Introduction to Quantum Computation

Åsa Hirvonen

Autumn 2018

Contents

Version of September 17, 2018

1 Introduction

Classical computation is concerned with bits - 0s and 1s - that are manipulated according to given rules to solve various problems. Working with quantum bits - or qubits - does not change the set of problems solvable, but due to its very different nature, allows one to formulate algorithms that solve problems faster than any known classical algorithm.

In this course we will look at the fundamentals of qubits and quantum computation, with the aim that the student gains an insight into the basic principles and main techniques underlying quantum algorithms. This material is largely based on the book Quantum Computation and Quantum Information by M. A. Nielsen and I. L. Chuang.

1.1 Quantum

One of the simplest striking examples of the counter intuitive world we enter when we go down to the quantum scale is the double slit experiment. In the early 19th century, Thomas Young invented it to prove that light is composed of waves. In the experiment a narrow beam of light is directed at a plate with two parallel slits and the emerging pattern of light is observed on a screen behind the plate. If either of the slits is covered, a single stripe of light can be observed. However, if both slits are open, the light waves will interfere producing alternating dark and light bands on the screen.

The experiment does not, however, exclude particle-like behaviour of light. In fact, light has wave-particle duality, it shows properties of both. As do, in fact, also matter. The experiment was later repeated with electrons, and even atoms and molecules, and the same interference phenomenon occurs. The same interference pattern will be observed even in the case where the particles used are sent through the slits one at a time - the particle will interfere with itself!

The story gets even more interesting with a modification to the experiment: If a detector is placed at the slits, showing which slit the particle took, the inference pattern disappears.

This version of the experiment illustrates the distinction between state and observable reality in quantum mechanics. Any system is modelled by a state, which is a complex linear combination of basic states. The basic states correspond to observable properties like 'the particle is in position x' or 'the electron has spin down'. A system that is not observed, may be in a 'superposition' state - in a combination of different observable states. But

observing, or measuring, the state, inevitably causes the system to enter a state corresponding to the observed value, with other possibilities discarded. The possibility of being in superposition is one of the key aspects of quantum computation. It is sometimes presented as the possibility to explore computation in several paths in parallel. But this phenomenon can not be used directly, as one can never observe all these possible paths. A key idea in quantum computation is to make use of constructive and destructive interference in a way that will 'pile up' probability to measure a certain state.

Another phenomenon in quantum information theory is entanglement. This stems from the possibility of having two parts of the same system in a superposition with 'missing possibilities'. The zero probability of certain measurement outcomes will cause the measurement of one part to collapse the superposition also of the other.

1.2 Quantum enters computation

One of the earliest ideas towards quantum computation was presented by Richard Feynman. In the early 1980's he pointed out that the computer simulation of a quantum system seems to need an exponential slowdown, and suggested one should use some sort of quantum computer instead, where one could use the quantum behaviour of one system to simulate the behaviour of another. The idea was developed further a couple of years later by Deutsch who defined a 'quantum Turing Machine' as a theoretical framework for quantum computation.

The model mainly used today for quantum computation, however, is circuit based. This was also developed by Deutsch, and was further developed by Yao, who also showed that the quantum circuit and quantum Turing machine model give equivalent models of computation. We will concentrate on the circuit model in this course.

The largest change of perspective quantum computation has brought about, is that computation is not an abstract notion, it is a physical process. And that process will limit or enable the methods you use. The pioneers of the theory of computation, such as Church, Gödel, Turing, and Kleene, developed their models of computation based on the metaphor of 'man and paper' - what one can do mechanically following certain procedures, with an unlimited supply of paper. The actual implementation does not matter.

For quantum computation one needs a new metaphor. As quantum computation is a physical process, but we don't want to go into the details of any specific implementation, we take as our background the basic principles

physicists currently agree on regarding quantum mechanics, formulated in a handful of postulates. From them we will extract the mathematical framework, and then work with only this abstract framework.

1.3 The postulates of quantum mechanics

Postulate 1 *A physical system at any point in time is completely described by a* state*. A state is modelled by a unit vector in a complex Hilbert space.*

This postulate separates the world we see from the structure used to model it. On one hand we have the world we see around us, with classrooms, blackboards, students, etc. On the other we have a unit vector describing all of this world. It can be seen as an information package containing information about all these teachers and students, and what they are doing. The point is, we cannot observe this state to extract all this information at once.

The special system we are interested in during this course is the qubit, which is a quantum analogue of a bit. As such, it needs to have two states, $|0\rangle$ and $|1\rangle$ which correspond to the bit values 0 and 1.[1] We need a model where the qubit can be 0 or 1 with absolute certainty, so the corresponding states will be orthogonal. Thus the smallest complex Hilbert space that can model the qubit is $\mathbb{C}^2$, and we will take this as our definition. Physical implementations for such a system can built using, e.g., electron spins, polarisation of light or nuclear magnetic resonance.

Of course, having a two-dimensional (complex) space, means the qubit can take as value also all length one linear combinations of the states $|0\rangle$ and $|1\rangle$, such as $\frac{1}{\sqrt{2}}|0\rangle + \frac{1}{\sqrt{2}}|1\rangle$.

Postulate 2 *The time evolution of a closed quantum system is described by the Schrödinger equation*

$$i\hbar\frac{d|\psi\rangle}{dt} = H|\psi\rangle,$$

where H is a so called Hamiltonian, an operator for energy.

If we don't look at the continuous evolution, but only consider what can happen in a fixed time interval, the postulate takes the following form:

[1] Note that $|0\rangle$ is *not* the zero vector of the Hilbert space, it is a unit vector corresponding to the bit value being 0.

Postulate 2' *The time evolution of a closed quantum system is described by* unitary transformations*, i.e., if the system is in state* $|\psi\rangle$ *at a given time* t_0*, then there is a unitary operator* U *such that the state* $|\psi_t\rangle$ *at time* $t_0 + t$ *is*

$$|\psi_t\rangle = U|\psi\rangle.$$

We will only look at discrete evolution (cf. computational steps), and take Postulate 2' as the grounds for what we can do to our qubits, i.e., we will demand that all computational actions we apply to our qubits are unitary operators. Thus we can, e.g., apply a `NOT` operator to the qubit using the matrix

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

This will map $|0\rangle$ to $|1\rangle$ and vice versa.

Postulate 3 *Quantum measurements are described by a set of measurement operators* $\{M_m : m \in M\}$ *satisfying the* completeness equation $\sum_m M_M^\dagger M_m = I$ *and such that a measurement of a state* ψ *gives outcome* $m \in M$ *with probability*

$$p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle$$

and if m *is the outcome, the state of the system will be*

$$\frac{M_m|\psi\rangle}{\sqrt{\langle\psi|M_m^\dagger M_m|\psi\rangle}}$$

after the measurement.

This is maybe the most striking part of quantum mechanics. It says that quantum measurements are non-deterministic processes. Knowing a state completely still does not tell what the measured outcome will be, it only gives the probability distribution of the outcomes. And, in fact, it also tells us, that in general we have no possibility of knowing a state completely. If we examine it in some way (i.e., perform a measurement), the state will change.

Note also that the postulate does not tell anything about the physical (or philosophical) nature of measurement. There are various interpretations of what it actually means to measure a state and thus destroy superposition. We will not go into those here. The takeaway points are: What can be observed of states are the outcomes of measurements. The probability distribution of the

outcomes is determined by a set of operators. Measurement changes the state irreversibly. For example, if we measure a qubit in state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ with respect to the computational basis $\{|0\rangle, |1\rangle\}$, i.e., with projections onto the basic elements as measurement operators, then we will observe a 0 with probability $|\alpha|^2$ and a 1 with probability $|\beta|^2$. We can *not* observe the coefficients α and β. And after measurement the qubit will be in the basis state corresponding to the measured value.

Postulate 4 *The state space of a composite system is the tensor product of the state spaces of the component systems.*

We will look closer at tensor products further on. As a quick glimpse, let's just look at what happens with three qubits. Each qubit corresponds to $\mathbb{C}^2$ with computational basis $\{|0\rangle, |1\rangle\}$. A tensor product of vector spaces has a basis that consists of formal products of the basic elements of the component spaces. Thus our 3-qubit space has a basis consisting of elements of the form $|i\rangle \otimes |j\rangle \otimes |k\rangle$, where $i, j, k \in \{0, 1\}$. These are usually written in a shorter way $|ijk\rangle$. Thus the state space of a three-qubit space has a basis consisting of all three-bit sequences. So the space has dimension 2^8.

In the following we leave the physical background to quantum computation. We only extract the mathematical model and work with that. Thus qubits are unit vectors in $\mathbb{C}^2$, several qubits are given by tensor products of this space. Computation happens by means of unitary transformations and measurements give non-deterministic outcomes with probabilities specified by a set of operators. When we measure with respect to an orthonormal basis, the probabilities can be calculated from the *amplitudes*, i.e., the coefficients of the basic elements.

2 The qubit

Our starting point is the quantum bit - the qubit - that have two basic states (corresponding to the classical bit values 0 and 1) spanning $\mathbb{C}^2$. We call these basic states $|0\rangle$ and $|1\rangle$, although a more familiar notation from linear algebra would be $|0\rangle = e_0 = (1, 0)$ and $|1\rangle = e_1 = (0, 1)$. Now every unit vector in the space $\mathbb{C}^2$ is a state vector, and thus we have states that are linear combinations, *superpositions*, of $|0\rangle$ and $|1\rangle$, of the form

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$$

[picture to be added]

Figure 1: The Bloch sphere.

where α and β are complex numbers satisfying

$$|\alpha|^2 + |\beta|^2 = 1.$$

We call the basis consisting of $|0\rangle$ and $|1\rangle$ the *computational basis*. The coefficients we refer to as *probability amplitudes*.

Qubit states are sometimes depicted using the *Bloch sphere*, the (real) three-dimensional unit sphere. Any unit vector in $\mathbb{C}^2$ can be written in the form

$$|\psi\rangle = e^{i\gamma} \cos\frac{\theta}{2}|0\rangle + e^{i\xi} \sin\frac{\theta}{2}|1\rangle.$$

Denoting $\varphi = \xi - \gamma$, we get the form

$$|\psi\rangle = e^{i\gamma}\left(\cos\frac{\theta}{2}|0\rangle + e^{i\varphi} \sin\frac{\theta}{2}|1\rangle\right).$$

It turns out, that the *global phase* $e^{i\gamma}$ cannot be observed. Thus physically this state is indistinguishable from the state

$$|\psi'\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\varphi} \sin\frac{\theta}{2}|1\rangle.$$

which corresponds to a unique point on the surface of the Bloch sphere (see Figure 2).

2.1 Dirac notation

In physics, one often uses a notation developed by Dirac. Here any vector is written as a *ket*-vector, $|a\rangle$, where a is any label. Often the label is chosen to convey essential information of the vector, e.g., we denote the computational basis vectors by the bit values they represent $|0\rangle$ and $|1\rangle$, when considering spin up and spin down, the corresponding vectors are denoted $|\uparrow\rangle$ and $|\downarrow\rangle$, eigenvectors of an operator are often labelled by the corresponding eigenvalue, etc.

To every vector $|v\rangle$ in a vector space V there is a corresponding functional[2] $\langle v|$. This is called a *bra*-vector, but it is really an operator $\langle v| : V \to \mathbb{C}$ defined by the formula

$$\langle v|(|u\rangle) = (|v\rangle, |u\rangle).$$

[2] A *functional* is a linear operator from a vector space to its scalar field.

The application of a "bra-vector" functional $\langle v||u\rangle$ is often written in the shortened form $\langle v|u\rangle$, reminiscent of the inner product of the v and u vectors giving the resulting value.

Using the Dirac notation one has short notations for several operators. E.g., the *outer product* $|w\rangle\langle b|$ is the linear operator defined by

$$(|w\rangle\langle v|)(|u\rangle) = |w\rangle\langle v|u\rangle = \langle v|u\rangle|w\rangle.$$

In the special case where $|v\rangle$ is a unit vector the outer product $|v\rangle\langle v|$ is the orthogonal projection onto the space spanned by $|v\rangle$. More generally we can define *projectors*:

Definition 2.1. If V is an inner product space and A is a subspace of V with an orthonormal basis $|v_1\rangle, \ldots, |v_n\rangle$, we define the *projector* onto A by

$$P_A = \sum_i |v_i\rangle\langle v_i|.$$

Exercise 2.1. Show that the definition of projector is independent of the choice of orthonormal basis.

Projectors are a special case of the more general concept of projection.

Definition 2.2. A *projection* is a linear operator P satisfying $P^2 = P$.

Exercise 2.2. Show that a projector is a projection.

2.2 Adjoint and Hermitian operators

In quantum mechanics observable properties correspond to self-adjoint operators on the state space. We will not look closer at this general theory, but we need the definition of the adjoint of an operator to understand the measurement postulate.

Fact 2.3. *If V is a Hilbert space and A is a linear operator on V, there is a unique linear operator $A^\dagger$, the* adjoint *or* Hermitian conjugate *of A, such that for all $|v\rangle, |w\rangle \in V$*

$$(|v\rangle, A|w\rangle) = (A^\dagger|v\rangle, |w\rangle).$$

If A is described by a matrix, the adjoint of A is the conjugate-transpose of A, $A^\dagger = (A^*)^T$. However adjoints exist in any Hilbert space.[3]

[3]For a proof, see the course material of Functional analysis.

Definition 2.4. An operator A is *self-adjoint*, if $A^\dagger = A$.

The following properties are rather straightforward to verify:

Exercise 2.3.

1. $(AB)^\dagger = B^\dagger A^\dagger$,
2. interpreting c as the operator 'scalar multiplication with c', $c^\dagger = c^*$,
3. $(A^\dagger)^\dagger = A$,
4. denoting $|v\rangle^\dagger := \langle v|$, we have $(A|v\rangle)^\dagger = \langle v|A^\dagger$, and thus $(A|v\rangle, B|w\rangle) = \langle v|A^\dagger B|w\rangle$,
5. $(|w\rangle\langle v|)^\dagger = |w\rangle\langle w|$.

Exercise 2.4. Projectors are self-adjoint.

Using the notion of adjoint we can define unitary operators.

Definition 2.5. An operator U is *unitary* if and only if $U^\dagger U = I$.

Exercise 2.5. An operator is unitary if and only if it is a bijection that preserves inner products.

2.3 Measurements

According to the third postulate, a quantum measurement is described by a collection of operators M_m such that

- if a state $|\psi\rangle$ is measured, the outcome m will occur with probability

$$p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle,$$

- if outcome m occurred when ψ was measured, the state of the system after the measurement is

$$\frac{M_m|\psi\rangle}{\sqrt{\langle\psi|M_m^\dagger M_m|\psi\rangle}},$$

- the measurement operators satisfy the *completeness equation*

$$\sum_m M_m^\dagger M_m = I.$$

A special case of measurements are the *projective measurements*. These are described by a self-adjoint operator M, an *observable*. In the finite-dimensional case the eigenvectors of such an operator will span the whole state space giving rise to the spectral decomposition

$$M = \sum_m m P_m$$

where P_m is the projector onto the eigenspace of M with eigenvalue m. Here the measurement operator corresponding to an outcome m is the corresponding projector. Thus the probability of getting outcome m when measuring a state ψ is (using Exercises 2.2 and 2.4)

$$p(m) = \langle\psi|P_m^\dagger P_m|\psi\rangle = \langle\psi|P_m P_m|\psi\rangle = \langle\psi|P_m|\psi\rangle.$$

Example 2.6. Consider the state $\psi = \frac{1}{\sqrt{3}}|0\rangle + \sqrt{\frac{2}{3}}|1\rangle$. Measuring with respect to the computational basis corresponds to the projective measurement using the observable $0P_0 + 1P_1 = P_1$. The measurement operators here are $P_0 = |0\rangle\langle 0|$ and $P_1 = |1\rangle\langle 1|$, and the probabilities of measuring 0 and 1, respectively, are

$$p(0) = \langle\psi|P_0|\psi\rangle = \frac{1}{3}$$

and

$$p(1) = \langle\psi|P_1|\psi\rangle = \frac{2}{3}.$$

Exercise 2.6. What are the probabilities if the state ψ above is measured with respect to the basis $\{|+\rangle, |-\rangle\} := \{\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\}$? Can you give a corresponding observable?

2.4 Global and relative phase

When presenting the Bloch sphere we claimed that the states ψ and $\psi' = e^{i\theta}\psi$ are indistinguishable. This is because all we can see of the state are the possible measurement outcomes, and these two states have the same probability distribution for them:

$$\begin{aligned} p(m) &= \langle\psi'|M_m^\dagger M_m|\psi'\rangle \\ &= \langle\psi|e^{-i\theta}M_m^\dagger M_m e^{i\theta}|\psi\rangle \\ &= \langle\psi|M_m^\dagger M_m|\psi\rangle. \end{aligned}$$

A unit length coefficient of the whole state like this is called a *global phase*.

A *relative phase*, however, can be distinguished. This is a coefficient of part of the state. An example is the '-1' of a relative phase is the '-1' when comparing $|+\rangle$ and $|-\rangle$. Although a measurement of these vectors in the computational basis will give the same probability distribution, another choice of basis (e.g. the vectors themselves), will certainly show a difference.

3 Multiple qubits - tensor products

According to the fourth postulate composite systems are modelled by tensor products.

Definition 3.1. Let V and W be vector spaces over some field K. We define the tensor product $V \otimes W$ of V and W as follows: First consider the K-vector space A with a basis consisting of all pairs $(v, w) \in V \times W$. (This is also called the *free vector space on* $V \times W$.) Let B be the linear subspace of A generated by all vectors in A of the form

- $(v, w_1 + w_2) - (v, w_1) - (v, w_2)$
- $(v_1 + v_2, w) - (v_1, w), -(v_2, w)$
- $\lambda(v, w) - (\lambda v, w)$
- $\lambda(v, w) - (v, \lambda w)$

with $v, v_i \in V$, $w, w_i \in W$, $\lambda \in K$. Then $V \otimes W$ is the quotient vector space A/B.

Fact 3.2. *The tensor product is the 'freest' possible vector space for which there is a bilinear map* $V \times w \to V \otimes W$.

Definition 3.3. A map $f : V \times W \to U$ is *K-bilinear* if

- $f(v_1 + v_2, w) = f(v_1, w) + f(v_2, w)$,
- $f(v, w_1 + w_2) = f(v, w_1) + f(v, w_2)$,
- $f(\lambda v, w) = \lambda f(v, w)$,
- $f(v, \lambda w) = \lambda f(v, w)$.

In practice, if V and W are two K vector spaces with bases $\{v_i :\leq i \leq m\}$ and $\{w_j : 1 \leq j \leq n\}$, respectively, then $V \otimes W$ is a vector space with basis $\{v_i \otimes w_j : 1 \leq i \leq m, 1 \leq j \leq n\}$ such that

$$(\sum_i a_i v_i) \otimes (\sum_j b_j w_j) = \sum_{i,j} a_i b_j \, v_k \otimes w_j.$$

Example 3.4. For two qubits, we get a basis

$$\{|0\rangle \otimes |0\rangle, |0\rangle \otimes |1\rangle, |1\rangle \otimes |0\rangle, |1\rangle \otimes |1\rangle\}.$$

The vector $|0\rangle \otimes |0\rangle$ is often written in the shorter form $|0\rangle|0\rangle$, or even $|00\rangle$ (and similarly for the other bit combinations). Generalising this we see that the computational basis for n qubits has the form

$$\{|\eta\rangle : \eta \in \{0,1\}^n\}.$$

The state $|+\rangle \otimes |+\rangle$ can thus be written as a linear combination of basic elements from the computational basis:

$$\begin{aligned} |+\rangle \otimes |+\rangle &= \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \\ &= \frac{1}{2}|0\rangle \otimes |0\rangle + \frac{1}{2}|0\rangle \otimes |1\rangle + \frac{1}{2}|1\rangle \otimes |0\rangle + \frac{1}{2}|1\rangle \otimes |1\rangle \\ &= \frac{1}{2}(|00\rangle + |01\rangle + |10\rangle + |11\rangle). \end{aligned}$$

Exercise 3.1.

1. $(A \otimes B)(C \otimes D) = AC \otimes BD$,
2. $(A + B) \otimes C = A \otimes C + B \otimes C$ and $A \otimes (C + D) = A \otimes C + A \otimes D$,
3. $(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger$,
4. if A and B are unitary, then so is $A \otimes B$.

Definition 3.5. If V and W are inner product spaces, we can define an inner product on $V \otimes W$ by

$$(\sum_i a_i v_i \otimes w_i, \sum_j b_j v'_j \otimes w'_j) = \sum_{i,j} a_i^* b_j (v_i, v'_j)(w_i, w'_j).$$

Exercise 3.2. Show that the definition above gives a well-defined inner product on $V \otimes W$.

3.1 Operators on $V \otimes W$

Definition 3.6. If V, V', W and W' are K-vector spaces and $A : V \to V'$ and $B : W \to W'$ are linear operators, we define the operator $A \otimes B : V \otimes W \to V' \otimes W'$ by

$$(A \otimes B)(v \otimes w) = Av \otimes Bw$$

for all $v \otimes w \in V \otimes W$.

It can be shown that this give a well defined linear operator on $V \otimes W$ (using the universality of the tensor, i.e., that $V \otimes W$ is the freest possible combination one gets while demanding bilinearity).

If we agree on an ordering on the basic vectors of a tensor product, we can write vectors in coordinate form and get matrix representations for operators of the form $A \otimes B$. We usually order the vectors so that the string of indexes of the vectors in the tensor product, are ordered alphabetically, e.g, for the two-qubit computational basis we get the following column vector representations:

$$|00\rangle = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, |01\rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, |10\rangle = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, |11\rangle = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Then if the linear operators A and B are represented by a $m \times n$ and $p \times q$ matrix, respectively, then we can form their *Kronecker product*, representing the operator $A \otimes B$:

$$A \otimes B = \begin{bmatrix} A_{11}B & A_{12}B & \cdots & A_{1n}B \\ A_{21}B & A_{22}B & \cdots & A_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1}B & A_{m2}B & \cdots & A_{mn}B \end{bmatrix}$$

This representation gives the matrix in blocks of size $p \times q$, where the blocks are formed by scaling the matrix B with the given coefficient from A.

Example 3.7. Consider the two-qubit state

$$|+\rangle \otimes |0\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \cdot 1 \\ 1 \cdot 0 \\ 1 \cdot 1 \\ 1 \cdot 0 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

We can apply the operator $H \otimes I$ to this state. This corresponds to the matrix

$$H \otimes I = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \cdot I & 1 \cdot I \\ 1 \cdot I & -1 \cdot I \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}.$$

Applying this matrix to the state $|+\rangle \otimes |0\rangle$ gives $H|+\rangle \otimes I|0\rangle = |00\rangle$:

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1+1 \\ 0 \\ 1-1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

Definition 3.8. Multiple qubit states that are tensor products of single qubit states are called *separable*. States that are not separable are *entangled*.

One of the most crucial parts of quantum computation is the existence of entangled states.

Lemma 3.9. *The state* $|\psi\rangle = \frac{1}{\sqrt{2}}|00\rangle + \frac{1}{\sqrt{2}}|11\rangle$ *is entangled.*

Proof. Suppose towards a contradiction that $|\psi\rangle$ is separable. Then it can be written as a tensor product $|\psi\rangle = |u\rangle \otimes |v\rangle$, where $|u\rangle = u_1|0\rangle + u_2|1\rangle$ and $|v\rangle = v_1|0\rangle + v_2|1\rangle$. Now

$$|u\rangle \otimes |v\rangle = u_1v_1|00\rangle + u_1v_2|01\rangle + u_2v_2|10\rangle + u_2v_2|11\rangle$$

so

$$u_1v_1 = \frac{1}{\sqrt{2}},\ u_1v_2 = 0,\ u_2v_1 = 0,\ \text{and} u_2v_2 = \frac{1}{\sqrt{2}},$$

which implies none of the elements u_1, v_1, u_2, v_2 is zero, but one of u_1 and v_2 is (as well as one of u_2 and v_1). This is a contradiction. □

The special thing about entangled states is their behaviour when measured. One can, namely, measure just one qubit out of a multi-qubit system. But if the state is entangled, the subsequent measurements of the other qubits will be affected by this. We will look at this phenomenon next.

3.2 Measuring multi-qubit systems

Measuring one qubit with respect to the computational bases, meant using the projectors P_0 and P_1 as measurement operators. In an n-qubit system, one can consider the projectors $P_\eta = |\eta\rangle\langle\eta|$ onto the subspaces spanned by single basic states $|\eta\rangle$, for $\eta \in \{0,1\}^n$. But using tensors one can also form measurements of single qubits in multi-qubit systems.

Now $\{P_0 \otimes I, P_1 \otimes I\}$ describes measurement of the first qubit (wrt the computational basis) in a multi-qubit system. Here I is the identity operator on all the other qubits (i.e., a tensor of identity operators, which you can easily check is just an identity operator in a larger dimension). To make sure this fits into our demands of a measurement, one needs to check the completeness relation, i.e., that $\sum_m M_m^\dagger M_m = I$:

$$\begin{aligned}(P_0 \otimes I_{2^{n-1}})^\dagger(P_0 \otimes I_{2^{n-1}}) + (P_1 \otimes I_{2^{n-1}})^\dagger(P_1 \otimes I_{2^{n-1}}) \\ &= (P_0^\dagger \otimes I_{2^{n-1}}^\dagger)(P_0 \otimes I_{2^{n-1}}) + (P_1^\dagger \otimes I_{2^{n-1}}^\dagger)(P_1 \otimes I_{2^{n-1}}) \\ &= (P_0 \otimes I_{2^{n-1}})^2 + (P_1 \otimes I_{2^{n-1}})^2 \\ &= (P_0^2 \otimes I_{2^{n-1}}^2) + (P_1^2 \otimes I_{2^{n-1}}^2) \\ &= (P_0 + P_1) \otimes I_{2^{n-1}} \\ &= I_2 \otimes I_{2^{n-1}} = I_{2^n}.\end{aligned}$$

Now, e.t., if we measure the first qubit of the two-qubit state

$$|\psi\rangle = a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle,$$

we have

$$(P_0 \otimes I)(a|00\rangle + b|01\rangle + c|10\rangle + d|11\rangle) = a|00\rangle + b|10\rangle.$$

So the probability of observing a 0, is $\langle\psi|P_0 \otimes I|\psi\rangle = a^2 + b^2$, and the state after observing a 0 is

$$\frac{a|00\rangle + b|10\rangle}{\sqrt{a^2 + b^2}}.$$

Similarly the probability of observing a 1, is $c^2 + d^2$, after which the state will be

$$\frac{c|10\rangle + d|11\rangle}{\sqrt{c^2 + d^2}}.$$

Now let us look at the so called *Bell state* $|\beta_{00}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$. If we measure its first qubit, we will end up in either the state $|00\rangle$ or the state $|11\rangle$. Thus looking at the first qubit (and knowing which state we were in to begin with), we will know with certainty the value we observe if we measure the second qubit.

Exercise 3.3. Show that if we measure the first qubit of the Bell state $|\beta_{00}\rangle$ and observe a 0, then after this the probability of observing a 0 when measuring the second qubit is 1.

Exercise 3.4. There are four Bell states

$$|\beta_{00}\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle),$$
$$|\beta_{01}\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle),$$
$$|\beta_{10}\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle),$$
$$|\beta_{11}\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle).$$

Show that the state β_{xy} is obtained from the two-qubit basic state $|xy\rangle$ by first applying a Hadamard operator to the first qubit, and then a controlled `NOT` operator to the qubits. The controlled `NOT` does nothing if the first qubit is 0, and flips the second qubit if the first is 1.