Chapter 9

Distance measures for quantum information

What does it mean to say that two items of information are similar?
What does it mean to say that information is preserved by some process?

- static measure : concerned with two broad classes of distance measure
- dynamic measure : quantify how well information has been preserved during a dynamic process

→ trace distance & fidelity

9.1 Distance measures for classical information

The *Hamming distance* : a way of quantifying the distance between classical codes. Unfortunately, the Humming distance between two objects is simply a matter of labeling, and *a priori* there aren't any labels in the Hilbert space arena of quantum mechanics.

A much better place to launch the study of distance measures for quantum information is with the comparison of classical probability distributions.

The relative frequency difference of information sources as probability distributions over some alphabet encourages us to concentrate on the comparison of probability distributions in our search for measures of distance.

What does it mean to say that tow probability distributions $\{p_x\}$ and $\{q_x\}$ over the same index set, x, are similar to one another? (there's no unique correct answer)The widely used two measures are the trace distance and the fidelity.

The *trace distance* is defined by the equation:

$$D(p_x, q_x) \equiv \frac{1}{2}\sum_x |p_x - q_x|. \tag{9.1}$$

This quantity is sometimes known as the L_1 *distance* or *Kolmogorov distance*.

The trace distance is a metric on probability distributions.

1. $D(x, x) = 0$

2. $D(x, y) = D(y, x)$
3. $D(x, z) \leq D(x, y) + D(y, z)$

The second measure of distance between probability distributions, the *fidelity* of the probability distributions $\{p_x\}$ and $\{q_x\}$, is defined by

$$F(p_x, q_x) \equiv \sum_x \sqrt{p_x q_x}. \tag{9.2}$$

The fidelity is a very different way of measuring distance between probability distributions that is the trace distance. It is not a metric ($F(p_x, p_x) = \sum_x p_x = 1$) A better geometric understanding of the fidelity is $F(p, q) = \sqrt{p} \cdot \sqrt{q} = \cos \sigma$.

The trace distance and fidelity are mathematically useful means of defining the notion of a distance between two probability distributions.

Do these measures have physically motivated operational meaning? In the case of trace distance, the answer to this question is yes.

$$D(p_x, q_x) = \max_S |p(S) - q(S)| = \max_S \left| \sum_{x \in S} p_x - \sum_{x \in S} q_x \right|, \tag{9.3}$$

where the maximization is over all subsets S of the index set $\{x\}$. The quantity being maximized is the difference between the probability that the event S occurs, according to the distribution $\{p_x\}$, and the probability that the event S occurs, according to the distribution $\{q_x\}$. The event S is thus in some sense the optimal event to examine when trying to distinguish the distance $\{p_x\}$ and $\{q_x\}$, with the trace distance governing how well it is possible to make this distribution.

Unfortunately, a similarly clear interpretation for the fidelity is not known. (open questions?)

It turns out that there are close connections between the fidelity and the trace distance, so properties of one quantity can often be used to deduce properties of the other.

The trace distance and fidelity are static measures of distance for comparing two fixed probability distributions.

A *dynamic measure of distance* : measures how well information is preserved by some physical process.

X : a input
Y : a output
(to form a Markov process : $X \to Y$.

This dynamic measure of distance can be understood as a special case of the static trace distance!

1. A copy of the input X : $\tilde{X} = X$

2. X now passes through the noisy channel, leaving as output the random variable Y.

How close is the initial perfectly correlated pair, $(\tilde{X}, X)$, to the final pair, $(\tilde{X}, Y)$? Using the trace distance as our measure of 'closeness', we obtain with some simple algebra,

$$\begin{aligned}
D((\tilde{X}, X), (X, Y)) &= \frac{1}{2} \sum_{xx'} |\delta_{xx'} p(X = x) - p(\tilde{X} = x, Y = x')| && (9.4)\\
&= \frac{1}{2} \sum_{x \neq x'} p(\tilde{X} = x, Y = x') + \frac{1}{2} \sum_x |p(X = x) - p(\tilde{X} = x, Y = x')| && (9.5)\\
&= \frac{1}{2} \sum_{x \neq x'} p(\tilde{X} = x, Y = x') + \frac{1}{2} \sum_x (p(X = x) - p(\tilde{X} = x, Y = x')) && (9.6)
\end{aligned}$$

$$= \frac{p(\tilde{X} \neq Y) + 1 - p(\tilde{X} = Y}{2} \tag{9.7}$$

$$= \frac{p(X \neq Y) + p(\tilde{X} \neq Y)}{2} \tag{9.8}$$

$$= p(X \neq Y). \tag{9.9}$$

Thus the probability of an error in the channel is equal to the trace distance between the probability distribution for $(\tilde{X}, X)$ and $(\tilde{X}, Y)$. This is an important construction, since it will be the basis for analogous quantum constructions. This is necessary because there is no *direct* quantum analogue of the probability $p(X \neq Y)$, since there is no notion in quantum mechanics analogous to the joint probability distribution for variables X and Y that *exist at different times*