & 0 & b & 0 & 0 & 0 \\ 0 & 0 & b & 0 & 0 & b & 0 & 0 \\ 0 & 0 & 0 & b & 0 & 0 & b & 0 \\ 0 & b & 0 & 0 & \frac{1+b}{2} & 0 & 0 & \frac{\sqrt{1-b^2}}{2} \\ 0 & 0 & b & 0 & 0 & b & 0 & 0 \\ 0 & 0 & 0 & b & 0 & 0 & b & 0 \\ 0 & 0 & 0 & 0 & \frac{\sqrt{1-b^2}}{2} & 0 & 0 & \frac{1+b}{2} \end{pmatrix}. \tag{3.23}$$

The eigenvalues of ρ^{pt} are

$$0, 0, 0, \frac{b}{7b+1}, \frac{2b}{7b+1}, \frac{2b}{7b+1}$$

$$\frac{1 + 14b^2 + 9b - \sqrt{98b^4 - 70b^3 + 23b^2 + 12b + 1}}{2\,(49b^2 + 14b + 1)},$$

$$\frac{1 + 14b^2 + 9b + \sqrt{98b^4 - 70b^3 + 23b^2 + 12b + 1}}{2\,(49b^2 + 14b + 1)}.$$

It can be shown that the seventh eigenvalue takes the maximum value $(25 - 2\sqrt{10})/130 \sim 0.144$ at $b = (47 - 10\sqrt{10})/31 \sim 0.496$ and the minimum value 0 at $b = 0$, and hence all the eigenvalues are non-negative for $0 \le b \le 1$ in spite of inseparability of ρ.

Note that the set of separable states form a convex subset in the set of bipartite states. Using basic matrix theory (functional analysis), for every inseparable state $\rho \in \mathbf{M}_m \otimes \mathbf{M}_n$, there is linear functional f on Hermitian matrices in $\mathbf{M}_{mn}$ such that $f(\rho) > 0 \ge f(\sigma_1 \otimes \sigma_2)$ for all $\sigma_1 \in \mathbf{D}_m, \sigma_2 \in \mathbf{D}_n$. Since every linear functional on Hermitian matrices in $\mathbf{M}_{mn}$ takes the form $f(X) = \mathrm{Tr}\,(FX)$ for some Hermitian matrix $F \in \mathbf{M}_{mn}$, we can regard F as an observable on the bipartite system. The linear functional f or the observable F associated with it is called the **entanglement witness** of ρ. So, we have the following.

THEOREM 3.5.3. *Let $\rho \in \mathbf{M}_m \otimes \mathbf{M}_n$. Then ρ is inseparble if and only if there is an entanglement witness F such that*

$$\mathrm{Tr}\,(F\rho) > 0 \geq \mathrm{Tr}\,(F(\sigma_1 \otimes \sigma_2)) \qquad \textit{for all } \sigma_1 \in \mathbf{D}_m, \sigma_2 \in \mathbf{D}_n.$$

It should be remarked that finding an entanglement witness of an inseparable state or showing the nonexistence could be a challenging problem.

3.6 Fidelity

It often happens that one has to compare two density matrices and tell how much they differ from each other. For instance, an experimentalist may conduct an experiment and then compare the resulting quantum state with a certain existing quantum state. A good measure for this purpose is the **fidelity** defined as follows; see [16].

DEFINITION 3.6.1. *Let $\rho_1, \rho_2 \in \mathbf{D}_n$. Then the fidelity is defined by*

$$F(\rho_1, \rho_2) = \left\{ \mathrm{Tr}\left(\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}} \right) \right\}^2. \tag{3.24}$$

Here, $\sqrt{\rho_1}$ is the positive semi-definite square root of ρ_1, and $\sqrt{\rho_1}\rho_2\sqrt{\rho_1}$ is positive semi-definite so that we can take its positive semi-definite square root.

We are going to prove some basic properties of fidelity. To do that we need to introduce some notation and terminology which will be useful in our subsequent discussion.

For $X \in \mathbf{M}_{m,n}$, let $vec(X) \in \mathbb{C}^{mn}$ be the vector obtained by stacking the columns of X with the first column on the top, and the last column at the bottom. Clearly, $X \mapsto vec(X)$ is an invertible linear map. One can define the inverse map from $\mathbb{C}^{mn}$ to $\mathbf{M}_{m,n}$ so that $|v\rangle \mapsto [v]_{m,n}$, where the first m entries of $|v\rangle$ form the first column of $[v]_{m,n}$, then next m entries of $|v\rangle$ form the second column of $[v]_{m,n}$, etc. Moreover, if we define the inner product

$$\langle X, Y \rangle = \mathrm{Tr}\,(X^\dagger Y)$$

on $\mathbf{M}_{m,n}$, the map $|v\rangle \mapsto [|v\rangle]_{m,n}$ satisfies $\langle v_1 | v_2 \rangle = \langle [|v_1\rangle]_{m,n}, [|v_2\rangle]_{m,n} \rangle$.

One may define the inner product norm for $A \in \mathbf{M}_{m,n}$ by $\|A\|_F = \sqrt{\langle A, A \rangle}$,[§] and the following Cauchy-Schwartz inequality holds (see Section 1.8.3:

$$|\langle A, B \rangle| \leq \|A\|_F \|B\|_F \qquad \text{for any } A, B \in \mathbf{M}_{m,n}.$$

[§] The norm $\|\cdot\|_F$ is also known as the Euclidean norm or Frobenius norm.

THEOREM 3.6.2. *Let $\rho_1, \rho_2 \in \mathbf{D}_n$. If $\rho_1^{1/2}\rho_2^{1/2}$ has singular values $s_1 \geq \cdots \geq s_n$, then*

$$F(\rho_1, \rho_2) = F(\rho_2, \rho_1) = \left[\sum_{j=1}^{n} s_j\right]^2,$$

and the following conditions hold.

(1) For any unitary U, $F(U\rho_1 U^\dagger, U\rho_2 U^\dagger) = F(\rho_1, \rho_2)$.

(2) If ρ_1 or ρ_2 is a pure state, then $F(\rho_1, \rho_2) = \mathrm{Tr}\,(\rho_1\rho_2)$.

(3) We have

$$F(\rho_1, \rho_2) \in [0, 1].$$

The equality $F(\rho_1, \rho_2) = 1$ holds if and only if $\rho_1 = \rho_2$.

The equality $F(\rho_1, \rho_2) = 0$ holds if and only if $\mathrm{Tr}\,(\rho_1\rho_2) = 0$, equivalently, $\sigma_1^r\sigma_2^s = 0$ for any positive numbers r, s.

Proof. Let $A = \rho_1^{1/2}\rho_2^{1/2}$ have a singular value decomposition $A = XDY^\dagger$ where $X, Y \in \mathbf{M}_n$ are unitary and D is the diagonal matrix with diagonal entries $s_1 \geq \cdots \geq s_n \geq 0$. Then

$$F(\rho_1, \rho_2) = \left[\mathrm{Tr}\left(\sqrt{\sqrt{\rho_1}\rho_2\sqrt{\rho_1}}\right)\right]^2 = [\mathrm{Tr}\,\sqrt{AA^\dagger}]^2 = [\mathrm{Tr}\,(XDX^\dagger)]^2 = \left[\sum_{j=1}^{n} s_j\right]^2,$$

and

$$F(\rho_2, \rho_1) = \left[\mathrm{Tr}\left(\sqrt{\sqrt{\rho_2}\rho_1\sqrt{\rho_2}}\right)\right]^2 = \left[\mathrm{Tr}\,\sqrt{A^\dagger A} = \mathrm{Tr}\,(YDY^\dagger)\right]^2 = \left[\sum_{j=1}^{n} s_j\right]^2.$$

It is clear that $\left(U\rho_1 U^\dagger\right)^{1/2}\left(U\rho_2 U^\dagger\right)^{1/2} = U\left(\rho_1^{1/2}\rho_2^{1/2}\right)U^\dagger$ and A have the same singular values. Thus (1) holds.

To prove (2), we may assume that ρ_1 is a pure state as $F(\rho_1, \rho_2) = F(\rho_2, \rho_1)$. By (1), to compute $F(\rho_1, \rho_2)$, we may replace (ρ_1, ρ_2) by $(U\rho_1 U^\dagger, U\rho_2 U^\dagger)$ and assume that ρ_1 is the diagonal matrix with diagonal entries $1, 0, \ldots, 0$. If ρ_2 has $(1, 1)$ entry μ, then

$$\begin{aligned} F(\rho_1, \rho_2) &= \left[\mathrm{Tr}\,\sqrt{\rho_1^{1/2}\rho_2\rho_1^{1/2}}\right]^2 = \sqrt{\mu}^2 = \mu \\ &= \mathrm{Tr}\,(\rho_1\rho_2\rho_1) = \mathrm{Tr}\,(\rho_1^2\rho_2) = \mathrm{Tr}\,(\rho_1\rho_2). \end{aligned}$$

Now, we turn to (3). Suppose $A = \rho_1^{1/2}\rho_2^{1/2}$ has singular value decomposition $UDV^\dagger$, where U, V are unitary matrices, and D is the matrix of singular values. Let $Z = VU^\dagger$. Then

$$F(\rho_1, \rho_2) = |\mathrm{Tr}\,(AZ)|^2 = |\mathrm{Tr}\,(\rho_1^{1/2}\rho_2^{1/2}Z)|^2$$

$$\leq \langle \rho_1^{1/2}, \rho_1^{1/2}\rangle\langle \rho_2^{1/2}Z, \rho_2^{1/2}Z\rangle = (\operatorname{Tr}\rho_1)(\operatorname{Tr}\rho_2) = 1.$$

The inequality follows from the Cauchy-Schwartz inequality. The equality holds if and only if $\rho_1 = \mu\rho_2$ for some $\mu \in \mathbb{C}$. Comparing the traces on both sides, one sees that $\mu = 1$.

Now, $0 = F(\rho_1, \rho_2) = (\sum_{j=1}^n s_j)^2$ if and only if $s_1 = \cdots = s_n = 0$, i.e., $A = \rho_1^{1/2}\rho_2^{1/2} = 0$. It follows that $AA^\dagger = 0$, and hence

$$0 = \operatorname{Tr}(AA^\dagger) = \operatorname{Tr}(\rho_1^{1/2}\rho_2\rho_1^{1/2}) = \operatorname{Tr}(\rho_1\rho_2).$$

Since $P = \rho_1^{1/2}$ and $Q = \rho_2^{1/2}$ are positive semi-definite, $PQ = 0$ is equivalent to $P^rQ^s = 0$ for any $r, s > 0$. To see this, suppose $P = UD_1U^\dagger$ where $U \in \mathbf{M}_n$ is unitary and D_1 has diagonal entries $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$ such that $\lambda_k > 0 = \lambda_{k+1}$. Then $0 = U^\dagger PQU = U^\dagger PUU^\dagger QU = D_1(U^\dagger QU)$ implies that the Hermitian matrix $U^\dagger QU$ has the form $0_k \oplus C$. Thus, $U^\dagger(P^rQ^s)U = D_1^r(0_k \oplus C^s) = 0$, and hence $P^rQ^s = 0$. Conversely, if $P^rQ^s = 0$, we can apply the above argument to show that $PQ = (P^r)^{1/r}(Q^s)^{1/s} = 0$. ▮

One may see the above properties proved in [16]. In [6], the authors suggested using properties (1) – (3) in the above theorem as the basic requirements for any generalization of fidelity measures for comparing mixed states.

We remarked that there are other comparisons between the difference of two quantum states ρ_1, ρ_2. We name a few examples in the following.

- (The **Bures distance**) $D_B(\rho_1, \rho_2) = \sqrt{2(1 - F(\rho_1, \rho_2)^{1/2})}$.
- (The **trace distance**) $\|\rho - \sigma\|_{\mathrm{Tr}} = \operatorname{Tr}|\rho_1 - \rho_2|$, where $|X| = (X^\dagger X)^{1/2}$ for any square matrix X.

Note that $D_B(\rho_1, \rho_2)$ arises from the quantity

$$\min_{U,V\in \mathrm{U}(n)} \|\rho_1^{1/2}U - \rho_2^{1/2}V\|_F^2 = \min_{W\in\mathrm{U}(n)} \operatorname{Tr}(\rho_1 + \rho_2 - \rho_1^{1/2}W\rho_2^{1/2} - \rho_2^{1/2}W^\dagger\rho_1^{1/2}),$$

which equals $2(1 - F(\rho_1, \rho_2)) = D_B(\rho_1, \rho_2)$. That is why $\sqrt{2}$ is in the formula. One can show

$$1 - \sqrt{F(\rho_1, \rho_2)} \leq \|\rho_1 - \rho_2\|_{\mathrm{Tr}} \leq \sqrt{1 - F(\rho_1, \rho_2)}.$$

3.7 Entropies

Information carried by a physical system is quantified by various entropies, both in classical and quantum information theories. Let us start with the Shannon entropy, which was proposed to quantify classical information.

3.7.1 Shannon Entropy

Suppose an event x, such as being hit by a car in 24 hours, happens with probability $p(x)$. The **information content** $I(x)$ is defined by

$$I(x) = -\log_2 p(x). \tag{3.25}$$

Note that the base of the log function is 2. We will omit 2 in the rest of this section to simplify the notation. We use ln for $\log_e$. For $q(x) = 1/p(x)$, $-\log p(x) = \log q(x)$ is approximately the number of digits required to represent $q(x)$ in binary number. Therefore the information content $I(x)$ is large if $p(x)$ is small.

Car accident does not happen very often and we are surprised if we hear our friend was hit by a car. The information we get by this news is large ($I(x) \gg 1$), which is often associated with a big surprise. In contrast, the probability of watching a dog in 24 hours is close to 1 in usual circumstances and we do not get new information by this ($I(x) \sim \log 1 \sim 0$) and no one will be surprised by someone watched a dog. Thus, $I(x)$ quantifies the information we get as well as our surprise at the information.

Let X be a random variable which takes values in $\mathcal{X} = \{x_1, x_2, \ldots, x_n\}$ and let $p(x_k)$ be the probability, with which X takes value x_k. We only consider the case where $\mathcal{X}$ is a finite set. The probabilities sum up to one: $\sum_{k=1}^n p(x_k) = 1$. The Shannon entropy $H(p(x))$ of this random variable X is defined as the average of the information contents over $\mathcal{X}$ as

$$H(p(x)) = -\sum_{k=1}^{n} p(x_k) \log p(x_k). \tag{3.26}$$

H is positive-semidefinite since $0 \le p(x_k) \le 1$ and $-\log p(x_k) \ge 0$. We assume $x \log x$ is continuous at $x = 0$ and put

$$x \log x|_{x=0} = \lim_{x \to 0} x \log x = 0.$$

$H(p(x))$ is the average number of bits to store the information.

Let us work out two extreme cases.

(i) Let $p(x_1) = 1$ and $p(x_k) = 0$ $(k \neq 1)$. Then $H(p(x)) = -1 \log 1 = 0$. This is the smallest possible value of H since it is positive-semidefinite.

(ii) Suppose there are n possible values of x_k and all these values appear with the same probability $p(x_k) = 1/n$. Then $H(p(x)) = -n\frac{1}{n} \log \frac{1}{n} = \log n$.

Let us show this is the maximal possible value of H. We maximize H under the constraint $\sum_k p(x_k) = 1$. Namely we consider the extremum of

$$-\sum_{k=1}^{n} p(x_k) \log p(x_k) - \lambda \left(\sum_{k=1}^{n} p(x_k) - 1 \right),$$

where λ is the Lagrange multiplier. Under the variation $\delta p(x_k)$, we require $\delta H = -\sum_k (\log p(x_k) + 1 + \lambda)\delta p(x_k) = 0$, from which we find $p(x_k) = 2^{-1-\lambda}$ is independent of k. From the constraint $\sum_k p(x_k) = 1$, we obtain $\lambda = \log n - 1$ and hence $p(x_k) = 1/n$. We obtain $0 \le H(p(x)) \le \log n$ for a general $\{p(x_k)\}$.

We guess from the above two examples that the Shannon entropy is large if X is more random and small if X is less random.

EXAMPLE 3.7.1. *(1) In coin tossing of a uniform coin, probabilities of getting heads (H) and tails (T) are* $p_{\mathrm{H}} = p_{\mathrm{T}} = 1/2$*. Then* $\mathcal{X} = \{H, T\}$ *and the entropy of this random process is*

$$H(p(x)) = -2 \times \frac{1}{2} \log \frac{1}{2} = \log 2 = 1,$$

which means one bit is required to store the information of H *or* T*.*
(2) Suppose the probability of getting H *is* p *and* T *is* $1-p$ *for a non-uniform coin. Then*

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

Note that $H(p(x)) = 0$ *if* $p = 0$ *or* $p = 1$*, namely if there is no randomness in coin tossing, and* $H(X)$ *takes maximum value* $\log 2 = 1$ *at* $p = 1/2$*, i.e., the most random case.*
(3) Probability of getting any of $1, 2, \ldots, 6$ *is* $1/6$ *on a roll of the uniform dice. The entropy of this process is*

$$H(p(x)) = -6\frac{1}{6} \log \frac{1}{6} = \log 6 \sim 2.58.$$

It requires 3 bits to store the number on average.

Entropy is additive. Let X, Y be two *independent* random variables taking values in $\mathcal{X} = \{x_j\}$ and $\mathcal{Y} = \{y_k\}$ and let $p(x_j)$ and $q(y_k)$ be their probability distributions. Suppose X and Y are measured and values x_j and y_k are obtained. This event takes place with probability $p(x_j)q(y_k)$. Then the entropy of this process is

$$\begin{aligned} H(p(x), q(y)) &= -\sum_{j,k} p(x_j)q(y_k) \log(p(x_j)q(y_k)) \\ &= -\sum_{j,k} p(x_j)q(y_k)[\log p(x_j) + \log q(y_k)] \\ &= -\sum_{j} p(x_j) \log p(x_j) - \sum_{k} q(y_k) \log q(y_k) \\ &= H(p(x)) + H(q(y)), \end{aligned} \tag{3.27}$$

showing entropy is additive for independent random variables.

3.7.2 Classical Relative Entropy

Let X be a classical random variables and $\mathcal{X} = \{x_k\}$ be the set of its values. Suppose there are two probability distributions $\{p(x_k)\}$ and $\{q(x_k)\}$ on the same variable set $\mathcal{X}$. The classical **relative entropy** of $\{p(x_k)\}$ to $\{q(x_k)\}$ is defined by

$$H(p(x)||q(x)) = \sum_k p(x_k)(\log p(x_k) - \log q(x_k)). \tag{3.28}$$

PROPOSITION 3.7.2. *$H(p(x)||q(x)) \geq 0$, where equality holds if and only if $p(x_k) = q(x_k)$ for all k.*

Proof. Use the well known inequality $\ln 2 \log x = \ln x \leq x - 1$ *for* $x > 0$ *to show*

$$\begin{aligned} H(p(x)||q(x)) &= \sum_k p(x_k) \log \frac{p(x_k)}{q(x_k)} \geq \frac{1}{\ln 2} \sum_k p(x_k) \left[1 - \frac{q(x_k)}{p(x_k)}\right] \\ &= \frac{1}{\ln 2} \sum_k [p(x_k) - q(x_k)] = 0. \end{aligned}$$

Clearly the inequality above is saturated if and only if $q(x_k)/p(x_k) = 1$ for all k, that is $p(x_k) = q(x_k)$ for all k.

Relative entropy measures how close two distributions are. In this sense, relative entropy is also known as the Kullback-Leibler distance although it does not satisfy the axioms of distance.

3.7.3 von Neumann Entropy

Quantum counterpart of the Shannon entropy is the von Neumann entropy. Let us consider a system with the n-dimensional Hilbert space $\mathcal{H}$. Suppose the system is in a mixed state ρ. Then the **von Neumann entropy** is defined as

$$S(\rho) = -\mathrm{Tr}\,(\rho \log \rho), \tag{3.29}$$

where log stands for $\log_2$ as before.

Let $\rho = \sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i|$ be the spectral decomposition of ρ, where $0 \leq \lambda_i \leq 1$ and $\sum_i \lambda_i = 1$. Now $S(\rho)$ is represented as $S(\rho) = -\sum_i \lambda_i \log \lambda_i$, from which we find $S(\rho) \geq 0$. Since a unitary transformation does not change the eigenvalues of ρ, von Neumann entropy is invariant under unitary transformation;

$$S(U\rho U^\dagger) = S(\rho), \quad U \in \mathrm{U}(n). \tag{3.30}$$

Let us consider two extreme cases as before.

(i) Consider a pure state $\rho = |\psi\rangle\langle\psi|$. If we take $|\psi\rangle$ as the first basis vector of $\mathcal{H}$, its component expression is $|\psi\rangle = (1, 0, \ldots, 0)^t$ and $\rho = \mathrm{diag}(1, 0, \ldots, 0)$, from which we obtain $S(\rho) = 0$.

(ii) Let us consider next the uniformly mixed state $\rho = I_n/n$, where I_n is the identity matrix in $\mathbf{M}_n$. Then

$$S(\rho) = -n\frac{1}{n}\log\frac{1}{n} = \log n,$$

which is the maximal possible value of S. To prove this, let us extremize $\tilde{S}(\rho) = -\mathrm{Tr}\,(\rho\log\rho) - \mu(\mathrm{Tr}\,\rho - 1)$, where μ is the Lagrange multiplier that takes care of the constraint $\mathrm{Tr}\,\rho = 1$. We impose $\delta\tilde{S}(\rho) = -\mathrm{Tr}\,[(\log\rho + 1 + \mu)\delta\rho] = 0$ under the variation $\delta\rho$, from which we obtain $\rho = 2^{-1-\mu}I_n$. The unit trace condition gives $\mu = \log n - 1$ as before, from which we obtain $\rho = I_n/n$ and $S(\rho) = \log n$.

The above observation shows that $0 \le S(\rho) \le \log n$ for a general ρ. $S(\rho)$ measures how much ρ differs from pure states. Note that $S(\rho) = 0$ if and only if ρ is a pure state. In fact, $S(\rho)$ has been shown to have only one maximum at $\rho = I_n/n$ above, namely at $\lambda_i = 1/n,\ \forall i$, where $\{\lambda_i\}$ is the set of eigenvalues of ρ. Then the minima are found at the edges, where $\lambda_k = 1$ and $\lambda_j = 0$ $(j \neq k)$, for $1 \le k \le n$. These edge points are pure states.

3.7.4 Entanglement Entropy

Consider a bipartite system with the associated Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ with $\dim\mathcal{H}_A = n_A$ and $\dim\mathcal{H}_B = n_B$. Take a pure state $|\psi\rangle \in \mathcal{H}_{AB}$ and $\rho = |\psi\rangle\langle\psi|$ and define the reduced density matrices $\rho_A = \mathrm{Tr}\,_B\rho$ and $\rho_B = \mathrm{Tr}\,_A\rho$.

PROPOSITION 3.7.3. *Suppose there is a pure state $|\psi\rangle \in \mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ and let ρ_A and ρ_B be the reduced density matrices. Then*
(i) $S(\rho_A) = S(\rho_B)$.
(ii) $S(\rho_A) = S(\rho_B) = 0$ if and only if $|\psi\rangle$ is a tensor product state.

Proof. (i) Let $|\psi\rangle = \sum_{i=1}^r \sqrt{s_i}|f_i^{(A)}\rangle|f_i^{(B)}\rangle$ be the Schmidt decomposition of $|\psi\rangle$, where r is the Schmidt number of $|\psi\rangle$. Extend $\{|f_1^{(a)}\rangle, \ldots, |f_r^{(a)}\rangle\}$ to an orthonormal basis $\{|f_1^{(a)}\rangle, \ldots, |f_r^{(a)}\rangle, |e_{r+1}^{(a)}\rangle, \ldots |e_{n_a}^{(a)}\rangle\}$ for $\mathcal{H}_a$, $a = A$ and B. Partial traces of $|\psi\rangle\langle\psi|$ with respect these bases yield

$$\rho_A = \mathrm{Tr}\,_B|\psi\rangle\langle\psi| = \sum_{i=1}^r s_i|f_i^{(A)}\rangle\langle f_i^{(A)}| \text{ and } \rho_B = \mathrm{Tr}\,_A|\psi\rangle\langle\psi| = \sum_{i=1}^r s_i|f_i^{(B)}\rangle\langle f_i^{(B)}|,$$

from which we obtain $S(\rho_A) = S(\rho_B) = -\sum_{i=1}^r s_i\log s_i$.
(ii) If $|\psi\rangle = |\psi_A\rangle|\psi_B\rangle$ be a tensor product state, then $|\psi\rangle\langle\psi| = |\psi_A\rangle\langle\psi_A| \otimes |\psi_B\rangle\langle\psi_B|$. It follows $\rho_A = |\psi_A\rangle\langle\psi_A|$ and $\rho_B = |\psi_B\rangle\langle\psi_B|$, and $S(\rho_A) = S(\rho_B) = 0$.

Conversely, suppose $S(\rho_A) = S(\rho_B) = 0$. Then both ρ_A and ρ_B are pure states as we have observed previously. Let us write $\rho_A = |\psi_A\rangle\langle\psi_A|$ and $\rho_B = |\psi_B\rangle\langle\psi_B|$. Since we assume the bipartite state is pure, it must take a form $|\psi\rangle = |\psi_A\rangle|\psi_B\rangle$. ▮

Entanglement entropy vanishes when a pure bipartite state is a tensor product state. Finiteness of entanglement entropy signals entanglement of the bipartite system. In this sense, entanglement entropy works as a measure of purity or entanglement.

EXAMPLE 3.7.4. *In $\mathbb{C}^2 \otimes \mathbb{C}^3$, take a vector*

$$|\psi\rangle = \frac{1}{2}(i|0\rangle|0\rangle + |1\rangle|2\rangle + i|1\rangle|0\rangle + |2\rangle|2\rangle).$$

Let us find the entanglement entropies of this state. It is easy to find the Schmidt decomposition of $|\psi\rangle$ as

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|f_1^{(1)}\rangle|f_1^{(2)}\rangle + |f_2^{(1)}\rangle|f_2^{(2)}\rangle)$$

We do not need the explicit form of $|f_i^{(a)}\rangle$ here. We find

$$\rho_1 = \mathrm{Tr}_2|\psi\rangle\langle\psi| = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \ \rho_2 = \mathrm{Tr}_1|\psi\rangle\langle\psi| = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

where we have taken $\{|f_i^{(a)}\rangle\}$ as bases. We obtain $S(\rho_1) = S(\rho_2) = 2 \times 1/2 \times \log_2 2 = 1$.

In contrast with ordinary thermodynamic entropy, entanglement entropy does not depend on the system size. $S(\rho_A) = S(\rho_B)$ holds even if A is much larger than B.

It should be noted that the equality $S(\rho_A) = S(\rho_B)$ is valid only for pure bipartite states and does not hold for general mixed states. Suppose $\rho_{AB} = \rho_A \otimes \rho_B$ for example, where ρ_A and ρ_B are arbitrary mixed states of subsystems A and B, respectively. Then $S(\rho_A) = -\mathrm{Tr}\,\rho_A \log \rho_A$ and $S(\rho_B) = -\mathrm{Tr}\,\rho_B \log \rho_B$ are different in general.

PROPOSITION 3.7.5. *The von Neumann entropy and its entanglement entropy satisfy*

(i) $S(\rho_{AB}) \le S(\rho_A) + S(\rho_B)$ (Subadditivity)

(ii) $|S(\rho_A) - S(\rho_B)| \le S(\rho_{AB})$ (Araki-Lieb inequality)

(iii) $S(\rho_{AB}) + S(\rho_{BC}) \ge S(\rho_{ABC}) + S(\rho_B)$ (Strong additivity).

The proof of the above inequalities is found in [NC], for example.

3.7.5 Quantum Relative Entropy

DEFINITION 3.7.6. *Let $\rho, \sigma \in \mathcal{S}(\mathcal{H})$. The* **relative entropy** *of ρ to σ is defined by*

$$S(\rho||\sigma) = \mathrm{Tr}\,(\rho \log \rho) - \mathrm{Tr}\,(\rho \log \sigma). \tag{3.31}$$

Note that the relative entropy is well defined only when $\ker \rho \supseteq \ker \sigma$ so that $\log 0$ is multiplied by 0. Otherwise $S(\rho||\sigma)$ is divergent.

PROPOSITION 3.7.7. *Let $\rho, \sigma \in \mathcal{S}(\mathcal{H})$. Then*
(i)

$$S(\rho||\sigma) \geq 0 \qquad (\textbf{Klein's inequality}) \tag{3.32}$$

The equality holds if and only if $\rho = \sigma$.
(ii)

$$S(\rho||\sigma) \geq \frac{1}{2}\|\rho - \sigma\|_1^2, \qquad (\textbf{Pinsker's inequality}) \tag{3.33}$$

where $\|O\|_1 = \mathrm{Tr}\sqrt{O^\dagger O}$.

(iii) Suppose ρ_{AB} and σ_{AB} are states of a bipartite system AB and ρ_A and σ_A are reduced density matrices of ρ_{AB} and σ_{AB}, respectively. Then

$$S(\rho_A||\sigma_A) \leq S(\rho_{AB}||\sigma_{AB}) \qquad (\textit{Monotonicity}). \tag{3.34}$$

The proof is found in [NC], for example. $\|\rho - \sigma\|_1$ is a measure of distinguishability of two states ρ and σ. It is hard to distinguish them if $\|\rho - \sigma\|_1$ is small. The relative entropy gives an upper bound of their distinguishability. As a result, it is harder to distinguish two states if the part B of the bipartite system is ignored by partial trace, which justifies the monotonicity inequality.

Now we are ready to proceed to the world of quantum information and quantum computation. Variations on the themes introduced here and in the previous chapter will appear repeatedly in the following chapters.

3.8 Notes and Open problems

Theorem 3.4.1 followed from a result of Choi [7] as a consequence of completely positive maps in the context of C^*-algebra. There has been new approach to general quantum operations that do not assume that the initial state of the principal system and the environment system is a product state to begin with. For example, see [15].

Using the materials discussed in the first three chapters, we can describe the general ideas on quantum computing and quantum information research.

For example, quantum information theory concerns the use of quantum state to store and transmit data securely and efficiently. For example, using the no-cloning theorem and the fact that measuring quantum states will alter them, we can design secure quantum encryption system as shown in the following chapter. Since quantum states always interact with the environment leading to degradation changes of the data (known as decoherence), we need to design of quantum error correction schemes to fight against such effects. If there is an ensemble of identical quantum state, one may apply measurement to these identical states, and get an estimate of the density matrix representing them. This leads to the tomography problem. Similarly, if one can determine the input and output states of a quantum channel many times, one may get an estimate of the error operators in the operator sum representation of the quantum channel. This is the subject of quantum tomography. In particular, the quantum channel tomography is needed for the design quantum error correction schemes.

One may give a general description of quantum algorithms for computing in terms of vector states as follows.

1. Set up the problem using a quantum sate $|\psi\rangle$, say, with n qubits.
2. Construct a bipartite system by setting up $|\phi\rangle \otimes |\psi\rangle$.
3. Apply a suitable quantum operation to get $|\Psi\rangle = U(|\phi\rangle \otimes |\psi\rangle)$.
4. Apply a suitable measurement to $|\Psi\rangle$ to obtain useful information.

Researchers have also applied quantum effects to other branches of sciences such as photosynthesis, evolution, image processing, quantum machine learning, quantum neural network, quantum cognition, etc. For example, see [16, 17, 18] and their references. In general, one would use a quantum state to encode the data needed to be processed in the problem. Then apply a suitable quantum operation to manipulate quantum state leading to the solution of the problem. For example, for image recognition problem, suppose one wants to decide whether a given image corresponds to the number $0, 1, \ldots, 9$. One can encode $0, 1 \ldots, 9$ as orthonormal vector states $|\psi_0\rangle, \ldots, |\psi_9\rangle$. Then design a quantum operation that will identify correctly a given state $|\psi\rangle$ (corresponding to an imperfect image) to a state in the set $\{|\psi_j : 0 \le j \le 9\}$, which $|\psi\rangle$ meant to be.

Here are some open problems.

1. Let $\rho_1 \in \mathbf{D}_m, \rho_2 \in \mathbf{D}_n$. Determine the set
$$\mathcal{S}(\rho_1, \rho_2) = \{\rho \in \mathbf{D}_{mn} : \mathrm{Tr}\,_1(\rho) = \rho_1, \mathrm{Tr}\,_2(\rho) = \rho_1\}.$$
2. Determine $\rho \in \mathcal{S}(A, B)$ with maximum rank and minimum rank.
3. Determine $\rho \in \mathcal{S}(A, B)$ with maximum von Neumann entropy $S(\rho) = \mathrm{Tr}\,(-\rho \ln \rho)$.
4. More generally, one may consider tripartite system with states in $\mathbf{D}_{n_1n_2n_3}$ and determine $\mathcal{S}(\rho_{23}, \rho_{13}) = \{\rho \in \mathbf{D}_{n_1n_2n_3} : \mathrm{Tr}\,_1(\rho) = \rho_{23}, \mathrm{Tr}\,_2(\rho) = \rho_{13}\}$, where $\rho_{23} \in \mathbf{D}_{n_2n_3}$ and $\rho_{12} \in \mathbf{D}_{n_1n_3}$ are two given states.
5. Given quantum states $\rho_1, \ldots, \rho_k \in \mathbf{D}_n$, $\sigma_1, \ldots, \sigma_k \in \mathbf{D}_m$, determine quantum operation Φ such that $\Phi(\rho_j) = \sigma_j$ for $j = 1, \ldots, k$.

Exercises for Chapter 3

EXERCISE 3.1. *Let $\rho = (P_{ij}) \in \mathbf{D}_{mn}$ with $P_{ij} \in \mathbf{M}_n$. Suppose $\sigma_2 = P_{jj}/\mathrm{Tr}\,(P_{jj}) \in \mathbf{M}_n$ for some $P_{jj} \neq 0$. Then ρ is an uncorrelated state if and only if $\rho_{rs} = a_{rs}\sigma_2$ with $a_{rs} \in \mathbb{C}$ for all $1 \leq r, s \leq m$. If a_{rs} do exist for all $1 \leq r, s \leq m$, then $\rho = \sigma_1 \otimes \sigma_2$ with $\sigma_1 = (a_{rs}) \in M_m$.*

Note: You need to argue why σ_1, σ_2 are density matrices.

EXERCISE 3.2. *Verify that*

$$\rho_1 = \begin{pmatrix} \frac{1+p}{4} & 0 & 0 & \frac{p}{2} \\ 0 & \frac{1-p}{4} & 0 & 0 \\ 0 & 0 & \frac{1-p}{4} & 0 \\ \frac{p}{2} & 0 & 0 & \frac{1+p}{4} \end{pmatrix} \qquad (0 \leq p \leq 1) \tag{3.35}$$

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

EXERCISE 3.3. *Verify that*

$$\rho_2 = \begin{pmatrix} \frac{p}{2} & 0 & 0 & \frac{p}{2} \\ 0 & \frac{1-p}{2} & \frac{1-p}{2} & 0 \\ 0 & \frac{1-p}{2} & \frac{1-p}{2} & 0 \\ \frac{p}{2} & 0 & 0 & \frac{p}{2} \end{pmatrix} \qquad (0 \leq p \leq 1) \tag{3.36}$$

is a density matrix. Show that the negativity vanishes only for $p = 1/2$.

EXERCISE 3.4. *Let*

$$|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle|1\rangle - |1\rangle|0\rangle).$$

Find the corresponding density matrix. Then partial-trace it over the first Hilbert space to find the reduced density matrix of the second system.

EXERCISE 3.5. *Let*

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

be a density matrix with a basis $\{|0\rangle, |1\rangle\}$. Find a purification of ρ_1.

EXERCISE 3.6. *Let*

$$|\Psi\rangle = \sum_{k=1}^{s} \sqrt{p_k}|\psi_k\rangle \otimes |\phi_k\rangle$$

be a purification of $\rho_1 = \sum_{k=1}^{s} p_k|\psi_k\rangle\langle\psi_k| \in \mathbf{D}_m$, where $\{|\phi_1\rangle, \ldots, |\phi_s\rangle\}$ is an orthonormal basis for $\mathbb{C}^s$. Show that

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

is another purification of ρ_1 for any $U \in \mathrm{U}(n)$.

EXERCISE 3.7. *Let U be a unitary operator acting on ρ_1 and ρ_2. Show that*

$$F(U\rho_1 U^\dagger, U\rho_2 U^\dagger) = F(\rho_1, \rho_2). \tag{3.37}$$

EXERCISE 3.8. *Let*

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

Find the fidelity $F(\rho_1, \rho_2)$.

EXERCISE 3.9. *Show that every Hermitian matrix $H \in \mathbf{M}_{mn}$ can be written as a linear combination of density matrices of the form $\rho_1 \otimes \rho_2$ with $\rho_1 \in \mathbf{D}_m$ and $\rho_2 \in \mathbf{D}_n$ using the following steps.*

(1) Let $\{|e_1\rangle, \ldots, |e_m\rangle\}$ be the standard basis for $\mathbb{C}^m$. Consider the set $\mathcal{B}_1$ consisting of matrices in $\mathbf{D}_m$ of the form: $D_j = |e_j\rangle\langle e_j|$ for $1 \le j \le m$, or

$$X_{rs} = |e_r + e_s\rangle\langle e_r + e_s|/2, \ Y_{rs} = |e_r - ie_s\rangle\langle e_r + ie_s|/2 \ \text{ for } 1 \le r < s \le m.$$

If $H = (h_{rs})$ is Hermitian and $h_{rs} = x_{rs} + iy_{rs}$ with $x_{rs}, y_{rs} \in \mathbb{R}$, show that $H - \sum_{r<s}(x_{rs}X_{rs} + y_{rs}Y_{rs})$ is a real diagonal matrix, and hence is a linear combination of $D_1, \ldots, D_m$. Conclude that H is a real linear combination of the elements in $\mathcal{B}_1$.

Similarly, if $\{|f_1\rangle, \ldots, |f_n\rangle\}$ is the standard basis for $\mathbb{C}^n$, then every Hermitian matrix $G = (g_{rs}) \in \mathbf{M}_n$ is a real linear combination of the elements in the set $\mathcal{B}_2$ consisting of elements in $\mathbf{D}_n$ of the form $|f_j\rangle\langle f_j|$ for $1 \le j \le n$, or

$$|f_r + f_s\rangle\langle f_r + f_s|/2, \ |f_r - if_s\rangle\langle f_r + if_s|/2 \ \text{ for } 1 \le r < s \le n.$$

(2) Let $\mathcal{B} = \{\rho_1 \otimes \rho_2 : \rho_1 \in \mathcal{B}_1, \rho_2 \in \mathcal{B}_2\}$. For any Hermitian $T = (T_{rs})_{1 \le r,s \le m}$ with $T_{rs} \in \mathbf{M}_n$, we have $T_{jj} \in \mathbf{M}_n$ is Hermitian, and $T_{rs} = H_{rs} + iG_{rs}$ with Hermitian matrices $H_{rs} = (T_{rs} + T_{rs})/2, G_{rs} = i(T_{ij}^\dagger - T_{ij}) \in \mathbf{M}_n$ for $r < s$. Show that $T_1 = \sum_{r<s}(X_{rs} \otimes H_{rs} + Y_{rs} \otimes G_{rs})$ is a real linear combination of matrices in $\mathcal{B}$, and deduce that $T - T_1$ is a real linear combination of matrices in $\mathcal{B}$, and so is T.

EXERCISE 3.10. *Let $\rho \in \mathbf{D}_{mn}$.*

(a) Suppose ρ is a pure state. Then ρ is separable if and only if $\rho = \rho_1 \otimes \rho_2$ for pure states $\rho_1 \in \mathbf{D}_m, \rho_2 \in \mathbf{D}_n$.

(b) Suppose ρ is a separable. Then ρ is a convex combination of product states $\rho_1 \otimes \rho_2$ such that $\rho_1 \in \mathbf{D}_m, \rho_2 \in \mathbf{D}_n$ are pure states.

EXERCISE 3.11. *Suppose $\Phi : M_n \to M_m$ is a quantum operation with the operator sum representation (3.15)* *(a) Show that* Φ *is trace preserving, i.e.,* $\operatorname{Tr}\Phi(A) = \operatorname{Tr} A$ *for all* $A \in M_n$*.*

[It suffices to check $\Phi(E_{ij}) = \delta_{ij}$*, where* $\{E_{11}, E_{12}, \dots, E_{nn}\}$ *is the standard basis for* $\mathbf{M}_n$*.*

(b) Show that Φ *maps positive semi-definite matrices to positive semi-definite matrices.*

(c) Show that Φ *is* k*-positive for every positive integer* k*, i.e., if* $A = (A_{ij}) \in M_k(M_n)$*, where* $A_{ij} \in M_n$ *for all* $1 \le i, j \le k$*, is positive semi-definite, then so is* $[\Phi(A_{ij})]$*.*

[Note that $[\Phi(A_{ij})] = \sum_{j=1}^{r}(F_j \otimes I_k)A(F_j^\dagger \otimes I_k)$*.]*

EXERCISE 3.12. *Prove the following uncertainty principle for a mixed state* $\rho \in \mathbf{M}_n$*. Let* $A, B \in \mathbf{M}_n$ *be Hermitian matrices,* $(\alpha, \beta) = (\operatorname{Tr}(A\rho), \operatorname{Tr}(B\rho))$*,* $\Delta(A) = \sqrt{\operatorname{Tr}[(A-\alpha I)^2\rho]}$ *and* $\Delta(B) = \sqrt{\operatorname{Tr}[(B-\beta I)^2\rho]}$*. Then*

$$\Delta(A)\Delta(B) \ge |\operatorname{Tr}([A,B]\rho)|/2.$$

The equality holds if and only if there is $\theta \in [0, 2\pi)$ *such that*

$$\cos\theta(A-\alpha I)\rho^r + i\sin\theta(B-\beta I)\rho^r = 0 \tag{3.38}$$

for any positive number r*.*

References

[1] E. Riefell and W. Polak, Quantum Computing: A Gentle Introduction, MIT Press (2011).

[2] M. A. Neilsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press (2000).

[3] D. M. Greenberger, M. A. Horne and A. Zeilinger, in *'Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, ed. M. Kafatos, Kluwer, Dordrecht (1989). Also avilable as arXiv:0712.0921 [quant-ph].

[4] W. Dür, G. Vidal and J. I. Cirac, Phys. Rev. A **62**, 062314 (2000).

[5] A. Einstein, B. Podolsky, N. Rosen, Phys. Rev. **41**, 777 (1935).

[6] Y.-C. Liang, Y.-H. Yeh, P.E.M.F. Mendonça, R.Y. Teh, M.D. Reid, and P.D. Drummond, Quantum fidelity measures for mixed states, Reports on Progress in Physics, Volume 82, Number 7, (2019).

[7] M.-D. Choi, Completely Positive Linear Maps on Complex Matrices, Linear Algebra and its Applications, 10, 285–290 (1975).

[8] K. Kraus, States, Effects and Operations: Fundamental Notions of Quantum Theory, Springer Verlag, (1983)

[9] K. Hornberger, e-print quant-ph/0612118 (2006).

[10] H. Barnum, M. A. Nielsen and B. Schumacher, Phys. Rev. A **57**, 4153 (1998).

[11] Y. Kondo, *et al.*, J. Phys. Soc. Jpn. **76** (2007) 074002.

[12] G. Lindblad, Commun. Math. Phys. **48**, 119 (1976).

[13] V. Gorini, A. Kossakowski and E. C. G. Sudarshan, J. Math. Phys., **17**, 821 (1976).

[14] Gurvits, L., Classical deterministic complexity of Edmonds' problem and quantum entanglement, in Proceedings of the 35th ACM Symposium on Theory of Computing, ACM Press, New York, 2003.

[15] H. Hayashi, G. Kimura, and Y. Ota,. Kraus representation in the presence of initial correlations. Physical Review A - Atomic, Molecular, and Optical Physics, 67(6), 2003. 621091-621095, 2003.

[16] M. Mohseni, Y. Omar, G. Engel, M.B. Plenio, Quantum effects in biology, Cambridge University Press, Cambridge, 2014.

[17] M. Asano, A. Khrennikov, M. Ohya, Y. Tanaka, I. Yamato, Quantum Adaptivity in Biology: From Genetics to Cognition, Springer, New York, 2020.

[18] F. Yan and S.E. Venegas-Andraca, Quantum Image Processing, Springer, New York, 2010.