Ph219/CS219: Quantum Computation

Winter 2006

Solutions to Problem Set 6

Problem 6.1

a) As we saw in the lecture, we can estimate the eigenvalue λ of U by repeating a procedure in which we prepare the state $|+\rangle \otimes |\lambda\rangle$ where $U|\lambda\rangle = \lambda|\lambda\rangle$, we apply $\Lambda(U)$ and then measure the first qubit in the eigenbasis of σ_{x} or σ_{y} [1]. Each time we repeat this procedure, a measurement of σ_{x} (resp. σ_{y}) gives us information about the real (resp. imaginary) part of λ, so that after k repetitions we can determine λ to accuracy $O(1/\sqrt{k})$.

Now, let the input state to $\Lambda(U)$ be

$$\rho_{\mathrm{in}} = |+\rangle\langle+| \otimes \frac{I_n}{2^n} = \frac{1}{2^{n+1}} \begin{bmatrix} I_n & I_n \\ I_n & I_n \end{bmatrix} , \tag{1}$$

where I_n is the $2^n \times 2^n$ identity matrix. That is, instead of preparing an eigenstate $|\lambda\rangle$ of U on the target register, we prepare the maximally mixed state $I_n/2^n$. After $\Lambda(U)$ acts on ρ_{in}, the state becomes

$$\begin{aligned} \rho_{\mathrm{out}} &= \Lambda(U)\left(|+\rangle\langle+| \otimes I/2^n\right)\Lambda(U)^\dagger \\ &= \tfrac{1}{2^{n+1}} \begin{bmatrix} I_n & 0 \\ 0 & U \end{bmatrix} \begin{bmatrix} I_n & I_n \\ I_n & I_n \end{bmatrix} \begin{bmatrix} I_n & 0 \\ 0 & U^\dagger \end{bmatrix} = \tfrac{1}{2^{n+1}} \begin{bmatrix} I_n & U^\dagger \\ U & I_n \end{bmatrix} . \end{aligned} \tag{2}$$

Suppose we subsequently measure the first qubit along the eigenbasis of σ_{x}. The projectors onto the $|+\rangle$ and $|-\rangle$ eigenstates are

$$\Pi_{\sigma_{\mathrm{x}},\pm 1} = |\pm\rangle\langle\pm| \otimes I_n = \frac{1}{2} \begin{bmatrix} I_n & \pm I_n \\ \pm I_n & I_n \end{bmatrix} , \tag{3}$$

so that the probability of obtaining the eigenvalue ±1 is

$$\begin{aligned} P(\sigma_{\mathrm{x}}, \pm 1) &= \mathrm{Tr}\left(\Pi_{\mathrm{x},\pm 1}\rho_{\mathrm{out}}\right) \\ &= \tfrac{1}{2^{n+2}}\mathrm{Tr} \begin{bmatrix} I_n \pm U & \pm I_n + U^\dagger \\ \pm I_n + U & I_n \pm U^\dagger \end{bmatrix} \\ &= \tfrac{1}{2}\left(1 \pm \tfrac{Re(\mathrm{Tr}U)}{2^n}\right) . \end{aligned} \tag{4}$$

[1] We use the notation $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ and $|\pm i\rangle = (|0\rangle \pm i|1\rangle)/\sqrt{2}$ for the eigenvectors of σ_{x} and σ_{y} respectively.

If we instead measure the first qubit in the eigenbasis of σ_{y}, then the projectors corresponding to the ±1 outcomes take the form

$$\Pi_{\sigma_{\mathrm{y}},\pm 1} = |\pm i\rangle\langle\pm i| \otimes I_n = \frac{1}{2}\left[\begin{array}{cc} I_n & \mp i I_n \\ \pm i I_n & I_n \end{array}\right] , \tag{5}$$

and we can compute $P(\sigma_{\mathrm{y}}, \pm 1) = \frac{1}{2}\left(1 \pm \frac{Im(\mathrm{Tr}U)}{2^n}\right)$.

Suppose we repeat the procedure in which the ancilla qubit is measured in the σ_{x} eigenbasis k times and we average over all measurement outcomes. Since each time we repeat the procedure new input states are prepared, there are no correlations between the measurement outcomes and so our problem is equivalent to estimating the bias of a coin if we toss it k times. For large k, we can use the central limit theorem to show that the average over the k trials will be a random variable following a Gaussian distribution with average $P(\sigma_{\mathrm{x}}, +1)(+1)+P(\sigma_{\mathrm{x}}, -1)(-1) = \mathrm{Re}(\mathrm{Tr}U)/2^n$ and standard deviation which decreases with k as $1/\sqrt{k}$. Therefore our accuracy in estimating the normalized trace of U after k trials is $O(1/\sqrt{k})$.

b) From part (a), we can estimate $\mathrm{Tr}U/2^n$ to accuracy $\delta = O(1/\sqrt{k})$ by repeating k times the procedure (i) prepare $|+\rangle\langle+| \otimes I/2^n$, (ii) apply the unitary gate $\Lambda(U)$, and (iii) measure the control qubit in the eigenbasis of σ_{x} or σ_{y}. Therefore to achieve some fixed accuracy δ_0 we need a number of repetitions $k = \Omega(1/\delta_0^2)$ which does not depend on the number of qubits n.

In each repetition, steps (i) and (iii) involve $O(n)$ operations on a quantum computer. It remains to estimate the complexity of step (ii), i.e. the complexity of simulating $\Lambda(U)$ given that U can be realized by a $poly(n)$ size quantum circuit using gates from some universal gate set. Since $\Lambda(U)$ can be executed by controlling each gate in the circuit realizing U from the same ancilla qubit, there are $poly(n)$ gates in $\Lambda(U)$ that we must approximate. Since errors add linearly, we require each gate in $\Lambda(U)$ to be approximated to accuracy ϵ such that $k\, poly(n)\ \epsilon < \delta_0$. That is, we require that the error, ϵ, in each individual gate multiplied with the number of gates in the whole circuit, $k\, poly(n)$, does not exceed our desired fixed accuracy δ_0. From this we obtain that $\epsilon = O(1/n^b)$ for some integer b. Finally, from the Kitaev-Solovay theorem we know that such an approximation can be achieved with overhead $polylog(1/\epsilon)$. This establishes that our algorithm has size $poly(n\, logn)$ and is therefore efficient.

Problem 6.2

As in Simon's original problem, we begin with the state $|0\rangle^{\otimes n} \otimes |0\rangle$ and apply

Hadamard gates on the first n qubits to obtain the state

$$(H^{\otimes n} \otimes I)|0\rangle^{\otimes n} \otimes |0\rangle = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle \otimes |0\rangle \ . \tag{6}$$

We then query the oracle to produce the state

$$\mapsto \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle \otimes |f(x)\rangle \ , \tag{7}$$

and finally measure the second register obtaining some specific $f(x_0)$. The post-measurement state on the first n qubits is then

$$\mapsto |x_o H\rangle = \frac{1}{\sqrt{|H|}} \sum_{x:f(x)=f(x_0)} |x\rangle \ , \tag{8}$$

which can also be written as

$$|x_o H\rangle = \frac{1}{\sqrt{|H|}} \sum_{\vec{b}\in\mathbb{Z}_2^k} |x_0 + b \cdot a\rangle \ , \tag{9}$$

where a is the vector with "components" the generators a_i of H, and b is a $k-$bit binary vector which, when dotted onto a, gives an arbitrary element of H (which is in itself a $n-$bit vector).

The next step is to perform a Hadamard transformation on each of these n qubits to produce the state

$$H^{\otimes n}|x_o H\rangle = \frac{1}{\sqrt{|H|}} \frac{1}{\sqrt{2^n}} \sum_{b\in\mathbb{Z}_2^k} \sum_{y\in\mathbb{Z}_2^n} (-1)^{(x_0+b\cdot a)y} |y\rangle \ , \tag{10}$$

For each fixed y, we can consider the sum over all values of b, $\sum_{b\in\mathbb{Z}_2^k}(-1)^{(b\cdot a)\,y}$. For the case $k = 1$ (so that $b = 0, 1$) the sum gives the familiar $1+(-1)^{ay}$ term in Simon's original problem. For $k = 2$, the sum equals $1 + (-1)^{a_1 y} + (-1)^{a_2 y} + (-1)^{a_1 y}(-1)^{a_2 y}$, which we can factorize as $(1 + (-1)^{a_1 y})(1 + (-1)^{a_2 y})$. In general, $\sum_{b\in\mathbb{Z}_2^k}(-1)^{(b\cdot a)\,y} = \sum_{i=1}^{k}(1 + (-1)^{a_i y})$. Clearly then, the only y's that survive in the sum of Eq. (10) are these for which $a_i y = 0 \ (mod\ 2), \forall i = 1, 2, \ldots, k$. This shows that a measurement in the computation basis will yield a vector y which will be orthogonal to the vectors in the subgroup H. Since H is $k-$dimensional, we can repeat this procedure $n-k$ times in order to obtain $n-k$ linearly independent vectors orthogonal to H which will uniquely determine it.

In fact, if we perform exactly $n-k$ queries we are not guaranteed to obtain $n-k$ linearly independent vectors y. Let y_1 be the first vector we measure, which is nontrivial ($y_1 \neq 0$) with probability no smaller than $1 - 1/2^{n-k}$. Since in Eq. (10) the various y's which are orthogonal to H occur with the same amplitude, the probability the second vector y_2 is linearly independent from y_1 is no smaller than $1-2/2^{n-k}$. Similarly, the probability that the third vector y_3 is linearly independent from y_1 and y_2 is no smaller than $1 - 2^2/2^{n-k}$. In general, the probability that the vectors $y_1, y_2, \ldots, y_{n-k}$ will be linearly independent is no smaller than

$$\begin{aligned} P_{\text{success}} &\geq (1 - \tfrac{1}{2^{n-k}})(1 - \tfrac{2}{2^{n-k}})(1 - \tfrac{2^2}{2^{n-k}})(1 - \tfrac{2^3}{2^{n-k}}) \cdots (1 - \tfrac{2^{n-k-1}}{2^{n-k}}) \\ &\geq (1 - \tfrac{1}{2})(1 - (\tfrac{1}{2^2} + \tfrac{1}{2^3} + \cdots + \tfrac{1}{2^{n-k-1}} + \tfrac{1}{2^{n-k}})) \\ &\geq \tfrac{1}{2}(1 - \tfrac{1}{2}) = \tfrac{1}{4}\ , \end{aligned} \tag{11}$$

since $\frac{1}{2^2} + \frac{1}{2^3} + \cdots + \frac{1}{2^{n-k-1}} + \frac{1}{2^{n-k}} = \frac{1}{2} - \frac{1}{2^{n-k}} \leq \frac{1}{2}$.

Problem 6.3

a) The intersection of two subgroups H and H' of a group G is clearly a subgroup of G itself [2]. Now, assume the intersection $g'H' \cap gH$ is non-empty. Then there must exist $h \in H$ and $h' \in H'$ such that $g'h' = gh$, or $g' = gh(h')^{-1}$. Therefore $g'H' \cap gH = gh(h')^{-1}H' \cap gH = ghH' \cap gH = g(hH' \cap H) = g(hH' \cap hH) = gh(H' \cap H)$ and so

$$\left|g'H' \cap gH\right| = \left|H' \cap H\right|\ , \tag{12}$$

where we have used the fact that[3] $hH = H$ when $h \in H$.

b) $P_{H'}$ is the projector onto the linear span of states $|g'H'\rangle$ with $g' \in G$. First, imagine one such coset state $|g'H'\rangle$ that has a non-trivial overlap with the fixed state $|gH\rangle$. From part (a), $g'H' \cap gH$ contains $|H \cap H'|$ elements and so

$$\langle g'H'|gH\rangle = \frac{1}{\sqrt{|H'|\,|H|}}\ |H \cap H'|\ . \tag{13}$$

Now consider the intersection $g'H' \cap gH$. As in part (a), we can write $g'H' \cap gH = gh(H' \cap H)$ for some $h \in H$. Since g is fixed, two different choices of h will lead to two disjoint sets $g'H' \cap gH$ unless they differ by an element in $H' \cap H$. This implies that the number of distinct sets $g'H' \cap gH$ for g fixed and varying g' is $|H|/|H \cap H'|$. The projector $P_{H'}$ is therefore a sum over $|H|/|H \cap H'|$ mutually orthogonal coset

[2] Indeed, if $h_1, h_2 \in H \cap H'$ then $h_1h_2 \in H$ and $h_1h_2 \in H'$ which implies $h_1h_2 \in H \cap H'$. Also, for any $h \in H \cap H'$ there exists a *unique* inverse h^{-1} such that $h^{-1} \in H$ and $h^{-1} \in H'$ which implies $h^{-1} \in H \cap H'$. Hence $H \cap H'$ is a group.

[3] That is, left-multiplying the elements in H with $h \in H$ just permutes them without changing the group.

states each with non-trivial overlap with $|gH\rangle$ given by Eq. (13). Thus

$$||P_{H'}|gH\rangle||^2 = \frac{|H|}{|H\cap H'|}\left(\frac{|H\cap H'|}{\sqrt{|H'|\,|H|}}\right)^2 = \frac{|H\cap H'|}{|H'|}\ . \tag{14}$$

If H' is a subgroup of H then $H\cap H' = H'$ and so Eq. (14) implies $P_{H'}|gH\rangle = |gH\rangle$. Alternatively, if H' is not a subgroup of H then, by Lagrange's theorem, $H\cap H'$ contains at most half the elements of H' (since, from part (a), $H\cap H'$ is a group and is also contained in H'). Then $||P_{H'}|gH\rangle|| \le 1/\sqrt{2}$.

c) Imagine first the hidden subgroup is the first one in our trial list ($r=1$). Then

$$P^{(r=1)}_{\text{success}} = ||P_{H_1}|\psi\rangle||^2 = 1\ . \tag{15}$$

If instead the hidden subgroup is second in our list ($r=2$) then

$$\begin{aligned}
P^{(r=2)}_{\text{success}} &= ||P_{H_2}(I-P_{H_1})|\psi\rangle||^2\\
&= ||(P_{H_2}(I-P_{H_1}) - P_{H_2} + P_{H_2})|\psi\rangle||^2\\
&\ge (||P_{H_2}|\psi\rangle|| - ||(P_{H_2}(I-P_{H_1}) - P_{H_2})|\psi\rangle||)^2\\
&= (1 - ||P_{H_2}P_{H_1}|\psi\rangle||)^2\\
&\ge (1 - ||P_{H_1}|\psi\rangle||)^2\\
&\ge \left(1 - \tfrac{1}{2^{k/2}}\right)^2\ .
\end{aligned} \tag{16}$$

where, in going from the second to the third line, we used the triangle inequality of the sup norm.

In general, we can define $S_t = P_{H_r}(I-P_{H_{r-1}})\cdots(I-P_{H_{r-t}})$ for $0\le t\le r-1$. Then $S_0 = P_{H_r}$ and $S_{r-1} = P_{H_r}(I-P_{H_{r-1}})\cdots(I-P_{H_2})(I-P_{H_1})$. Also note that $||S_t|| - ||S_{t+1}|| \ge 0$, since $||S_{t+1}|| = ||S_t(I-P_{H_{r-t-1}})|| \le ||S_t||\cdot||I-P_{H_{r-t-1}}|| \le ||S_t||$.

Then

$$||S_0|| - ||S_1|| \le ||S_1 - S_0|| = ||-P_{H_r}P_{H_{r-1}}|| \le ||P_{H_{r-1}}||\ , \tag{17}$$

$$||S_1|| - ||S_2|| \le ||S_2 - S_1|| = ||-S_1(-P_{H_{r-2}})|| \le ||P_{H_{r-2}}||\ , \tag{18}$$

and in general

$$||S_{t-1}|| - ||S_t|| \le ||S_t - S_{t-1}|| = ||-S_{t-1}(-P_{H_{r-t}})|| \le ||P_{H_{r-t}}||\ . \tag{19}$$

Therefore

$$||S_0|| - ||S_t|| \le ||P_{H_{r-1}}|| + ||P_{H_{r-2}}|| + \cdots + ||P_{H_{r-t}}|| \le \frac{t}{2^{k/2}}\ . \tag{20}$$

and so

$$P^{(r)}_{\text{success}} = ||S_{r-1}||^2 \geq \left(1 - \frac{r-1}{2^{k/2}}\right)^2 \ , \tag{21}$$

since $||S_0|| = ||P_{H_r}|| = 1$.

Problem 6.4

a) A problem of size n is in BQP if there exists a polynomial-size uniform circuit family that depends on n, where for each such circuit the state for the first output qubit is $|0\rangle$ (resp. $|1\rangle$) with probability at least 2/3 if the correct answer is 0 (resp. 1). Recall that, according to our model, the input qubits to each circuit are prepared in the standard state $|0\rangle^{\otimes n}$ which we will denote as $|0\rangle$ for simplicity. If U is the unitary implemented by the quantum circuit, the output state is therefore $U|0\rangle$. Let P_0 be the projector onto the state $(|0\rangle\langle 0|)_1$ on the first qubit. Then, the probability of obtaining the outcome 0 after measuring the first output qubit is

$$P(0) = \langle 0|U^\dagger(P_0 \otimes I_{n-1})U|0\rangle \ . \tag{22}$$

where I_{n-1} is the identity operator acting on all qubits except for the first.

We can now view $V = U^\dagger(P_0 \otimes I_{n-1})U$ as a new matrix, so that the classical simulation of the quantum computation executed by U reduces to the calculation of the matrix element $\langle 0|V|0\rangle$. We note that V consists first of the unitary U, then of the projector onto the $|0\rangle$ state on the first qubit, and finally of the inverse of the quantum circuit U^{-1}. This is reminiscent of the idea used by Bennett to show that reversible classical computation (i.e., computation in which no "junk" bits are produced and which can be performed with zero thermodynamic cost) is possible.

b) We can write $V = V_L \ldots V_2 V_1$ where each V_i is either a Hadamard gate (H), or a Toffoli gate ($\Lambda(X)$), or a projection P_0 tensored with the identity on all qubits on which it acts trivially. Therefore, our classical simulation will need to compute the matrix element $\langle 0|V_L \ldots V_2 V_1|0\rangle$.

For every Toffoli gate in this expression, we can insert the identity in the form $I = H^2$ following it on its target qubit. This way we obtain a new V, say V', of size $S \leq 2L$, and let us denote by h the number of Hadamard gates it contains. Then we need to compute

$$P(0) = \frac{1}{\sqrt{2^h}}\langle 0|V'_S \ldots V'_2 V'_1|0\rangle \ , \tag{23}$$

where for every Hadamard gate in V' we have collected the $1/\sqrt{2}$ normalization in front of the expression and replaced the corresponding gate $V'_a = H$ by the gate $V'_a = \sqrt{2}H$.

We now note that $\sqrt{2}H$ acts in the computation basis as

$$\sqrt{2}H \ : \ |i\rangle \mapsto \sum_{j=0}^{1}(-1)^{ij}|j\rangle \ . \tag{24}$$

Similarly, for the Toffoli gate

$$\Lambda(X) \ : \ |i,j,k\rangle \mapsto |i,j,k\oplus ij\rangle \ , \tag{25}$$

which, when combined with the two Hadamard gates following it by the construction of V', becomes

$$\begin{aligned} (I\otimes I\otimes H)\Lambda(X) \quad &: \quad |i,j,k\rangle \mapsto \sum_{l=0}^{1}(-1)^{(k\oplus ij)l}|i,j,l\rangle \ , \\ (I\otimes I\otimes H^2)\Lambda(X) \quad &: \quad |i,j,k\rangle \mapsto \sum_{l=0}^{1}\sum_{m=0}^{1}(-1)^{(k\oplus ij\oplus m)l}|i,j,m\rangle \ . \end{aligned} \tag{26}$$

The result now follows. The initial state is a state in the computation basis (i.e., all qubits in the state $|0\rangle$). For every Hadamard gate initially in V, we can use Eq. (24) to generate a new index and express the output state in the computation basis. For every $\Lambda(X)$ gate initially in V, we can group it with the two Hadamard gates we inserted to obtains V' and use Eq. (26) to generate two new indices. At the end of the computation we "project" onto the $|0\rangle$ state, so that all free indices at that point are forced to take the value 0. Thus we have reduced the calculation of $P(0)$ in Eq. (23) to the calculation of a sum over $h \leq 2L$ intermediate binary indices of the form

$$P(0) = \frac{1}{\sqrt{2^h}}\sum_{x}(-1)^{\phi(x)} \ , \tag{27}$$

where x is a binary vector with the h intermediate indices as components and $\phi(x)$ is a polynomial of degree at most three in the intermediate indices. If N_0 (resp. N_1) is the number of values of x for which $\phi(x) = 0$ (resp. $\phi(x) = 1$), it follows that

$$P(0) = \frac{1}{\sqrt{2^h}}(N_0 - N_1) \ . \tag{28}$$