MASSACHUSETTS INSTITUE OF TECHNOLOGY

Department of Physics, EECS, and Department of Applied Math

MIT 6.443J / 8.371J / 18.409 / MAS.865

Quantum Information Science

March 16, 2006

Problem Set #4

(due in class, 06-Apr-06)

Lecture Topics (3/16, 3/21, 3/23, 3/4): quantum algorithms; entanglement; typical subspaces

Recommended Reading: Nielsen and Chuang, Sections 5.4 and 12.1 - 12.4

Problems:

P1: (**Quantum factoring as a feedback process**) Shor's quantum factoring algorithm was independently (re-)discovered by Alexi Kitaev, in Russia. Kitaev's formulation allows for an interesting observation of how quantum factoring can be viewed as a feedback process, involving quantum control and optimal estimation, as we explore in this problem.

Let N be a composite number we wish to factor, and choose some y coprime to N. Define the unitary transform U to be

$$U|m\rangle = |my \bmod N\rangle\,, \tag{1}$$

where the state lives in an N dimensional Hilbert space (for example, of $n = \lceil \log_2 N \rceil$ qubits).

(a) Show that the eigenstates of U are

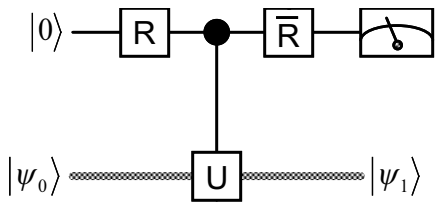
$$|\lambda_k\rangle = \frac{1}{\sqrt{r}} \sum_{l=0}^{r-1} e^{2\pi i l \phi_k} |y^l \bmod N\rangle\,, \tag{2}$$

where $\phi_k = k/r$, and r is the order of y, i.e. the smallest integer such that $y^r \bmod N = 1$. Also show that

$$U|\lambda_k\rangle = e^{-2\pi i \phi_k} |\lambda_k\rangle\,. \tag{3}$$

It is a fact from number theory that once r is known, with probability greater than 50%, a factor of N can be found. Fatoring N is thus equivalent to finding r. The calculation here indicates that finding r is equivalent to finding an eigenvalue of U. We consider next a circuit by which this may be accomplished.

(b) Consider this quantum circuit:



This is one step of a Kitaev factoring algorithm, in which the top wire carries an ancilla qubit, and the bottom (thick grey) wire carries the main n qubit state. Let the initial input state into the controlled-U gate be $|\psi_0\rangle = |\lambda_k\rangle$. The R gate acting on the ancilla qubit is the Hadamard transform

$$R = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} . \tag{4}$$

Following the initial state through the circuit, and show that the ancilla is measured to be 0 with probability

$$p_0 = \cos^2(\pi\phi_k) , \tag{5}$$

and independent of the measurement result, the final state $|\psi_1\rangle = |\lambda_k\rangle$ for this example. Note that therefore, it may be reused.

The interesting observation is that after repeated trials, we are able to estimate p_0 and thus determine the eigenvalue ϕ_k. If we may repeat the procedure with powers of U, i.e., U^{2^j}, then we may estimate ϕ_k efficiently (in a number of trials polynomial in $\log N$).

(c) Unfortunately, the above scheme would not be very useful if we already knew enough to be able to generate an eigenstate at the outset to feed into the system! What happens if we do not start with an eigenstate, and instead have the input state

$$|\psi_0\rangle = |1\rangle = \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} |\lambda_k\rangle , \tag{6}$$

which is an equally weighted superposition of eigenstates?

Note that $|1\rangle$ is simple to generate. It is thus convenient to define the ancilla state

$$|\eta_k\rangle = \frac{1}{2} \left[(1 + e^{-2\pi i \phi_k})|0\rangle + (1 - e^{-2\pi i \phi_k})|1\rangle \right] . \tag{7}$$

Compute the output state after one trial, and show that it is given by

$$|\psi_1\rangle = \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} |\eta_k\rangle|\lambda_k\rangle . \tag{8}$$

(d) Compute the output state after t trials, $|\psi_t\rangle$, where the output of each trial is fed back as the input to the next iteration of the circuit.

(e) Each measurement of an ancilla qubit $|\eta_k\rangle$ gives either 0 or 1, and by symmetry the order of results doesn't matter, so the only important quantity is the total number of zeros measured, n_0 out of the t trials. Let us try to understand what a-posteriori state results for a given n_0 by considering the joint probability distribution $p(k, n_0)$, where $n_1 = t - n_0$ is the number of one's which resulted.

This distribution is what one would obtain if a projective measurement were carried out on the $|\psi\rangle$ state in the $|\lambda_k\rangle$ basis. Give an expression for $p(k, n_0)$.

(f) The interesting thing is that to a very good approximation, $p(n_0) \approx 1/2$, and so the *conditional* probability for getting some k, given n_0, is

$$p(k|n_0) \approx \frac{2}{r}\binom{t}{n_0} \cos^{2n_0}\left[\frac{\pi k}{r}\right] \sin^{2n_1}\left[\frac{\pi k}{r}\right] . \tag{9}$$

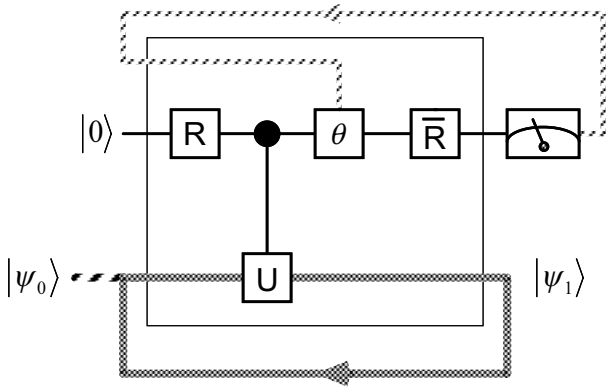
Verify this expression and plot the distribution; show that it has two peaks with widths which decrease as $\mathcal{O}(1/t)$.

This shows that with each successive increase of t, the state $|\psi_t\rangle$ increasingly converges into a superposition of two eigenstates of U, and moreover, knowledge of n_0 increasingly determines k.

(g) Helpful insight is gained by a numerical example. Try running this algorithm for $N = 143$, $y = 5$, and $r = 20$, and plot $p(k|n_0)$ for a sequence of values of t.

The critical quantity is the convergence rate of our knowledge of the eigenvalue.

(h) One of the inefficiencies of the scheme derived above is the fact that even after the system has converged into a perfect eigenstate, the measurement result from each iteration can still vary quite randomly. That is, once k has converged to a fixed value, we still obtain a zero with probability $\cos^2(\pi\phi_k)$, which can be significant. Ideally, we would like arrange the output distribution so as to maximize the mutual information between each measurement and the unknown eigenvalue ϕ_k. We can take a step in that direction by modifying the above quantum circuit to become:



Note that an additional component is added in the control path, a θ box, which implements the transform

$$\begin{bmatrix} 1 & 0 \\ 0 & e^{2\pi i\theta} \end{bmatrix} \tag{10}$$

on the control bit, where θ is a classically determined angle, provided by a classical control apparatus. Operation of this circuit is very similar to the previous scenario: an initial state $|\psi_0\rangle$ is prepared and fed into the lower loop. This state will continually circulate, and eventually converges into an eigenstate of the system.

The difference now is that depending on the accumulated sequence of measurement results, we can estimate the state of the system and change θ accordingly to bias future measurement outputs so that they have low entropy.

Analyze how this circuit works in detail, by following the state around one iteration of the loop, assuming it starts initially in with an eigenstate input, $|0\rangle|\lambda_k\rangle$. Show that if we choose $\theta = \phi_k$ then $p_0 = 1$. This is good, because then a measurement of 1, being an unlikely event (if our estimator is correct), would give us a relatively large amount of information about the error $|\theta - \phi_k|$.

(i) Coming up with a good estimator model is nontrivial, especially since the system changes non-deterministically each time a measurement is performed. In particular, when feedback is performed, Eq.(9) is no longer a good estimate of the state, since the Hamiltonian now becomes a function of the record of prior measurement results!

Construct an algorithm for updating θ based on the measurement record obtained, using the idea that $\{f_0, f_1, \ldots\}$ is a model (series of functions of ϕ) of what we expect the system's conditional probability distribution for ϕ_k to look like, approximating $p(n_0) \approx 1/2$ (this is not very good at late times). Append new multiplicative terms to this function after each iteration, depending on the measurement results obtained.

Evaluate your algorithm, for example, using a trial run with parameters $N = 143$, $y = 5$, $r = 20$.

(j) [optional] The procedure suggested in the last step is somewhat unstable in practice, because the estimator for θ is very bad at early times. An improved solution would be to estimate θ based on a running average of ϕ, or from the frequency of occurance of 0. Ideally, you would want something like a Kalman filter. Try to derive an optimal estimation procedure for this feedback based quantum factoring algorithm, and compare your result with Shor's algorithm. What θ update rule would you need to be able to obtain the quantum Fourier transform circuit?

P2: (**Quantum search by continuous-time simulation**) Grover's quantum search algorithm can also be constructed as a continuous time quantum algorithm involving the simulation of a particular Hamiltonian. Consider the Hamiltonian

$$H = |x\rangle\langle x| + |\psi\rangle\langle\psi|\,, \tag{11}$$

where $|\psi\rangle$ is the initial state of the system, and $|x\rangle$ is the solution state (with an unknown x). Suppose you are given an oracle which you can call, which implements $U_x(\Delta t) = \exp(-i|x\rangle\langle x|\Delta t)$ for a specified value of Δt (your choice). Moreover, you also have available $U_\psi(\Delta t) = \exp(-i|\psi\rangle\langle\psi|\Delta t)$ (you can perform this yourself, since $|\psi\rangle$ is known).

(a) Show that $U(\Delta t) = U_\psi(\Delta t)U_x(\Delta t)$ can be expressed as (up to an unimportant global phase factor)

$$\begin{aligned} U(\Delta t) = &\left(\cos^2\left(\frac{\Delta t}{2}\right) - \sin^2\left(\frac{\Delta t}{2}\right)\vec{\psi}\cdot\hat{z}\right) I \\ &-2i\sin\left(\frac{\Delta t}{2}\right)\left(\cos\left(\frac{\Delta t}{2}\right)\frac{\vec{\psi}+\hat{z}}{2} + \sin\left(\frac{\Delta t}{2}\right)\frac{\vec{\psi}\times\hat{z}}{2}\right)\cdot\vec{\sigma}\,, \end{aligned} \tag{12}$$

using the Bloch vector representation, $|x\rangle\langle x| = (I+Z)/2 = (I+\hat{z}\cdot\vec{\sigma})/2$, with $\hat{z} \equiv (0,0,1)$ being the unit vector in the z direction, and $|\psi\rangle\langle\psi| = (I+\vec{\psi}\cdot\vec{\sigma})/2$. We may choose $\vec{\psi} = (2\alpha\beta, 0, (\alpha^2-\beta^2))$ by recognizing that H acts only in a two-dimensional space spanned by $|x\rangle$ and $|y\rangle = |\psi\rangle - \langle x|\psi\rangle|x\rangle$ (un-normalized).

(b) Show that by choosing $\Delta t = \pi$, the operations U_ψ and U_x are identical to the operations used in the quantum simulation algorithm.

(c) How can Δt be chosen such that we obtain a quantum search algorithm which uses $O(\sqrt{N})$ queries, and for which the final state is $|x\rangle$ *exactly*, that is, the algorithm works with probability 1, rather than with some smaller probability?

P3: (**Measures of pure state entanglement**) Entanglement is *a property of a composite quantum system that cannot be changed by local operations and classical communications.* How do we mathematically determine if a given state is entangled or not? And if a state is entangled, how entangled is it?

(a) Recall that by virtue of the Schmidt decomposition (book, page 109), a pure state $|\psi\rangle$ in the Hilbert space of systems A and B can be written as

$$|\psi\rangle = \sum_k \lambda_k |k_A\rangle |k_B\rangle \,, \tag{13}$$

where $|k_A\rangle$ and $|k_B\rangle$ are orthonormal states of systems A and B, respectively, and $\sum_k \lambda_k^2 = 1$. The *Schmidt number* is the number of nonzero λ_k. Prove that $|\psi\rangle$ is a product state, that is $|\psi\rangle = |\psi_A\rangle|\psi_B\rangle$, if and only if the Schmidt number of $|\psi\rangle$ is 1.

(b) Prove that the Schmidt number cannot be changed by local unitary transforms and classical communication. The Schmidt number is one measure of how entangled a state is.

(c) Give the Schmidt numbers for each of the following states:

$$|\phi_1\rangle = \frac{|00\rangle + |11\rangle + |22\rangle}{\sqrt{3}} \tag{14}$$

$$|\phi_2\rangle = \frac{|00\rangle + |01\rangle + |10\rangle + |11\rangle}{2} \tag{15}$$

$$|\phi_3\rangle = \frac{|00\rangle + |01\rangle + |10\rangle - |11\rangle}{2} \tag{16}$$

$$|\phi_4\rangle = \frac{|00\rangle + |01\rangle + |11\rangle}{\sqrt{3}} \,. \tag{17}$$

P4: (**Typical sequences (computational)**) Let $X_1, X_2, X_3, \ldots$ be an i.i.d sequence of random variables $\mathcal{X}$ with range $\{a, b, c\}$ and probability mass function $p(a) = 0.8$, $p(b) = p(c) = 0.1$.

(a) Calculate the entropy rate $H(\mathcal{X}) = H(X_1)$.

(b) The set of ϵ-typical sequences of length n, $A_\epsilon^{(n)}$, consists of sequences for which the number n_a of occurrences of the value a is close to the expected value $0.8n$. Find inequalities that tell when a sequence is ϵ-typical in terms of ϵ, n, and n_a.

(c) Let $A_{0.1}^{(100)}$ be the set of 0.1-typical sequences of length 100. Compute $\Pr(A_{0.1}^{(100)})$.

(d) Compute $|A_{0.1}^{(100)}|$, the number of typical sequences, and the number of bits needed to represent all typical sequences.