QUANTUM INFORMATION & COMPUTATION

Nilanjana Datta, DAMTP Cambridge

1 Introduction: why *quantum* computation and information?

Let us begin by asking: "*What is information?*" Well, intuitively we "get information" when we acquire knowledge of an alternative that we did not know before (and consequently, there is a "reduction in uncertainty"). In classical computing and communication, information is conventionally represented by a *classical bit* (or string of bits) viz. a Boolean (i.e. binary) variable that can take values 0 or 1. It represents the "elementary unit of information" giving the result of a single binary decision e.g. a yes/no question. More elaborate kinds of information are represented by bit strings to provide distinct labels for more than two *a priori* possible answers.

Having our notion of information we can then go on to introduce *computation* as *information processing* i.e. the updating of a bit string by a prescribed sequence of steps (the "program"). These steps (at the basic bit string level) are normally taken to be local Boolean gates – in each step one or two of the bits (at prescribed locations in the string) are updated by applying a prescribed Boolean function, or gate, to them. Typical examples of such gates are the 1-bit NOT operation and the 2-bit operations called AND and OR; and it can be shown that this small basic set is "universal" in the sense that it suffices for the construction of any Boolean function (with inputs and outputs being bit strings of any length).

All the above will probably be unsurprising to many readers but we can go further – if information is represented by bits then: "*What is a bit?*" Above we have associated it with a Boolean variable, an abstract mathematical concept. But this cannot be our answer because when we acquire information, we need to actually receive "something real", not just entertain an abstract mathematical concept. The key point here is well expressed by the quote (R. Landauer 1996) "Information is not a disembodied abstract entity; it is always tied to a physical representation". Indeed the Boolean values 0 and 1 serve only to provide two recognisably different labels. So our answer to "What is a bit?" is: a bit is given by any two different physical states (of some physical system) that can be reliably distinguished (by a physical measurement). The Boolean values 0 and 1 are just two distinguishable patterns of physical ink on a page; when we ask a question and hear 'yes' or 'no', we are just using our ears as a physical device to distinguish between two different soundwave forms in the air; and in a computer memory, bits can be represented by two different voltage levels in a material etc. The key message here is: "No information without representation!"

Consequently, if information is represented in physical states or degrees of freedom of

some physical system, then any possible act of computation, or information processing, must correspond to a physical evolution of that physical system. For example, any actual computer is always a physical device whose operation must obey the laws of physics.

More generally, the fact to keep in mind is:
the possibilities and limitations of information storage, processing (i.e. computation) and communication must all rest on the laws of physics and cannot be determined by abstract thought or mathematics alone!
Hence, there must be a deep and fundamental connection between physics and computation, and *that* is why we need *quantum* information and computation (QIC), since, as Feynman put it more succinctly: because "Nature isn't classical, dammit...". And indeed it can be argued that our conventional, generally accepted model of computation (in any of its equivalent forms, e.g. based on Turing machines or viewing computations in terms of Boolean gates etc.) amounts to the computational possibilities allowed by the laws of classical physics.

Quantum physics differs dramatically from classical physics in the way it represents the physical world and the kinds of processes that it allows (as we will see in detail shortly). QIC is the study of the possible applications and exploitation of these novel quantum features in issues of information storage, computation, computational complexity, cryptography and communication. The subject is a remarkable synthesis of mathematics, theoretical computer science, classical information theory and cryptography with quantum physics, promising a series of benefits of much practical significance, beyond the remit (even in principle) of conventional (classical) computing and information technology.

The subject emerged in the mid-1980's (even though there had already been some significant contribution made by Alexander Holevo in Russia in the 1970's) and it is currently one of the most active areas of all scientific research internationally. Here we will just briefly highlight some of the key issues of relevance in discussing its benefits:

Computing power - computational complexity issues.
As we will see later in the course, a quantum computer cannot compute any computational task that is not already computable in principle on a classical computer. However, the key issue here is not 'computability in principle' but 'computability in practice' i.e. that of computational *hardness*. In computational complexity theory the 'hardness' of a computational task is measured by the amount of computational resources needed to compute it; the resources considered are 'time' i.e. the number of computational steps, and 'space' i.e. the amount of computer memory workspace needed.

For example, think of the task of factoring a given integer with n digits, and how hard this is to do, as a function of n. Here n is called the 'input size' for the computation. For any proposed algorithm we ask: does the time (i.e. number of steps) grow polynomially with n (so-called poly-time algorithm) or exponentially with n (faster than any polynomial growth), as n gets larger and larger. Poly-time algorithms are regarded as "feasible in practice" whereas a task with only exponential-time algorithms, while computable in principle, is regarded as infeasible, or effectively uncomputable, in practice – because for relatively modest input sizes, the time required would exceed any reasonably available

period (e.g. exceeding the age of the universe).

This is where quantum computing has a major impact: we will see that the formalism of quantum theory leads to new kinds of "non-classical" modes of computation (new kinds of computational steps for information processing), providing remarkable new possibilities for computational algorithms. In some cases these possibilities are able to cross the boundary between poly-time and exponential-time algorithms i.e. there are some computational tasks for which no known classical poly-time algorithm exists but which can be solved in poly-time on a quantum computer i.e. these tasks, which are effectively uncomputable in practice on a classical computer, become computable in practice on a quantum computer.

The most famous example is the computational task of *integer factorisation*. In classical computation there is no known algorithm that runs in polynomial time (in the number of digits) but in 1994 Peter Shor discovered a poly-time *quantum* algorithm for factorisation. We emphasise that this exponential speedup in time is achieved not by an increase in clock speed of steps on the computer, but by exploiting entirely new (quantum) kinds of computational steps (and needing exponentially fewer of them) that are simply not available to classical computers.

We will see a variety of examples of such quantum computational benefits (including factoring) in the second half of the course.

Communication and security issues - quantum states as information carriers.
Intrinsically quantum (i.e. non-classical) features of quantum states (including the possibilities of quantum superposition, entanglement and principles of quantum measurement theory) can be exploited to provide novel possibilities (beyond what is achievable with classical physics) for information communication and security. These include the so-called processes of *quantum teleportation and superdense coding*, and a variety of important cryptographic issues such as the ability to *implement provably secure communication*. In this course we will discuss quantum teleportation, superdense coding, the Bennett-Brassard quantum scheme for secure communication, and some further features of quantum states when viewed as information carriers.

Technological issues.
Historically in computer science (before the advent of quantum computing) there was a phenomenon known as Moore's law viz. that since 1965 there has been a steady rate of miniaturisation of computer components, by approximately a factor of 4 every 3.5 years. With this trend we have now effectively reached the atomic scale where classical physics fails completely and quantum effects are dominant – components begin to malfunction in 'bizarre' quantum ways. To deal with this, we could either aim to re-design our components to stamp out the new effects to provide the same functionality as before, or else we could embrace the new quantum effects, aiming to exploit them in new kinds of computational ways. Our discussion of computational complexity above shows that the latter is surely the way to go!

However this involves immense technological challenges: it turns out that quantum states and processes are intrinsically more fragile and difficult to control cleanly, than their

classical counterparts. Inspired by the theoretical technological possibilities on offer, in recent years there has been a huge effort devoted to developing the needed quantum technological capability. To date, some quantum cryptographic protocols have been implemented (including secure communication), even to the level of being commercially available. However these require quantum processing of only relatively small systems (thus within our current quantum technological capabilities) and similarly some quantum algorithms on very small input instances have been demonstrated. But the ultimate 'holy grail' of a working scalable universal quantum computer is currently beyond our quantum technological capability. Some of the world's leading information technology companies (including IBM, Google, Microsoft) have mounted huge research and development efforts with just that aim.

In fact, there has been a recent breakthrough by Google (October 2019). Google claims to have reached "quantum supremacy" with an array of 53 qubits. Quantum supremacy is the goal of demonstrating that a programmable quantum device can solve a problem in practice that a classical computer cannot. *Discussion in the lecture.*

2 Quantum Mechanics: Mathematical Preliminaries

Quantum Mechanics is a physical theory which replaces Newtonian Mechanics at atomic and sub–atomic levels. The basic principles of quantum mechanics can be summarized in four postulates. These postulates provide a connection between the physical world and the mathematical formalism of Quantum Mechanics We will begin by setting out these postulates (QM1) - (QM4), while simultaneously introducing and explaining the formalism of Dirac notation, which we will use to express their mathematical content.

Dirac notation is nothing more than an alternative notation for basic linear algebra which is widely used in quantum mechanics. We will use it for essentially all aspects of this course so it will be important for you to master it at the outset!

2.1 Mathematical Preliminaries and the Dirac bra-ket notation

Recall: A *complex, inner product space* $\mathcal{V}$ is a vector space over the field $\mathbb{C}$ of complex numbers, with an additional structure called the *inner product*. We shall use Dirac's bra-ket notation to denote vectors in $\mathcal{V}$ and their inner products.

1. A vector in $\mathcal{V}$ is denoted by the ket-vector (or simply *ket*) $|\psi\rangle$.
2. Its conjugate transpose is denoted by the bra-vector (or simply *bra*): $\langle\psi| \equiv |\psi\rangle^\dagger$.
3. If $\mathcal{V} \equiv \mathbb{C}^n$ (i.e. n-dimensional complex vector space), then any $|\psi\rangle \in \mathcal{V}$ is given by a column vector of length n and complex entries.
4. $\langle\psi|$ is given by a row vector of length n whose elements are complex conjugates of the elements of $|\psi\rangle$.

5. The inner product (or scalar product) of two vectors $|\psi\rangle, |\phi\rangle \in \mathcal{V}$ is denoted as

$$(|\psi\rangle, |\phi\rangle) \equiv \langle\psi|\phi\rangle$$

.

Properties of the inner product

1. Positivity: $\langle\psi|\psi\rangle \geq 0$ with equality if and only if $|\psi\rangle = 0$.
2. Linearity (in the second argument):

$$\langle\phi|a\psi_1 + b\psi_2\rangle = a\langle\phi|\psi_1\rangle + b\langle\phi|\psi_2\rangle, \quad a, b \in \mathbf{C}$$

3. Anti-linearity in the first argument: if $|\phi\rangle = \alpha\,|\phi_1\rangle + \beta\,|\phi_2\rangle$, then

$$\langle\phi|\psi\rangle = \alpha^*\langle\phi_1|\psi\rangle + \beta^*\langle\phi_2|\psi\rangle.$$

4. Skew-symmetry (or conjugate symmetry):

$$\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*$$

5. The norm of a vector $|\psi\rangle \in \mathcal{V}$ is given by

$$||\psi|| = \sqrt{\langle\psi|\psi\rangle}.$$

The special case of $\mathcal{V} \equiv \mathbb{C}^2$: In this course we will often work with a 2-dimensional complex vector space $\mathcal{V} \equiv \mathbb{C}^2$ with a chosen orthonormal basis denoted $\{|0\rangle, |1\rangle\}$ i.e. the basis vectors are labelled by the bit values 0 and 1, and this basis will be called the *computational basis or standard basis*[1]. In this course we adopt the following, standard convention:

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

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

Hence, a ket vector $|v\rangle = a\,|0\rangle + b\,|1\rangle \in \mathbb{C}^2$ is given (in component form) by the *column* vector

$$|v\rangle = \begin{pmatrix} a \\ b \end{pmatrix}.$$

The corresponding *bra vector* $\langle v| \equiv |v\rangle^\dagger$ is given (in component form) by the row vector

$$\langle v| = |v\rangle^\dagger = a^*\,\langle 0| + b^*\,\langle 1| = (a^*\ b^*).$$

[1] All of our constructions and formulae can be easily generalised to arbitrary finite dimensional spaces.

If $|w\rangle = c\,|0\rangle + d\,|1\rangle$, then the inner product of $|v\rangle$ and $|w\rangle$ is given by

$$\langle v|w\rangle = |v\rangle^\dagger\,|w\rangle = (a^*\ b^*)\begin{pmatrix} c \\ d \end{pmatrix} = a^*c + b^*d.$$

Outer product: The outer product of two vectors $|\psi\rangle, |\phi\rangle \in \mathcal{V}$ is given by $|\psi\rangle\langle\phi|$. If $\mathcal{V} = \mathbb{C}^n$, then $|\psi\rangle\langle\phi|$ is given by an $n \times n$ matrix.

Let us denote a complete orthonormal basis (o.n.b.) of $\mathcal{V} \equiv \mathbb{C}^n$ by $\{|i\rangle\}_{i=1}^n$, with $|i\rangle \in \mathcal{V}$ for all $i = 1, 2, \ldots n$: we have $\langle i|j\rangle = \delta_{ij}$ (orthonormality) and

$$|i\rangle\langle i| = I, \quad \text{completeness relation} \tag{1}$$

where I denotes the $n \times n$ identity matrix. Any vector $|\psi\rangle \in \mathcal{V}$ can be expressed in this basis as: $|\psi\rangle = \sum_{i=1}^n c_i\,|i\rangle, \quad c_i \in \mathbb{C}$. If the ket $|\psi\rangle$ is chosen to be normalized (i.e. $\langle\psi|\psi\rangle = 1$), then we have

$$1 = \langle\psi|\psi\rangle = \sum_{i=1}^n |c_i|^2.$$

Since $|c_i|^2 \geq 0$ and $\sum_{i=1}^n |c_i|^2 = 1$, it follows that $\{|c_i|^2\}_{i=1}^n$ forms a probability distribution, in this case. For $\mathcal{V} \equiv \mathbb{C}^2$, $\{|0\rangle, |1\rangle\}$ is an example of a complete orthonormal basis, with the completeness relation $|0\rangle\langle 0| + |1\rangle\langle 1| = I$, where I is the 2×2 identity matrix.

Remark: Indeed the whole Dirac notation formalism is motivated by the bracket notation $(\underline{v}, \underline{w})$ for inner products commonly used in mathematics, hence the terms "bra" and "ket" vectors, giving the inner product as a "bra-ket". In more abstract terms, bra vectors are a notation for elements of the dual space $\mathcal{V}^*$, i.e. $\langle v|$ is the linear functional whose value on any ket $|w\rangle$ is the inner product $\langle v|w\rangle$.

Tensor Products

The *tensor product* is a way of putting vector spaces together to form larger vector spaces. If $\mathcal{V}$ and $\mathcal{W}$ are vector spaces of dimensions m and n with bases $\{|e_1\rangle, \ldots, |e_m\rangle\}$ and $\{|f_1\rangle, \ldots, |f_n\rangle\}$ respectively, then the tensor product space $\mathcal{V} \otimes \mathcal{W}$ has dimension mn and can be regarded as consisting of all formal linear combinations of the symbols $|e_i\rangle \otimes |f_j\rangle$ for $i = 1, \ldots, m$ and $j = 1, \ldots, n$. There is a natural bilinear embedding $\mathcal{V} \times \mathcal{W} \to \mathcal{V} \otimes \mathcal{W}$ defined as follows. If $|\alpha\rangle = \sum_i a_i |e_i\rangle$ and $|\beta\rangle = \sum_j b_j |f_j\rangle$ are general vectors in $\mathcal{V}$ and $\mathcal{W}$ respectively then

$$(|\alpha\rangle, |\beta\rangle) \mapsto |\alpha\rangle \otimes |\beta\rangle = \sum_{ij} a_i b_j |e_i\rangle \otimes |f_j\rangle \qquad (*)$$

In other words, the elements of $\mathcal{V} \otimes \mathcal{W}$ are linear combinations of tensor products of elements of $\mathcal{V}$ and $\mathcal{W}$, and $\{|e_i\rangle \otimes |f_j\rangle\}$ is an orthonormal basis of $\mathcal{V} \otimes \mathcal{W}$.

Dirac notation: tensor products of vectors

Product vectors and entangled vectors
Any vector of the form $|u\rangle \otimes |v\rangle$ in $\mathcal{V} \otimes \mathcal{W}$ is called a *product vector*. We often write the product vector $|u\rangle \otimes |v\rangle$ simply as $|u\rangle |v\rangle$ (omitting the $\otimes$). The mapping in eqn. (*) above is not surjective – vectors in $\mathcal{V} \otimes \mathcal{W}$ that are not product vectors are called *entangled vectors*. We will have detailed discussions about *entanglement*, which plays a key role in quantum information theory, later.

Focus: We will be mostly concerned with tensor products of the two-dimensional space $\mathcal{V} \equiv \mathbb{C}^2$ with itself (multiple times). For the k-fold tensor power we write $(\mathbb{C}^2)^{\otimes k} \equiv \otimes^k \mathbb{C}^2 = \mathbb{C}^2 \otimes \ldots \otimes \mathbb{C}^2$ which is a space of dimension 2^k with basis $\{|i_1\rangle \otimes \ldots \otimes |i_k\rangle \; ; \; i_1, \ldots i_k \in \{0, 1\}\}$ labelled by the 2^k k-bit strings $i_1 \ldots i_k$. We often write $|i_1\rangle \otimes \ldots \otimes |i_k\rangle$ simply as $|i_1 \ldots i_k\rangle$. This basis is also called the *computational (or standard) basis* of $(\mathbb{C}^2)^{\otimes k}$.

Example. For $k = 2$, if $|v\rangle = a|0\rangle + b|1\rangle$ and $|w\rangle = c|0\rangle + d|1\rangle$, we have $|v\rangle |w\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$. By formal multiplication we get (omitting the symbol $\otimes$)

$$|v\rangle \otimes |w\rangle = (a|0\rangle + b|1\rangle)(c|0\rangle + d|1\rangle) = ac|00\rangle + ad|01\rangle + bc|10\rangle + bd|11\rangle$$

and in terms of components we have

$$|v\rangle |w\rangle \equiv |v\rangle \otimes |w\rangle = \begin{pmatrix} a \\ b \end{pmatrix} \otimes \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} ac \\ ad \\ bc \\ bd \end{pmatrix} = \begin{pmatrix} a \begin{pmatrix} c \\ d \end{pmatrix} \\ b \begin{pmatrix} c \\ d \end{pmatrix} \end{pmatrix}.$$

Note how the last expression gives the pattern for how to get the final tensor product components from those of the individual vectors: take each numerical component of the first vector in turn and "expand it up (doubling it to two numbers)" by multiplying it by the components of the second vector taken in order, and then list all these in order

in a column. Note also that this illustrates that the tensor product is not commutative i.e. $|v\rangle \otimes |w\rangle \neq |w\rangle \otimes |v\rangle$ in general. $\square$

Example. The vector

$$|\Phi\rangle := \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \in \mathbb{C}^2 \otimes \mathbb{C}^2$$

is entangled i.e. it is not a product vector. To see this in an elementary way, suppose that it is a product vector, that is,

$$\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) = (a\,|0\rangle + b\,|1\rangle)(c\,|0\rangle + d\,|1\rangle) = ac\,|00\rangle + ad\,|01\rangle + bc\,|10\rangle + bd\,|11\rangle\,,$$

for some a, b, c, d. Then comparing the first and last expressions, we must have $ad = 0$ (and also $bc = 0$), so either $a = 0$ or $d = 0$. Thus respectively either $|00\rangle$ or $|11\rangle$ has coefficient zero too, which is a contradiction.

This argument may be generalised to show that an arbitrary vector $\alpha\,|00\rangle + \beta\,|01\rangle + \gamma\,|10\rangle + \delta\,|11\rangle \in \mathbb{C}^2 \otimes \mathbb{C}^2$ is entangled if and only if $\alpha\delta - \beta\gamma \neq 0$ (see Exercise Sheet 1). But beware: this simple single-equation characterisation no longer suffices if the component spaces have dimension greater than 2. Indeed in that general case, a vector $\sum_{i,j} A_{ij}\,|i\rangle\,|j\rangle$ is a product vector if and only if the matrix $(A_{ij})_{i,j}$ of coefficients has rank one. $\square$

Inner product in $\mathcal{V} \otimes \mathcal{W}$

The inner products on $\mathcal{V}$ and $\mathcal{W}$ give a natural inner product on $\mathcal{V} \otimes \mathcal{W}$ defined "slot-wise" (the slots being the component spaces). Thus for product vectors we have the inner product of $|\alpha_1\rangle\,|\beta_1\rangle$ with $|\alpha_2\rangle\,|\beta_2\rangle$ being $\langle\alpha_1|\alpha_2\rangle\langle\beta_1|\beta_2\rangle$. This extends to general (entangled) vectors by linearity since general vectors are always linear combinations of product vectors (e.g. of the product basis vectors $|e_i\rangle\,|f_j\rangle$). More explicitly, (and here using the summation convention for repeated indices) if $|A\rangle = a_{ij}\,|e_i\rangle\,|f_j\rangle$ and $|B\rangle = b_{ij}\,|e_i\rangle\,|f_j\rangle$ are two vectors in $\mathcal{V} \otimes \mathcal{W}$ then

$$\langle A|B\rangle = \left(\, a^*_{ij}\,\langle e_i|\,\langle f_j|\,\right)\,\left(\, b_{pq}\,|e_p\rangle\,|f_q\rangle\,\right) = a^*_{ij} b_{pq}\,\langle e_i|e_p\rangle\,\langle f_j|f_q\rangle = a^*_{ij} b_{ij}$$

where we have used the basis orthonormality relations $\langle e_i|e_p\rangle = \delta_{ip}$ and $\langle f_j|f_q\rangle = \delta_{jq}$.

Thus inner products are calculated by contracting indices on the vector components (relative to an orthonormal product basis), also with a complex conjugation for the left vector

We often write the bra vector of a product ket $|\alpha\rangle\,|\beta\rangle \in \mathcal{V} \otimes \mathcal{W}$ as $\langle\beta|\,\langle\alpha|$ ("reflecting" the symbols) with order of spaces reversed. However, irrespective of how we write it, it will generally be important to remain aware of which vector space corresponds to which slot. If needed, we can make this explicit by using subscripts to denote the names of the spaces e.g. writing the bra vector of $|\alpha\rangle_{\mathcal{V}}\,|\beta\rangle_{\mathcal{W}}$ as ${}_{\mathcal{W}}\langle\beta|\;{}_{\mathcal{V}}\langle\alpha|$ or ${}_{\mathcal{V}}\langle\alpha|\;{}_{\mathcal{W}}\langle\beta|$.