Part II — Quantum Information and Computation

Based on lectures by Dr Nilanjana Datta and notes by thirdsgames.co.uk

Lent 2023

Contents

§1 Mathematical background

§1.1 Motivation

In classical computation, the elementary unit of information is the **bit**, which takes a value in $\{0, 1\}$. This gives the result of a single binary decision problem, where the zero and one correspond to different answers to the problem. Binary strings of length greater than one are used to provide more than 2 answers to a problem; if we have n bits, we can encode 2^n different messages.

Classical computation is understood to be the processing of information: taking an initial bit string and and updating it by a prescribed sequence of steps. The steps are taken to be the action of local Boolean logic gates, such as conjunction, disjunction, or negation. At each step, a small number of bits in prescribed locations are edited.

Information in the real world must be tied to a physical representation. For example, bits in a processor are often represented by different voltages of specific components. Importantly, there is no information **without** representation. Performing a computation classically must therefore involve the evolution of a physical system over time, which is covered by the laws of classical physics.

However, nature does not abide by classical physics at subatomic levels, and we must use quantum mechanics to accurately model such behaviours. One such behaviour modelled by quantum mechanics is the superposition principle, that the corresponding quantum analog of the bit need not be in precisely one state. Quantum entanglement is the phenomenon where particles can be linked in such a way that their states can be manipulated even at a distance. Quantum measurement is probabilistic and alters the underlying system.

Quantum information and computation therefore exploits these features of quantum mechanics to address issues of information storage, communication, computation, and cryptography. The features of quantum mechanics seem to allow us benefits which are beyond the limits of classical information and computation, even in principle. Note that a quantum computer cannot perform any task that cannot in principle be performed classically. We only hope that quantum techniques allow a reduction in the complexity of certain algorithms.

§1.2 Benefits of quantum information and computation

In complexity theory, we study the **hardness** of a certain computational task. One must consider the resources required for the task; which in classical computation are normally limited to time (measured in number of computational steps) and space (amount of memory required).

If an algorithm takes time bounded by a polynomial function in the input size n, we say the algorithm is **polynomial-time**. Otherwise, we say it is an **exponential-time** algorithm. Polynomial-time algorithms are typically taken to be computable in practice, but exponential-time algorithms are usually considered only computable in principle. Quantum mechanical techniques can provide polynomial-time algorithms that have only exponential-time classical versions. One example is Shor's integer factorisation algorithm.

Quantum states of physical sytems can be used to encode information, such as spin states of electrons. There are certain tasks possible with such quantum states which are impossible in classical physics; one example is quantum teleportation.

There are also some technological issues with classical physics. Components of processors have become minified to atomic scale, and therefore they cannot be shrunk much further without dealing with the effects of quantum mechanics. Conversely, there are technological challenges with quantum physics. Quantum systems are very fragile, and modern quantum computers typically require temperatures close to absolute zero to reduce noise.

Quantum supremacy refers to the hypothetical moment at which a programmable quantum computer can first solve a problem in practice that a classical computer cannot. At the time of writing, there is no concensus that quantum supremacy has been achieved.

§1.3 Hilbert spaces

Every quantum mechanical system is associated with a Hilbert space $\mathcal{V}$, a complex inner product space that is a complete metric space with respect to the distance function induced by the inner product. We use Dirac's **bra-ket** notation: a vector is represented by $|v\rangle \in \mathcal{V}$, and its conjugate transpose is denoted $\langle v| \in \mathcal{V}^\star$. If $\mathcal{V} = \mathbb{C}^n$, we write

$$|\psi\rangle = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}; \quad \langle\psi| = \begin{pmatrix} a_1^\star & \cdots & a_n^\star \end{pmatrix}$$

The inner product of ψ and φ is written $\langle\psi|\varphi\rangle$. Recall that an inner product satisfies

- $\langle\psi|\psi\rangle \geq 0$, and equal to zero if and only if $|\psi\rangle = 0$;
- linearity in the second argument, so $\langle\psi|a\varphi_1 + b\varphi_2\rangle = a\langle\psi|\varphi_1\rangle + b\langle\psi|\varphi_2\rangle$;
- antilinearity in the first argument, so $\langle a\psi_1 + b\psi_2|\varphi\rangle = a^\star\langle\psi_1|\varphi\rangle + b^\star\langle\psi_2|\varphi\rangle$;
- skew-symmetry, so $\langle\psi|\varphi\rangle^\star = \langle\varphi|\psi\rangle$;

and induces a norm $\|\psi\| = \||\psi\rangle\| = \sqrt{\langle\psi|\psi\rangle}$. In this course, we will often consider $\mathcal{V} = \mathbb{C}^2$ and define

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

For an arbitrary $|v\rangle \in \mathbb{C}^2$, we can write $|v\rangle = a|0\rangle + b|1\rangle$, giving

$$|v\rangle = \begin{pmatrix} a \\ b \end{pmatrix}; \quad \langle v| = \begin{pmatrix} a^\star & b^\star \end{pmatrix}$$

If $|w\rangle = c|0\rangle + d|1\rangle$, then $\langle v|w\rangle = a^\star c + b^\star d$.

We can also compute the **outer product** of two vectors, defined to be $|\psi\rangle\langle\varphi| = |\psi\rangle\langle\varphi|$. If $\mathcal{V} = \mathbb{C}^n$, the outer product is an $n \times n$ matrix. An orthonormal basis $(|i\rangle)_{i=1}^n$ for $\mathcal{V}$ is called **complete** if $\sum_{i=1}^n |i\rangle\langle i|$ is the identity matrix.

If $\mathcal{V}$ has a complete orthonormal basis, we can write $|\psi\rangle = \sum_{i=1}^n c_i |i\rangle$ for some c_i. If $\langle\psi|\psi\rangle = 1$, we say $|\psi\rangle$ is **normalised**. In this case, $\sum |c_i|^2 = 1$, and the $|c_i|^2$ form a discrete probability distribution. We call the c_i the **probability amplitudes**.

Let $\mathcal{V}, \mathcal{W}$ be vector spaces, where $\dim \mathcal{V} = n, \dim \mathcal{W} = m$. Let $|v\rangle \in \mathcal{V}, |w\rangle \in \mathcal{W}$. Suppose $|v\rangle = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix}^\intercal$, and $|w\rangle = \begin{pmatrix} b_1 & \cdots & b_m \end{pmatrix}^\intercal$. Then, $|v\rangle \otimes |w\rangle$ is the **tensor product** of $|v\rangle$ and $|w\rangle$, defined by

$$|v\rangle \otimes |w\rangle = \begin{pmatrix} a_1 b_1 \\ \vdots \\ a_1 b_m \\ a_2 b_1 \\ \vdots \\ a_n b_m \end{pmatrix} \in \mathcal{V} \otimes \mathcal{W}$$

If $(|e_i\rangle)_{i=1}^n$ is a complete orthonormal basis for $\mathcal{V}$ and $(|f_j\rangle)_{j=1}^m$ is a complete orthonormal basis for $\mathcal{W}$, then $(|e_i\rangle \otimes |f_j\rangle)_{i,j=1}^{n,m}$ is a complete orthonormal basis for $\mathcal{V} \otimes \mathcal{W}$. We sometimes write $|v\rangle \otimes |w\rangle$ as $|v\rangle|w\rangle$ or $|vw\rangle$. Note that this is not commutative

If $|\alpha\rangle \in \mathcal{V}$, we can write $|\alpha\rangle = \sum a_i |e_i\rangle$, and similarly if $|\beta\rangle \in \mathcal{W}$, we can write $|\beta\rangle = \sum b_j |f_j\rangle$. Then, $|\alpha\beta\rangle = \sum a_i c_j |e_i f_j\rangle$.

We say $|\Psi\rangle \in \mathcal{V} \otimes \mathcal{W}$ is a **product vector** if $|\Psi\rangle = |\psi\rangle \otimes |\varphi\rangle$ for some ψ, φ. Vectors that are not product vectors are called **entangled vectors**.

Let $\mathcal{V} = \mathbb{C}^2 = \mathcal{W}$. Define $|\varphi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Suppose $|\varphi^+\rangle = |\psi\rangle \otimes |\varphi\rangle = (a|0\rangle + b|1\rangle) \otimes (c|0\rangle + d|1\rangle)$. Then, $|\varphi^+\rangle = ac|00\rangle + ad|01\rangle + bc|10\rangle + bd|11\rangle$. So one of a and d, and one of b and c is equal to zero, contradicting the assumption, so $|\varphi^+\rangle$ is entangled.

We define the inner product on the product space by defining

$$\langle\varphi_1|\psi_2\rangle = (\langle\alpha_1|\langle\beta_1|)(|\beta_2\rangle|\alpha_2\rangle) = \langle\alpha_1|\alpha_2\rangle\langle\beta_1|\beta_2\rangle$$

where $|\psi_i\rangle = |\alpha_i\rangle |\beta_i\rangle$. In the general case, $|A\rangle = \sum a_{ij} |e_i\rangle |f_j\rangle, |B\rangle = \sum b_{ij} |e_i\rangle |f_j\rangle$, and we define

$$\langle A|B\rangle = \left(\sum a_{ij}^\star \langle e_i| \langle f_j|\right)\left(\sum b_{ij} |e_i\rangle |f_j\rangle\right) = \sum a_{ij}^\star b_{ij} \delta_{ii'} \delta_{jj'} = \sum a_{ij}^\star b_{ij}$$

where δ is the Kronecker δ symbol.

We define the k-fold **tensor power** of a vector space $\mathcal{V}$ by

$$\mathcal{V}^{\otimes n} = \underbrace{\mathcal{V} \otimes \cdots \otimes \mathcal{V}}_{n \text{ times}}$$

If $\mathcal{V} = \mathbb{C}^2$, this has dimension 2^k, and complete orthonormal basis $|i_1 \dots i_k\rangle$ for $i_j \in \{0, 1\}$. Note that $|v\rangle |w\rangle \neq |w\rangle |v\rangle$.

§1.4 First postulate: quantum states

In this course, we will restrict our attention to finite-dimensional vector spaces, and finite time evolution. We describe the **postulates** for quantum mechanics that we will work under.

The first postulate is that, given an isolated quantum mechanical system S, we can associate a finite-dimensional vector space $\mathcal{V}$. The physical state of the system is given by a unit vector $|\psi\rangle$ in $\mathcal{V}$. More precisely, the state is given by a **ray**, an equivalence class of vectors $e^{i\theta} |\psi\rangle$ for $\theta \in \mathbb{R}$. No measurements can distinguish states in a given equivalence class. Note that states $a |\psi_1\rangle + b |\psi_2\rangle$ and $a |\psi_1\rangle + be^{i\theta} |\psi_2\rangle$ can be distinguished by measurement, since the phase difference is relative, not global.

Example 1.1

Let $\mathcal{V} = \mathbb{C}^2$ with (complete orthonormal) basis $|0\rangle, |1\rangle$. The elementary unit of quantum information is known as the **qubit**, which is any quantum system with $\mathcal{V} = \mathbb{C}^2$. The spin of an electron, which is some superposition of spin-up and spin-down, can be modelled by $\mathbb{C}^2$. A property of the polarisation of a photon, such as vertical or horizontal, or right-circular or left-circular, can also be modelled in this way.

Define $|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ and $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$. This is another complete orthonormal basis for $\mathcal{V}$, sometimes called the **conjugate basis**.

§1.5 Second postulate: composite systems

The second postulate of quantum mechanics is that two quantum systems S_1, S_2 with associated vector spaces $\mathcal{V}_1, \mathcal{V}_2$ can be composed into the **composite system** with vector space $\mathcal{V}_1 \otimes \mathcal{V}_2$.

Example 1.2

Consider $\mathcal{V}^{\otimes n}$, the space of n qubits. An orthonormal basis is $|i_1 \dots i_n\rangle$ where $i_j \in \{0,1\}$. A vector in $\mathcal{V}^{\otimes n}$ can be written $\sum a_{i_1 \dots i_n} |i_1 \dots i_n\rangle$. There are 2^n different amplitudes $a_{i_1 \dots i_n}$, providing exponential growth in information. However, in a product state, we obtain only linear growth in information.

§1.6 Observables

An **observable** is a property of a physical system which can, in theory, be measured. Mathematically, these are modelled by linear self-adjoint (or Hermitian) operators.

The action of a linear operator A on a state space $\mathcal{V}$ is a written $A|\psi\rangle$. By linearity, we have $A(a|\psi\rangle + b|\varphi\rangle) = aA|\psi\rangle + bA|\varphi\rangle$ for $a, b \in \mathbb{C}$. For any operator A acting on $\mathcal{V}$, there is a unique linear operator $A^\dagger$ such that $\langle v|Aw\rangle = \left\langle A^\dagger v \middle| w \right\rangle$, called the **adjoint** of A; operators equal to their adjoints are called **self-adjoint**.

We can easily show that $(AB)^\dagger = B^\dagger A^\dagger$. By convention, we define $|\psi\rangle^\dagger = \langle\psi|$, so for a self-adjoint operator A, we have $(A|\psi\rangle)^\dagger = \langle\psi| A$. There are four important operators which act on the single-qubit space $\mathbb{C}^2$.

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}; \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

σ_0 is the identity matrix, and $\sigma_x, \sigma_y, \sigma_z$ are called the **Pauli matrices**. The actions of these matrices on the basis vectors $|0\rangle$ and $|1\rangle$ are

$$\sigma_0|0\rangle = |0\rangle; \quad \sigma_0|1\rangle = |1\rangle; \quad \sigma_x|0\rangle = |1\rangle; \quad \sigma_x|1\rangle = |0\rangle;$$

$$\sigma_y|0\rangle = i|1\rangle; \quad \sigma_y|1\rangle = -i|0\rangle; \quad \sigma_z|0\rangle = |0\rangle; \quad \sigma_z|1\rangle = -|1\rangle$$

Note that

$$\sigma_x\sigma_y = i\sigma_z; \quad \sigma_y\sigma_z = i\sigma_x; \quad \sigma_z\sigma_x = i\sigma_y$$

Intuitively, σ_x is a bit flip, σ_y is a phase flip, and σ_z is a combined bit and phase flip.

§1.7 Dirac notation for linear operators

Let $|v\rangle = a|0\rangle + b|1\rangle$, and $|w\rangle = c|0\rangle + d|1\rangle$. The outer product is

$$M = |v\rangle\langle w| = \begin{pmatrix} a \\ b \end{pmatrix} \begin{pmatrix} c^\star & d^\star \end{pmatrix} = \begin{pmatrix} ac^\star & ad^\star \\ bc^\star & bd^\star \end{pmatrix}$$

which is a linear map on $\mathcal{V} = \mathbb{C}^2$. One can show that $M|x\rangle = (|v\rangle\langle w|)|x\rangle = |v\rangle\langle w|x\rangle$, which is the scalar product of the vector $|v\rangle$ with the inner product $\langle w|x\rangle$. Such outer products yield the linear maps from $\mathbb{C}^2$ to $\mathbb{C}^2$ that have rank 1, and the kernel of M is the subspace of vectors orthogonal to $|w\rangle$. Note that

$$|0\rangle\langle 0| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}; \quad |0\rangle\langle 1| = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}; \quad |1\rangle\langle 0| = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}; \quad |1\rangle\langle 1| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

Hence, we can write

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \implies A = a\,|0\rangle\langle 0| + b\,|0\rangle\langle 1| + c\,|1\rangle\langle 0| + d\,|1\rangle\langle 1|$$

In particular, $|0\rangle\langle 0|, |0\rangle\langle 1|, |1\rangle\langle 0|, |1\rangle\langle 1|$ forms a basis for the vector space $\mathcal{V} \otimes \mathcal{V}^\star$ of linear maps on $\mathcal{V}$. Note also that $\langle w|v\rangle = \operatorname{Tr} |v\rangle\langle w|$.

§1.8 Projection operators

Suppose that $|v\rangle$ is a normalised vector, so $\langle v|v\rangle = 1$. Then, $\Pi_v = |v\rangle\langle v|$ is the **projection operator** onto v, satisfying $\Pi_v \Pi_v = \Pi_v$ and $\Pi_v^\dagger = \Pi_v$. In Dirac notation, one can see that

$$\Pi_v \Pi_v = |v\rangle\langle v|\,|v\rangle\langle v| = |v\rangle\,\langle v|v\rangle\,\langle v| = |v\rangle\langle v| = \Pi_v$$

If $|a\rangle$ is orthogonal to $|v\rangle$, then $\Pi_v\,|a\rangle = |v\rangle\,\langle v|a\rangle = 0$. Therefore, $\Pi_v\,|x\rangle$ is the vector obtained by projection of $|x\rangle$ onto the one-dimensional subspace of $\mathcal{V}$ spanned by $|v\rangle$.

Now suppose $\mathcal{E}$ is any linear subspace of some vector space $\mathcal{V}$, and $|e_1\rangle, \dots, |e_d\rangle$ is any orthonormal basis of $\mathcal{E}$. Then,

$$\Pi_{\mathcal{E}} = |e_1\rangle\langle e_1| + \dots + |e_d\rangle\langle e_d|$$

is the projection operator into $\mathcal{E}$. This property can be checked by extending $|e_1\rangle, \dots, |e_d\rangle$ into an orthonormal basis of $\mathcal{V}$.

Note that if $|x\rangle = A\,|v\rangle$, then $\langle x| = (A\,|v\rangle)^\dagger = |v\rangle^\dagger A^\dagger = \langle v| A^\dagger$. Therefore, when constructing inner products, we can write $\langle a|M|b\rangle$ as $\langle a|x\rangle$ or $\langle y|b\rangle$ where $|x\rangle = M\,|b\rangle$ or $|y\rangle = M^\dagger\,|a\rangle$ (so that we have $\langle y| = \langle a|\,M$).

§1.9 Tensor products of linear maps

Suppose A, B are linear maps $\mathbb{C}^2 \to \mathbb{C}^2$. Then, we define $A \otimes B \colon \mathbb{C}^2 \otimes \mathbb{C}^2 \to \mathbb{C}^2 \otimes \mathbb{C}^2$ by its action on the basis $(A \otimes B)\,|i\rangle\,|j\rangle = A\,|i\rangle\,B\,|j\rangle$. In particular, for product vectors we obtain $(A \otimes B)(|v\rangle\,|w\rangle) = A\,|v\rangle \otimes B\,|w\rangle$.

The 4×4 matrix of components of $A \otimes B$ has a simple block form, which can be seen by writing down its action on basis states.

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}; \quad B = \begin{pmatrix} p & q \\ r & s \end{pmatrix} \implies A \otimes B = \begin{pmatrix} aB & bB \\ cB & dB \end{pmatrix} = \begin{pmatrix} ap & aq & bp & bq \\ ar & as & br & bs \\ cp & cq & dp & dq \\ cr & cs & dr & ds \end{pmatrix}$$

Note that $A \otimes I$ and $I \otimes A$ can be thought of as acting only on one of the subspaces. Consider $|\Phi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, and define A as above. Then,

$$\begin{aligned}
(A \otimes I)|\Phi\rangle &= \frac{1}{\sqrt{2}}[(A|0\rangle)|0\rangle + (A|1\rangle)|1\rangle] \\
&= \frac{1}{\sqrt{2}}[(a|0\rangle + c|1\rangle)|0\rangle + (b|0\rangle + d|1\rangle)|1\rangle] \\
&= \frac{1}{\sqrt{2}}[a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle] \\
(I \otimes A)|\Phi\rangle &= \frac{1}{\sqrt{2}}[|0\rangle(A|0\rangle) + |1\rangle(A|1\rangle)] \\
&= \frac{1}{\sqrt{2}}[|0\rangle(a|0\rangle + c|1\rangle) + |1\rangle(b|0\rangle + d|1\rangle)] \\
&= \frac{1}{\sqrt{2}}[a|00\rangle + c|01\rangle + b|10\rangle + d|11\rangle]
\end{aligned}$$

§1.10 Third postulate: physical evolution of quantum systems

The third postulate of quantum mechanics is that any physical finite-time evolution of a closed quantum system is represented by a unitary operation on the corresponding vector space of states. Recall that the following are equivalent for a linear operator U:

- U is unitary, so $U^{-1} = U^\dagger$;
- U maps an orthonormal basis to an orthonormal set of vectors;
- the columns (or rows) of U form an orthonormal set of vectors.

If a system is in a state $|\psi(t_1)\rangle$ at a time t_1 and later in a state $|\psi(t_2)\rangle$ at a time t_2, then $|\psi(t_2)\rangle = U(t_1, t_2)|\psi(t_1)\rangle$ for some unitary map $U(t_1, t_2)$ which depends only on t_1, t_2. This operator is derived from the **Schrödinger equation**, which is

$$i\hbar \frac{\partial}{\partial t}|\psi(t)\rangle = H|\psi(t)\rangle$$

where H is a self-adjoint operator known as the **Hamiltonian**. In particular, if H is time-independent, we have

$$U(t_1, t_2) = e^{-\frac{i}{\hbar}H(t_2 - t_1)}$$

In the more general case,

$$U(t_1, t_2) = e^{-\frac{i}{\hbar}\int_{t_1}^{t_2} H(t)\mathrm{d}t}$$

The unitary evolution of a closed system is deterministic.

§1.11 Partial inner products

A vector $|v\rangle \in \mathcal{V}$ defines a linear map $\mathcal{V} \otimes \mathcal{W} \to \mathcal{W}$ called the **partial inner product** with $|v\rangle$, defined on the basis $|e_i\rangle\,|f_j\rangle$ of $\mathcal{V} \otimes \mathcal{W}$ by $|e_i\rangle\,|f_j\rangle \mapsto \langle v|e_i\rangle\,|f_j\rangle$. Similarly, for any $|w\rangle \in \mathcal{W}$, we obtain a partial inner product $\mathcal{V} \otimes \mathcal{W} \to \mathcal{V}$. If $\mathcal{V}, \mathcal{W}$ are isomorphic, we must specify which partial inner product is intended.

§1.12 Fourth postulate: quantum measurement

Consider a system S with state space $\mathcal{V}$, and let A be an observable. A can be written as its **spectral projection** $A = \sum_k a_k P_k$ where $A\,|\varphi_k\rangle = a_k\,|\varphi_k\rangle$. If a_k is nondegenerate, $P_k = |\varphi_k\rangle\langle\varphi_k|$. If a_k is degenerate of multiplicity m, then $P_k = \sum_{i=1}^m |\varphi_k^i\rangle\langle\varphi_k^i|$.

The fourth postulate is that when an observable is measured, the resulting measurement will be an eigenvalue a_j, with probability $p(a_j) = \langle\psi|P_j|\psi\rangle$. Then, $|\psi\rangle$ is replaced with the post-measurement state

$$\frac{P_j\,|\psi\rangle}{\sqrt{p(a_j)}}$$

This is known as **Born's rule**. Such a measurement is called a **projective measurement** (or sometimes a **von Neumann measurement**), since the post-measurement state is given by a projection operator.

Suppose A, B are operators that do not commute, so $[A, B] = AB - BA \neq 0$. Then, the measurement of A will influence the outcome probabilities of a subsequent measurement of B. For instance, suppose $|\psi\rangle = |+\rangle\,, A = \sigma_z, B = \sigma_x$.

§1.13 Complete and incomplete projective measurements

Let $|\psi\rangle \in \mathcal{V}$ be a state in a state space of dimension n. Let $\mathcal{B} = \{|e_i\rangle\}$ be a set of n orthogonal basis vectors for $\mathcal{V}$. Then $|\psi\rangle = \sum a_j\,|e_j\rangle$ where $a_k = \langle e_k|\psi\rangle$. If the outcomes of a measurement are the indices of the basis vectors $j = 1, \dots, n$, we have $p(j) = \langle\psi|P_j|\psi\rangle$ where $P_j = |e_j\rangle\langle e_j|$. Therefore, $p(j) = |\langle\psi|e_j\rangle|^2 = |a_j|^2$. If the outcome is j, the post-measurement state is

$$\frac{P_j\,|\psi\rangle}{\sqrt{p(j)}} = \frac{|e_j\rangle\,\langle e_j|\psi\rangle}{\sqrt{p(j)}} = |e_j\rangle$$

Hence the state collapses to a basis vector. Taking another measurement immediately in the same basis, we obtain the result j with probability 1. Such a measurement is called a **complete** projective measurement; it is called complete as all P_j are of rank 1. When we measure a state $|\psi\rangle$ in a basis, it is often helpful to consider an orthogonal decomposition of $\mathcal{V}$ using the basis vectors.

Conversely, an **incomplete** projective measurement corresponds to an arbitrary orthogonal decomposition of $\mathcal{V}$. Consider a decomposition of $\mathcal{V}$ into d mutually orthogonal subspaces $\mathcal{E}_1, \dots, \mathcal{E}_d$, so $\mathcal{V} = \mathcal{E}_1 \oplus \dots \oplus \mathcal{E}_d$, and $\dim \mathcal{V} = \sum \dim \mathcal{E}_j$. Let Π_i be a projection operator onto $\mathcal{E}_i$. Since the spaces are mutually orthogonal, $\Pi_i \Pi_j = \delta_{ij} \Pi_i$. Consider a measurement with outcomes $1, \dots, d$ representing a particular subspace. The probability of observing outcome i is $\langle\psi|\Pi_i|\psi\rangle$. If the outcome is i, $|\psi\rangle$ is replaced with $\frac{\Pi_i|\psi\rangle}{\sqrt{p(i)}}$. In this case, the Π_i are no longer rank 1 projection operators. If $\mathcal{E}_i$ has basis $\{|f_j\rangle\}$, we can write $\Pi_i = \sum |f_j\rangle\langle f_j|$.

Incomplete projective measurement is a generalisation of complete projective measurement. One can refine an incomplete measurement into a complete measurement by first considering a complete measurement, and then summing the relevant outcome probabiilities to obtain a description of the incomplete measurement probabilities. Let $\left\{\left|e_k^{(j)}\right\rangle\right\}_{k=1}^{d_j}$ be a basis for $\mathcal{E}_j$ for each j. Then $\mathcal{V} = \bigoplus_{i=1}^{d} \mathcal{E}_j$ has orthonormal basis $\left\{\left|e_k^{(j)}\right\rangle\right\}_{j,k}$. Then, $\left\langle e_i^{(k_1)} \middle| e_j^{(k_2)} \right\rangle = \delta_{ij}\delta_{k_1 k_2}$.

Consider a two-bit string $b_1 b_2$. The **parity** of this string is $b_1 \oplus b_2$, where $\oplus$ represents addition modulo 2. Consider the orthogonal decomposition of $\mathcal{V}$ into $\mathcal{E}_0 \oplus \mathcal{E}_1$, where $\mathcal{E}_0 = \operatorname{span}\{|00\rangle, |11\rangle\}$ is the even parity subspace, and $\mathcal{E}_1 = \operatorname{span}\{|01\rangle, |10\rangle\}$ is the odd parity subspace. The outcomes of an incomplete measurement are then the labels 0 and 1 of the subspaces $\mathcal{E}_0$ and $\mathcal{E}_1$. Note that $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$ is a complete orthonormal basis for $\mathcal{V}$, so we can consider the complete projective measurement. $\langle\psi|P_{00}|\psi\rangle$ is the probability of outcome 00 for the complete measurement, where $P_{00} = |00\rangle\langle 00|$. For the incomplete measurement, $p(0) = \langle\psi|\Pi_0|\psi\rangle$ is the probability of outcome 0, where $\Pi_0 = P_{00} + P_{11}$. So $p(0) = \langle\psi|P_{00}|\psi\rangle + \langle\psi|P_{11}|\psi\rangle$.

§1.14 Extended Born rule

Let S_1, S_2 be quantum systems with state spaces $\mathcal{V}, \mathcal{W}$ with dimensions m, n, and we consider the composite system $S_1 S_2$. Let $\{|e_i\rangle\}$ be a complete orthonormal basis of $\mathcal{V}$, and let $\{|f_j\rangle\}$ be a complete orthonormal basis of $\mathcal{W}$. Suppose the composite system is in an initial state $|\psi\rangle = \sum a_{ij} |e_i\rangle |f_j\rangle$. Suppose now that we want to measure $|\psi\rangle$ in the basis $\{|e_i\rangle\}$; this amounts to an incomplete measurement with subspaces $\mathcal{E}_i = \operatorname{span}\{|e_i\rangle \otimes |\varphi\rangle : |\varphi\rangle \in \mathcal{W}\}$ for $1 \leq i \leq m$. The outcomes of such a measurement are $\{1, \dots, m\}$, and the $\mathcal{E}_i$ are mutually orthogonal. The probability of a given outcome is

$p(k) = \langle\psi|P_k \otimes I|\psi\rangle$, where $P_k = |e_k\rangle\langle e_k|$. Hence,

$$p(k) = \left(\sum a^\star_{i'j'} \langle e'_i| \left\langle f'_j \right|\right)(|e_k\rangle\langle e_k| \otimes I)\left(\sum a_{ij} |e_i\rangle |f_j\rangle\right) = \sum_{j=1}^{n} a^\star_{kj} a_{kj}$$

If the outcome is k, then the post-measurement state is given by

$$|\psi_{\text{after}}\rangle = \frac{(P_k \otimes I) |\psi\rangle}{p(k)} = \frac{\sum_j a_{kj} |e_k\rangle |f_j\rangle}{\sqrt{\sum_j |a_{kj}|^2}}$$

Using partial inner products, one can show that $|\psi_{\text{after}}\rangle$ is normalised. These rules are referred to as the **extended Born rule**.

Consider a quantum system S with state space $\mathcal{V}$. A measurement relative to any basis $\mathcal{C}$ can be performed by first performing a unitary operator, then performing a measurement in a fixed basis $\mathcal{B}$. Let $\mathcal{B} = \{|e_i\rangle\}$, and $\mathcal{C} = \{|e'_i\rangle\}$. Let U be a unitary operator such that $|e'_i\rangle = U |e_i\rangle$. Then, $U^\dagger = U^{-1}$ has the property that $U^{-1} |e'_i\rangle = |e_i\rangle$. Suppose we have a state $|\psi\rangle \in \mathcal{V}$. Let $|\psi\rangle = \sum c_i |e'_i\rangle$. Applying U^{-1} to $|\psi\rangle$, we obtain $U^{-1} |\psi\rangle = \sum c_i |e_i\rangle$ by linearity. We can then measure $|\psi'\rangle = U^{-1} |\psi\rangle$ in the basis $\mathcal{B}$. By the Born rule, $p(i) = \langle\psi'|P_i|\psi'\rangle = \langle\psi| U P_i U^\dagger |\psi\rangle$ where $P_i = |e_i\rangle\langle e_i|$, as we are performing a complete projective measurement. If the outcome is i, then the post-measurement state is $|\psi'_{\text{after}}\rangle = \frac{P_i|\psi'\rangle}{p(i)}$.

§1.15 Standard measurement on multi-qubit systems

Consider a system of n qubits. The state space is $(\mathbb{C}^2)^{\otimes n}$. The **computational basis** or **standard basis** is $\mathcal{B} = \{|i_1 \dots i_n\rangle \mid i_j \in \{0, 1\}\}$. The labels of the elements of the standard basis are labelled by bit strings of length n.

Suppose we are measuring a subset of k qubits of the n-qubit system. Let $n = 3$, and let

$$|\psi\rangle = \frac{i}{2} |000\rangle + \frac{1+i}{2\sqrt{2}} |001\rangle - \frac{1}{2} |101\rangle + \frac{3}{10} |110\rangle - \frac{2i}{5} |111\rangle$$

The standard measurement of any of the three qubits will always have the outcome zero or one. Suppose we perform a standard measurement on the first qubit. By the extended Born rule, we obtain

$$p^{(1)}(1) = \langle\psi| P_1 \otimes I \otimes I |\psi\rangle = \langle\psi| (|1\rangle\langle 1| \otimes I \otimes I) |\psi\rangle = \frac{1}{4} + \frac{9}{100} + \frac{4}{25} = \frac{1}{2}$$

If we measure the outcome 1, the post-measurement state is $|\psi_{\text{after}}\rangle = \frac{(P_1 \otimes I \otimes I)|\psi\rangle}{\sqrt{p^{(1)}(1)}}$.

§1.16 Reliably distinguishing states

Note that the measurement postulate implies that states with guaranteed (with probability 1) different measurement outcomes always lie in mutually orthogonal subspaces. We say that two states are **reliably distinguishable** if there exists a measurement which outputs two distinct outcomes with probability 1 when applied to the two states. Therefore, two states $|\psi\rangle, |\varphi\rangle$ are reliably distinguishable if and only if they are orthogonal, so $\langle\psi|\varphi\rangle = 0$.

Let $|\psi\rangle$ and $|\varphi\rangle$ be orthogonal. Let $\mathcal{B} = \{|\psi\rangle, |f_1\rangle, \ldots, |f_{m-1}\rangle\}$ be a complete orthonormal basis for $\mathcal{V}$. Then $\langle\psi|f_j\rangle = 0$ and $\langle f_j|f_k\rangle = \delta_{jk}$. Measuring $|\psi\rangle$ in this basis, $p(1) = \langle\psi| P_1 |\psi\rangle$ where $P_1 = |\psi\rangle\langle\psi|$, so the probability is 1. Measuring $|\varphi\rangle$ in this basis, $p(1) = \langle\psi|\varphi\rangle \langle\varphi|\psi\rangle = 0$. This is an example of a measurement which can reliably distinguish $|\psi\rangle$ and $|\varphi\rangle$.

Vectors $|v\rangle = |\psi\rangle$ and $|v'\rangle = e^{i\theta} |\psi\rangle$ are not distinguishable. For any measurement, the probability of obtaining a particular outcome when measuring $|v\rangle$ is always the same as the probability when measuring $|v'\rangle$.

§2 Quantum states as information carriers

§2.1 Using higher Hilbert spaces

Quantum information is encoded in the states of a quantum system. Classical information is encoded in states chosen from an orthonormal set, since all distinct classical messages can be distinguished. Given a quantum system S and a quantum state $|\psi\rangle$, we can perform this sequence of operations.

- (ancilla) Consider an auxiliary system A in a fixed state $|A\rangle \in \mathcal{V}_A$. The composite system SA has vector space $\mathcal{V}_S \otimes \mathcal{V}_A$. The initial joint state is $|\psi\rangle\,|A\rangle$. This results in an embedding of quantum information in a higher dimensional space.
- (unitary) Consider the action of a unitary operator U on SA (or on S), modelling the time evolution of the quantum system.
- (measure) We can perform measurements on SA (or on S). The post-measurement state of S is retained, and the auxiliary system A is discarded.

This process is sometimes known as 'going to the church of the higher Hilbert space'. The presence of the ancilla allows for entanglement with other quantum systems.

§2.2 No-cloning theorem

Classically, information can be easily copied by measuring all relevant information and reproducing it. Quantum copying involves three systems:

- a system A containing some quantum information to be copied;
- a system B with $\mathcal{V}_B \simeq \mathcal{V}_A$ initially in some fixed state $|0\rangle$ where the information is to be copied;
- a system M which represents any physical machinery in some 'ready' state $|M_0\rangle$ required for performing the copy.

The initial state of this composite system ABM is $|\psi\rangle\,|0\rangle\,|M_0\rangle$. Note that the $|\psi\rangle$ and $|0\rangle\,|M_0\rangle$ are **uncorrelated** in this state, as we are using the tensor product to combine them. Suppose that the cloning process is performed using some unitary operator U, so $U\,|\psi_A\rangle\,|0\rangle\,|M_0\rangle = |\psi_A\rangle\,|\psi_B\rangle\,|M_\psi\rangle$. Note $|\psi_A\rangle = |\psi\rangle_A = |\psi\rangle$. This cloning process may be required to work either for all states of A, or for some subset of A.

Theorem 2.1

Let $\mathcal{S}$ be any set of states of the system A that contains at least one pair of distinct[a] non-orthogonal states. Then $\nexists$ any unitary operator U that clones all states in $\mathcal{S}$.

[a] $\langle\xi|\psi\rangle \neq 1$ iff ξ, ψ distinct states.

Proof. Let $|\xi\rangle, |\eta\rangle$ be distinct non-orthogonal states in $\mathcal{S}$, so $\langle\xi|\eta\rangle \neq 0$. Suppose such a unitary operator U exists. Then, we must have

$$U|\xi_A\rangle|0_B\rangle|M_0\rangle = |\xi_A\rangle|\xi_B\rangle|M_\xi\rangle\,; \quad U|\eta_A\rangle|0_B\rangle|M_0\rangle = |\eta_A\rangle|\eta_B\rangle|M_\eta\rangle$$

Unitary operators preserve inner products. Hence,

$$\langle\xi_A|\eta_A\rangle\langle 0_B|0_B\rangle\langle M_0|M_0\rangle = \langle\xi_A|\eta_A\rangle\langle\xi_B|\eta_B\rangle\langle M_\xi|M_\eta\rangle$$

Hence, $\langle\xi|\eta\rangle = (\langle\xi|\eta\rangle)^2\langle M_\xi|M_\eta\rangle$[a]. By taking the absolute value, $|\langle\xi|\eta\rangle| = |\langle\xi|\eta\rangle|^2|\langle M_\xi|M_\eta\rangle|$. Since $\xi \neq \eta$, we must have $0 < |\langle\xi|\eta\rangle| < 1$, and $0 \leq |\langle M_\xi|M_\eta\rangle| \leq 1$. Therefore, $1 = |\langle\xi|\eta\rangle||\langle M_\xi|M_\eta\rangle| < 1$, which is a contradiction. $\square$

[a] As we assume states are normalised so $\langle M_0|M_0\rangle = \langle 0_B|0_B\rangle = 1$.

If quantum cloning were possible, superluminal (indeed, instantaneous) communication would also be possible. Suppose we have a state $\left|\psi^+_{AB}\right\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A|0\rangle_B + |1\rangle_A|1\rangle_B) \in \mathbb{C}^2 \otimes \mathbb{C}^2$. Let A, B be the entangled parts of this quantum state, and suppose that we send qubit A to Alice and B to Bob, far apart from each other.

If we want to send the bit 'yes' from Alice to Bob, we measure the qubit A in the basis $\{|0\rangle, |1\rangle\}$, which gives outcomes $0, 1$ with probability $\frac{1}{2}$. If the outcome is 0, the final state of B is $|0\rangle$, and if the outcome is 1, the final state of B is $|1\rangle$. If we want to send 'no', we instead measure A in the basis $\{|+\rangle, |-\rangle\}$, which gives the outcomes $+, -$ with probability $\frac{1}{2}$. Similarly, the final state of B is $|+\rangle$ or $|-\rangle$.

We claim that these 'yes' $(|0\rangle, |1\rangle)$ and 'no' $(|+\rangle, |-\rangle)$ **preparations** of qubit B are indistinguishable by Bob with any local action on the qubit. That is, they each give exactly the same probability distribution of outcomes of any measurement so no superluminal communication yet. In fact, the distribution matches the prior distribution before qubit A was measured.

Let Π_i be the projection operator for outcome i on qubit B. Suppose that 'yes' was sent. Then,

$$p_{\text{yes}}(i) = \frac{1}{2}\langle 0|\Pi_i|0\rangle + \frac{1}{2}\langle 1|\Pi_i|1\rangle = \frac{1}{2}\operatorname{Tr}[\Pi_i(|0\rangle\langle 0| + |1\rangle\langle 1|)] = \frac{1}{2}\operatorname{Tr}\Pi_i$$

In the 'no' case,

$$p_{\text{no}}(i) = \frac{1}{2}\langle +|\Pi_i|+\rangle + \frac{1}{2}\langle -|\Pi_i|-\rangle = \frac{1}{2}\operatorname{Tr}[\Pi_i(|+\rangle\langle +| + |-\rangle\langle -|)] = \frac{1}{2}\operatorname{Tr}\Pi_i$$

These probability distributions match.

Suppose that cloning were possible. We clone the qubit B multiple times after the message was sent, to produce one of the states $|0\rangle \dots |0\rangle, |1\rangle \dots |1\rangle, |+\rangle \dots |+\rangle, |-\rangle \dots |-\rangle$. We now measure each qubit in the basis $|0\rangle, |1\rangle$ separately. If the 'yes' message was sent, all measurements will result in 0 or 1. If 'no' was sent, it is possible that two measurements would differ. In expectation, half of the measurements would result in the outcome 0 and half would result in the outcome 1. Therefore, the 'yes' and 'no' errors can be distinguished with probability of error 2^{-N+1} if we make N copies of B.

§2.3 Distinguishing non-orthogonal states

Suppose you know a state $|\psi\rangle$ has state $|\alpha_0\rangle$ or $|\alpha_1\rangle$ with probability $\frac{1}{2}$, where $\langle\alpha_0|\alpha_1\rangle \neq 0$. Since the states are non-orthogonal, we cannot perfectly distinguish the states, but must allow some error rate. The simplest possibility is to not make a measurement and guess randomly; in which case, the guess is correct with probability $\frac{1}{2}$.

Suppose we append an auxiliary system $|A\rangle$ to $|\alpha_i\rangle$. Note that $\langle A| \langle\alpha_i|\alpha_i\rangle |A\rangle = \langle\alpha_i|\alpha_i\rangle$ as $|A\rangle$ is normalised. If we apply a unitary operator U to $|\alpha_i\rangle$ then perform a projective measurement in the basis $\{\Pi_0, \Pi_1\}$, our action corresponds to simply performing a measurement $\Pi_0' = U^\dagger \Pi_0 U$ or $\Pi_1' = U^\dagger \Pi_1 U$, which leads to the same probabilities of outcomes. Indeed,

$$p(i) = \langle U\xi|\Pi_i|U\xi\rangle = \langle\xi|U^\dagger \Pi_i U|\xi\rangle = \langle\xi|\Pi_i'|\xi\rangle$$

Therefore, in this particular problem, we gain no benefit from moving to a larger Hilbert space or applying unitary operators.

We now describe the **state estimation** or **state discrimination** process. We will consider a two-outcome measurement $\{\Pi_0, \Pi_1\}$, where $\Pi_0 + \Pi_1 = I$. The average success probability is

$$\begin{aligned} p_S(\Pi_0, \Pi_1) &= \frac{1}{2}\mathbb{P}\left(0 \mid |\psi\rangle = |\alpha_0\rangle\right) + \frac{1}{2}\mathbb{P}\left(1 \mid |\psi\rangle = |\alpha_1\rangle\right) \\ &= \frac{1}{2}\langle\alpha_0|\Pi_0|\alpha_0\rangle + \frac{1}{2}\langle\alpha_1| \underbrace{\Pi_1}_{I-\Pi_0} |\alpha_1\rangle \\ &= \frac{1}{2} + \frac{1}{2}\operatorname{Tr}\left[\Pi_0(|\alpha_0\rangle\langle\alpha_0| - |\alpha_1\rangle\langle\alpha_1|)\right] \end{aligned}$$

as $\operatorname{Tr}(A|\psi\rangle\langle\psi|) = \langle\alpha|A|\alpha\rangle$. The optimal choice of measurement maximises the average success probability p_S. Note that $\Delta = |\alpha_0\rangle\langle\alpha_0| - |\alpha_1\rangle\langle\alpha_1|$ is self-adjoint, and we can write $p_S = \frac{1}{2} + \frac{1}{2}\operatorname{Tr}(\Pi_0\Delta)$. Therefore, the eigenvalues of Δ are real, and the eigenvectors form an orthonormal basis. For a state $|\beta\rangle$ orthogonal to both $|\alpha_0\rangle$ and $|\alpha_1\rangle$, we have $\Delta|\beta\rangle = 0$. Therefore, Δ acts nontrivially only in the vector space spanned by $|\alpha_0\rangle$ and $|\alpha_1\rangle$, and hence has at most two nonzero eigenvalues, and its eigenvectors lie in $\operatorname{span}\{|\alpha_0\rangle, |\alpha_1\rangle\}$.

Now, $\operatorname{Tr}\Delta = 0$ so the eigenvalues are δ and $-\delta$ for some $\delta \in \mathbb{R}$. Let $|p\rangle$ be the eigenvector for δ, and $|m\rangle$ be the eigenvector for $-\delta$, so $\langle p|m\rangle = 0$. We can write Δ in its spectral decomposition, giving $\Delta = \delta\,|p\rangle\!\langle p| - \delta\,|m\rangle\!\langle m|$.

Let $\left|\alpha_0^\perp\right\rangle \in \operatorname{span}\{|\alpha_0\rangle, |\alpha_1\rangle\}$ be a normalised vector such that $\left\langle\alpha_0^\perp\middle|\alpha_0\right\rangle = 0$. Then, $\left\{|\alpha_0\rangle, \left|\alpha_0^\perp\right\rangle\right\}$ is an orthonormal basis. Hence, we can write $|\alpha_1\rangle = c_0\,|\alpha_0\rangle + c_1\left|\alpha_0^\perp\right\rangle$. In this basis,

$$\Delta = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} -|c_0|^2 & -c_0c_1^\star \\ -c_0^\star c_1 & -|c_1|^2 \end{pmatrix} = \begin{pmatrix} 1-|c_0|^2 & -c_0c_1^\star \\ -c_0^\star c_1 & -|c_1|^2 \end{pmatrix} = \begin{pmatrix} |c_1|^2 & -c_0c_1^\star \\ -c_0^\star c_1 & -|c_1|^2 \end{pmatrix}$$

which has eigenvalues $\delta = |c_1|, -\delta = -|c_1|$. Since $|c_0| = |\langle\alpha_0|\alpha_1\rangle| = \cos\theta$ where $\theta \geq 0$, we have $\delta = \sin\theta$. Then,

$$\begin{aligned} p_S(\Pi_0, \Pi_1) &= \frac{1}{2} + \frac{1}{2}\operatorname{Tr}(\Pi_0\Delta) \\ &= \frac{1}{2} + \frac{1}{2}\operatorname{Tr}(\Pi_0[\sin\theta\,|p\rangle\!\langle p| - \sin\theta\,|m\rangle\!\langle m|]) \\ &= \frac{1}{2} + \frac{\sin\theta}{2}[\,\langle p|\Pi_0|p\rangle - \langle m|\Pi_0|m\rangle] \end{aligned}$$

Note that for any $|\varphi\rangle$, we have $0 \leq \langle\varphi|\Pi|\varphi\rangle \leq 1$, so the measurement is maximised when $\langle p|\Pi_0|p\rangle = 1$ and $\langle m|\Pi_0|m\rangle = 0$. We therefore define $\Pi_0 = |p\rangle\!\langle p|$. Then, the optimal average success probability is

$$p_S^\star = \frac{1}{2} + \frac{\sin\theta}{2}$$

Theorem 2.2 (Holevo–Helstrom theorem for pure states)
Let $|\alpha_0\rangle, |\alpha_1\rangle$ be equally likely states, with $|\langle\alpha_0|\alpha_1\rangle| = \cos\theta$, $\theta \geq 0$. Then, the probability p_S of correctly identifying the state by any quantum measurement satisfies

$$p_S \leq \frac{1}{2} + \frac{\sin\theta}{2}$$

and this bound can be attained.

In the case of orthogonal states, the theorem implies that $p_S \leq 1$ and the bound can be attained, which was shown before.

§2.4 No-signalling principle

Suppose we have a possibly entangled state $|\varphi_{AB}\rangle \in \mathcal{V}_A \otimes \mathcal{V}_B$ shared between two agents Alice (A) and Bob (B). Suppose we perform a complete projective measurement on

$|\varphi_A\rangle$. By the extended Born rule, each measurement outcome will lead to an instantaneous change of $|\varphi_B\rangle$. If this change in state could be detected by measuring $|\varphi_B\rangle$, instantaneous communication between A and B would be possible.

Consider $\left|\varphi_{AB}^+\right\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. Suppose qubit A is measured in the standard basis $\{|0\rangle, |1\rangle\}$.

outcome	probability	post-measurement state	final state of B
0	$\frac{1}{2}$	$\|00\rangle$	$\|0\rangle$
1	$\frac{1}{2}$	$\|11\rangle$	$\|1\rangle$

Suppose qubit B is subsequently measured in $\{|b_0\rangle, |b_1\rangle\}$. If B is in the state $|0\rangle$, we can write $|0\rangle = c_0|b_0\rangle + c_1|b_1\rangle$, and $p_{|0\rangle}(i) = |c_i|^2 = |\langle b_i|0\rangle|^2$. If B is in the state $|1\rangle$, we write $|1\rangle = d_0|b_0\rangle + d_1|b_1\rangle$, and $p_{|1\rangle}(i) = |d_i|^2 = |\langle b_i|1\rangle|^2$. Therefore, $p(i) = \frac{1}{2}|\langle b_i|0\rangle|^2 + \frac{1}{2}|\langle b_i|1\rangle|^2 = \frac{1}{2}$. The two outcomes for this measurement are equally likely, regardless of the choice of complete orthonormal basis $\{|b_0\rangle, |b_1\rangle\}$.

Suppose instead A is not measured, but we perform the same measurement on B. The initial state is $\left|\varphi_{AB}^+\right\rangle$, so by the extended Born rule, $p(i) = \left\langle\varphi_{AB}^+\middle|(I_A \otimes |b_i\rangle\langle b_i|)\middle|\varphi_{AB}^+\right\rangle = \frac{1}{2}$. We can therefore not detect through measuring B whether a measurement was performed at A. This is the no-signalling principle.

We now prove the more general case. Let $|\varphi_{AB}\rangle \in \mathcal{V}_A \otimes \mathcal{V}_B$ be an arbitrary possibly entangled state.

Suppose we measure B in a complete orthonormal basis $\{|b\rangle\}_{b=1}^{\dim \mathcal{V}_B}$, which is a complete projective measurement on B. Let $\{|a\rangle\}_{a=1}^{\dim \mathcal{V}_A}$ be a complete orthonormal basis for $\mathcal{V}_A$. Then, expanding $|\varphi_{AB}\rangle$, in this basis, we can write $|\varphi_{AB}\rangle = \sum_{a,b} c_{ab}|a\rangle|b\rangle$. We obtain outcome b with probability $p(b) = \langle\varphi_{AB}|(I_A \otimes P_b)|\varphi_{AB}\rangle = \sum_{a=1}^{\dim \mathcal{V}_A}|c_{ab}|^2$. The post-measurement state is $|\varphi'_{AB}\rangle$.

Suppose that we first measure A in a complete orthonormal basis $\{|a\rangle\}_{a=1}^{\dim \mathcal{V}_A}$, and then perform the measurement $\{|b\rangle\}_{b=1}^{\dim \mathcal{V}_B}$ on B. The outcome of the first measurement is a with probability $p(a) = \langle\varphi_{AB}|(P_a \otimes I_B)|\varphi_{AB}\rangle = \sum_{b=1}^{\dim \mathcal{V}_B}|c_{ab}|^2$. We denote the post-measurement state of the joint system by $|\varphi''_{AB}\rangle = \frac{(P_a \otimes I_B)|\varphi_{AB}\rangle}{\sqrt{p(a)}}$. Then, the outcome of the second measurement is b with probability

$$\begin{aligned}
p(b \mid a) &= \langle\varphi''_{AB}|(I_A \otimes P_b)|\varphi''_{AB}\rangle \\
&= \frac{1}{p(a)}\langle\varphi_{AB}|\underbrace{(P_a \otimes I_B)(I_A \otimes P_b)}_{\text{As projections: }(P_a \otimes P_B)}(P_a \otimes I_B)|\varphi_{AB}\rangle \\
&= \frac{1}{p(a)}\langle\varphi_{AB}|(P_a \otimes P_b)|\varphi_{AB}\rangle \\
p(a,b) = p(a)p(b \mid a) &= \langle\varphi_{AB}|(P_a \otimes P_b)|\varphi_{AB}\rangle = |c_{ab}|^2
\end{aligned}$$

Hence $p(b) = \sum_{a=1}^{\dim \mathcal{V}_A} |c_{ab}|^2$, which is exactly the distribution we obtained when no measurement on A was performed. This proves the no-signalling principle.

§2.5 The Bell basis

Let $\mathbb{C}^2 \otimes \mathbb{C}^2$ model a quantum system representing the spins of two electrons. Consider $\left|\varphi^+_{AB}\right\rangle = \frac{1}{2}(|00\rangle + |11\rangle) \in \mathbb{C}^2 \otimes \mathbb{C}^2$. This is a **maximally entangled state**; we have information about the whole system, but no information about the individual states.

$$\left|\varphi^\pm_{AB}\right\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle); \quad \left|\psi^\pm_{AB}\right\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle)$$

$\left\{\left|\varphi^\pm_{AB}\right\rangle, \left|\psi^\pm_{AB}\right\rangle\right\}$ forms a complete orthonormal basis of $\mathbb{C}^2 \otimes \mathbb{C}^2$. This is called the **Bell basis**. The basis vectors are sometimes known as **EPR states**, after Einstein, Podolsky, and Rosen.

One bit of classical information can be encoded in a single qubit, and two bits can be encoded in a pair of qubits in the Bell basis. The Bell states have a **parity** 0 or 1, representing parallel $\{|\varphi^\pm\rangle\}$ or antiparallel $\{|\psi^\pm\rangle\}$ spins. The states also have a **phase**, which can be positive $\{|\varphi^+\rangle, |\psi^+\rangle\}$ or negative $\{|\varphi^-\rangle, |\psi^-\rangle\}$. For example, we can encode the classical message 01 using the state $|\varphi^-\rangle$.

We can perform a complete projective measurement on both qubits in the Bell basis to recover the encoded information with certainty. For instance, $P_{00} = |\varphi^+\rangle\langle\varphi^+|$. If we prepare a pair of electrons $|\varphi\rangle$ in the state $|\varphi^-\rangle$ for example, we obtain $p(00) = p(10) = p(11) = 0$ and $p(01) = 1$.

Note $(A \otimes I)|\varphi^+\rangle = (I \otimes A^\intercal)|\varphi^+\rangle$.

§2.6 Superdense coding

Suppose Alice wants to send a classical message to Bob. Two bits of classical information can be sent reliably via a single qubit, provided that Alice and Bob share an entangled state, using **superdense coding** or **quantum dense coding**. Let

$$X = \sigma_x; \quad Z = \sigma_z; \quad Y = i\sigma_y = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

One can check that the Bell basis vectors satisfy

$$\begin{aligned}
\left|\varphi^+\right\rangle &= (I \otimes I)\left|\varphi^+\right\rangle = (I \otimes I)\left|\varphi^+\right\rangle \\
|\varphi^-\rangle &= (Z \otimes I)\left|\varphi^+\right\rangle = (I \otimes Z)\left|\varphi^+\right\rangle \\
\left|\psi^+\right\rangle &= (X \otimes I)\left|\varphi^+\right\rangle = (I \otimes X)\left|\varphi^+\right\rangle
\end{aligned}$$

$$\left|\psi^-\right\rangle = (Y \otimes I)\left|\varphi^+\right\rangle = -(I \otimes Y)\left|\varphi^+\right\rangle$$

Suppose we have shared the entangled Bell state $\left|\varphi^+_{AB}\right\rangle$ between Alice and Bob. The superdense coding protocol is

Alice's message	local action on A	final state of AB
00	I	$\|\varphi^+\rangle$
01	Z	$\|\varphi^-\rangle$
10	X	$\|\psi^+\rangle$
11	Y	$\|\psi^-\rangle$

Then, Alice sends qubit A to Bob, so Bob has the entire state AB. Bob performs a Bell measurement, which distinguishes between the four Bell states, thus recovering Alice's message. Since the state is maximally entangled, an eavesdropper who may intercept Alice's transmission cannot recover any part of the message.

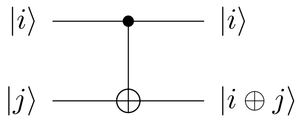
§2.7 Quantum gates

A quantum gate is given by a unitary operator acting on some qubits. Such gates have matrix representations in the computational basis.

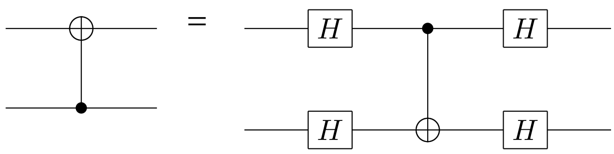
1. The **Hadamard gate** is
$$H = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$
One can show that
$$H\left|0\right\rangle = \left|+\right\rangle; \quad H\left|1\right\rangle = \left|-\right\rangle; \quad H\left|+\right\rangle = \left|0\right\rangle; \quad H\left|-\right\rangle = \left|1\right\rangle$$
Note that $H^\intercal = H^\dagger = H$ and $H^2 = I$. As an orthogonal transformation in $\mathbb{R}^2$, it acts as a reflection by an angle of $\frac{\pi}{8}$ to the positive x axis. This gate is drawn
$$—\boxed{H}—$$
In general, by linearity we obtain
$$a\left|0\right\rangle + b\left|1\right\rangle \quad —\boxed{H}— \quad a\left|+\right\rangle + b\left|-\right\rangle$$
2. The X, Z gates are given by
$$X\left|k\right\rangle = \left|k \oplus 1\right\rangle; \quad Z\left|k\right\rangle = (-1)^k\left|k\right\rangle$$
where $\oplus$ denotes addition modulo 2. They X, Z, Y gates are drawn
$$—\boxed{X}—; \quad —\boxed{Z}—; \quad —\boxed{X}—\boxed{Z}—$$

3. The **phase gate** is
$$P_\theta = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix}$$
Note that $Z = P_\pi$.

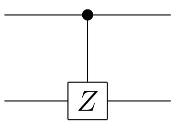
4. The **controlled-X** gate, also called a **CNOT** gate, is
$$CX = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & X \end{pmatrix}$$
Note that $CX\,|i\rangle\,|j\rangle = |i\rangle\,|i \oplus j\rangle$. The first qubit is called the **control** qubit, and the second is called the **target** qubit. If $i = 0$, there is no action on the second qubit. If $i = 1$, X is performed on the second qubit. In general, $CX\,|0\rangle\,|\psi\rangle = |0\rangle\,|\psi\rangle$, and $CX\,|1\rangle\,|\psi\rangle = |1\rangle\,(X\,|\psi\rangle)$. The circuit diagram is as follows.



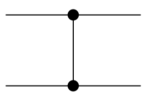
One can show that



5. The **controlled-Z** gate, also called a **CZ** gate, is
$$CZ = \begin{pmatrix} I & 0 \\ 0 & Z \end{pmatrix}$$
So $CZ\,|0\rangle\,|\psi\rangle = |0\rangle\,|\psi\rangle$ and $CZ\,|1\rangle\,|\psi\rangle = |1\rangle\,(Z\,|\psi\rangle)$. CZ is symmetric in its action on the two qubits; for example, $CZ_{12}\,|0\rangle\,|1\rangle = CZ_{21}\,|0\rangle\,|1\rangle$. This gate is drawn



or



§2.8 Quantum teleportation

Suppose Alice and Bob share the Bell state $|\varphi^+\rangle_{AB}$, and that Alice wants to send the state of qubit $|\psi\rangle_C$ to Bob, but only classical communication between them is possible.

It is possible to transfer the information about the state of $|\psi\rangle_C$ without physically transferring qubit C to Bob. This state transfer can be accomplished in such a way that is unaffected by any physical process in the space between Alice and Bob, since it relies only on classical communication.

The initial state of CAB is $|\Psi\rangle = |\psi\rangle_C \otimes |\varphi^+\rangle_{AB}$, assuming $|\psi\rangle_C$ is uncorrelated with $|\varphi^+\rangle_{AB}$. Let $|\psi\rangle_C = a\,|0\rangle_C + b\,|1\rangle_C$, so

$$|\Psi\rangle = |\psi\rangle_C \otimes \left|\varphi^+\right\rangle_{AB} = \frac{1}{\sqrt{2}}[a\,|000\rangle + a\,|011\rangle + b\,|100\rangle + b\,|111\rangle]$$

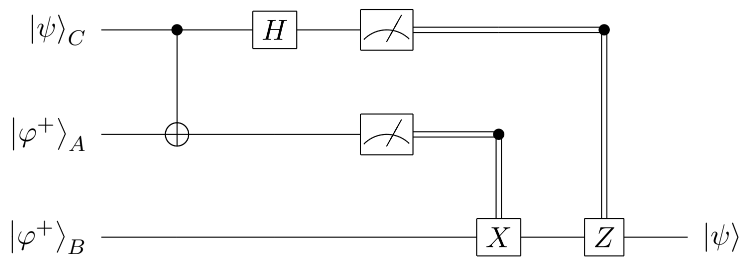
Alice sends C and A through a CX gate. Now,

$$|\Psi\rangle = |\varphi_1\rangle = \frac{1}{\sqrt{2}}[a\,|000\rangle + a\,|011\rangle + b\,|110\rangle + b\,|101\rangle]$$

She now sends C through a Hadamard gate.

$$\begin{aligned} |\Psi\rangle = |\varphi_2\rangle &= \frac{1}{\sqrt{2}}[a\,|{+00}\rangle + a\,|{+11}\rangle + b\,|{-10}\rangle + b\,|{-01}\rangle] \\ &= \frac{1}{2}\left[\,|00\rangle\,|\psi\rangle + |01\rangle\,(X\,|\psi\rangle) + |10\rangle\,(Z\,|\psi\rangle) + |11\rangle\,(-Y\,|\psi\rangle)\right] \end{aligned}$$

Alice now measures CA in the computational basis of $\mathbb{C}^2 \otimes \mathbb{C}^2$. The probability of each outcome is $\frac{1}{4}$, irrespective of the values of a and b and hence of $|\psi\rangle$. She then sends the result of her measurement to Bob. If Alice measures outcome ij, B is in state $X^j Z^i\,|\psi\rangle$. Then, Bob can act on B using $Z^i X^j$, as X and Z are involutive, giving $|\psi\rangle$ as desired. This process can be represented with the following diagram, where double-struck wires are classical, and the meter symbol denotes a measurement of the quantum state.



Note that after the measurement of CA, the entanglement between CA and B is broken. No-cloning is not violated, as the original state $|\psi\rangle_C$ is destroyed.

Note that the first steps of this process including Alice's measurement correspond to performing a Bell measurement on CA. This is because the action of CX_{CA} then H_C corresponds to a rotation of the Bell basis to the standard basis.

§3 Quantum cryptography

§3.1 One-time pads

We can use quantum information theory to securely transmit messages between agents Alice and Bob, who may be in distant locations, without the possibility that an eavesdropper Eve can recover the message that was sent.

We will assume that Alice and Bob have an authenticated classical channel through which they can send classical information; Alice and Bob can verify that any particular message on the channel came from a particular sender. We also assume that Eve cannot block the channel or modify any messages transmitted, but she can monitor the channel freely. Hence, Alice and Bob can receive messages from each other without error.

In the classical setting, there exists a provably secure classical scheme for private communications, called the **one-time pad**. This requires that Alice and Bob share a private key K, which is a binary string. K must have been created beforehand, and must be chosen uniformly at random from the set of binary strings of the same length as the message M. Suppose $M, K \in \{0,1\}^n$.

The protocol is as follows. First, Alice computes the encrypted message $C = M \oplus K$. She then sends C to Bob through the classical channel. Bob can then compute $C \oplus K = M \oplus K \oplus K = M$ to obtain the message that was sent by Alice. Eve cannot learn any information about the message (apart from its length), as she has no knowledge of K. In general, the probability that a particular K was chosen is 2^{-n}. This scheme cannot be broken.

Suppose that Alice and Bob use the same key K to send two messages M_1, M_2. Eve can obtain $M_1 \oplus K$ and $M_2 \oplus K$, and can therefore compute $(M_1 \oplus K) \oplus (M_2 \oplus K) = M_1 \oplus M_2$, which gives some information about the messages that were sent. Any key must only be used once, so the one-time pad protocol is inefficient. To solve this problem, we will construct methods for distributing keys, using techniques from quantum information theory.

§3.2 The BB84 protocol

Quantum key distribution allows Alice and Bob to generate a private key without needing to physically meet. This key can then be used to send messages over the one-time pad protocol. In addition to a classical channel, we assume that Alice and Bob also have access to a quantum channel through which they can send qubits. We will show that Eve cannot gain information about the key that Alice and Bob generate without being detected.

Consider the bases $\mathcal{B}_0 = \{|0\rangle, |1\rangle\}, \mathcal{B}_1 = \{|+\rangle, |-\rangle\}$. These are examples of **mutually unbiased bases**; a pair of bases such that if any basis vector is measured relative to the other basis, all outcomes are equally likely. For example, measuring $|+\rangle$ relative to $\mathcal{B}_0$ gives probability $\frac{1}{2}$ for outcomes 0 and 1.

First, Alice generates two m-bit strings $x = x_1 \dots x_m \in \{0,1\}^m, y = y_1 \dots y_m \in \{0,1\}^m$ uniformly at random. She then prepares the m-qubit state $|\psi_{xy}\rangle = |\psi_{x_1 y_1}\rangle \otimes \cdots \otimes |\psi_{x_m y_m}\rangle$ where

$$|\psi_{x_i y_i}\rangle = \begin{cases} |0\rangle & x_i = 0; y_i = 0 \\ |1\rangle & x_i = 1; y_i = 0 \\ |+\rangle & x_i = 0; y_i = 1 \\ |-\rangle & x_i = 1; y_i = 1 \end{cases}$$

Alice sends the qubits $|\psi_{xy}\rangle$ to Bob with m uses of the quantum channel. The qubits received are not necessarily in the state $|\psi_{xy}\rangle$ due to noise or malicious manipulation of the channel. Bob then generates an m-bit string $y' = y'_1 \dots y'_m \in \{0,1\}^m$ uniformly at random. If $y'_i = 0$, he measures the ith qubit in the basis $\mathcal{B}_0 = \{|0\rangle, |1\rangle\}$. If $y'_i = 1$, he acts on the ith qubit by the Hadamard gate and then measures in $\mathcal{B}_0$. Equivalently, he measures the ith qubit in the basis $\mathcal{B}_1 = \{|+\rangle, |-\rangle\}$. Let the sequence of outcomes be $x' = x'_1 \dots x'_m \in \{0,1\}^m$.

If $y'_i = y_i$, we have $x'_i = x_i$. Indeed, suppose $y'_i = 0 = y_i$. Then $|\pi_{x_i y_i}\rangle \in \mathcal{B}_0$, and Bob measures in basis $\mathcal{B}_0$, so he can determine x_i with probability 1. If $y'_i = 1 = y_i$, $|\pi_{x_i y_i}\rangle \in \mathcal{B}_1$, and Bob measures in basis $\mathcal{B}_1$.

Now, Alice and Bob compare their values of y and y' over the classical channel, and discard all x_i and x'_i for which $y_i \neq y'_i$. The remaining x_i and x'_i match, given that Bob receives $|\psi_{xy}\rangle$ exactly, and this forms the shared private key $\widetilde{x} = \widetilde{x}'$. The average length of $\widetilde{x}$ is $\frac{m}{2}$.

In the case $m = 8$, suppose $x = 01110100$ and $y = 11010001$. Alice prepares $|\psi_{xy}\rangle$ and sends the qubits to Bob. Suppose that Bob receives $|\psi_{xy}\rangle$ exactly, and he generates $y' = 01110110$. Bob measures qubit 1 in the basis $\mathcal{B}_0$, but the qubit is in state $|+\rangle$, so he obtains both outcomes for x'_1 with equal probability. He measures qubit 2 in the basis $\mathcal{B}_1$, and the qubit is in state $|-\rangle$, so after applying H and measuring, he obtains the correct outcome $x'_2 = 1$ with probability 1. After discarding mismatched y_i, the obtained private key is $\widetilde{x} = 110$.

In the general case, however, there may be noise or malicious activity on the channel. We therefore include the further step of **information reconciliation** at the end of the BB84 protocol. Alice and Bob want to estimate the **bit error rate**, which is the proportion of bits in $\widetilde{x}$ and $\widetilde{x}'$ that differ. They can publicly compare a random sample of their bits, and discard the bits used in the test. They assume that the bit error rate in the sample is approximately the same as the bit error rate of $\widetilde{x}$ and $\widetilde{x}'$.

Suppose that Alice and Bob have estimated the bit error rate to be $\frac{1}{7}$, and now have strings a, b of length 7. They can use classical error correcting code techniques to fix any remaining errors. They publicly agree to act on a, b by a matrix

$$\widetilde{H} = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}$$

which is the check matrix of a Hamming code. Alice computes the **syndrome** for a, given by $s^A = (s_1^A, s_2^A, s_3^A)^\intercal = \widetilde{H} a^\intercal$, and sends this to Bob on the public channel. Bob computes the syndrome s^B for b, and calculates $s = s^B - s^A$. There is a unique bit string v with at most one nonzero entry such that $\widetilde{H} v^\intercal = s$; he can therefore recover a.

The estimation of the bit error rate and the transmission of the syndrome can reveal some information on the public channel. Alice and Bob want to estimate the maximum amount of information that an eavesdropper could gain about the remaining bits, using **privacy amplification**. This depends on the choice of action that Eve takes.

As an example, suppose $a^\star = (a_1, a_2, a_3) \in \{0,1\}^3$, and suppose Eve knows at most one bit of this string. Let $c = (a_1 \oplus a_3, a_2 \oplus a_3)$. We claim that Eve has no knowledge about c. Indeed, we can explicitly enumerate all possibilities of $a^\star$ and the corresponding values of c, and show that Eve's knowledge about any of the bits of $a^\star$ does not change the distribution of c.

One strategy for Eve, called the **intercept and resend** strategy, is to intercept the qubits as they are transferred to Bob, measure them, and retransmit the post-measurement state. The best possible measurement she can perform is in the **Breidbart basis** $\{|\alpha_0\rangle, |\alpha_1\rangle\}$ where

$$|\alpha_0\rangle = \cos\frac{\pi}{8}|0\rangle - \sin\frac{\pi}{8}|1\rangle; \quad |\alpha_1\rangle = \sin\frac{\pi}{8}|0\rangle + \cos\frac{\pi}{8}|1\rangle$$

Note that

$$|\langle\alpha_0|0\rangle|^2 = |\langle\alpha_0|+\rangle|^2 = \cos^2\frac{\pi}{8}; \quad |\langle\alpha_1|1\rangle|^2 = |\langle\alpha_1|-\rangle|^2 = \cos^2\frac{\pi}{8}$$

The $|\alpha_i\rangle$ provide the best possible simultaneous approximations of $|0\rangle, |+\rangle$ and $|1\rangle, |-\rangle$. Suppose $y_i' = y_i$, and suppose Eve intercepts the ith qubit and measures it in the Breidbart basis. Her outcomes are 0 or 1, and she learns the correct value of x_i with probability $\cos^2\frac{\pi}{8} \approx 0.85$. If she measures 0, she transmits $|\alpha_0\rangle$ to Bob, and if she measures 1, she transmits $|\alpha_1\rangle$ to Bob.

The probability that Bob makes an incorrect inference of the value of the ith bit after this manipulation is $\frac{1}{4}$, regardless of the state of the qubit transmitted by Alice. Suppose $|\psi_{x_i y_i}\rangle = |0\rangle$, so $x_i = 0, y_i = 0$. Then,

$$\begin{aligned}\mathbb{P}(x_i' \neq x_i) &= \mathbb{P}(B \text{ measures } 1 \mid A \text{ sent } |0\rangle) \\ &= \mathbb{P}(E \text{ sent } |\alpha_0\rangle \mid A \text{ sent } |0\rangle)\,\mathbb{P}(B \text{ measures } 1 \mid E \text{ sent } |\alpha_0\rangle) \\ &+ \mathbb{P}(E \text{ sent } |\alpha_1\rangle \mid A \text{ sent } |0\rangle)\,\mathbb{P}(B \text{ measures } 1 \mid E \text{ sent } |\alpha_1\rangle)\end{aligned}$$

$$
\begin{aligned}
&= |\langle \alpha_0|0\rangle|^2 |\langle \alpha_0|1\rangle|^2 + |\langle \alpha_1|0\rangle|^2 |\langle \alpha_1|1\rangle|^2 \\
&= \frac{1}{4}
\end{aligned}
$$