41076: Methods in Quantum Computing

`Measurements, norms and quantum channels`

Dr. Mária Kieferová based on the materials from Dr. Min-Hsiu Hsieh
Centre for Quantum Software & Information, Faculty of Engineering and Information Technology, University of Technology Sydney

Abstract

Contents to be covered in this lecture are

1. Measurement
2. Norms and distance measures
3. Quantum channels
4. Noise channels

1 Quantum Measurement

Quantum measurement is a process to observe the classical information within a quantum state. It can destroy the superposition property of a quantum state. The quantum measurement postulate evolves from *Born's rule* in his seminal paper in 1926, which states that "the probability density of finding a particle at a given point is proportional to the square of the magnitude of the particle's wave function at that point". Given the qubit state $|b\rangle = \alpha|0\rangle + \beta|1\rangle$, Born's rule says that we can observe this qubit in state $|0\rangle$ with probability $|\alpha|^2$ and in state $|1\rangle$ with probability $|\beta|^2$. Furthermore, after the measurement, the qubit state $|b\rangle$ will disappear and collapse to the observed state $|0\rangle$ or $|1\rangle$.

In general, a quantum measurement is mathematically described by a collection of $\Upsilon := \{M_i\}$, where each measurement operator $M_i \in \mathcal{L}(\mathcal{H})$ satisfies

$$\sum_i M_i = I \tag{1}$$

and each M_i is apositive semi-definite operator. We call these measurements positive operator-valued measure (POVM). The probability of obtaining an outcome i on a quantum state ρ is

$$p_i := \mathrm{Tr}(M_i\rho). \tag{2}$$

The state after measurement will be altered as

$$\rho_i := \frac{M_i\rho}{p_i}.$$

The normalised condition in Eq. (1) guarantees that

$$\begin{aligned}\sum_i p_i &= \sum_i \mathrm{Tr}(M_i\rho) \\ &= \mathrm{Tr}\left(\sum_i M_i\rho\right) \\ &= \mathrm{Tr}\,\rho = 1.\end{aligned} \tag{3}$$

Projective Measurement and Observables

A special instance of quantum measurements is the *projective* measurement. A projective measurement Υ is a collection of projectors $\{P_0, P_1, \cdots, P_{L-1}\}$ which sum to identity. Note that $P_iP_j = 0$ for $i \neq j$ and $P_i^2 = P_i$. When we measure a quantum state $|\phi\rangle$ with Υ, we will get the outcome j with probability

$$p_j := \mathrm{Tr}(P_j|\phi\rangle\langle\phi|)$$

and the resulting state

$$\frac{P_j|\phi\rangle}{\sqrt{p_j}}.$$

A projective measurement $\Upsilon = \{P_i\}$ with the corresponding measurement outcomes $\{\lambda_i\} \in \mathcal{R}$ can be efficiently represented by a Hermitian matrix $H = \sum_i \lambda_i P_i$. Such a matrix is called an *observable*. In physics, an observable is a physical quantity that can be measured. Examples of *observables* of a physical system include the position or momentum of a particle, among many others.

Measuring the observable H means that performing the projective measurement $\Upsilon = \{P_i\}$ on a quantum state $|\phi\rangle$. It follows that the expected value of the outcomes if we measure the state $|\phi\rangle$ with $\Upsilon = \{P_i\}$ is

$$\langle H\rangle := \sum_i \lambda_i \,\mathrm{Tr}\, P_i|\phi\rangle\langle\phi| = \langle\phi|H|\phi\rangle. \tag{4}$$

Exercise 1. *Show that every POVM can be constructed by a projective measurement on a larger Hilbert space.*

Given a measurement, we cannot always distinguish quantum states. For example, let us take $|+\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$, $|-\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}}$ and measurement statistics for $M = \{|0\rangle\langle 0|, |1\rangle\langle 1|\}$. However, we could distinguish them if we measured $M' = \{|+\rangle\langle +|, |-\rangle\langle -|\}$.

We can use quantum measurements to characterize quantum states. This process is known as **tomography**.

Exercise 2. *A single qubit is fully characterized by a vector $\vec{r}, |r| \leq 1$ such that*

$$\rho = \frac{1}{2}(I + r_0\sigma_x + r_1\sigma_y + r_2\sigma_z). \tag{5}$$

Take a set of operators

$$M = \left\{\frac{I+X}{6}, \frac{I-X}{6}, \frac{I+Y}{6}, \frac{I-Y}{6},, \frac{I+Z}{6}, \frac{I-Z}{6}\right\}. \tag{6}$$

Show that

1. *M is a POVM (operators are positive and sum to identity).*
2. *M is tomographically complete, i.e. measuring enough times will allow us to learn the vector r.*

2 Distance Measures

Now we understand how to talk about quantum states, operations and measurements. However, in practice, we won't be always able to prepare the exact right quantum state or apply the perfect desired operation. To talk about how close we are to the desired result, we need to introduce the notion of distance. We are already familiar with computing norms for pure states.

- Norm $\|\psi\| = \langle\psi|\psi\rangle$

We can extend the concept of norms to matrices.

2.1 Matrix Norm

We will introduce a few useful matrix norms in this section. First of all, every norm $\|\cdot\|$ must satisfy the following conditions.

- $\|A\| \geq 0$ with equality if and only if $A = 0$.
- $\|\alpha A\| = |\alpha|\|A\|$ for any $\alpha \in \mathbb{C}$.
- Triangle inequality: $\|A + B\| \leq \|A\| + \|B\|$.

Definition 3 (Schatten norm)**.** *For $p \in [1, \infty)$, the Shatten p-norm of a matrix $A \in \mathbb{C}^{m\times n}$ is defined as*

$$\|A\|_p := \mathrm{Tr}(|A|^p)^{\frac{1}{p}} \tag{7}$$

where $|A| := \sqrt{A^\dagger A}$. We extend $p \to \infty$ as follows

$$\|A\|_\infty := \max\{\|A\boldsymbol{x}\| : \forall \boldsymbol{x} \in \mathbb{C}^n,\ \|\boldsymbol{x}\| = 1\}. \tag{8}$$

Properties of Schatten p-norms are summarized below

1. The Schatten norms are unitarily invariant: for any unitary operators U and V

$$\|UAV\|_p = \|A\|_p \tag{9}$$

   for any $p \in [1, \infty]$.

2. The Schatten norms satisfy Hölder's inequality: for $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times \ell}$, it holds that

$$\|AB\|_1 \leq \|A\|_p\|B\|_q, \tag{10}$$

   where $p, q \geq 1$ and $\frac{1}{p} + \frac{1}{q} = 1$.

3. Sub-multiplicativity: for $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times \ell}$, it holds that

$$\|AB\|_p \leq \|A\|_p\|B\|_p. \tag{11}$$

4. Monotonicity: for $1 \leq p \leq q \leq \infty$, it holds that

$$\|A\|_1 \geq \|A\|_p \geq \|A\|_q \geq \|A\|_\infty. \tag{12}$$

Exercise 4. *Denote by $\sigma_i(A)$ the i-th (non-zero) singular value of A. Show that*

$$\|A\|_p = \left(\sum_i (\sigma_i(A))^p\right)^{\frac{1}{p}}. \tag{13}$$

There are important special cases of Schatten p-norm. Specifically, the Schatten 1-norm is commonly known as the *trace norm*, and will lead to the definition of trace distance in Sec. 2.3. The Schatten 2-norm is also known as the *Frobenius norm* whose explicit form is given below.

Definition 5 (Frobenuis norm)**.** *The Frobenius norm (or the Hilbert-Schmidt norm) of a matrix $A \in \mathbb{C}^{m\times n}$ is defined as*

$$\|A\|_2 \equiv \|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |A_{i,j}|^2}. \tag{14}$$

Finally, the Schatten ∞-norm is also called the *operator norm* or the *spectral norm* whose definition is given in Eq. (8).

2.2 Distance

A distance function is a function d which maps pairs of objects to real numbers and satisfies the following rules:

- The distance between an object and itself is always zero.
- The distance between distinct objects is always positive.
- Distance is symmetric: the distance from x to y is always the same as the distance from y to x.
- Triangle inequality: $d(x, y) \leq d(x, z) + d(z, y)$

We can compute the distance between to matrices as

$$\|A - B\|. \tag{15}$$

Exercise 6. *Verify that this definition satisfies the conditions of a distance.*

Depending on which norm is used, the distance can have different interpretations.

2.3 Trace Distance and Fidelity

We will introduce two commonly used distance measures in quantum information science; namely the trace distance and fidelity.

Definition 7 (Trace Distance)**.** *The trace distance between two operators A and B is given by*

$$Tr(A,B) = \frac{1}{2}\|A-B\|_1 := \frac{1}{2}\operatorname{Tr}|A-B|.$$

The trace distance of two density operators is an extension of total variation distance of probability measures:

$$T(P,Q) = \frac{1}{2}\sum_x |p(x)-q(x)|, \tag{16}$$

where P and Q are probability distributions with pdf $p(x)$ and $q(x)$, respectively.

Properties of the trace distance include

- $T(\rho,\sigma) = 0$ if and only if $\rho = \sigma$.
- Invariant under unitary operation: $T(U\rho U^\dagger, U\sigma U^\dagger) = T(\rho,\sigma)$
- Contraction: $T(\mathcal{N}(\rho),\mathcal{N}(\sigma)) \leq T(\rho,\sigma)$, where $\mathcal{N}$ is any trace-preserving and completely positive map.
- Convexity: $T(\sum_i p_i\rho_i,\sigma) \leq \sum_i p_i T(\rho_i,\sigma)$.

The trace distance is related to the maximum probability of distinguishing between two quantum states through **Helstrom measurement**

$$p_{success} = \frac{1}{2}(1+T(\rho,\sigma)). \tag{17}$$

Assuming that an unknown state is either ρ or σ with equal probabilities, $p_{success}$ is the maximum probability of distinguishing which state we have been given.

Definition 8 (Fidelity)**.** *For $\rho,\sigma \in \mathcal{D}(\mathcal{H})$, their fidelity is*

$$F(\rho,\sigma) := \left(\operatorname{Tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}\right)^2.$$

Note that fidelity is not a distance between quantum states, however, infidelity $I(\rho,\sigma) = 1 - F(\rho,\sigma)$ would be a distance measure.

Exercise 9. *Show that, fidelity has the following properties*

1. $0 \leq F(\rho,\sigma) \leq 1$.
2. $F(U\rho U^\dagger, U\sigma U^\dagger) = F(\rho,\sigma)$.
3. $F(|\psi_\rho\rangle, |\psi_\sigma\rangle) = |\langle\psi_\rho|\psi_\sigma\rangle|^2$.
4. *Symmetry:* $F(\rho,\sigma) = F(\sigma,\rho)$.

Exercise 10. *What is the fidelity between a maximally mixed state and any pure state?*

3 Quantum Channels

Recall that the most general operation on quantum states is a quantum channel, also known as a completely positive trace-preserving map (CPTP map). Any such channel can be written as

$$\Phi(\sigma) = \sum_i B_i \sigma B_i^\dagger \quad \text{where} \quad \sum_i B_i^\dagger B_i = \mathbf{1}. \tag{18}$$

This is known as the Kraus representation and the operators B_i as Kraus operators.

Examples

- Dephasing Channel:
$$\mathcal{N}(\rho) = (1-p)\rho + pZ\rho Z.$$
- Depolarizing Channel:
$$\mathcal{N}(\rho) = (1-p)\rho + p\pi,$$
where π is the completely mixed state.
- Pauli Channel:
$$\mathcal{N}(\sigma) = \sum_{i,j=0}^{1} p(i,j) Z^i X^j \sigma X^j Z^i$$
where we denote $X^0 = Z^0 = I$.
- Measure-and-prepare channel: For a POVM $\{\Lambda_i\}$ and a collection of quantum states $\{\sigma_i\}$, we can define
$$\mathcal{N}(\rho) = \sum_i \sigma_i \operatorname{Tr}(\Lambda_i \rho). \tag{19}$$
This channel is also known as an *entanglement-breaking* channel.

Formally, the average fidelity of a channel is defined with respect to the identity channel

$$F(\mathcal{E}) = \int d\psi \langle\psi|\mathcal{E}(\psi)|\psi\rangle \tag{20}$$

as an average over all state fidelities. To obtain the average, we must integrate over all the quantum states in a given Hilbert space with equal weightings and satisfy $\int d\psi = 1$. This is known as integration over Haar measure.

Exercise 11. *Compute the fidelity of a qubit depolarizing channel* $\mathcal{E}(\rho) = (1-p)|\psi\rangle\langle\psi| + p\frac{I}{d}$.

We can then use the definition for any channel by seen as a perfect channel followed by a noise on the identity channel. Computing average gate fidelities can be further simplified using Nielsen's formula [3]. In a special case when the channel is unitary, we can compute its fidelity (with respect to the identity channel) as

$$F(U) = \frac{d + |Tr(U)|^2}{d + d^2}. \tag{21}$$

Exercise 12. *1. Verify that $F(I)$=1 in* (21)*.*

2. The hottest quantum startup promises to do quantum computing by implementing Hadamard and Toffoli gates. However, they have a minor issue: their Toffoli gates are not working and they are simply doing nothing (i.e. identity gates). What is the fidelity of their "Toffoli" gate?

3. What if they replace all m-controlled-NOT gates with the identity?

Further Reading

A very good lecture note by Ronald de Wolf can be downloaded here [1].

For a better understanding of quantum channels, I would recommend [2].

References

[1] Ronald de Wolf, *Quantum computing: Lecture notes*, 2019.

[2] Vinayak Jagadish and Francesco Petruccione, *An invitation to quantum channels*, arXiv preprint arXiv:1902.00909 (2019).

[3] Michael A Nielsen, *A simple formula for the average gate fidelity of a quantum dynamical operation*, Physics Letters A **303** (2002), no. 4, 249–252.