MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Department of Physics, EECS, and Department of Applied Math

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

Quantum Information Science

April 10, 2008

Problem Set #4 Solutions

Problems:

P1: (**Quantum factoring as a feedback process**)

(a)

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

$$= \frac{1}{\sqrt{r}}\sum_{l=0}^{r-1} e^{2\pi i l k/r} U|y^{l+1} \bmod N\rangle \tag{2}$$

$$= \frac{1}{\sqrt{r}}\sum_{l'=0}^{r-1} e^{2\pi i (l'-1) k/r} U|y^{l'} \bmod N\rangle \tag{3}$$

$$= e^{-2\pi i k/r}|\lambda_k\rangle = e^{-2\pi i \phi_k}|\lambda_k\rangle \tag{4}$$

(b) Following the circuit,

$$\frac{1}{\sqrt{2}}(|0\rangle|\lambda_k\rangle + e^{-2\pi i\phi_k}|1\rangle|\lambda_k\rangle) = \frac{1}{2}((|0\rangle + |1\rangle)|\lambda_k\rangle + e^{-2\pi i\phi_k}(|0\rangle - |1\rangle)|\lambda_k\rangle) \tag{5}$$

$$= \frac{1}{2}((1 + e^{-2\pi i\phi_k})|0\rangle|\lambda_k\rangle + (1 - e^{-2\pi i\phi_k})|1\rangle|\lambda_k\rangle) \tag{6}$$

$$= e^{-\pi i\phi_k}(\cos(\pi\phi_k)|0\rangle + i\sin(\pi\phi_k)|1\rangle)|\lambda_k\rangle \tag{7}$$

Measurement of the ancilla gives 0 with probability $p(0) = \cos^2(\pi\phi_k)$.

(c) Following the circuit and letting the right register start in $|1\rangle = \frac{1}{\sqrt{r}}\sum_{k=0}^{r-1}|\lambda_k\rangle$, we find by linearity and the solution to part (b),

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

where

$$|\eta_k\rangle = \frac{1}{2}((1 + e^{-2\pi i\phi_k})|0\rangle + (1 - e^{-2\pi i\phi_k})|1\rangle). \tag{9}$$

(d) If we choose to postpone measuring the ancillas, the final state after t iterations is

$$|\psi_t\rangle = \frac{1}{\sqrt{r}}\sum_{k=0}^{r-1}|\eta_k\rangle^{\otimes t}|\lambda_k\rangle. \tag{10}$$

This expression shows that the state of the output is independent of the order in which the ancila are measured.

(e) We want to find

$$p(k, n_0) = \mathbb{P}(k \text{ obtained and } n_0 \text{ zeros obtained in } t \text{ trials}). \tag{11}$$

Let $c_{jk} := \frac{1}{\sqrt{2}}(1 + (-1)^j e^{-2\pi i \phi_k})$, then

$$p(k, n_0) = \frac{1}{r 2^t} \binom{t}{n_0} |c_{0k}|^{2n_0} |c_{1k}|^{2(t-n_0)} \tag{12}$$

$$= \frac{1}{r} \binom{t}{n_0} \cos^{2n_0}(\pi\phi_k) \sin^{2(t-n_0)}(\pi\phi_k). \tag{13}$$

(f) From the joint distribution $p(k, n_0)$, we can find the marginal distribution

$$p(n_0) = \sum_{k=0}^{r-1} p(k, n_0) = \frac{1}{r} \binom{t}{n_0} \sum_{k=0}^{r-1} \cos^{2n_0}(\pi\phi_k) \sin^{2(t-n_0)}(\pi\phi_k) \tag{14}$$

and the resulting conditional distribution $p(k|n_0) = p(k, n_0)/p(n_0)$ by Bayes rule. Define $\alpha = n_0/t$. Observe that α is approximately uniformly distributed in $[0, 1]$ and $p(n_0)$ depends only on the ratio α (Can you show this?).

Taking $p(n_0) \approx 1/2$, the conditional distribution is

$$p(k|n_0) \approx \frac{2}{r} \binom{t}{n_0} \left[\cos^2\left(\frac{k\pi}{r}\right)\right]^{n_0} \left[\sin^2\left(\frac{k\pi}{r}\right)\right]^{t-n_0}. \tag{15}$$

In order to find the maxima of the distribution $p(k|n_0)$, we opt to work with $\log p(k|n_0)$ which is a strictly increasing function of p. Accordingly, we have

$$\frac{\partial}{\partial k} \log p(k|n_0) = \frac{\partial p_0}{\partial k} \left(\frac{n_0}{p_0} - \frac{n_1}{1 - p_0}\right) \tag{16}$$

$$= 0. \tag{17}$$

The only relevant solutions are the ones corresponding to the solutions of $n_0/p_0 - n_1/(1 - p_0) = 0$, namely, $p_0 = n_0/t$, which gives

$$\cos^2\left(\frac{k\pi}{r}\right) = \frac{n_0}{t} \implies k_1 = \frac{r}{\pi} \text{Cos}^{-1} \sqrt{\frac{n_0}{t}}, \qquad k_2 = r\left(\frac{1}{\pi} \text{Cos}^{-1} \sqrt{-\frac{n_0}{t}}\right). \tag{18}$$

Consider the Taylor expansion of $\log p$ around $p_0^* = n_0/t$ to the first nonzero order:

$$\log p \approx \log p\Big|_{p_0^*} + \frac{\partial^2}{\partial p_0^2} \log p\Big|_{p_0^*} (p_0 - p_0^*). \tag{19}$$

Evaluating the term $\partial^2 \log p / \partial p_0^2$ at $p_0 = p_0^*$, shows that for large t, it is proportional to t. The expression for p would be of the form

$$p \approx p^* e^{-t(p_0 - p_0^*)} \tag{20}$$

where p^* denotes the value of the distribution for $p_0 = p_0^*$. This shows that the width over which the values of the distribution are considerable around the peak decreases as $\mathcal{O}(1/t)$.

(g) See Figure 1. We find that $\frac{6}{\pi^{-1}\text{Arccos}(\sqrt{.3})} \approx 19.01$ and $\frac{14}{\pi^{-1}\text{Arccos}(-\sqrt{.3})} \approx 20.45$.

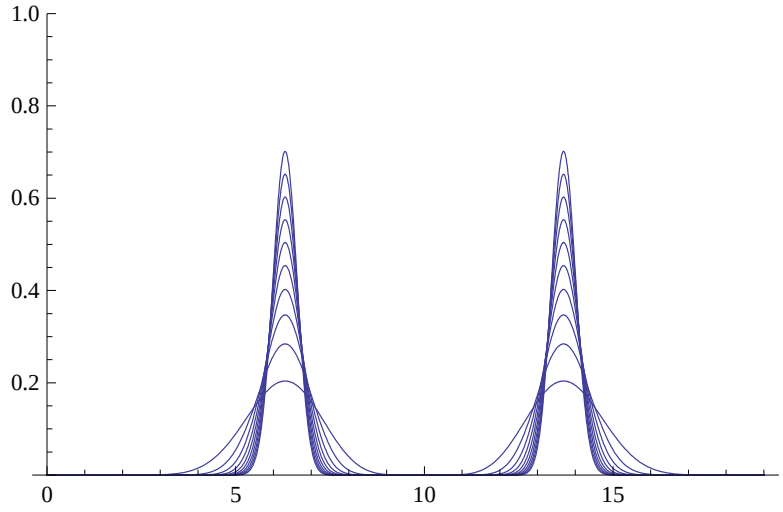

Figure 1: Conditional probability $p(k|\alpha)$ plotted for $\alpha = 0.3$ and t ranging from 10 to 100 in increments of 10. The peaks occur at around $k = 6, 14$.

(h) Following through the circuit,

$$|\psi_1\rangle = \frac{1}{2}\left((1 - e^{2\pi i(\theta-\phi_k)})|0\rangle|\lambda_k\rangle + (1 + e^{2\pi i(\theta-\phi_k)})|1\rangle|\lambda_k\rangle\right), \tag{21}$$

so when $\theta = \phi_k$, $|\psi_1\rangle = |0\rangle|\lambda_k\rangle$.

(i) Let $\theta = 0$, $f_0 = 1$, $n = 0$, $p = 0$, and for $U|m\rangle = |my^{2^p} \mod N\rangle$, $U|\eta_k\rangle = \phi_k^p|\eta_k\rangle$, $\phi_k^p = 2^p k/r$, $\eta_k^\theta = \frac{1}{2}\left[(1 + e^{2\pi i(\theta-\phi_k^p)}|0\rangle + (1 - e^{2\pi i(\theta-\phi_k^p)})|1\rangle\right]$, and $|\phi_0\rangle = |1\rangle$. Do the following:

(a) Run one cycle:

$$|\psi_n\rangle = \sum_n c_{k,n}|\lambda_k\rangle \tag{22}$$

$$\mapsto |0\rangle|\psi_n\rangle \tag{23}$$

$$\mapsto \sum_k c_{k,n}|\eta_k^\theta\rangle|\lambda_k\rangle \tag{24}$$

$$\mapsto c_{k,n+1}|\lambda_k\rangle = |\psi_{n+1}\rangle, \text{with outcome } b_n \tag{25}$$

(b) Update model: If $b_n = 0$ then $f_{n+1} = \cos^2(\phi - \theta)f_n$, else $f_{n+1} = \sin^2(\phi - \theta)f_n$. Let θ be the value of ϕ that maximizes f_{n+1}. If θ is sufficiently accurate, increment p.

(c) Increment n and go to the first step.

The asymptotic value of θ as $n \to \mathcal{O}(\log N)$ gives k/r.

P2: (**Entanglement distillation**) It is somewhat easier to take $|\phi_+\rangle$ as the state when no error occurs, so have Alice apply Z to each of her qubits and Bob apply X to his half. This maps $|\psi_-\rangle$ to $|\phi_+\rangle$. The

state $|\phi_+\rangle$ is mapped to each of the other Bell states by some single qubit Pauli: Z_A gives $|\phi_-\rangle$, $Z_A X_B$ gives $|\psi_-\rangle$, and X_B gives $|\psi_+\rangle$. Since $U \otimes I|\phi_+\rangle = I \otimes U^T|\phi_+\rangle$, we can move the Z_A errors to Bob's half. Now, the state Alice and Bob share is equivalent to Alice starting with $|\phi_+\rangle$ and sending Bob's qubits to him through a channel that acts on each qubit as $\mathcal{E}(\rho) = (1-3\epsilon)\rho + \epsilon(X\rho X + Y\rho Y + Z\rho Z)$.

If there are no errors, the initial stabilizer is generated by $X_i X_{i+9}$, $Z_i Z_{i+9}$ for $i = 1, 2, \ldots, 9$. Alice measures the generators of Shor's code. Each generator anticommutes with some element of the stabilizer, so the string of outcome bits is uniformly distributed. After the measurements, Alice and Bob share a logical Bell pair stabilized by the stabilizer of Shor's code on Alice's half (with syndrome $\vec{s}_A$) and Bob's half (also with syndrome $\vec{s}_A$), since, for a generator g_A of Shor's code on Alice's side, $g_A g_B \in S$, so $\pm g_A(g_A g_B) = \pm g_B \in S$. The logical Bell pair is stabilized by $\bar{X}_A \bar{X}_B$ and $\bar{Z}_A \bar{Z}_B$.

Bob can now measure the syndrome on his side to obtain the same syndrome $\vec{s}_B = \vec{s}_A$ as Alice, if there are no errors. If there are errors on Bob's half, some of the signs are flipped on the physical Bell pair generators, so for example $-g_A g_B \in S$ and $\pm g_A(-g_A g_B) = \mp g_B \in S$. So, if the errors are detectable by Shor's code, $\vec{s}_B \neq \vec{s}_A$ and the errors on Bob's half are indicated by the syndrome $\vec{s}_A \oplus \vec{s}_B$.

(a) Postselecting on $\vec{s}_A = \vec{s}_B = 0$ gives a state where no detectable errors have occurred. When Alice and Bob decode their halves, the only way the resulting state can be in error is if an element of the normalizer of Shor's code was applied to Bob's half. The lowest weight $\bar{X}$ errors are $X_i X_{i+1} X_{i+2}$ for $i = 1, 2, 3$ (i.e. in the same block), the lowest weight $\bar{Z}$ errors are $Z_i Z_j Z_k$ where $i \in \{1, 2, 3\}$, $j \in \{4, 5, 6\}$, and $k \in \{7, 8, 9\}$ (i.e. in different blocks), and the lowest weight $\bar{Y}$ errors have weight greater than three. Therefore,

$$\rho \approx (1 - 12\epsilon^3 - O(\epsilon^4))|\phi_+\rangle\langle\phi_+| + 3\epsilon^3|\psi_+\rangle\langle\psi_+| + 9\epsilon^3|\phi_-\rangle\langle\phi_-| + O(\epsilon^4)|\psi_-\rangle\langle\psi_-|. \tag{26}$$

(b) Postselecting on $\vec{s}_A = \vec{s}_B$ gives a state where, again, no detectable errors have occurred. After Alice and Bob send each other their syndromes, they know this, so they can each apply an element of the Pauli group to reset their syndromes to zero (not the usual corrections, since we want these corrections to commute with $\bar{X}$ and $\bar{Z}$). Alice and Bob decode and get the same result as (a).

(c) Suppose Alice and Bob find that $\vec{s}_A \neq \vec{s}_B$. Bob computes $\vec{s}_A \oplus \vec{s}_B$ and uses this new syndrome to correct the error on his half. Alice and Bob's syndromes agree after the correction. If the error is correctable, then Alice and Bob can continue as in (b) to obtain an undamaged state. However, if the error happens to be detectable but not correctable, Bob's correction maps the logical Bell pair to a different logical Bell pair. For Shor's code, all the single qubit errors are correctable. Furthermore, the following double errors are correctable: $X_i Z_j$ for $i \neq j$, $Z_i Z_j$ for i, j in the same block, $X_i X_j$ for i, j in different blocks, $Y_i X_j$ for i, j in different blocks, and $Y_i Z_j$ for i, j in the same block. The remaining double errors are detectable but not correctable: $X_i X_j$ for i, j in the same block ($\bar{X}$), $Z_i Z_j$ for i, j in different blocks ($\bar{Z}$), $Y_i Y_j$ for all i, j (either $\bar{X}$ or $\bar{Z}$ depending on whether they are in the same block or different blocks), $Y_i X_j$ for i, j in the same block ($\bar{X}$), and $Y_i Z_j$ for i, j in different blocks ($\bar{Z}$). The number of ways a pair of the same type of errors can be distributed into different blocks is $\binom{3}{2} \times 3 \times 3 = 27$ and into the same block is $3 \times \binom{3}{2} = 9$. The count doubles if the errors are of different types.

Therefore,

$$\rho \approx \left(1 - 144\epsilon^2 - O(\epsilon^3)\right) |\phi_+\rangle\langle\phi_+| + (36\epsilon^2 + O(\epsilon^3)|\psi_+\rangle\langle\psi_+| \tag{27}$$

$$+ (108\epsilon^2 + O(\epsilon^3))|\phi_-\rangle\langle\phi_-| + O(\epsilon^4)|\psi_-\rangle\langle\psi_-|. \tag{28}$$

P3: (**State identification and the pretty good measurement**)

(a) Let E_i, $i = 1, 2, \ldots, k$, $\sum_{i=1}^k E_i = I$, be POVM elements, then

$$\mathbb{P}(\text{guess right}) = \sum_{i=1}^{k} p_i \text{Tr}(E_i \rho_i). \tag{29}$$

Since $\tau - p_i\rho_i$ is positive and Hermitian for all i, Tr $E_i(\tau - p_i\rho_i) \geq 0$. Indeed, if A and B are positive and Hermitian, then each admits a positive square root. Therefore,

$$\text{Tr } AB = \text{Tr}(\sqrt{A}\sqrt{A}\sqrt{B}\sqrt{B}) \tag{30}$$

$$= \text{Tr}(\sqrt{B}\sqrt{A}\sqrt{A}\sqrt{B}) \tag{31}$$

$$= \text{Tr}((\sqrt{A}\sqrt{B})^\dagger(\sqrt{A}\sqrt{B})) \geq 0. \tag{32}$$

Taking the sum over elements, we have

$$\sum_{i=1}^{k} \text{Tr } E_i(\tau - p_i\rho_i) \geq 0 \implies \text{Tr } \tau \geq \mathbb{P}(\text{guess right}), \tag{33}$$

by linearity of the trace and completeness of the POVM elements.

(b) Suppose $k > 1$. For this ensemble, $M = I/2$ so $M^{-1/2} = \sqrt{2}I$. The PGM elements are

$$E_j = \frac{1}{k} \begin{pmatrix} 1 & e^{-2\pi ij/k} \\ e^{2\pi ij/k} & 1 \end{pmatrix}. \tag{34}$$

The probability of success is $2/k$. This can be proven optimal by taking $\tau = \sum_i p_i E_i \rho_i = \text{diag}(1/k, 1/k)$ since the eigenvalues of $\tau - p_i\rho_i$ are 0 and $1/k$ for all i.

(c) Pick two orthogonal states $|i\rangle$ and $|j\rangle$ and form the POVM with elements $E_0 = |i\rangle\langle i|$, $E_1 = |j\rangle\langle j|$, and $E_2 = I - E_1 - E_2$. This POVM guesses these two states without error, so the success probability is at least $2/k$. From (b) we know this is optimal.

(d) Consider the POVM to distinguish the two orthogonal states of the ensemble: $E_0 = I - E_1 - E_2$, $E_1 = |+\rangle\langle+|$, and $E_2 = |-\rangle\langle-|$. This POVM succeeds with probability $2/3$ for this ensemble. It can be proven optimal using $\tau = \text{diag}(1/3, 1/3)$ since $\tau - p_i\rho_i$ has eigenvalues $1/3$ and 0 for all i. The PGM has elements $E_0 = \begin{pmatrix} 1/2 & 0 \\ 0 & 0 \end{pmatrix}$, $E_1 = \begin{pmatrix} 1/4 & 1/(2\sqrt{2}) \\ 1/(2\sqrt{2}) & 1/2 \end{pmatrix}$, and $E_2 = \begin{pmatrix} 1/4 & -1/(2\sqrt{2}) \\ -1/(2\sqrt{2}) & 1/2 \end{pmatrix}$. The PGM succeeds with probability $5/12 + 1/(3\sqrt{2}) \approx 0.652 < 1/3$.

(e) The PGM elements are $E_0 = \begin{pmatrix} 2/3 & 0 \\ 0 & 0 \end{pmatrix}$, $E_1 = \begin{pmatrix} 1/6 & 1/(2\sqrt{3}) \\ 1/(2\sqrt{3}) & 1/2 \end{pmatrix}$, and $E_2 = \begin{pmatrix} 1/6 & -1/(2\sqrt{3}) \\ -1/(2\sqrt{3}) & 1/2 \end{pmatrix}$, and its success probability is $1/2 + 1/(4\sqrt{3}) \approx 0.644 > 1/2$. However, this is not optimal. Take

$\tau = \begin{pmatrix} 1/2 & 0 \\ 0 & 1/6 \end{pmatrix}$, then $\tau - p_i\rho_i \geq 0$ for all i and Tr $\tau = 2/3$. A POVM achieving this is $E_0 = (8/9)|0\rangle\langle 0|$, $E_1 = (1/18)\begin{pmatrix} 1 & 3 \\ 3 & 9 \end{pmatrix}$, and $E_2 = (1/18)\begin{pmatrix} 1 & -3 \\ -3 & 9 \end{pmatrix}$.

P4: (**Measures of pure state entanglement**)

(a) If the Schmidt number is one, then $|\psi\rangle = |k_A\rangle|k_B\rangle$ is a product state. Suppose $|\psi\rangle = |\psi_A\rangle|\psi_B\rangle$. Choose $|\psi_A\rangle$ as a basis vector of subsystem A and choose an arbitrary basis for $I_A - |\psi_A\rangle\langle\psi_A|$. Apply the Gram-Schmidt procedure to obtain an orthonormal basis of A containing $|\psi_A\rangle$. Repeat for subsystem B. The resulting bases give a Schmidt decomposition for $|\psi\rangle$ with Schmidt number one.

(b) Let $|\psi\rangle = \sum_k \lambda_k |k_A\rangle|k_B\rangle$. Consider $U = U_A \otimes U_B$ acting on this state,

$$U|\psi\rangle = \sum_k \lambda_k U_A|k_A\rangle \otimes U_B|k_B\rangle. \tag{35}$$

Choose a new Schmidt basis for $U|\psi\rangle$ to be $U_A|k_A\rangle$ and $U_B|k_B\rangle$. This implies that the Schmidt rank is unchanged by local unitaries.

(c) We have

$$|\phi_1\rangle = \sum_{k=1}^{3} \frac{1}{\sqrt{3}}|kk\rangle \qquad \Longrightarrow 3 \tag{36}$$

$$|\phi_2\rangle = |+\rangle|+\rangle \qquad \Longrightarrow 1 \tag{37}$$

$$|\phi_3\rangle = |0\rangle|+\rangle + |1\rangle|-\rangle \qquad \Longrightarrow 2. \tag{38}$$

For the last state, $|\phi_4\rangle$, we can apply the singular value decomposition to $\begin{pmatrix} 1/\sqrt{3} & 1/\sqrt{3} \\ 0 & 1/\sqrt{3} \end{pmatrix}$ to obtain $UDV^\dagger$ where $D = \text{diag}(\sqrt{(3+\sqrt{5})/6}, \sqrt{(3-\sqrt{5})/6})$, so the Schmidt number is 2.