QUANTUM INFORMATION & COMPUTATION

Nilanjana Datta, DAMTP Cambridge

1 Quantum states as information carriers

Recall from the introduction that information is represented by distinguishable states of a physical system. In classical physics it is axiomatic that any two different states are perfectly distinguishable by a measurement, but in quantum physics, quantum measurements can reliably (i.e. with probability 1) distinguish alternative states only if they lie in orthogonal subspaces of the state space. Hence two (or more) states cannot be reliably distinguished unless they are (pairwise) orthogonal. A quantum system with a d-dimensional state space, despite having infinitely many different (pure) states, i.e. kets, can represent at most d reliably distinguishable messages, and a qubit is the simplest quantum system that can be used to represent a classical bit.

The idea of quantum information
In quantum information theory, information is encoded in the states of a quantum system. For closed systems (considered in this lecture course) these are given by kets in the state space of the system. Such states (also called pure states) are the most definite kind of state that a quantum system can have, and we can reliably *prepare* any desired pure state. But if we receive such a state (of unknown identity) we cannot identify it with certainty yet we are still receiving a kind of definite signal or message, albeit unreadable in the classical sense. We use the term *quantum information* to refer to what we acquire when we receive a quantum state. This turns out to be a very useful concept with many intriguing properties (some of which we will soon see) if thought of as a kind of quantum analogue of familiar classical information. It is a useful intuition being the "stuff" that's processed in quantum computations and communicated over quantum channels. Quantum measurement provides a link back to classical information. A rich theory of its properties and applications has been developed, called quantum Shannon theory, inspired by the formalism and applications of classical information as given in Shannon's classical information theory.

Classical information can be thought of (or represented) as a special case of quantum information in which all quantum states are required to be drawn from a fixed orthonormal set (e.g. computational basis states of a qubit) and its tensor powers.

Given some quantum information in the form of an unknown quantum state $|\psi\rangle$ of a quantum system S, there are three basic kinds of operations that we can perform on them:

(Ancilla): we can take a fixed (known) quantum state $|A\rangle$ of a second quantum system A (called the *ancilla*) and adjoin it to $|\psi\rangle$ to obtain the state $|\psi\rangle\,|A\rangle$ of the composite system SA. This has the useful effect of embedding our quantum information into a

larger dimensional space, that of the joint system SA, having dimension $d_1 d_2$ where d_1 resp. d_2 are the dimensions of the state spaces of S and A.

(Unitary): we can apply a unitary operator U of our choice so that $|\psi\rangle$ becomes $U|\psi\rangle$. We can also apply a unitary operator on the joint state of the composite system SA and later discard the ancilla A.

(Measure): we can perform a measurement on (possibly only part of) the joint state of SA, record the result, and retain the post-measurement state for further processing. The output here will generally be a probabilistic mixture of the possible post-measurement states, with probabilities as given by the Born rule.

The most general quantum action we can apply to the state of the system S is then represented by a sequence of these three basic operations.

1.1 The no-cloning theorem

We now give our first tantalising property of quantum information which is not shared by classical information: quantum information cannot be copied or 'cloned'!

The copying of classical information is a very familiar process (e.g photocopying) and it is common to view the possibility of copying as obvious and unremarkable. In fundamental terms, given some classical information represented in the state of a classical physical system A (e.g. text ink and paper page) we can take another similar system B, initially in some fixed ("blank") state independent of A (blank sheet of paper). We can then carry out a physical operation (operation of the photocopier) on the joint system AB to reliably measure or "read" the state of A (without any corruption) and evolve the state of B to that measured state, to achieve cloning.

Now let us consider similar ideas for quantum information to establish what we'd mean by *quantum* copying or cloning. Our process will involve three quantum subsystems A, B and M. A will contain our quantum information to be copied; B is system (with state space of same dimension as A) which should finally contain the copy that we seek; M will represent any extra "machinery" or physical objects that are needed in the cloning process (like the photocopy machine for classical copying).

Initially A will contain $|\psi\rangle$ and B will contain some standard "blank" state denoted $|0\rangle$; M will also be in some fixed starting state, denoted $|M_0\rangle$ representing the initial "ready" state of the machine and any materials that it may require. A crucially important point here is that the initial state $|0\rangle|M_0\rangle$ of BM should be independent of $|\psi\rangle$.

The cloning process will be a fixed quantum physical evolution of ABM achieving the following transformation:

$$|\psi\rangle_A |0\rangle_B |M_0\rangle_M \longrightarrow |\psi\rangle_A |\psi\rangle_B |M_\psi\rangle_M$$

i.e. the quantum information is copied into B (while also remaining intact in A) and the final state of M is allowed to be arbitrary and even depend possibly on $|\psi\rangle$ (as indicated).

The cloning process may be required to work for all states $|\psi\rangle$ of A or alternatively only for some restricted subset of states.

No-cloning theorem. Let $\mathcal{S}$ be any set of states of A that contains at least one non-orthogonal pair of states. Then no unitary cloning process exists that achieves cloning for all states in $\mathcal{S}$. $\square$

Remark.
The no-cloning theorem actually remains true for *arbitrary* prospective cloning processes, not just unitary ones i.e. even if the further operations of (Ancilla) and (Measure) are allowed to be included. If (Measure) is used then all probabilistic branches are required to lead to perfect cloning. The use of these extra operations can in fact be reduced to the fully unitary case, and in this course we will prove only the unitary case.

Proof of the no-cloning theorem (for unitary processes)
Let $|\xi\rangle$ and $|\eta\rangle$ be two distinct non-orthogonal states in $\mathcal{S}$. Then the cloning process must do both the following evolutions:

$$|\xi\rangle_A |0\rangle_B |M_0\rangle_M \longrightarrow |\xi\rangle_A |\xi\rangle_B |M_\xi\rangle_M$$
$$|\eta\rangle_A |0\rangle_B |M_0\rangle_M \longrightarrow |\eta\rangle_A |\eta\rangle_B |M_\eta\rangle_M$$

(for some possibly different states $|M_\xi\rangle$ and $|M_\eta\rangle$ of M). Now, any unitary process preserves inner products so the inner product of the two initial states must equal that of the two final states:

$$\langle\xi|\eta\rangle \langle 0|0\rangle \langle M_0|M_0\rangle = \langle\xi|\eta\rangle \langle\xi|\eta\rangle \langle M_\xi|M_\eta\rangle \qquad (*)$$

Taking absolute values in eqn. (*) and using $\langle 0|0\rangle = \langle M_0|M_0\rangle = 1$ we get

$$|\langle\xi|\eta\rangle| = |\langle\xi|\eta\rangle|^2 |\langle M_\xi|M_\eta\rangle|.$$

Since $|\xi\rangle \neq |\eta\rangle$ and $|\xi\rangle$ is not orthogonal to $|\eta\rangle$ we have $0 \lneqq |\langle\xi|\eta\rangle| \lneqq 1$ and cancelling it gives

$$1 = |\langle\xi|\eta\rangle| \, |\langle M_\xi|M_\eta\rangle|$$

which is a contradiction since $|\langle\xi|\eta\rangle| \lneqq 1$ and $|\langle M_\xi|M_\eta\rangle| \leq 1$. $\square$

Example A. (cloning and superluminal signalling)
The no-cloning theorem was proved in 1982 independently by D. Dieks and by W. Wootters & W. Zurek. The theorem also appears (at least implicitly) in earlier work of D. Park of 1970 but this went completely unnoticed until recently. The 1982 work arose in response to a proposal by F. Herbert for a method of superluminal (in fact instantaneous) communication using quantum methods. If Herbert's result were correct it would have cast serious doubt on quantum theory as an acceptable physical theory! Fortunately there was an error in Herbert's argument – he had taken for granted without justification or discussion that quantum states could be cloned!

Herbert's method was as follows. Suppose Alice and Bob are distantly separated and they share the entangled state

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|0\rangle |0\rangle + |1\rangle |1\rangle)$$

i.e. they each hold one qubit of this 2-qubit state. Note that we can also write (as is easily directly checked):

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

where $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$ are the Pauli X eigenstates.

Alice wants to communicate a yes/no decision instantaneously to Bob at noon. She does the following.
Alice's action:
at noon, for 'yes' she measures her qubit in the standard (i.e. Pauli Z) basis $\{|0\rangle, |1\rangle\}$, and for 'no' she measures her qubit in the X basis $\{|+\rangle, |-\rangle\}$.
According to the Born rule, after her 'yes' action, Bob's qubit will be in state $|0\rangle$ or $|1\rangle$ with 50/50 probability; and after her 'no' action, Bob's qubit will be in state $|+\rangle$ or $|-\rangle$ with 50/50 probability (as is immediately clear from the second formula for $|\phi^+\rangle$ above).

Fact:
these 'yes' and 'no' preparations of Bob's qubit are completely indistinguishable by any local action (measurement) on Bob's qubit i.e. they each give exactly the same probability distribution of outcomes for any measurement (and also in fact the same as that for Alice having done no action at noon!) Indeed if Π_i is the projection operator for outcome i of a measurement by Bob, then in the 'yes' case, his probability of seeing i is

$$\mathrm{prob}_{\mathrm{yes}}(i) = \frac{1}{2}\langle 0|\,\Pi_i\,|0\rangle + \frac{1}{2}\langle 1|\,\Pi_i\,|1\rangle = \mathrm{Tr}\,\Pi_i\left(\frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2}\right)$$

(since $\langle A|B\rangle = \mathrm{Tr}\,(|B\rangle\langle A|)$.
On the other hand, in the 'no' case we similarly get

$$\mathrm{prob}_{\mathrm{no}}(i) = \mathrm{Tr}\,\Pi_i\left(\frac{|+\rangle\langle +| + |-\rangle\langle -|}{2}\right).$$

Now (as can easily be checked by computing the 2×2 matrices of components) we have

$$|0\rangle\langle 0| + |1\rangle\langle 1| \;=\; |+\rangle\langle +| + |-\rangle\langle -| = I \quad \text{(completeness relation)}$$

so Bob cannot detect any effect of Alice's attempted signalling!

But: Suppose Bob can clone quantum states. Then he proceeds as follows.

Bob's action:
Immediately after noon he clones his qubit to make many copies (say one million copies). Then he measures them all in the standard basis to obtain a bit string B of length one million, of the measurement results.

In the 'yes' case, all the qubits will be $|0\rangle$ or all $|1\rangle$ so B will be the string of all 0's or all 1's.

In the 'no' case all qubits will be $|+\rangle$ or all $|-\rangle$. However both of these give a 50/50 outcome of 0 and 1 upon standard basis measurement so B will now have a uniformly random bit string of length one million. This is easily distinguishable from the 'yes' case

except with negligibly small probability $2/2^{10^6}$, so with arbitrarily high probability Bob can instantaneously get Alice's message.

The 'Fact' above is actually a special case of the so-called quantum no-signalling principle which we will discuss later. □

Now moving on from no-cloning, we will consider another property of quantum information:

1.2 Distinguishing non-orthogonal states

Suppose we are given an unknown quantum state $|\psi\rangle$ that is promised to be one of two non-orthogonal states $|\alpha_i\rangle$ for $i = 0, 1$, and we wish to determine which one it is i.e. the value of subscript i. We have seen that this is impossible to do *with certainty* (which would then e.g. also provide a method for cloning the states!) but can we still obtain *some* information about i, and then, how much? Since quantum measurement outputs are generally probabilistic we ask if we can identify the state while allowing some probability of error in the answer, or failure of the process. For example we could just do nothing with the state and randomly guess $i = 0$ or 1, which would always be correct with probability half. But we can do better than this by performing a quantum measurement on $|\psi\rangle$ to guide our answer.

More formally we consider a state estimation process of the following general kind. Given $|\psi\rangle$ we first adjoin an ancillary system in some fixed state $|A\rangle$ (independent of $|\psi\rangle$) which has the effect of enlarging the state space that we can work in. Then we apply a unitary operation to the joint system and finally a measurement with two outcomes labelled 0 and 1 corresponding to our guess of $|\psi\rangle$ having been $|\alpha_0\rangle$ or $|\alpha_1\rangle$.

We can simplify the mathematical description of our process as follows. Adjoining the ancilla state $|A\rangle$ amounts to just converting the discrimination problem from $|\alpha_0\rangle$ vs. $|\alpha_1\rangle$ to $|\alpha_0\rangle|A\rangle$ vs. $|\alpha_1\rangle|A\rangle$ i.e. just another example of two non-orthogonal states, which in fact even have the same inner product. Secondly applying a unitary U to any state $|\xi\rangle$ before a measurement with orthogonal projectors Π_0 and Π_1, is equivalent to just performing only a measurement with U-rotated orthogonal projectors $\Pi'_i = U^\dagger \Pi_i U$ as these give the same outcome probabilities:

$$\text{prob}(i) = (\langle\xi| U^\dagger)(\Pi_i)(U|\xi\rangle) = \langle\xi|(U^\dagger \Pi_i U)|\xi\rangle .$$

Hence we can recast our state estimation process as just a single measurement: given one of two possible non-orthogonal states $|\alpha_0\rangle$ and $|\alpha_1\rangle$ (which are now the $|\alpha_i\rangle$'s with the ancilla adjoined), perform a single two outcome measurement with projectors Π_0 and Π_1 (which are now the U-rotated versions of the original measurement).

Some measurements will be better than others by providing the correct answers with higher probability. To formalise this we introduce a definition of success probability p_S for the process to quantify how good it is, and we'll seek to optimise this. In the absence of any prior knowledge about which of the two states $|\alpha_0\rangle$ or $|\alpha_1\rangle$ we will receive we

assume a prior probability of half for each. Then the success probability is defined by

$$\begin{aligned} p_S &= \tfrac{1}{2}\,\text{prob}\,\{\text{process outputs 0 given } |\alpha_0\rangle \text{ was sent}\} + \\ &\quad \tfrac{1}{2}\,\text{prob}\,\{\text{process outputs 1 given } |\alpha_1\rangle \text{ was sent}\} \end{aligned}$$

which by the Born rule becomes

$$p_S = \frac{1}{2}\left(\langle\alpha_0|\,\Pi_0\,|\alpha_0\rangle + \langle\alpha_1|\,\Pi_1\,|\alpha_1\rangle\right).$$

Since $\Pi_0 + \Pi_1 = I$ (the identity operator on the full space) we have $\Pi_1 = I - \Pi_0$ and so

$$p_S = \frac{1}{2} + \frac{1}{2}\text{Tr}\left(\,\Pi_0\left(|\alpha_0\rangle\langle\alpha_0| - |\alpha_1\rangle\langle\alpha_1|\right)\right). \qquad (*)$$

The optimal choice of measurement $\{\Pi_0, I - \Pi_0\}$ will be the one that maximises p_S (for the given known pair of states $|\alpha_i\rangle$).

To explicitly identify the optimal measurement let's look in detail at the operator

$$\Delta = |\alpha_0\rangle\langle\alpha_0| - |\alpha_1\rangle\langle\alpha_1|.$$

It has the following properties:
(i) it is self-adjoint so has a complete basis of eigenstates (and all eigenvalues real).
(ii) For any $|\beta\rangle$ orthogonal to both $|\alpha_0\rangle$ and $|\alpha_1\rangle$ we have $\Delta\,|\beta\rangle = 0$. Hence Δ has at most two non-zero eigenvalues, whose eigenvectors must lie in the span of $|\alpha_0\rangle$ and $|\alpha_1\rangle$.
(iii) $\text{Tr}\,\Delta = 0$ so the two non-zero eigenvalues sum to 0. Writing them as $+\delta$ and $-\delta$, and corresponding normalised eigenvectors as $|p\rangle$ and $|m\rangle$ respectively, we have

$$\Delta = \delta\,|p\rangle\langle p| - \delta\,|m\rangle\langle m|.$$

(iv) We can determine δ in terms of $|\alpha_0\rangle$ and $|\alpha_1\rangle$ as follows. Working in the 2 dimensional subspace spanned by $|\alpha_0\rangle$ and $|\alpha_1\rangle$, choose a unit vector $\left|\alpha_0^\perp\right\rangle$ orthogonal to $|\alpha_0\rangle$ and write $|\alpha_1\rangle = c_0\,|\alpha_0\rangle + c_1\left|\alpha_0^\perp\right\rangle$. Thus in components relative to the $\{|\alpha_0\rangle, \left|\alpha_0^\perp\right\rangle\}$ basis we have

$$|\alpha_0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |\alpha_1\rangle = \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} \quad \text{with } |c_0|^2 + |c_1|^2 = 1.$$

So

$$\Delta = \begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \end{pmatrix} - \begin{pmatrix} c_0 \\ c_1 \end{pmatrix}\begin{pmatrix} c_0^* & c_1^* \end{pmatrix} = \begin{pmatrix} |c_1|^2 & -c_0c_1^* \\ -c_1c_0^* & -|c_1|^2 \end{pmatrix}.$$

A straightforward calculation shows that $\det(\Delta - \delta I) = 0$ has solutions $\delta = \pm|c_1| = \pm\sin\theta$, where we have introduced θ defined by $\cos\theta = |\langle\alpha_0|\alpha_1\rangle|$.

Now, finally returning to our formula eq. (*) for p_S and substituting our expression for Δ, we get

$$\begin{aligned} p_S &= \tfrac{1}{2} + \tfrac{\delta}{2}\text{Tr}\left(\,\Pi_0\left(|p\rangle\langle p| - |m\rangle\langle m|\right)\right) \\ &= \tfrac{1}{2} + \tfrac{\delta}{2}\left(\langle p|\,\Pi_0\,|p\rangle - \langle m|\,\Pi_0\,|m\rangle\right). \end{aligned}$$

For any projector Π and state $|\xi\rangle$ we have $0 \le \langle\xi|\,\Pi\,|\xi\rangle \le 1$ (since $\langle\xi|\,\Pi\,|\xi\rangle$ is the squared length of the projected state $\Pi\,|\xi\rangle$). Thus we see that p_S achieves its maximum value of

$(1+\delta)/2$ if Π_0 is chosen to be any subspace that contains $|p\rangle$ (so $\Pi_0 |p\rangle = |p\rangle$) and is orthogonal to $|m\rangle$ (so $\Pi_0 |m\rangle = 0$); and then $\Pi_1 = I - \Pi_0$ will have $\Pi_1 |m\rangle = |m\rangle$. Such a choice of Π_0 is always possible since $|p\rangle$ and $|m\rangle$ are always orthogonal.

In particular for example if $|\alpha_0\rangle$ and $|\alpha_1\rangle$ are qubit states then we can just work entirely in their two dimensional space and an optimal measurement to discriminate them will be the measurement relative to the eigenbasis of the hermitian operator $\Delta = |\alpha_0\rangle \langle\alpha_0| - |\alpha_1\rangle \langle\alpha_1|$.

The achievable optimal success probability above is known as the Holevo-Helström bound for discriminating pure states.

In summary, we have proven

Theorem (Holevo-Helström bound for pure states): Given one of two equally likely states $|\alpha_0\rangle$ and $|\alpha_1\rangle$ with $|\langle\alpha_0|\alpha_1\rangle| = \cos\theta$, the probability p_S of correctly identifying the state by any quantum measurement process is bounded by $p_S \leq \frac{1}{2}(1+\sin\theta)$ and the bound is tight (i.e. achieved by a particular choice of measurement). $\square$

Remark (unambiguous state discrimination)
Finally we mention that by "changing the rules of the game" we can formulate other interesting kinds of state discrimination tasks, such as so-called *unambiguous state discrimination*. For this task we are again given an unknown one of two states $|\alpha_0\rangle$ and $|\alpha_1\rangle$ and we want to construct quantum measurement process with *three* outcomes called 0, 1 and 'fail' with the following properties:
(i) if measurement outcome 0 occurs then the state was certainly $|\alpha_0\rangle$;
(ii) if measurement outcome 1 occurs then the state was certainly $|\alpha_1\rangle$;
(iii) if measurement outcome 'fail' occurs then our process has failed and we have generally irretrievably lost all information about the given state.
In this scenario we would seek to minimise the average probability of obtaining the third outcome.

Both scenarios fall short of providing reliably perfect discrimination of non-orthogonal states albeit in interestingly different ways: in the first we always get an answer 0 or 1 with the caveat that it may be incorrect, whereas in the second the 0 and 1 answers are always certainly correct but the catch now is that the process sometimes fails, and destroys the state (so we cannot try again!) See exercise sheet 1 for an example of an actual unambiguous state discrimination process.

1.3 The no-signalling principle

Consider two parties Alice and Bob separated distantly in space, each holding their own local quantum systems A and B respectively. Suppose they possess a (generally) entangled) joint quantum state $|\phi_{AB}\rangle$ of the joint system (sometimes called a *bipartite state* since there are two subsystems). They each have access only to their respective part of the bipartite state, which they can manipulate locally by quantum actions. Suppose Alice performs a complete measurement on her subsystem A. According to the Born rule, for each measurement outcome the state of Bob's system will change instantaneously. If

Bob could notice this change then they would be able to communicate instantaneously!

Example B.
Suppose that the shared state of a bipartite system AB is the entangled state

$$\left|\Phi^+_{AB}\right\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right).$$

If Alice performs a standard basis measurement on A then Bob's system will be "collapsed" into pure state $|0\rangle$ or $|1\rangle$ corresponding to Alice's outcome, each occurring with probability half. Before the measurement however, B was not in a pure state (but was part of an entangled state). Can Bob notice this change by just local actions on his subsystem? Note also that he does not know Alice's outcome - acquisition of that knowledge would require communication from Alice!

Suppose Bob performs a complete measurement in basis $\{|b_i\rangle : i = 0, 1\}$. By the (extended) Born rule, after Alice's measurement he will get

$$\text{prob}(i) = \frac{1}{2}\left|\langle 0|b_i\rangle\right|^2 + \frac{1}{2}\left|\langle 1|b_i\rangle\right|^2 = \frac{1}{2}\langle b_i|b_i\rangle = \frac{1}{2}$$

(where we have averaged over Alice's two possible outcomes using their probabilities of half each). However before Alice's measurement he would have had

$$\text{prob}(i) = \left\langle\phi^+_{AB}\right| \left(I \otimes P^B_i\right) \left|\phi^+_{AB}\right\rangle = \left\langle\phi^+_{AB}\right| \left(I \otimes |b_i\rangle\langle b_i|\right) \left|\phi^+_{AB}\right\rangle = \frac{1}{2} \quad \text{too.}$$

Thus even though each individual outcome of Alice's measurement will give noticeably different probabilities of i for Bob (viz. $|\langle 0|b_i\rangle|^2$ and $|\langle 1|b_i\rangle|^2$), if Bob does not know Alice's outcome he must average over their probabilities and his ability to notice any change is lost! □

This turns out to be true in full generality: for any bipartite state $|\phi\rangle_{AB}$, no local actions by Alice can be noticed by Bob locally i.e. for any local measurement by Bob, the output probability distribution is always unaffected by any local action by Alice. This is the quantum no signalling principle. It seems bizarrely remarkable that quantum theory appears to involve non-local effects (at the level of state descriptions viz. post-measurement states arising from local actions on composite systems) yet the full quantum formalism conspires to prevent us from being able to harness this nonlocality for communication!

We now give a more formal formulation and proof of the no-signalling principle.

Local operations on a composite system

(loc-Unitary): a local unitary operation U by Alice resp. Bob on a bipartite system is mathematically represented as the operator $U_A \otimes I_B$ resp. $I_A \otimes U_B$ on the full state of AB (and here I is the identity operation). Note that any two local unitary operations on disjoint subsystems always commute (as $(U \otimes I)(I \otimes V) = (I \otimes V)(U \otimes I) = U \otimes V$)).

(loc-Ancilla): Alice and Bob can adjoin local ancillary systems A' and B' which simply enlarge their locally held systems.

(loc-Measure): Let $\mathcal{H}_A$, $\mathcal{H}_B$ and $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ denote the state spaces of A, B and AB respectively. If Alice performs a (generally incomplete) local measurement on A corresponding to the decomposition of $\mathcal{H}_A$ into the orthogonal subspaces $\mathcal{E}_a$'s labelled by outcomes a, then mathematically on the full state space this is represented by the measurement with the orthogonal decomposition $\mathcal{E}_a \otimes \mathcal{H}_B$'s of $\mathcal{H}_{AB}$. In particular even a complete measurement on A will be an incomplete measurement of the full system.

If $\mathcal{F}_b$'s are the orthogonal subspaces of a local measurement by Bob on B with outcomes b, then one can check from the Born rule that the joint probabilities $\text{prob}(a, b)$ obtained from performing both measurements is independent of whether A or B goes first, or whether they measure 'simultaneously', corresponding to the global measurement with orthogonal subspaces $\mathcal{E}_a \otimes \mathcal{F}_b$ for all pairs (a, b).

No-signalling theorem: Suppose Alice and Bob have access to subsystems A and B respectively of a joint state $|\phi_{AB}\rangle$. Then Alice cannot convey any information to Bob by performing local operations i.e. no local action by Alice can change the output probability distribution of any local quantum process by Bob.

Proof: Consider first the basic case of Bob performing a complete measurement on B relative to a basis $\{|b\rangle\}$ labelled by the outcomes b. We use this basis of B to express $|\phi_{AB}\rangle$ as

$$|\phi_{AB}\rangle = \sum_{a,b} c_{ab} |a\rangle |b\rangle . \qquad (*)$$

Here $\{|a\rangle\}$ denotes an orthonormal basis of the state space of the system A. By the Born rule, the probability of Bob seeing outcome b in the absence of any local operation by Alice is given by

$$\begin{aligned} p(b) \equiv \text{prob}(b) &= \langle\phi_{AB}| (I_A \otimes P_b) |\phi_{AB}\rangle \\ &= \langle\phi_{AB}| (I_A \otimes |b\rangle \langle b|) |\phi_{AB}\rangle \\ &= \sum_a |c_{ab}|^2 \end{aligned} \qquad (1)$$

The post-measurement state of AB if Bob's measurement outcome is b is given by

$$|\phi'_{AB}\rangle = \frac{(I_A \otimes P_b) |\phi_{AB}\rangle}{\sqrt{p(b)}}. \qquad (2)$$

Now suppose Alice first performs a complete measurement relative to the basis $\{|a\rangle\}$ and subsequently Bob performs his measurement above. If Alice gets an outcome a, which she gets with probability

$$p(a) = \sum_b |c_{ab}|^2, \qquad (3)$$

(as can be easily checked), the post-measurement state of AB is

$$|\phi''_{AB}\rangle = \frac{(P_a \otimes I_B) |\phi_{AB}\rangle}{\sqrt{p(a)}}, \qquad (4)$$

where $P_a = |a\rangle\langle a|$. Bob on doing his measurement now (that is after Alice gets an outcome a), gets an outcome b with probability

$$\begin{aligned} p(b|a) &= \langle \phi''_{AB}| \left(I_A \otimes |b\rangle\langle b|\right) |\phi''_{AB}\rangle \\ &= \frac{1}{p(a)} \langle \phi_{AB}| \left(P_a \otimes I_B\right)(I_A \otimes P_b) |\phi_{AB}\rangle . \end{aligned} \tag{5}$$

Then the joint probability of Alice getting outcome a and Bob getting outcome b is given by

$$p(a,b) = p(b|a)p(a) = \langle \phi_{AB}| \left(P_a \otimes P_b\right) |\phi_{AB}\rangle = |c_{ab}|^2.$$

(which actually holds regardless of which time order the local measurements are performed in). Then the marginal probability distribution for b i.e. the distribution that Bob will see, is

$$p(b) = \sum_a p(a,b) = \sum_a |c_{ab}|^2,$$

which is the same (cf above) as in the case of Alice not having done anything on her side[1].□

Remark (communication complexity)
The fact that quantum theory appears to include non-local effects has a long history going back at least to the iconically influential 1935 'EPR' paper by A. Einstein, B. Podolsky and N. Rosen. Later in the 1960's J. Bell introduced what are now known as the Bell inequalities, providing a much simplified and experimentally accessible way of demonstrating the non-local effects. But (perhaps because of the no-signalling property) these effects were largely ignored by "serious physicists" and viewed as just an awkward curiosity or inconvenience. Then early in the 1990's something remarkable happened: it was realised that if Alice and Bob shared entangled states *and were also allowed classical communication too*, then although entanglement *by itself* cannot provide communication (by no-signalling) it can nevertheless greatly *assist* (when used alongside classical communication) by greatly reducing the amount of classical communication needed to achieve some distributed tasks, involving inputs from both Alice and Bob. In some cases the amount of classical communication could be reduced by a massive exponential amount at the expense of a modest consumption of entangled states used alongside. As a result, a whole new research area, called quantum communication complexity was born.

The basic scenario is the following: Alice and Bob possess separate n-bit strings x and y and they wish to compute some joint function $f(x,y)$ of both strings. Clearly they'll need to communicate (e.g. at least the results of some intermediate local calculations) and n bits of communication each way always suffices (they just exchange the information of their strings, and each can then compute f locally). It is now known that for some f's, if Alice and Bob share some entanglement (e.g. some $|\phi^+\rangle$ states), then f can be computed using exponentially less classical communication than is possible by any method involving

[1]This argument may be readily generalised to Alice and/or Bob performing incomplete measurements and/or Alice first performing local unitaries on her system A. Moreover, if Alice and Bob include extra local ancillas this just has the effect of enlarging their local spaces and the above arguments go through unchanged in this albeit enlarged scenario.

just classical communication. (and for other f's entanglement provides no help at all). Thus a communication network having shared entanglement along its connections (a 'quantum internet') can solve some distributed computing tasks with exponentially less classical bit traffic across the network. Correspondingly entanglement is now recognised to be a preciously valuable communication resource. In so-called quantum teleportation (cf below) we'll see another communication use for entanglement, this time for the task of communicating *quantum* states. □

1.4 The Bell basis

The state of two qubits A and B given by

$$|\Phi^+_{AB}\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right) \tag{6}$$

which appeared in Example A and Example B is said to be a *maximally entangled state*. It has the following important feature: even though the state of the composite 2–qubit system is known exactly, we have no information about the states of each individual qubit. Say the two qubits correspond to two electrons. Then, if we perform any local measurement of A or B, e.g. if we measure spin of A along *any* axis, the result is completely random. There is a probability 1/2 of getting spin–up and a probability 1/2 of getting spin–down.

The state $|\Phi^+_{AB}\rangle$ belongs to a complete orthonormal basis of the state space $\mathbb{C}^2 \otimes \mathbb{C}^2$ of the two qubits. This basis consists of the following 4 maximally entangled states:

$$\begin{aligned} |\Phi^\pm_{AB}\rangle &= \frac{1}{\sqrt{2}}\left(|00\rangle \pm |11\rangle\right) \\ |\Psi^\pm_{AB}\rangle &= \frac{1}{\sqrt{2}}\left(|01\rangle \pm |10\rangle\right) \end{aligned} \tag{7}$$

These states are known as Bell states or EPR states after the various scientists (Bell, Einstein, Podolsky and Rosen) who first appreciated the significance of entanglement. The basis is often referred to as the *Bell basis*. Sometimes the labels A and B are suppressed (as in Example A).

In physical terms, if $|0\rangle$ and $|1\rangle$ are the spin $+1/2$ and spin $-1/2$ states of a spin half particle (say in the Z direction) then $\left|\Psi^-_{AB}\right\rangle$ is the spin-zero singlet state of two spin half particles and the other three Bell states span the 3-dimensional spin-1 triplet space.

The four Bell states can be characterized by 2 *classical bits* as follows (the subscript AB has been omitted for notational simplicity):

1. One bit is the *parity bit*.

   Choose 0 to denote that the spins are parallel, i.e., the state is a $\left|\Phi^\pm_{AB}\right\rangle$ state. Choose 1 to denote that the spins are antiparallel, i.e., the state is a $\left|\Psi^\pm_{AB}\right\rangle$ state.

2. The other bit is the *phase bit* (+ or −). It specifies what superposition of the two states of like parity the particular Bell state corresponds to, e.g. choose 0 to denote + and 1 to denote −.

e.g. The classical bit string 01 corresponds to the Bell state $|\Phi^-_{AB}\rangle$.

Hence we can encode two classical bits in the state of the two qubit system. This information can be recovered by performing a joint measurement (on the two qubits) which projects onto the Bell basis. This is called a *Bell measurement* and is a measurement in which the measurement operators are the projectors $|\Phi^\pm_{AB}\rangle\langle\Phi^\pm_{AB}|$ and $|\Psi^\pm_{AB}\rangle\langle\Psi^\pm_{AB}|$.

The outcomes of the Bell measurement are the binary sequences 00, 01, 10 and 11, and the corresponding projection operators are $P_{00} = \left|\Phi^+_{AB}\right\rangle\left\langle\Phi^+_{AB}\right|$, $P_{01} = \left|\Phi^-_{AB}\right\rangle\left\langle\Phi^-_{AB}\right|$, $P_{10} = \left|\Psi^+_{AB}\right\rangle\left\langle\Psi^+_{AB}\right|$, and $P_{11} = \left|\Psi^-_{AB}\right\rangle\left\langle\Psi^-_{AB}\right|$.

Information recovery is possible, however, only if the two qubits are in the same location so that joint measurement is possible.

1.5 Superdense coding

We have seen that an individual qubit (e.g. if received as a quantum message) can reliably encode only a single classical bit, corresponding to having a maximum of two mutually orthogonal states. Superdense coding is a way of doubling this information capacity: a receiver (say Bob) can reliably extract two classical bits from a single qubit which he receives from a sender (say, Alice) *provided* Alice and Bob initially share a Bell state.

Note that in terms of the operators $X := \sigma_X$, $Y := i\sigma_y$ and $Z := \sigma_z$ we have the following:

$$\begin{array}{rcl}
\left|\Phi^+_{AB}\right\rangle & = & (I\otimes I)\left|\Phi^+_{AB}\right\rangle \\
\left|\Phi^-_{AB}\right\rangle & = & (Z\otimes I)\left|\Phi^+_{AB}\right\rangle \quad = \quad (I\otimes Z)\left|\Phi^+_{AB}\right\rangle \\
\left|\Psi^+_{AB}\right\rangle & = & (X\otimes I)\left|\Phi^+_{AB}\right\rangle \quad = \quad (I\otimes X)\left|\Phi^+_{AB}\right\rangle \\
\left|\Psi^-_{AB}\right\rangle & = & (Y\otimes I)\left|\Phi^+_{AB}\right\rangle \quad = \quad -(I\otimes Y))\left|\Phi^+_{AB}\right\rangle
\end{array}$$

(Note: from the action of the Pauli operators from the states $|0\rangle, |1\rangle$, it follows that $Y|0\rangle = -|1\rangle$ and $Y|1\rangle = |0\rangle$.)

The superdense coding protocol
Alice and Bob (distantly separated in space) each possess one qubit of a $\left|\Phi^+_{AB}\right\rangle$ state. In order to reliably communicate two classical bits to Bob by sending him only a single qubit, Alice first locally applies the operation I, Z, X or Y to her qubit, to represent the messages 00, 01, 10 or 11 respectively, and then sends her qubit over to Bob. On receiving Alice's qubit Bob simply performs a Bell measurement on the two qubits which he now holds, to reliably read out Alice's 2-bit message. *Details in lecture.*

We will see the Bell measurement again below when we discuss quantum teleportation.