Part II: Quantum Information and Computation - Revision

Lectures by Richard Jozsa, notes by James Moore

1 Formalism

1.1 Entanglement

Definition: Let V and W be vector spaces, and $V \otimes W$ be their tensor product. If $|\xi\rangle \in V \otimes W$ can be written in the form $|\xi\rangle = |a\rangle |b\rangle$, it is called a *product vector*; otherwise it is called *entangled*.

Theorem: In $V_2 \otimes V_2$, the state

$$|v\rangle = \alpha |00\rangle + \beta |01\rangle + \gamma |10\rangle + \delta |11\rangle$$

is entangled if and only if $\alpha\delta - \beta\gamma \neq 0$.

Proof: Write as a product state and compare coefficients. □

1.2 Schmidt form

Theorem: Let $|\psi\rangle_{AB}$ be a state of a composite system $A \otimes B$ where $\dim(A) = m$ and $\dim(B) = n$. Let $d = \min(m, n)$. Then there are orthonormal bases $\{|\alpha_1\rangle, |\alpha_2\rangle, ..., |\alpha_m\rangle\}$ and $\{|\beta_1\rangle, |\beta_2\rangle, ..., |\beta_n\rangle\}$ of A and B and non-negative real numbers $\lambda_1, \lambda_2, ..., \lambda_d$ such that

$$|\psi\rangle_{AB} = \sum_{i=1}^{d} \lambda_i |\alpha_i\rangle |\beta_i\rangle .$$

Proof: Let

$$|\psi\rangle_{AB} = \sum_{i,j} a_{ij} |i\rangle |j\rangle .$$

Express a_{ij} as its singular value decomposition

$$a_{ij} = [UDV^\dagger]_{ij} = \sum_{k,l} u_{ik} d_{kl} v^*_{jl},$$

where U, V are unitary and D is diagonal. Then

$$|\psi\rangle_{AB} = \sum_{i,j,k,l} d_{kl}(u_{ik}|i\rangle)(v^*_{jl}|j\rangle) = \sum_{i,j,k} d_{kk}(u_{ik}|i\rangle)(v^*_{jk}|j\rangle).$$

Write $|\alpha_k\rangle = \sum u_{ik}|i\rangle$ and $|\beta_k\rangle = \sum v^*_{jk}|j\rangle$. This gives the required Schmidt form (can easily show $\langle\alpha_l|\alpha_k\rangle = \langle\beta_l|\beta_k\rangle = \delta_{lk}$). □

2 Quantum information properties

2.1 Operations on quantum information

Given some quantum information $|\psi\rangle$, there are three possible operations we can perform:

1. Adjoin an *ancilla*, i.e. adjoin a fixed known quantum state $|A\rangle$ to the system, and consider the combined state $|\tilde{\psi}\rangle = |\psi\rangle |A\rangle$.
2. Apply any unitary operation to $|\psi\rangle$.
3. Measure the system, giving us both classical information from the result of the measurement, and a post-measurement state which we can further manipulate.

Theorem: Operations on quantum information can always be reduced to the sequence (i) adjoin an ancilla; (ii) apply a single unitary operation; (iii) measure the system once.

Proof: Non-examinable. □

2.2 The basic quantum gates

Definition: A unitary operation on one or two qubits is called a *quantum gate*.

The main examples of quantum gates are:

The Hadamard gate:

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

The Pauli gates:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The controlled not: Note CX_{12} means 1 is the control qubit, 2 is the target qubit.

$$CX = CX_{12} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

2.3 The no-cloning theorem

Definition: A *quantum cloning operation* is an operation of the form:

$$|\alpha\rangle\,|0\rangle\,|M\rangle \mapsto |\alpha\rangle\,|\alpha\rangle\,|M_\alpha\rangle\,.$$

Theorem: There does not exist a unitary cloning operation.

Proof: Let $|\xi\rangle$ and $|\eta\rangle$ be two distinct non-orthogonal states. Then the cloning process must achieve both:

$$|\xi\rangle\,|0\rangle\,|M\rangle \mapsto |\xi\rangle\,|0\rangle\,|M_\xi\rangle\,, \qquad |\eta\rangle\,|0\rangle\,|M\rangle \mapsto |\eta\rangle\,|\eta\rangle\,|M_\eta\rangle\,.$$

Since unitary operations preserve inner products, we have

$$\langle\xi|\eta\rangle = \langle\xi|\eta\rangle\,\langle\xi|\eta\rangle\,\langle M_\xi|M_\eta\rangle \quad\Rightarrow\quad 1 = \langle\xi|\eta\rangle\,\langle M_\xi|M_\eta\rangle\,.$$

But since the states are distinct, the right hand side is less than 1. Contradiction. □

2.4 Superluminal communication

If cloning *is* allowed, we can signal faster than light.

Herbert's method: Let Alice and Bob share an entangled Bell state $|\phi^+\rangle$. To signal yes, Alice measures her qubit in the basis $\{|0\rangle\,,|1\rangle\}$ and to signal no, she measures her qubit in the basis $\{|+\rangle\,,|-\rangle\}$.

Alice and Bob agree on a time Alice will signal Bob. As soon as Alice signals, Bob clones his qubit a million times. He then measures his million qubits in the computational basis. He gets all 0's or all 1's if Alice is signalling yes, and on average, 50% 0's and 50% 1's if Alice is signalling no. So he can tell what she is saying.

2.5 The no-deleting principle

Lemma: Let $|\xi_i\rangle$ and $|\eta_i\rangle$ for $i = 0, 1$ be states with $\langle\xi_0|\xi_1\rangle = \langle\eta_0|\eta_1\rangle$. Then there is a unitary operation U with $U\,|\xi_i\rangle = |\eta_i\rangle$.

Proof: Assume all states normalised. If $|\xi_0\rangle$ and $|\xi_1\rangle$ are equal, then the $|\eta_i\rangle$ are equal. So trivially exists such a U.

Otherwise, $|\xi_0\rangle \neq |\xi_1\rangle \Rightarrow |\eta_0\rangle \neq |\eta_1\rangle$ and $\{|\xi_0\rangle\,,|\xi_1\rangle\}$, $\{|\eta_0\rangle\,,|\eta_1\rangle\}$ form bases of V_2. Use

$$U = \left(|\eta_0\rangle \quad |\eta_1\rangle\right)\,.$$

in the $\{|\xi_0\rangle\,,|\xi_1\rangle\}$ basis. Write $|v\rangle = a\,|\xi_0\rangle + b\,|\xi_1\rangle$. We can easily show $||U\,|v\rangle\,||^2 = ||\,|v\rangle\,||^2$, hence U is norm-preserving, so is unitary. □

Definition: A *deleting operation* for the states $|\alpha_i\rangle$ ($i = 0, 1$, non-orthogonal and distinct) is a process effecting the following:

$$|\alpha_i\rangle\,|\alpha_i\rangle\,|M\rangle \mapsto |\alpha_i\rangle\,|0\rangle\,|M_i\rangle\,.$$

Theorem: For any unitary deleting process, the state $|\alpha_i\rangle$ can be reconstituted from $|M_i\rangle$ alone.

Proof: Assuming we need to delete both $|\alpha_1\rangle$ and $|\alpha_2\rangle$, we can take the inner product before and after deletion to get $\langle\alpha_0|\alpha_1\rangle^2 = \langle\alpha_0|\alpha_1\rangle\,\langle M_0|M_1\rangle$. Hence $\langle\alpha_0|\alpha_1\rangle = \langle M_0|M_1\rangle$. So by the Lemma, there exists a unitary map U such that $U\,|M_0\rangle = |\alpha_0\rangle\,|N\rangle$, $U\,|M_1\rangle = |\alpha_1\rangle\,|N\rangle$ (for some $|N\rangle$ arbitrary). □

2.6 Distinguishing non-orthogonal states

Theorem: Suppose we receive an unknown quantum state $|\psi\rangle$ which is one of two known, distinct quantum states $|\alpha_0\rangle$, $|\alpha_1\rangle$. Suppose that $|\,\langle\alpha_0|\alpha_1\rangle\,| = \cos(\theta)$. Then the probability P_S of correctly identifying the given state via a quantum process is bounded by

$$P_S \leq \frac{1}{2}(1 + \sin(\theta)).$$

Furthermore, this bound is tight.

Proof: Adjoining an ancilla means we need to distinguish $|\alpha_0\rangle\,|A\rangle$ and $|\alpha_1\rangle\,|A\rangle$, so problem hasn't changed.

Performing a unitary operation U followed by a measurement with projectors π_i is equivalent to measuring with respect to projectors $\tilde{\pi}_i = U^\dagger \pi_i U$ (this is easy to show by comparing the probabilities of getting i in both cases).

Hence WLOG the quantum process we must use to distinguish the states is a single measurement. Let the projectors be π_0 and π_1. Then the probability P_S is

$$P_S = \frac{1}{2}\underbrace{||\pi_0\,|\alpha_0\rangle\,||^2}_{\substack{\text{prob of } 0\\ \text{if } |\alpha_0\rangle}} + \frac{1}{2}\underbrace{||\pi_1\,|\alpha_1\rangle\,||^2}_{\substack{\text{prob of } 1\\ \text{if } |\alpha_1\rangle}} = \frac{1}{2}(\langle\alpha_0|\pi_0|\alpha_0\rangle + \langle\alpha_1|\pi_1|\alpha_1\rangle).$$

Using $\pi_0 + \pi_1 = I$, rewrite this as

$$P_S = \frac{1}{2} + \frac{1}{2}\mathrm{tr}(\pi_0(|\alpha_0\rangle\,\langle\alpha_0| - |\alpha_1\rangle\,\langle\alpha_1|)) = \frac{1}{2} + \frac{1}{2}\mathrm{tr}(\pi_0 D).$$

We note D is Hermitian, so has real eigenvalues, and orthogonal eigenstates for distinct eigenvalues. If $|\beta\rangle$ is orthogonal to $|\alpha_0\rangle$, $|\alpha_1\rangle$, then $D\,|\beta\rangle = 0$ so D has at most 2 non-zero eigenvalues. D has trace 0, so write $\pm\delta$ for the eigenvalues of D and $|p\rangle$ and $|m\rangle$ as the corresponding eigenstates.

Just need to determine δ. Let $|\alpha_\perp\rangle$ be perpendicular to $|\alpha_0\rangle$ and write $|\alpha_1\rangle$ in the basis $\{|\alpha_0\rangle, |\alpha_\perp\rangle\}$ as (c_0, c_1), where $|c_0|^2 + |c_1|^2 = 1$ and $c_0 = \langle\alpha_0|\alpha_1\rangle$ by definition. Thus $|c_0| = \cos(\theta)$ and $|c_1| = \sin(\theta)$. Then

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

Find eigenvalues: $\pm\delta = \pm|c_1| = \pm\sin(\theta)$. Inserting this back into the expression for P_S, we have:

$$P_S = \frac{1}{2} + \frac{1}{2}\sin(\theta)(\langle p|\pi_0|p\rangle - \langle m|\pi_0|m\rangle).$$

To get the bound, use $\langle p|\pi_0|p\rangle = ||\pi_0\,|p\rangle\,||^2 \leq 1$ (since probability) and $\langle m|\pi_0|m\rangle = ||\pi_0\,|m\rangle\,||^2 \geq 0$. The bound is tight when the projector π_0 is the projection onto $|p\rangle$.

Definition: An *unambiguous state discrimination process* is a state discrimination process with three outputs 0, 1 and 'fail', such that if 0 is obtained, the state is definitely $|\alpha_0\rangle$, if 1 is obtained the state is definitely $|\alpha_1\rangle$ and if 'fail' is obtained, then all information about the state has been lost.

2.7 The no-signalling theorem

Theorem: Suppose Alice and Bob have access to subsystems A and B respectively of a joint state $|\phi\rangle_{AB}$. Then no local action by Alice can change the output probability distribution of any local process by Bob.

Proof: Suppose Bob measures with respect to a basis $\{|b\rangle\}$ and Alice does nothing. Write

$$|\phi\rangle_{AB} = \sum_b |\xi_b\rangle_A\,|b\rangle_B\,.$$

On measurement, Bob gets b with probability $\mathrm{Prob}(b) = ||_B\langle b|\,|\xi_b\rangle_A\,|b\rangle_B\,||^2 = \langle\xi_b|\xi_b\rangle\,.$

Instead, assume Alice goes first and she measures with respect to a basis $\{|a\rangle\}$. The probability Alice gets a is $\mathrm{Prob}(a) = ||_A\langle a|\phi\rangle_{AB}\,||^2$, and the post-measurement state of system B is

$$|\eta_a\rangle_B = \frac{_A\langle a|\phi\rangle_{AB}}{\sqrt{\mathrm{Prob}(a)}}.$$

Hence, given that Alice got a, the conditional probability Bob gets b is

$$\mathrm{Prob}(b|a) = \frac{|_B\langle b|\,_A\langle a|\phi\rangle_{AB}\,|^2}{\mathrm{Prob}(a)} = \frac{|_A\langle a|\xi_b\rangle_A\,|^2}{\mathrm{Prob}(a)}.$$

Recall $\mathrm{Prob}(a,b) = \mathrm{Prob}(a)\mathrm{Prob}(b|a)$. Summing $\mathrm{Prob}(a,b)$ over a gives the marginal distribution for b:

$$\mathrm{Prob}(b) = \sum_a |_A\langle a|\xi_b\rangle_A\,|^2 = \sum_a \langle\xi_b|a\rangle\,\langle a|\xi_b\rangle = \langle\xi_b|\xi_b\rangle\,.$$

Hence the probabilities are unchanged. The method easily generalises to incomplete measurements.

Other local operations that Alice and Bob can perform are adjoining an ancilla and performing a unitary operation. Adjoining an ancilla just enlarges the local state space, so doesn't affect anything at the other end.

If Alice performs a local unitary operation before measuring, then all that happens is

$$|\phi\rangle_{AB} = \sum_b |\xi_b\rangle_A\,|b\rangle_B \mapsto |\phi\rangle_{AB} = \sum_b (U_A\,|\xi_b\rangle_A)\,|b\rangle_B\,,$$

which does not affect probabilities since $||U_A\,|\xi_b\rangle_A\,||^2 = ||\,|\xi_b\rangle_A\,||^2$. Hence the proof goes through as before. □

3 Dense coding and teleportation

3.1 Bell states

Definition: The *Bell states* are

$$|\phi^\pm\rangle = \frac{1}{\sqrt{2}}(|00\rangle \pm |11\rangle)\,,$$
$$|\psi^\pm\rangle = \frac{1}{\sqrt{2}}(|01\rangle \pm |10\rangle)\,.$$

A measurement with respect to the Bell states is called a *Bell measurement*.

Property: We can construct all Bell states from $|\phi^+\rangle$ with just a single one-qubit gate applied to the first qubit. We have:

$$|\phi^+\rangle = (I\otimes I)\,|\phi^+\rangle$$
$$|\phi^-\rangle = (Z\otimes I)\,|\phi^+\rangle$$
$$|\psi^+\rangle = (X\otimes I)\,|\phi^+\rangle$$
$$|\psi^-\rangle = (Y\otimes I)\,|\phi^+\rangle\,.$$

3.2 Dense coding

Protocol: Suppose Alice and Bob are distantly separated and share a $|\phi^+\rangle$ state.

If Alice wants to send 00, 01, 10 or 11 to Bob, she applies I, Z, X or Y to her qubit, respectively. She then sends her qubit to Bob who now holds either $|\phi^+\rangle$, $|\phi^-\rangle$, $|\psi^+\rangle$, $|\psi^-\rangle$.

Bob then performs a Bell measurement, and depending on what Alice sent, he gets either ϕ^+, ϕ^-, ψ^+ or ψ^- with certainty. He interprets these results as 00, 01, 10 and 11 respectively.

3.3 Quantum teleportation

Protocol: Suppose that Alice holds an unknown qubit $|\alpha\rangle$ and also shares a $|\phi^+\rangle$ state with Bob. The joint system is thus $|\alpha\rangle_A |\phi^+\rangle_{AB}$. Alice can teleport $|\phi^+\rangle$ to Bob as follows.

1. Alice performs CX_{12} followed by $H \otimes I$ on her two held qubits. She then measures with respect to the computational basis to see a result ij.

2. Alice sends the result ij over a classical channel to Bob.

3. Bob applies the unitary operation $Z^i X^j$ to his held qubit. The final joint state is $|ij\rangle_{AA} |\alpha\rangle_B$.

To see this works, write $|\alpha\rangle = a|0\rangle + b|1\rangle$ and write $|\phi^+\rangle$ out in full.

A circuit diagram for the procedure is:

4 Quantum cryptography

4.1 The BB84 quantum key distribution

Protocol: Suppose Alice and Bob can communicate via a classical and quantum channel and want to publicly establish a secret key. We define the states $|\psi_{00}\rangle = |0\rangle$, $|\psi_{01}\rangle = |+\rangle$, $|\psi_{10}\rangle = |1\rangle$ and $|\psi_{11}\rangle = |-\rangle$.

1. Alice generates two uniformly random bit strings of length m: $X = x_1 x_2 ... x_m$ and $Y = y_1 y_2 ... y_m$. She then sends the states $|\psi_{x_i y_i}\rangle$ (in order) to Bob over the quantum channel.

2. Bob generates a random 'guess' for Y, say $Y' = y_1' y_2' ... y_m'$. He then measures $|\psi_{x_i y_i}\rangle$ in the basis $B_{y_i'}$ where $B_0 = \{|0\rangle, |1\rangle\}$ and $B_1 = \{|+\rangle, |-\rangle\}$. The result is a string $X' = x_1' x_2' ... x_m'$. If $y_i = y_i'$, Bob gets $x_i' = x_i$ with certainty.

3. Alice and Bob publicly announce Y and Y', and then remove all bits in X and X' where their corresponding Y, Y' bits differ. This leaves the shorter strings $\tilde{X}$ and $\tilde{X}'$.

4. If the channel was noisy, or there was eavesdropping, $\tilde{X} \neq \tilde{X}'$. Alice and Bob choose a random sample of bit positions in X and $\tilde{X}'$ and compare them to estimate the bit error rate of the entire string. They discard the bits they announced.

5. (*Non-examinable*): Using the estimated bit error rate, it is possible to use classical techniques of privacy amplification to get shorter strings about which the eavesdropper can have practically no knowledge at all, with very high probability.

4.2 The intercept-resend attack

The intercept-resend attack involves an eavesdropper, Eve, who captures the states Alice sends, measures them in some basis, and sends the results onto Bob.

Theorem: Suppose Eve performs an intercept-resend attack, measuring in the *Breidbart basis*

$$\begin{aligned} |\alpha_0\rangle &= \cos(\pi/8)|0\rangle + \sin(\pi/8)|1\rangle \\ |\alpha_1\rangle &= -\sin(\pi/8)|0\rangle + \cos(\pi/8)|1\rangle . \end{aligned}$$

Eve thinks Alice's bit was a 0 if she gets $|\alpha_0\rangle$ and 1 if she gets $|\alpha_1\rangle$. Then the chance Eve will correctly identify any bit will be $\cos^2(\pi/8)$, and the bit error rate in $\tilde{X}, \tilde{X}'$ will be $1/4$.

Proof: If Alice sent $|0\rangle$, then measurement in the Breidbart basis gives $|\alpha_0\rangle$ with probability $\cos^2(\pi/8)$. Similarly for all other possible states Alice sends.

To estimate the bit error rate, we draw a probability tree diagram:

If Alice sends $|0\rangle$, then Eve gets $|\alpha_0\rangle$ with probability $\cos^2(\pi/8)$ and $|\alpha_1\rangle$ with probability $\sin^2(\pi/8)$. Bob measures with respect to $|0\rangle$, $|1\rangle$ (as we are comparing the tilde strings), and from there we see that he gets the wrong answer with probability

$$2\cos^2(\pi/8)\sin^2(\pi/8) = \frac{1}{4}.$$

Similar for all other qubits Alice could send. □

5 Models of computation

5.1 Computational tasks

Write $B = B_1 = \{0, 1\}$ and B_n for the set of all n-bit strings. Write B^* for the set of all bit-strings of finite length, i.e.

$$B^* = \bigcup_n B_n.$$

Definition: A *computational task* consists of:

(i) an *input bit string* $x = i_1 i_2 ... i_n$ with *size* n;

(ii) a *language* $L \subset B^*$;

(iii) a *decision problem*; that is, given any $x \in B^*$, is $x \in L$?

(iv) a 1-bit *output*, 1 for ‘yes’ and 0 for ‘no’.

5.2 Classical circuit model

The *circuit model* for classical computation is described as follows.

For input $x = i_1 i_2 ... i_n$, extend with extra zeros to get $i_1 ... i_n 00...0$. A basic computational step is a specified Boolean gate: AND, OR or NOT, applied to specified bits in the list $i_1 ... i_n 00...0$. A computation is a prescribed sequence of these steps, C_n, for each input size n, depending only on n and not on x. The output is the value of some designated bit after the final step.

C_n is called a *Boolean circuit* and $(C_1, C_2, C_3 ...)$ is called a *circuit family*.

We can also consider *probabilistic* classical computation. This works as above, but we initially extend $i_1 i_2 ... i_n$ by random bits $r_1 r_2 ... r_k$, and then by zeroes $00...0$. The output of our circuit is then *probabilistic*.

This formulation includes the possibility of random gates; we simply use three-bit control gates with a random bit r_j as the control bit.

5.3 Complexity

Definition: The *time complexity* of a task is the function $T(n)$, which is the number of gates in the circuit C_n.

Definition: The *space complexity* of a task is the function $S(n)$, which is the length of the extended input string $i_1 ... i_n r_1 ... r_k 00...0$.

Definition: If there exists n_0, c and k such that $T(n) < cn^k$ for all $n \geq n_0$, then we write $T(n) = O(\text{poly}(n))$. The task is called *polynomial time*.

5.4 Complexity classes

We define the following classes:

Definition: The complexity class P is the class of decision problems having deterministic polynomial time algorithms.

Definition: The complexity class BPP is the class of decision problems having probabilistic polynomial time algorithms such that for every input x, the probability the answer is correct is greater than $2/3$. (BPP stands for bounded-error probabilistic polynomial time.)

5.5 Quantum circuit model

The *quantum circuit model* for quantum computation is described as follows.

For input $x = i_1 ... i_n \in B_n$, start with qubits $|i_1\rangle |i_2\rangle ... |i_n\rangle$ in the computational basis, and extend to $|i_1\rangle |i_2\rangle ... |i_n\rangle |0\rangle |0\rangle ... |0\rangle$. Random qubits are never needed, since quantum measurements provide this.

The computational steps are quantum gates on designated qubits. The output is the outcome of a quantum measurement in the basis $\{|0\rangle, |1\rangle\}$ on specified qubits. We need a fixed measurement, otherwise you could hide the solution to difficult problems in the choice of measurement.

A computation is a prescribed sequence of quantum gates, C_n, for each input size n, depending only on n and not on x.

5.6 Universal sets of gates

A general unitary gate U on n qubits has continuous parameters, whereas classical gates are discrete. So no finite quantum gate set can be universal - the number of finite circuits from a finite set of gates is only countable infinite, hence we cannot get a continuous infinite of U. Instead, we want:

Definition: A gate set G is called *approximately universal* if for all $\epsilon > 0$ and any gate W on n qubits, there is a circuit $\tilde{W}$ of gates from G such that $||\tilde{W} - W|| < \epsilon$ (where $|| \cdot ||$ denotes the operator norm).

Theorem (Solovay-Kitaev): For each fixed n, there is a polynomial p such that for all gates W on n qubits, we can choose $\tilde{W}$ obeying $||W - \tilde{W}|| < \epsilon$ with size bounded by $p(\log(1/\epsilon))$.

Proof: Non-examinable. □

5.7 The class BQP

Definition: The class BQP is the class of all decision problems that can be solved with polynomial-sized quantum circuits, getting the answer correct with probability greater than $2/3$.

Theorem: Suppose we are only allowed an approximately universal set of quantum gates. Then BQP is independent of approximately universal gate set chosen.

Proof: Suppose that $||U_1 - V_1|| < \epsilon$ and $||U_2 - V_2|| < \epsilon$. Then we have:

$$\begin{aligned} ||U_2U_1 - V_2V_1|| &= ||U_2(U_1 - V_1) + V_1(U_2 - V_2)|| \\ &\leq ||U_2|| \cdot ||U_1 - V_1|| + ||V_1|| \cdot ||U_2 - V_2|| < 2\epsilon, \end{aligned}$$

using $||AB|| \leq ||A|| \cdot ||B||$ and $||U|| = 1$ for U unitary. By induction, if $||U_i - V_i|| < \epsilon$ for $i = 1, ..., n$ we have:

$$||U_n...U_1 - V_n...V_1|| < n\epsilon.$$

Now suppose that G and H are approximately universal gate sets. Let D be in BQP with circuit from elements G. Then for each input of size n, there is a quantum circuit C_n of size $\mathrm{pol}(n)$ comprising gates from G, returning the correct answer for the decision problem with probability at least $2/3$.

Write the circuit C_n as the sequence U_{r_1}, U_{r_2} ... $U_{r_{\max}}$, taken from $G = \{U_1, U_2, ...U_n\}$. Since C_n has size $\mathrm{poly}(n)$, we know $r_{\max}(n) = \mathrm{poly}(n)$.

Now, for each U_i, there exists a circuit $\mathrm{circ}_i(H)$ formed from gates in H such that

$$||U_i - \mathrm{circ}_i(H)|| < \frac{\epsilon}{r_{\max}},$$

where $\mathrm{circ}_i(H)$ uses $\mathrm{poly}(r_{\max}/\epsilon)$ gates, by the Solovay-Kitaev theorem. Hence by the earlier work,

$$||U_{r_1}...U_{r_{\max}} - \mathrm{circ}_{r_1}(H)...\mathrm{circ}_{r_{\max}}(H)|| < \epsilon.$$

Thus $\mathrm{circ}_{r_1}(H)...\mathrm{circ}_{r_{\max}}(H)$ approximates the original circuit to arbitrary high accuracy, hence gives the same result with arbitrarily high accuracy. The number of gates needed in this new circuit is

$$r_{\max} \cdot \mathrm{poly}(r_{\max}/\epsilon) = \mathrm{poly}(n),$$

since $r_{\max} = \mathrm{poly}(n)$. □

6 Quantum algorithms

6.1 Boolean functions as unitary gates

Let $f : B_m \to B_n$ and let $x \mapsto y = f(x)$. Consider $\tilde{f} : B_{m+n} \to B_{m+n}$ define by $(x, y) \mapsto (x, y \oplus f(x))$. Then $\tilde{f}$ is clearly invertible (it's self-inverse), and hence is a permutation of $m + n$-bit strings.

Definition: For any $f : B_m \to B_n$, define the quantum operation

$$U_f \underbrace{|x\rangle}_{m \text{ bits}} \underbrace{|y\rangle}_{n \text{ bits}} = |x\rangle \, |y \oplus f(x)\rangle \,,$$

on the computational basis and extend by linearity. This must be unitary, since by the above it just permutes $m + n$-bit strings, so its columns are orthonormal.

6.2 Promise problems

Instead of an input $x = i_1i_2...i_n \in B_n$ for a computational task, assume we have an *oracle* U_f that computes some Boolean function $f : B_m \to B_n$. Each use of the oracle counts as one computational step.

Definition: A *promise problem* is defined as follows. Let f be a function with a *promise* on its form, e.g. f is zero for half of its outputs. Then we must determine some property of f either with certainty or good probability based only on the output of the oracle.

Definition: The *query complexity* of a promise problem is the number of uses of the oracle needed to solve the task. The *total time complexity* is the total size of the circuit used to solve the task, where using the oracle counts as one step.

6.3 The Deutsch-Jozsa algorithm

Problem: The *balanced versus constant problem* is a promise problem with:

(i) $f : B_n \to B_1$ is either constant ($f(x) \equiv 1$ or $f(x) \equiv 0$), or *balanced*, i.e. exactly half of the 2^n possible $f(x)$ values are 0 and the other half are 1;

(ii) we must determine with certainty whether f is balanced or constant.

Theorem: Classically, the balanced versus constant problem has query complexity at least $2^n/2 + 1$.

Proof: Suppose there is an algorithm with $2^n/2$ queries. We can reply 0 to all of these queries, and then the $(2^n/2 + 1)$th query could be answer 0 or 1. □

Theorem: Quantumly, the balanced versus constant problem has query complexity 1.

We show this is true via the *Deutsch-Jozsa algorithm*.

Protocol: The Deutsch-Jozsa algorithm operates as follows:

1. Set the output register to
$$\frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) = HX|0\rangle =: |\alpha\rangle .$$

2. Note that for any $|x\rangle$, $x \in B_n$, we have
$$|x\rangle\left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) \underbrace{\mapsto}_{U_f} |x\rangle\left(\frac{|f(x)\rangle - |1 \oplus f(x)\rangle}{\sqrt{2}}\right)$$
$$= \begin{cases} |x\rangle\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right) & \text{if } f(x) = 0, \\ |x\rangle\left(\frac{|1\rangle-|0\rangle}{\sqrt{2}}\right) & \text{if } f(x) = 1. \end{cases}$$
Hence $|x\rangle|\alpha\rangle \mapsto (-1)^{f(x)}|x\rangle|\alpha\rangle$.

3. Do the above in superposition over all $|x\rangle$:
$$\frac{1}{\sqrt{2^n}}\sum_{x\in B_n}|x\rangle|\alpha\rangle \underbrace{\mapsto}_{U_f} \frac{1}{\sqrt{2^n}}\left(\sum_{x\in B_n}(-1)^{f(x)}|x\rangle\right)|\alpha\rangle .$$
Discard $|\alpha\rangle$ as it is no longer needed.

4. Now notice that for f constant we get the state
$$|\xi_{\text{constant}}\rangle = \pm\frac{1}{\sqrt{2^n}}\sum_{x\in B_n}|x\rangle ,$$
and for f balanced we get the state $|\xi_{\text{balanced}}\rangle$, for which exactly half the signs are $\pm$. These states are orthogonal. We also note that
$$|00...0\rangle \underbrace{\mapsto}_{H\otimes H\otimes ...\otimes H} \frac{1}{\sqrt{2^n}}\sum_{x\in B_n}|x\rangle = \pm|\xi_{\text{constant}}\rangle .$$

5. Hence apply the rotation $H\otimes H\otimes ...\otimes H$ (which is self-inverse) to the state given by the above steps. Write $|\eta_f\rangle = H\otimes H\otimes ...\otimes H|\xi_f\rangle$. Then $|\eta_{\text{constant}}\rangle = |00...0\rangle$ and $|\eta_{\text{balanced}}\rangle$ is a sum of states not including $|00...0\rangle$.

   Thus measure in the computational basis. If we get all zeroes, we know f was constant. Otherwise, f was balanced.

The circuit diagram for the Deutsch-Jozsa algorithm is given below.

6.4 The quantum Fourier transform

Definition: The *quantum Fourier transform* modulo N, written QFT_N is a unitary operator on an N-dimensional states space with basis $\{|0\rangle, ...|N-1\rangle\}$ labelled by $\mathbb{Z}_N$ given by
$$QFT_N|a\rangle = \frac{1}{\sqrt{N}}\sum_{b=0}^{N-1}e^{2\pi iab/N}|b\rangle$$
on the basis states, and extended by linearity.

6.5 The periodicity-finding algorithm

Problem: Let $f : \mathbb{Z}_N \to \mathbb{Z}_M$. Define the *periodicity problem* as the promise problem with:

(i) we promise that f is periodic, i.e. $f(x+r) = f(x)$ for some least $r > 0$, for all x, and that f is one-to-one in each period;

(ii) the problem is to determine r with any constant level of probability $1-\epsilon$, $\epsilon > 0$.

Protocol: The *periodicity determination algorithm* is described as follows:

1. Make the uniform superposition
$$\frac{1}{\sqrt{N}}\sum_{x=0}^{N-1}|x\rangle .$$

2. Query the oracle using this state in the first register to get
$$|f\rangle = \frac{1}{\sqrt{N}}\sum_{x\in B_n}|x\rangle|f(x)\rangle .$$

3. Measure the second register with respect to $\{|0\rangle, |1\rangle, ...|M-1\rangle\}$. We see some value $y \in \mathbb{Z}_M$. Let x_0 be the least x_0 with $f(x_0) = y$. Then the post-measurement state is
$$|\text{per}\rangle = \frac{1}{\sqrt{A}}\sum_{j=0}^{A-1}|x_0 + jr\rangle ,$$
where $Ar = N$.

4. Apply QFT_N to $|\text{per}\rangle$; the result is

$$QFT_N\,|\text{per}\rangle = \frac{1}{\sqrt{N}} \sum_{j=0}^{A-1} \left(\frac{1}{\sqrt{A}} \sum_{y=0}^{N-1} \omega^{(x_0+jr)y}\,|y\rangle \right),$$

   where $\omega = e^{2\pi i/N}$ is the Nth root of unity. Switch the order of summation to get:

$$QFT_N\,|\text{per}\rangle = \frac{1}{\sqrt{NA}} \left(\sum_{y=0}^{N-1} \omega^{x_0 y} \left(\sum_{j=0}^{A-1} \omega^{jry} \right) |y\rangle \right),$$

   and after performing the interior sum (which is non-zero iff ry is a multiple of N), we are left with

$$QFT_N\,|\text{per}\rangle = \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} \omega^{x_0 kN/r}\,|\frac{kN}{r}\rangle\,.$$

5. Now measure with respect to $\{|0\rangle, |1\rangle \dots |N-1\rangle\}$. We will obtain some c defined by

$$c = \frac{k_0 N}{r},$$

   with $0 \leq k_0 \leq r-1$. Rewrite as

$$\frac{c}{N} = \frac{k_0}{r}.$$

   If k_0 and r are coprime, then r is the denominator of c/N when it is cancelled down (c and N are both known). The chance of this happening is given by...

6. Use the *coprimality theorem* from number theory. This theorem states that the number of integers less than r which are coprime to r grows as $O(r/\log(\log(r)))$.

   Thus, since k_0 is uniformly random on $0 \leq k_0 \leq r-1$, the chance it is coprime to r is $O(1/\log(\log(r))) > O(1/\log(\log(N))$. Hence repeat $O(\log(\log(N))$ times and check the answer each time.

---

7 Quantum search problems

7.1 The unstructured search problem

Problem: We are given an unstructured database with $N = 2^n$ items containing a *unique* good item. The problem is to find the good item with constant probability $1-\epsilon$.

Classically, it is clear that $O(N)$ operations are necessary and sufficient, where the constant depends on ϵ.

Quantumly, we represent the database by an oracle for a Boolean function $f : B_n \to B_1$ with a promise that there is a unique x_0 such that $f(x_0) = 1$ and $f(x) = 0$ for all $x \neq x_0$. The problem is to find x_0.

7.2 Grover's algorithm

First we replace the quantum oracle $U_f\,|x\rangle\,|y\rangle = |x\rangle\,|y \oplus f(x)\rangle$ with an oracle I_{x_0} on n qubits defined by

$$I_{x_0}\,|x\rangle = \begin{cases} |x\rangle & \text{if } x \neq x_0, \\ -\,|x\rangle & \text{if } x = x_0. \end{cases}$$

We can implement I_{x_0} with one query to U_f by supplying U_f with the state $|x\rangle\,|-\rangle$ (see circuit diagram below).

Clearly, a formula for I_{x_0} is $I_{x_0} = I - 2\,|x_0\rangle\,\langle x_0|$. More generally, we define:

Definition: Define the *reflection in the mirror hyperplane* perpendicular to $|\alpha\rangle$ by $I_{|\alpha\rangle} = I - 2\,|\alpha\rangle\,\langle\alpha|$. For computational basis states, write $I_{|x_0\rangle} = I_{x_0}$ and $I_{|00\dots0\rangle} = I_0$, say.

Protocol: *Grover's algorithm* works as follows.

1. Define the Grover iteration operator on n qubits by $Q = -H_n I_0 H_n I_{x_0}$ where $H_n = H \otimes H \otimes \dots \otimes H$. Let $\mathcal{P}(x_0)$ be the plane spanned by $|x_0\rangle$ and the uniform superposition states $|\psi_0\rangle$. We will show in the next section that Q is a rotation through an angle 2α, where

$$\sin(\alpha) = \frac{1}{\sqrt{N}},$$

   when Q acts in the plane $\mathcal{P}(x_0)$. Outside of the plane $\mathcal{P}(x_0)$, Q acts as $-I$.

2. Repeatedly apply Q to $|\psi_0\rangle$ to rotate it near to $|x_0\rangle$. Let $\cos(\beta) = \langle x_0|\psi_0\rangle = 1/\sqrt{N}$, which is independent of the unknown x_0. Then the number of iterations required is

$$\frac{\beta}{2\alpha} = \frac{\arccos(1/\sqrt{N})}{2\text{arcsin}(1/\sqrt{N})}.$$

   For N very large, $\beta \approx \pi/2$, and $\alpha \approx 1/\sqrt{N}$. Hence the number of iterations required is

$$\frac{\beta}{2\alpha} \approx \frac{\pi}{4}\sqrt{N}.$$

   In particular, this shows that the query complexity is $O(\sqrt{N})$.

7.3 The action of Q: proofs

We now prove the claims we made above about the action of the Grover iteration operator Q.

Theorem: Q preserves the plane $\mathcal{P}(x_0)$.

Proof: First we note that for any unitary U,

$$UI_{|\psi\rangle}U^\dagger = UIU^\dagger - 2U|\psi\rangle\langle\psi|U^\dagger = I_U|\psi\rangle.$$

Since $H = H^\dagger$, we can thus write the Grover iteration operator as $Q = -I_{|\psi_0\rangle}I_{|x_0\rangle}$, where $|\psi_0\rangle = H_n\,|00...0\rangle$ is the uniform superposition. So Q is the composition of two reflections.

Now note that for any $|\xi\rangle$, $|\psi\rangle$, we have $I_{|\psi\rangle}\,|\xi\rangle = |\xi\rangle - 2\,|\psi\rangle\,\langle\psi|\xi\rangle$. So $I_{|\psi\rangle}$ modifies $|\xi\rangle$ by a multiple of $|\psi\rangle$. Hence Q modifies $|\xi\rangle$ by a multiplies of $|x_0\rangle$ and then a multiple of $|\psi_0\rangle$. So Q preserves the plane $\mathcal{P}(x_0)$. □

Theorem: Q acts as an anticlockwise rotation by 2α in the plane $\mathcal{P}(x_0)$, where $\sin(\alpha) = 1/\sqrt{N}$.

Proof: We use two facts from 2D Euclidean geometry:

1. We have $-I_{|v\rangle} = I_{|v^\perp\rangle}$ in a plane. This follows by writing $u = a\,|v\rangle + b\,|v^\perp\rangle$; then $I_{|v\rangle}$ reverses the sign of a and $I_{|v^\perp\rangle}$ reverses the sign of b.
2. Reflection in a mirror lines M_1, then in a mirror line M_2, is the same thing as rotation through the angle 2θ, where θ is the angle between the mirror lines.

Then $Q = -I_{|\psi_0\rangle}I_{|x_0\rangle} = I_{|\psi_0\rangle}I_{|x_0^\perp\rangle}$. So Q is reflection in the mirror line along $|x_0\rangle$ followed by a reflection in the mirror line along $|\psi_0^\perp\rangle$, which is equivalent to a rotation through twice the angle between $|x_0\rangle$ and $|\psi_0^\perp\rangle$.

From the below diagram, this angle is 2α, where

$$\sin(\alpha) = \langle\psi_0|x_0\rangle = \frac{1}{\sqrt{N}}. \quad \square$$

Theorem: In $\mathcal{P}(x_0)^\perp$, $Q = -I$.

Proof: Let $|\xi\rangle \in \mathcal{P}(x_0)^\perp$. Then $|\xi\rangle$ is perpendicular to both $|x_0\rangle$ and $|\psi_0\rangle$. So $Q = -I_{|\psi_0\rangle}I_{|x_0\rangle}$ must act as $-I \cdot I = -I$. □