Quantum Computing and Information - Problem Set 2 Solutions

Exercise 1. **Constructing a Toffoli gate from CNOT and single-qubit gates** *This exercise will prove that two-qubit unitary gates are universal. For a single-qubit unitary U, define the controlled-U operation to be $C_U := |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes U$. Note that* CNOT $= C_X$*. To indicate the systems that these gates act on we use the notation $[C_X]_{i,j}$ to mean a controlled-U operation where qubit i is the control and qubit j is the target.*

a) *Show that*

$$[C_U]_{1,3}[C_X]_{2,1}[C_U^\dagger]_{13}[C_X]_{21}[C_U]_{23}$$

implements a doubly-controlled U^2: i.e. applies $[U^2]_3$ only if qubits 1 and 2 are both in the $|1\rangle$ state. Thus if $U = e^{i\varphi}\sqrt{X}$ for some φ then this implements gate that is related to the Toffoli gate.

b) *Now we need to construct a controlled-$\sqrt{X}$ gate from CNOTs and single-qubit gates. Show that*

$$[V]_3[C_X]_{1,3}[V^\dagger]_3[W]_3[C_X]_{1,3}[W^\dagger]_3 = [C_U]_{1,3},$$

where $U = VXV^\dagger WXW^\dagger$.

c) *Find V, W such that $VXV^\dagger WXW^\dagger = e^{i\varphi}\sqrt{X}$ for some φ. (Part (d) of question 2 on problem set 1 may help here, although you will need to calculate $(\vec{v}\cdot\vec{\sigma})\cdot(\vec{w}\cdot\vec{\sigma})$ rather than the commutator.)*

a) Note that $[C_X]_{2,1} = \sum_{a,b\in\{0,1\}} |a\rangle\langle a|_2 \otimes |b\oplus a\rangle\langle b|_1$. Now we calculate

$$[C_U]_{2,3} = \sum_{a,b\in\{0,1\}} |a\rangle\langle a| \otimes |b\rangle\langle b| \otimes U^b$$

$$[C_X]_{2,1}[C_U]_{2,3} = \sum_{a,b\in\{0,1\}} |b\oplus a\rangle\langle a| \otimes |b\rangle\langle b| \otimes U^b$$

$$[C_U^\dagger]_{1,3}[C_X]_{2,1}[C_U]_{2,3} = \sum_{a,b\in\{0,1\}} |b\oplus a\rangle\langle a| \otimes |b\rangle\langle b| \otimes U^{b-(b\oplus a)}$$

$$[C_X]_{2,1}[C_U^\dagger]_{1,3}[C_X]_{2,1}[C_U]_{2,3} = \sum_{a,b\in\{0,1\}} |a\rangle\langle a| \otimes |b\rangle\langle b| \otimes U^{b-(b\oplus a)}$$

$$[C_U]_{1,3}[C_X]_{2,1}[C_U^\dagger]_{1,3}[C_X]_{2,1}[C_U]_{2,3} = \sum_{a,b\in\{0,1\}} |a\rangle\langle a| \otimes |b\rangle\langle b| \otimes U^{a+b-(b\oplus a)}$$

Finally, we observe that for $a, b \in \{0,1\}$, $a + b - (a\oplus b) = ab$.

b) If qubit 1 is in the $|0\rangle$ state then we can ignore the $[C_X]_{1,3}$ gates, and we are left with $VV^\dagger WW^\dagger = I$ acting on qubit 3. On the other hand, if qubit 1 is in the $|1\rangle$ state, then the $[C_X]_{1,3}$ gates act as $[X]_3$ gates, and we obtain $VXV^\dagger WXW^\dagger$ acting on qubit 3. This is equivalent to the claimed $[C_U]_{1,3}$ behavior.

c) Note that $X = e^{i\frac{\pi}{2}X}$, so $\sqrt{X} = e^{i\frac{\pi}{4}X} = (I + iX)/\sqrt{2}$. Define $\vec{v}, \vec{w}$ such that $\vec{v}\cdot\vec{\sigma} = VXV^\dagger$ and $\vec{w}\cdot\vec{\sigma} = WXW^\dagger$. We claim that varying over all unitary V is equivalent to varying over all unit vectors $\vec{v}$ (and similarly for $W, \vec{w}$). Why? First, according to the spectral theorem, the set $\{VXV^\dagger : V \in \mathcal{U}_2\}$ equals the set of Hermitian matrices with eigenvalues $\{1, -1\}$. Second, any traceless 2×2 Hermitian matrix can be written in the form $\vec{v}\cdot\vec{\sigma}$ for some not-necessarily-unit vector $\vec{v}$. Third, $(\vec{v}\cdot\vec{\sigma})^2 = \|\vec{v}\|^2 I$, implying that $(\vec{v}\cdot\vec{\sigma})$ has eigenvalues $\pm\|\vec{v}\|$. Thus if $\vec{v}$ is a unit vector then $\vec{v}\cdot\vec{\sigma}$ has eigenvalues ± 1 and therefore can be written as $VXV^\dagger$ for $V \in \mathcal{U}_2$; and conversely, for any $V \in \mathcal{U}_2$, $VXV^\dagger$ has eigenvalues ± 1 and therefore equals $\vec{v}\cdot\vec{sigma}$ for some unit vector $\vec{v}$.

We now return to the problem at hand. From 2d of the last problem set plus a small calculation, we find that

$$(\vec{v}\cdot\vec{\sigma})(\vec{w}\cdot\vec{\sigma}) = (\vec{v}\cdot\vec{w})I + i(\vec{v}\times\vec{w})\cdot\sigma.$$

Thus we need to choose unit vectors $\vec{v}, \vec{w}$ satisfying $\vec{v} \cdot \vec{w} = 1/\sqrt{2}$ (so the angle between the vectors is $\pi/4$) and $\vec{v} \cdot \vec{w} = (1,0,0)/\sqrt{2}$. Thus, the vectors should be in the y-z plane. One choice that works is $\vec{v} = (0,1,1)/\sqrt{2}, \vec{w} = (0,0,1)$.

Finally, we need to find the corresponding V, W whose existence is guaranteed by the spectral theorem. Using the spectral theorem, we should choose V to map the eigenbasis of X to the eigenbasis of $\vec{v}\cdot\vec{\sigma}$, and similarly should choose W to map the eigenbasis of X to the eigenbasis of $\vec{w}\cdot\vec{\sigma}$. This can be done with matlab, or by using problem 2g of the last problem set to observe that since $\vec{v}$ has polar coordinates $\theta = \pi/4$, $\phi = \pi/2$, we have $\vec{v}\cdot\vec{\sigma} = 2|\alpha\rangle\langle\alpha| - I = |\alpha\rangle\langle\alpha| - |\beta\rangle\langle\beta|$ for $|\alpha\rangle = \cos(\pi/8)e^{-i\pi/4}|0\rangle + \sin(\pi/8)e^{i\pi/4}|1\rangle$ and $|\beta\rangle = \sin(\pi/8)e^{-i\pi/4}|0\rangle - \cos(\pi/8)e^{i\pi/4}|1\rangle$. Thus, we can take $V = |\alpha\rangle\langle+| + |\beta\rangle\langle-|$. We can do something similar for W, or just notice that $W = H$ works, for H the Hadamard matrix.

An alternate solution (due to Kamil) for V is to define $T = \exp(i\frac{\pi}{8}\sigma_z)$, observe that $XTX = T^{-1}$ and that $T^4 = Z$. Thus, $TXT^{-1} = T^2X$ and $(TH)^\dagger X(TH) = HT^\dagger XTH = HXT^2H = ZHT^2H$. We take $V = H$ and $W = (TH)^\dagger = HT^\dagger$ and find $VXV^\dagger WXW^\dagger = (HXH)\cdot(ZHT^2H) = Z\cdot ZHT^2H = HT^2H = H\sqrt{Z}H = \sqrt{X}$.

Exercise 2. **The hybrid argument** *The* operator norm *is defined as follows. If M is a matrix, then define*

$$\|M\| := \max |\langle\alpha| M |\beta\rangle|,$$

where the max is taken over all unit vectors $|\alpha\rangle$ and $|\beta\rangle$.

a) *Show that the operator norm obeys the triangle inequality: $\|A+B\| \leq \|A\| + \|B\|$.*

b) *Show that the norm is right and left unitarily-invariant. That is, for any unitary U and any matrix M, $\|M\| = \|MU\| = \|UM\|$.*

c) *Suppose that we would like to perform a quantum circuit $U_{(T)} := U_1U_2\cdots U_T$ but only are able to apply each gate approximately. Thus, we instead perform $\tilde{U}_{(T)} := \tilde{U}_1\cdots\tilde{U}_T$ for some unitaries $\tilde{U}_1, \ldots, \tilde{U}_T$ satisfying $\|U_i - \tilde{U}_i\| \leq \epsilon_i$ for $i = 1, \ldots, T$. Prove that $\|U_{(T)} - \tilde{U}_{(T)}\| \leq \epsilon_{(T)} := \sum_{i=1}^T \epsilon_i$.*

a) Let unit vectors $\langle\alpha|$ and $|\beta\rangle$ satisfy $\langle\alpha| M |\beta\rangle = M$. Then $\|A\| \geq |\langle\alpha| A |\beta\rangle|$ and $\|B\| \geq |\langle\alpha| B |\beta\rangle|$ by the definitions of the operator norm, and thus

$$\begin{aligned}\|A\| + \|B\| &\geq |\langle\alpha| A |\beta\rangle| + |\langle\alpha| B |\beta\rangle| \\ &\geq \langle\alpha| (A+B) |\beta\rangle = \|A+B\| \qquad \text{by the triangle inequality for } \mathbb{C}\end{aligned}$$

b) Since U is a bijection on the set of unit vectors, maximizing over $|\beta\rangle$ is the same as maximizing over $U|\beta\rangle$. Similarly, maximizing over $\langle\alpha|$ is the same as maximizing over $\langle\alpha|U$.

c) We prove the claim by induction on T. The base case ($T = 1$) is immediate. Now assume that $\|U_{(T-1)} - \tilde{U}_{(T-1)}\| \leq \epsilon_1 + \ldots + \epsilon_{T-1}$. Use first the right invariance of the operator norm and then the triangle inequality to obtain

$$\begin{aligned}\|U_{(T)} - \tilde{U}_{(T)}\| &= \|U_{(T-1)}U_T - U_{(T-1)}\tilde{U}_T + U_{(T-1)}\tilde{U}_T - \tilde{U}_{(T-1)}\tilde{U}_T\| &(1)\\ &\leq \|U_{(T-1)}U_T - U_{(T-1)}\tilde{U}_T\| + \|U_{(T-1)}\tilde{U}_T - \tilde{U}_{(T-1)}\tilde{U}_T\| &(2)\\ &\leq \|U_T - \tilde{U}_T\| + \|U_{(T-1)} - \tilde{U}_{(T-1)}\| &(3)\\ &\epsilon_T + \sum_{i=1}^{T-1}\epsilon_i &(4)\end{aligned}$$

In Eq. (3), we have used the right and left unitary invariance of the operator norm, and in the final equation we used the induction hypothesis.

Exercise 3. **A lazier Quantum Fourier Transform (QFT)**
When implementing the QFT, a lot of time is spent on $R_k = \exp(\frac{2\pi i|1\rangle\langle1|}{2^k})$ rotations that, for large values of k, are very close to I. Suppose we replace R_k with the identity whenever $k \geq k_0$ for some cut-off value k_0.

a) *The standard QFT uses $O(n^2)$ gates. Give an asymptotic estimate for the number of gates in the lazy QFT described here, noting that identity gates don't count.*

b) *Give an upper bound on the error in the resulting approximate QFT.*

c) *How many gates suffice to achieve an error that scales as $1/n^{100}$?*

a) Each qubit is now involved in $\leq k_0$ controlled rotations, so the total number of gates is $O(nk_0)$. In fact, this is not much of an overestimate, since only k_0 qubits are involved in fewer than k_0 gates.

b) $\|R_k - I\| = |e^{2\pi i/2^k} - 1| = sin(\pi/2^k) \leq \pi/2^k$ using the fact that $\sin(x) \leq |x|$ for all x. The total error is $\leq \sum_{j=0}^{n-k_0} \pi(n-k_0-j)2^{-k_0-j} \leq \pi n 2^{-k_0} \sum_{j=0}^{\infty} 2^{-j} = 2\pi n 2^{-k_0} = O(n2^{-k_0})$.

c) $101\log(n)$.

Exercise 4. **Phase estimation**

a) *Suppose we start with the state*

$$\frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle, \tag{5}$$

apply the conditional phase $\sum_{x=0}^{2^n-1} e^{2\pi i\varphi x}|x\rangle\langle x|$ and then the inverse QFT $\frac{1}{\sqrt{2^n}} \sum_{x,y=0}^{N-1} e^{-\frac{2\pi ixy}{N}} |x\rangle \langle y|$. Finally we measure the state in the computational basis and obtain outcome y. Calculate $Pr[y]$.

b) *Assume that $0 \leq \varphi \leq 1$. Define $\Delta := y/N - \varphi$ and $\delta = \min(|\Delta|, 1-|\Delta|)$. This definition is meant to express the idea that δ is the error in the phase estimation procedure. Show that there exists a constant $c > 0$ such that*

$$Pr\left[\delta \geq \frac{k}{N}\right] \leq \frac{c}{k},$$

for any positive integer k. Hint: For $\alpha \geq 0$, it may be helpful to use the bounds $\alpha - \alpha^3/6 \leq \sin\alpha \leq \alpha$.

c) Optional: *Now suppose we do the same procedure but replace the state in Eq. (5) with*

$$\frac{1}{\sqrt{2^{n-1}}} \sum_{x=0}^{2^n-1} \sin \frac{\pi(x+\frac{1}{2})}{2^n} |x\rangle. \tag{6}$$

Check that this state is normalized, calculate $Pr[y]$ for this strategy, and show that it satisfies

$$Pr\left[\delta \geq \frac{k}{N}\right] \leq \frac{c}{k^3},$$

for any positive integer k and for a possibly different constant c. Thus, while the width of this distribution cannot be substantially improved, the tails can be made to drop off faster. This question relates to the construction of optimal quantum clocks.

a) Use the expression for a finite geometric series, valid for all $x \neq 1$: $\sum_{j=0}^{N-1} x^j = (1-x^N)/(1-x)$. Then we obtain:

$$\Pr[y] = \left|\frac{1}{N}\sum_{x=0}^{N-1} e^{2\pi i x\Delta/N}\right|^2 = \left|\frac{1-e^{2\pi i\Delta}}{N(1-e^{2\pi i\Delta/N})}\right|^2 = \frac{\sin^2(\pi\Delta)}{N^2\sin^2(\pi\Delta/N)} = \frac{\sin^2(\pi\delta)}{N^2\sin^2(\pi\delta/N)}$$

b) Suppose $|\delta| \leq N/\pi$. Then

$$\begin{aligned}\sin^2(\pi\delta/N) &\geq \left(\frac{\pi\delta}{N}\left(1-\frac{1}{6}\left(\frac{\pi\delta}{N}\right)^2\right)\right)^2 \\ &\geq \left(\frac{5\pi}{6}\frac{\delta}{N}\right)^2 \geq \delta^2/N^2.\end{aligned}$$

Using $\sin^2(\pi\delta) \le 1$, we find that $\Pr[y] \le 1/\delta^2$.

On the other hand, if $|N\delta| > 1/\pi$, then we also have $|N\delta| < 1/2$ by the definition of δ. Thus $\sin^2(\pi\delta/N) \ge \sin^2(1) \ge 0.7$. We conclude that $\Pr[y] \ge 2/\delta^2$. Finally, we can sum over $|\delta| \ge k$ to obtain $\Pr[|\delta| \ge k] \le 4/\delta$.

c) This calculation is in appendix A.3 of arXiv:0811.3171. The proof there has (at least) one mistake: the δ^2 at the end should be δ^4.

Exercise 5. **Collision detection**
Suppose we are given a black-box function $f : \{0,1\}^n \to \{0,1\}^{n-1}$ that is 2-to-1: i.e. exactly two inputs go to each output. Our goal is to find $x, y \in \{0,1\}^n$ such that $f(x) = f(y)$. However, unlike in Simon's algorithm, we now have no promise about any periodicity of f. As a result it turns out that quantum computers cannot achieve exponential speedups in this case. Define $N = 2^n$.

a) *Give a classical algorithm that finds a collision with high probability ($\ge 1/2$) using only $O(\sqrt{N})$ queries to f.*

b) *Suppose now that only $O(N^{1/3}\log(N))$ bits of memory are available. (Note that $\log(N)$ bits can store one integer between 1 and N.) Now describe a classical algorithm that finds a collision with high probability that uses $O(N^{2/3})$ queries.*

c) *Give a quantum algorithm that finds a collision in $O(\sqrt{N})$ queries and uses $O(\log(N))$ space.* Hint: Use Grover's algorithm.

d) *Give a quantum algorithm that finds a collision in $O(M + \sqrt{N/M})$ queries and uses $O(M\log(N))$ space for any choice of M. Choosing $M = N^{1/3}$ will then yield a $\tilde{O}(N^{1/3})$-query algorithm, where $\tilde{O}$ neglects log factors.* Hint: combine parts (b) and (c).

a) Query a random subset $S \subset \{0,1\}^n$ of size $c\sqrt{N} + 1$ and check for collisions. Suppose that after k queries, we haven't yet seen a collision. Then the probability of seeing a collision on the $k + 1^{\text{st}}$ query is $k/(N-k) \ge k/N$. Thus, the probability of *failing* to see a collision on the $k+1^{\text{st}}$ query is $\le 1 - k/N \le e^{-k/N}$. The probability that no collision is found after $c\sqrt{N} + 1$ queries is $\le \prod_{i=1}^{c\sqrt{N}}(1 - i/N) \le \exp(-\sum_{i=1}^{c\sqrt{N}} i/N) \le e^{-c^2}$. Taking $c = \sqrt{\ln(2)}$ then suffices.

An alternate approach is to observe that there are $N(N-1)\cdots(N-t+1)$ subsets of $[N]$ of size t, but only $N(N-2)\cdots(N-2(t+1))$ of these are collision-free. We then bound $(1-2j/N)/(1-j/N) \le e^{-j/N}$ by comparing the powers of j/N on each side, and then the proof proceeds as above.

b) Choose a random subset S of size $N^{1/3}$ and query f on those positions. Storing the answer takes $N^{1/3}\log(N)$ bits of memory. If there is already a collision, then we are done. If not, then query $cN^{2/3}$ random positions in $\{0,1\}^n - S$ and check for collisions with S. If the function is 2-1, then each query has a $1 - N^{-2/3}$ chance of finding a collision. Thus, a collision is found with probability $1 - e^{-c}$. Taking $c = \ln(2)$, we find a collision with probability $\ge 1/2$.

c) Query $f(0)$, store the answer, and then Grover search for $i \neq 0$ s.t. $f(i) = f(0)$.

d) Choose a random subset S of size M and query f on those positions. This takes M queries. Grover search for $i \in \{0,1\}^n - S$ s.t. $f(i) \in f(S)$. Assuming that f is 1-1 on S (and again, if this is not true, then we are done), there are M targets in a search space of size $N - M$. Thus, Grover search takes $O(\sqrt{\frac{N-M}{M}}) \le O(\sqrt{N/M})$ queries. The total number of queries is $O(M + \sqrt{N/M})$.

Exercise 6. *Optional, but recommended:* **Quantum counting**
We are given a black-box function $f : \{0,1\}^n \to \{0,1\}$ and would like to estimate $|f^{-1}(1)|$: that is, the number of $x \in \{0,1\}^n$ such that $f(x) = 1$. Let $M = |f^{-1}(1)|$ and $N = 2^n$.

a) *Suppose we are given access to* $U_f = \sum_{x\in\{0,1\}^n}\sum_{y\in\{0,1\}} |x\rangle\langle x| \otimes |y\oplus f(x)\rangle\langle y|$. *We would like to use* U_f *to apply the phase* $(-1)^{f(x)}$ *conditioned on an additional qubit. This operation is defined as*

$$V_f = I \otimes |0\rangle\langle 0| + \sum_{x\in\{0,1\}^n} (-1)^{f(x)}|x\rangle\langle x| \otimes |1\rangle\langle 1|.$$

Show how we can use U_f *to implement* V_f.

b) *Define* $|s\rangle = \frac{1}{\sqrt{2^n}}\sum_{x=0}^{N-1}|x\rangle$. *Define the Grover iteration*

$$G = (I - 2|s\rangle\langle s|)\cdot \sum_{x\in\{0,1\}^n} (-1)^{f(x)}|x\rangle\langle x|.$$

Find the eigenvalues of G.

c) *Show how the construction of part (a) can be used to perform*

$$\sum_{t=0}^{T-1} |t\rangle\langle t| \otimes G^t$$

using $T-1$ *queries to* U_f.

d) *Assume that* M *divides* N. *Show that quantum phase estimation can be used to determine* M/N *up to accuracy* $O(\sqrt{M/N}/T)$ *with high probability. How large does* T *have to be in order to have a* $\geq 1/2$ *probability of determining* M *exactly? How many queries are necessary to achieve this classically?*

a) Apply Hadamards to the last qubit before and after U_f.

b) Let $\Pi = \sum_{x\in f^{-1}(1)}|x\rangle\langle x|$. Let $|s_1\rangle = \sum_{x\in f^{-1}(1)}|x\rangle/\sqrt{M}$ and $|s_2\rangle = \sum_{x\in f^{-1}(0)}|x\rangle/\sqrt{N-M}$. If we define $p = M/N$, then note that $|s\rangle = \sqrt{M/N}\,|s_1\rangle + \sqrt{1-M/N}\,|s_2\rangle$. Also G acts trivially on the subspace orthogonal to $\{|s_1\rangle, |s_2\rangle\}$. On the $\{|s_1\rangle, |s_2\rangle\}$ subspace, G acts as

$$\begin{pmatrix} 1-2p & -2\sqrt{p(1-p)} \\ -2\sqrt{p(1-p)} & -1+2p \end{pmatrix} \cdot \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} = -\begin{pmatrix} 1-2p & 2\sqrt{p(1-p)} \\ -2\sqrt{p(1-p)} & 1-2p \end{pmatrix}$$

which has eigenvalues $e^{\pi i\theta}$ where $\theta = \sin^{-1}(2\sqrt{p(1-p)})$.

c) Write t in unary (i.e. $T-1$ bits, of which t are equal to 1 and $T-1-t$ equal to zero). Then apply V_f $T-1$ times, with the same first register and with the control register stepping through the $T-1$ bits.

d) Apply phase estimation to $|s\rangle$ and we learn either θ or $-\theta$ to accuracy $1/T$. To translate this into the error in p, we observe that $2\sqrt{p(1-p)} = \sin(\theta)$. Assume that $0 \leq p \leq 1/2$, so $\sqrt{p} \leq 2\sqrt{p(1-p)} \leq 2\sqrt{p}$. Thus, $p \sim \sin^2(\theta)$.

Suppose now phase estimation returns $\theta + \epsilon$ instead of θ. Then our estimate for p will be off by $\sim \epsilon\sin(\theta)\cos(\theta) \sim \epsilon\sqrt{p}$.

Substituting $\epsilon \sim 1/T$, we find that the algorithm outputs $p \pm O(\sqrt{p}/T)$ with high probability. Thus, to learn M/N exactly, we need $\sqrt{p}/T \ll 1/N$, and therefore need $T \gg N\sqrt{p} = \sqrt{MN}$. By contrast, learning M exactly classically requires $\Omega(N)$ queries, even if we allow a probability of error.

If $\frac{1}{2} < p \leq 1$, then the above bounds hold, but we can do better in the $p \approx 1$ regime by estimating $|f^{-1}(0)|$ instead of $|f^{-1}(1)|$.