Solutions to Hw 4
December 2013

Problem 4.1

(a) Simple trigonometry tells us

\[
\operatorname{Prob}(y) = \frac{1}{NA} \left( \frac{\sin^2 \pi Ar \delta}{\sin^2 \pi r \delta} \right) \leq \frac{1}{NA} \left( \frac{1}{\sin^2 \pi r \delta} \right) \tag{1}
\]

Further, note that

\[
\frac{2}{\pi} \leq \frac{\sin x}{x} \leq 1, \quad \forall x \in [0, \frac{\pi}{2}]
\]

\[
\Rightarrow \frac{\sin^2 x}{x^2} \geq \frac{4}{\pi^2}, \quad \forall x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \tag{2}
\]

Therefore,

\[
\sin^2(\pi r \delta) \geq (\pi r \delta)^2 \frac{4}{\pi^2}, \quad \text{since } \pi r \delta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \tag{3}
\]

Thus we have,

\[
\operatorname{Prob}(y) \leq \frac{1}{NA} \left( \frac{1}{4r^2 \delta^2} \right) = \frac{1}{4NA r^2 \delta^2} \tag{4}
\]

(b) Let us denote by \(y_\delta\), the measurement outcome \(y\) whose distance to the nearest multiple of \(1/r\) is exactly equal to \(\delta\), ie.

\[
\delta = \frac{y_\delta}{N} - \frac{k}{r} \tag{5}
\]

And let us define \(y_{\max}\) as,

\[
\delta_{\max} = \frac{y_{\max}}{N} - \frac{k}{r} = \frac{1}{2r} \tag{6}
\]

Now,

\[
\operatorname{Prob}(> \delta) \leq \int_{y_\delta}^{y_{\max}} \frac{1}{4NA r^2 \left( \frac{y}{N} - \frac{k}{r} \right)^2} \leq \frac{2N}{4NA r^2} \left( \frac{1}{\frac{2N}{r} - \frac{k}{r}} - \frac{1}{\frac{2\max}{N} - \frac{k}{r}} \right)
\]

\[
= \frac{1}{2Ar^2 \delta} - \frac{1}{Ar}
\]

\[
< \frac{1}{Nr \delta} \leq \frac{1}{N\delta} \tag{7}
\]

\(^1\text{This simply follows from the fact that } \frac{\sin x}{x} \text{ is a monotonically decreasing function in the interval } [0, \frac{\pi}{2}].\)
since $2Ar < N$, because by definition, $A > \frac{N}{r} > \frac{N}{2^r}$ and $1 \ll r \ll 2^n$.

**Problem 4.2**

(a) Recall from the lecture notes that we can estimate the eigenvalue $\lambda$ of $U$ by repeating the following procedure:

First, we prepare the state $|+\rangle \otimes |\lambda\rangle$ where $U |\lambda\rangle = \lambda |\lambda\rangle$, then apply $\Lambda(U)$ and finally, measure the first qubit in the eigenbasis of $\sigma_x$ or $\sigma_y$.

Each time we repeat this procedure, a measurement of $\sigma_x$ and $\sigma_y$ gives us information about the real and imaginary parts, respectively, of $\lambda$, so that after $k$ repetitions we can determine $\lambda$ to accuracy $O(1/\sqrt{k})$.

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

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

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 $\rho_\text{in}$, the state becomes

$$\rho_\text{out} = \Lambda(U) \left( |+\rangle \langle +| \otimes \frac{I_n}{2^n} \right) \Lambda(U)^\dagger = \frac{1}{2^n+1} \begin{pmatrix} I_n & 0 \\ 0 & U \end{pmatrix} \begin{pmatrix} I_n & I_n \\ I_n & I_n \end{pmatrix} \begin{pmatrix} I_n & 0 \\ 0 & U^\dagger \end{pmatrix} = \frac{1}{2^n+1} \begin{pmatrix} I_n & U^\dagger \\ U & I_n \end{pmatrix}.$$

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

$$\Pi_{\sigma_x,\pm 1} = |\pm\rangle \langle | \otimes I_n = \frac{1}{2} \begin{pmatrix} I_n & \pm I_n \\ \pm I_n & I_n \end{pmatrix},$$

so that the probability of obtaining the eigenvalue $\pm 1$ is

$$P(\sigma_x, \pm 1) = \text{Tr} \left( \Pi_{\sigma_x,\pm 1} \rho_\text{out} \right) = \frac{1}{2^n+2} \text{Tr} \left( \pm I_n + U \pm I_n + U^\dagger \right) = \frac{1}{2} \left( 1 \pm \frac{\text{Re}(\text{Tr}(U))}{2^n} \right) . \tag{11}$$

If we instead measure the first qubit in the eigenbasis of $\sigma_y$, then the projectors corresponding to the $\pm 1$ outcomes take the form

$$\Pi_{\sigma_y,\pm 1} = |\pm i\rangle \langle | \otimes I_n = \frac{1}{2} \begin{pmatrix} I_n & \pm i I_n \\ \pm i I_n & I_n \end{pmatrix},$$

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

Suppose we repeat the procedure in which the ancilla qubit is measured in the $\sigma_x$ eigenbasis $k$ times and we average over all measurement outcomes. Since new input states are prepared each time we repeat the procedure, there are no correlations between the measurement

\footnote{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 $\sigma_x$ and $\sigma_y$ respectively.}
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 an average value:

\[ P(\sigma_x, +1)(+1) + P(\sigma_x, -1)(-1) = \text{Re}(\text{Tr}U)/2^n \]

and a 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 \(\text{Tr}U/2^n\) to accuracy \(\delta = O(1/\sqrt{k})\) by repeating the following procedure k times - (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 \(\sigma_x\) or \(\sigma_y\). To achieve a fixed accuracy \(\delta_0\), we therefore 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 \(\text{poly}(n)\) 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 \(\text{poly}(n)\) gates in \(\Lambda(U)\) that we must approximate. Since errors add linearly, we require each gate in \(\Lambda(U)\) to be approximated to an accuracy \(\epsilon\) such that, the error in each individual gate multiplied with the number of gates in the whole circuit ( \(k\text{poly}(n)\)), does not exceed our desired fixed accuracy, i.e. \(\delta_0 \text{poly}(n) \epsilon < \delta_0\). 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 \(\text{polylog}(1/\epsilon)\). This establishes that our algorithm has size \(\text{poly}(n \log n)\) and is therefore efficient.

**Problem 4.3**

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

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

(13)

We then query the oracle to produce the state

\[
\frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle \otimes |0\rangle \rightarrow \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle \otimes |f(x)\rangle,
\]

(14)

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

\[
\frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle \otimes |f(x)\rangle \rightarrow |x_0H\rangle = \frac{1}{\sqrt{|H|}} \sum_{x:f(x)=f(x_0)} |x\rangle,
\]  

(15)

3
which can also be written as
\[ |x_oH⟩ = \frac{1}{\sqrt{|H|}} \sum_{b ∈ \mathbb{Z}_2^k} |x_0 + b ⋅ a⟩, \]

where \( a \) is the vector with the generators \( a_i \) of \( H \) as its “components”, 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
\[ H^⊗n |x_oH⟩ = \frac{1}{\sqrt{|H|}} \sqrt{2^n} \sum_{y ∈ \mathbb{Z}_2^n} (−1)^{y ⋅ a} |y⟩, \]

For each fixed \( y \), we can consider the sum over all values of \( b \), \( \sum_{b ∈ \mathbb{Z}_2^k} (−1)^{(b ⋅ a) y} \). For the case \( k = 1 \) (so that \( b = 0, 1 \)) the sum gives the familiar \( 1 + (−1)^y \) 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, \( \sum_{b ∈ \mathbb{Z}_2^k} (−1)^{(b ⋅ a) y} = \prod_{i=1}^k (1 + (−1)^{a_i y}) \). Clearly, the only \( y \)’s that survive in the sum of Eq. (24) are those for which \( a_i y = 0 \) \( (mod 2) \), \( ∀i = 1, 2, \ldots, k \). 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.

However, 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. It will be nontrivial \( (y_1 \neq 0) \) with a probability no smaller than \( 1 − 1/2^{n−k} \). Since in Eq. (24) 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
\[ P_{\text{success}} ≥ \left(1 - \frac{1}{2^n}\right)\left(1 - \frac{2}{2^n}\right)\left(1 - \frac{2^2}{2^n}\right)\cdots\left(1 - \frac{2^{n−k−1}}{2^n}\right) \]
\[ ≥ \left(1 - \frac{1}{2}\right)\left(1 - \left(\frac{1}{2} + \frac{1}{2^n}\right) + \cdots + \frac{1}{2^{n−k−1}} + \frac{1}{2^n}\right) \]
\[ ≥ \frac{1}{2} \left(1 - \frac{1}{2}\right) = \frac{1}{4}, \]

since
\[ \frac{1}{2^2} + \frac{1}{2^3} + \cdots + \frac{1}{2^{n−k−1}} + \frac{1}{2^n−k} = \frac{1}{2} - \frac{1}{2^n−k} ≤ \frac{1}{2}. \]

**Problem 4.4**

(a) Let us choose \( c\sqrt{N} \) distinct inputs \( x_i \) at random, where \( c \) is a constant. For each of these inputs, we can compute \( f(x_i) \) and store the \( c\sqrt{N} \) pairs \((x_i, f(x_i))\) in memory. This requires space \( O(\sqrt{N}) \) and can be accomplished with \( O(\sqrt{N}) \) oracle queries. We now compute the probability that we fail to find a single collision among these \( c\sqrt{N} \) randomly chosen inputs.
We first pick the input \( x_1 \). Let \( A_2 \) denote the event that the second input \( x_2 \) does not collide with \( x_1 \) (i.e., \( f(x_1) \neq f(x_2) \)). Event \( A_2 \) occurs with probability

\[
P(A_2) = \sum_{x_1} P(x_1)P(A_2|x_1) = N \left( \frac{1}{N} \right) \left( \frac{N-1}{N} \right) = \frac{N-1}{N}
\]

Similarly, let \( A_3 \) be the event that there is no collision among \( x_1, x_2 \) and our third input \( x_3 \). Then

\[
P(A_3) = P(A_3|A_2)P(A_2) = \left( \frac{N-2}{N} \right) \left( \frac{N-1}{N} \right)
\]

After choosing \( k \) distinct inputs \( x_i \), the probability that we fail to find a collision is

\[
P(A_k) = P(A_k|A_{k-1})P(A_{k-1}|A_{k-2}) \cdots P(A_3|A_2)P(A_2)
\]

\[
= \left( 1 - \frac{k-1}{N} \right) \left( 1 - \frac{k-2}{N} \right) \cdots \left( 1 - \frac{2}{N} \right) \left( 1 - \frac{1}{N} \right)
\]

\[
\leq \exp \left( -\frac{1}{N} \sum_{i=1}^{k-1} i \right)
\]

\[
= \exp \left( -\frac{k(k-1)}{2N} \right)
\]

where we have used \( 1 + x \leq e^x \).

Substituting \( k = c\sqrt{N} \) and using \( N >> 1 \), Eq.(3) gives us the probability of failure to find a collision after \( c\sqrt{N} \) queries - \( P(A_k) \leq e^{-c^2/2} \). Thus we can upper bound the failure probability by any constant \( 0 < \delta \leq 1 \) if we choose \( c \) to be sufficiently large.

(b) We now choose \( cN^{1/3} \) distinct inputs \( x_i \) at random for some constant \( c \). Again we compute \( f(x_i) \) for all of them and store the \( cN^{1/3} \) pairs \( (x_i, f(x_i)) \) as before. This requires space \( O(N^{1/3}) \) and \( O(N^{1/3}) \) oracle queries. Let \( X = \bigcup_i \{ x_i \} \). We can check whether there is a collision within \( X \) and declare success if we find one. If we are unsuccessful, we know that every \( x_i \in X \) collides with a unique \( y \in S \setminus X \), where \( S = \{0, 1\}^n \) is the total input space.

We define a new function \( g : S \setminus X \to \{0, 1\} \) such that \( g(y) = 1 \) if \( f(y) = f(x) \) for some \( x \in X \) and \( g(y) = 0 \) otherwise. Thus \( g \) takes the value 1 at exactly \( |X| = cN^{1/3} \) inputs and our goal is to find one such input. This can be done in \( O(N^{2/3}) \) queries since \( g \) is defined in \( |S \setminus X| < N \) inputs and has \( O(N^{1/3}) \) marked inputs (i.e., inputs for which \( g \) takes the value 1).

In particular, the probability that after \( k \) trials \( y_i \in S \setminus X \) we fail to find a single collision is

\[
P_{\text{fail}} = \left( \frac{|S \setminus X| - |X|}{|S \setminus X|} \right)^k \leq \exp \left( -\frac{k|X|}{|S \setminus X|} \right).
\]

Since \( |S \setminus X| = N - cN^{1/3} \) and taking \( N >> 1 \), the failure probability can be made smaller than any constant \( 0 < \delta \leq 1 \) by making \( k = c'N^{2/3} \) for some large enough constant \( c' \).

Overall, we succeed with high probability using \( k + cN^{1/3} = O(N^{2/3}) \) oracle queries and space.
O(N^{1/3}).

(c) We first pick a random input \( x_0 \in S \). Then we consider the new function \( h : S \setminus \{ x_0 \} \to \{ 0, 1 \} \) such that \( h(y) = 1 \) if \( f(y) = f(x_0) \) and \( h(y) = 0 \) otherwise. Using Grover’s algorithm for the function \( h \), we can find a collision (i.e., the unique \( y \) such that \( f(y) = f(x_0) \)) in \( O(\sqrt{N}) \) oracle queries, using space \( O(1) \), since we just need to store the pair \((x_0, f(x_0))\).

(d) We now choose \( M \) distinct inputs \( x_i \) at random and compute the \( M \) pairs \((x_i, f(x_i))\) as before. This requires space \( O(M) \) and \( O(M) \) oracle queries. Using the function \( g \) of part (b), we can perform Grover’s algorithm to find one of the \( M \) marked inputs of \( g \), in \( O(\sqrt{N/M}) \) oracle queries, with high probability. We then query the oracle one additional time to learn the value of \( f \) for this input and compare with the \( M \) pairs \((x_i, f(x_i))\) to find the \( x_i \) with which it collides. Overall, choosing \( M = N^{1/3} \), we can find a collision using space \( O(N^{1/3}) \) and oracle queries \( O(N^{1/3} + \sqrt{N/N^{1/3}}) = O(N^{1/3}) \).

Problem 4.5

(a) Observe that if we feed \(|x\rangle \otimes |−\rangle\) to the black box performing \( U_f \), we obtain, as the output, \((-1)^{f(x)} |x\rangle \otimes |−\rangle\). On the other hand, if the input is in the state \(|x\rangle \otimes |+\rangle\), then the black box does nothing to that state. Thus, to implement \( \Lambda(U_f) \) (the last qubit is the controlled qubit!) one has to apply the Hadamard gate to the last qubit, then the black box \( U_f \), and finally the Hadamard gate to the last qubit.

(b) We can use a \( m \)-bit register for \( t \) such that \( t = \sum_{k=0}^{m-1} t_k 2^k \). Now we control the unitary \((U_{\text{Grover}})^{2^k}\) from the \( k \)-th bit. Thus, overall, we use \( 2^0 + 2^1 + 2^2 + \cdots + 2^{m-1} = 2^m - 1 = T - 1 \) oracle queries.

(c) We start with a uniform superposition over all counter values and all input values,

\[
|\Psi_{\text{initial}}\rangle = \frac{1}{\sqrt{T}} \sum_{t=0}^{T-1} |t\rangle \otimes \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} |j\rangle
\]

We then apply \( V \) to obtain the state

\[
V |\Psi_{\text{initial}}\rangle = \frac{1}{\sqrt{T}} \sum_{t=0}^{T-1} |t\rangle \otimes \left( \cos((2t + 1)\theta) |\Psi_X^\perp\rangle + \sin((2t + 1)\theta) |\Psi_X\rangle \right)
\]

since the initial state \(|s\rangle = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} |j\rangle = \cos(\theta) |\Psi_X^\perp\rangle + \sin(\theta) |\Psi_X\rangle \) and each Grover iteration rotates \(|s\rangle \) closer to \(|\Psi_X\rangle \) by an angle \( 2\theta \).

Finally, we want to apply the QFT on the counter register and then measure in the computation.
basis. Applying the QFT on $V\Psi_{\text{initial}}$ gives

$$
|\Psi_{\text{QFT}}\rangle = \frac{1}{\sqrt{T}} \sum_{l=0}^{T-1} e^{\frac{2\pi i}{T} l T} |l\rangle \otimes (\cos(\theta l + 1) |\Psi_X\rangle + \sin(\theta l + 1) |\Psi_X\rangle)
$$

$$
= \frac{1}{\sqrt{T}} \sum_{l=0}^{T-1} |l\rangle \otimes \left( \frac{a_1 + b_1}{2} |\Psi_X\rangle + \frac{a_1 - b_1}{2T} |\Psi_X\rangle \right),
$$

where $a = \sum_{l=0}^{T-1} \exp\left(\frac{2\pi i (l + \theta l)}{T} \right) + \theta)$, $b = \sum_{l=0}^{T-1} \exp\left(\frac{2\pi i (l - \theta l)}{T} \right) - \theta$.

Therefore, the probability of measuring the outcome $l$ is given by

$$
P(l) = \frac{1}{4T^2} \left( |a_1 + b_1|^2 + |a_1 - b_1|^2 \right) = \frac{1}{2T^2} \left( |a_1|^2 + |b_1|^2 \right).$$

(i) First we consider the case when $T\theta/\pi \in \mathbb{Z}$. Then, the only values of $l$ that survive in Eq. (25) are $l = \pm T\theta/\pi$ since $a = T\delta_{l,-T\theta/\pi}$ and $b = T\delta_{l,T\theta/\pi}$. Measuring the counter in the computation basis will therefore yield the integers $T\theta/\pi$ or $T - T\theta/\pi$ with equal probability 1/2, from which we can learn $\theta$ with accuracy $O(1/T)$.

(ii) However, in general $T\theta/\pi \notin \mathbb{Z}$. Therefore, we consider next, the case $0 < T\theta/\pi < 1$. Then, for $0 < T\theta/\pi < 1$, success means measuring $l = 0$ which is the closest integer to $T\theta/\pi$ and we have

$$
P(l = 0) = \frac{1}{2T^2} \left( |a_{l=0}|^2 + |b_{l=0}|^2 \right) \geq \frac{1}{2T^2} \left( \left( \frac{T}{2} \right)^2 + \left( \frac{T}{2} \right)^2 \right) = \frac{4}{\pi^2}.\tag{27}
$$

since $|a_{l=0}| = \left| \frac{e^{i2T\theta} - 1}{e^{i2\theta} - 1} \right| \geq \frac{2e^{i2\theta}}{2\theta} = T^2 \pi$, and $|b_{l=0}| = \left| \frac{e^{i2T\theta} + 1}{e^{i2\theta} - 1} \right| = |a_{l=0}|$. When $1/2 < T\theta/\pi < 1$, success means measuring $l = \pm 1$ (i.e., $l = 1$ or $l = T - 1 = -1 (\text{mod}T)$) since 1 is now the closest integer to $T\theta/\pi$. We calculate

$$
P(l = 1) = \frac{1}{2T^2} \left( |a_{l=1}|^2 + |b_{l=1}|^2 \right) \geq \frac{|a_{l=1}|^2}{2T^2} \geq \frac{2}{\pi^2},
$$

$$
P(l = T - 1) = \frac{1}{2T^2} \left( |a_{l=T-1}|^2 + |b_{l=T-1}|^2 \right) \geq \frac{|a_{l=T-1}|^2}{2T^2} \geq \frac{2}{\pi^2},
$$

since $|b_{l=1}| = \left| \frac{e^{i2(T\theta)} - 1}{e^{i2(2\theta)} - 1} \right| \geq \frac{2e^{i2\theta}}{2\theta} = T^2 \pi$, and $|a_{l=T-1}| = \left| \frac{e^{i2T\theta} - 1}{e^{i2(2\theta)} - 1} \right| = |b_{l=1}|$. Therefore, the success probability is again lower bounded by $4/\pi^2$.

(iii) Next, we look at the case when $1 < T\theta/\pi < \frac{T}{2} - 1$. Let $f^- = \lfloor \frac{T\theta}{\pi} \rfloor$ and $f^+ = \lceil \frac{T\theta}{\pi} \rceil$. Then, if $T\theta/\pi - f^- \leq \frac{1}{2}$, success means measuring $l = \pm f^-$ (i.e., $l = f^-$ or $l = T - f^-$), and we can calculate

$$
P(l = f^-) = \frac{1}{2T^2} \left( |a_{l=f^-}|^2 + |b_{l=f^-}|^2 \right) \geq \frac{|a_{l=f^-}|^2}{2T^2} \geq \frac{2}{\pi^2},
$$

$$
P(l = T - f^-) = \frac{1}{2T^2} \left( |a_{l=T-f^-}|^2 + |b_{l=T-f^-}|^2 \right) \geq \frac{|a_{l=T-f^-}|^2}{2T^2} \geq \frac{2}{\pi^2}.
$$

since, defining $\phi = \frac{T\theta}{\pi} - f^-$, we have, $|b_{l=f^-}| = \left| \frac{e^{i2\phi} - 1}{e^{i2\phi} + 1} \right| \geq \frac{2e^{i2\phi}}{2\phi} = T^2 \pi$, and $|a_{l=T-f^-}| = \left| \frac{e^{i2\phi}}{e^{i2\phi} + 1} \right| = |b_{l=f^-}|$. When $f^+ - \frac{T\theta}{\pi} < \frac{1}{2}$, success means measuring $l = \pm f^+$ (i.e., $l = f^+$ or $f^+ - \frac{T\theta}{\pi}$ for $|\phi| \leq \pi$.

---

*In what follows we make repetitive use of the bounds (a) $|e^{i\theta} - 1| \leq |\phi|$ for $|\phi| < 1$, and (b) $|e^{i\theta} - 1| \geq 2\phi^\pi$ for $|\phi| \leq \pi$.*
\( l = T - f^+ \), and we have

\[
\begin{align*}
P(l = f^+) &= \frac{1}{2T^2} \left( |a_{l=f^+}|^2 + |b_{l=f^+}|^2 \right) \geq \frac{|b_{l=f^+}|^2}{2T^2} \geq \frac{2}{\pi^2}, \\
P(l = T - f^+) &= \frac{1}{2T^2} \left( |a_{l=T-f^+}|^2 + |b_{l=T-f^+}|^2 \right) \geq \frac{|a_{l=T-f^+}|^2}{2T^2} \geq \frac{2}{\pi^2},
\end{align*}
\]

(30)
since, defining \( \phi' = f^+ - \frac{T\theta}{\pi} \), we have, \( |b_{l=f^+}| = \left| \frac{e^{i2\pi\phi'}}{e^{i2\pi\phi'/T} - 1} \right| \geq T\frac{2}{\pi} \) and \( |a_{l=T-f^+}| = \left| \frac{e^{-i2\pi\phi'}}{e^{i2\pi\phi'/T} - 1} \right| \). The success probability is again lower bounded by \( 4/\pi^2 \).

(iv) The final case is when \( \frac{T}{2} - 1 < T\theta/\pi < \frac{T}{2} \). First, for \( T\theta/\pi - \left( \frac{T}{2} - 1 \right) \leq \frac{1}{2} \), success means measuring \( l = \pm \frac{T}{2} - 1 \) (i.e., \( l = \frac{T}{2} - 1 \) or \( l = \frac{T}{2} + 1 \)). Setting \( f^- = \frac{T}{2} - 1 \), it follows from Eq. (29) that the success probability is at least \( 4/\pi^2 \). For \( \frac{T}{2} - T\theta/\pi < \frac{1}{2} \), success means measuring \( l = \frac{T}{2} \). We can calculate

\[
P(l = T/2) = \frac{1}{2T^2} \left( |a_{l=T/2}|^2 + |b_{l=T/2}|^2 \right) \geq \frac{1}{2T^2} \left( T^2 \pi \right)^2 = \frac{4}{\pi^2},
\]

(31)
since, defining \( \phi'' = \frac{T}{2} - \frac{T\theta}{\pi} \), we have, \( |a_{l=T/2}| = \left| \frac{e^{-i2\pi\phi''}}{e^{-i2\pi\phi''/T} - 1} \right| \geq T^2 \frac{2}{\pi} \), and \( |b_{l=T/2}| = \left| \frac{e^{i2\pi\phi''}}{e^{-i2\pi\phi''/T} - 1} \right| = |a_{l=T/2}| \).

Therefore, for all cases above, measuring the counter in the computation basis reveals the closest integer to \( \frac{T\theta}{\pi} \) with a \textit{constant} probability of success which is at least \( 4/\pi^2 \), so that \( \theta \) can be determined to accuracy \( O(1/T) \).

We also have \( \theta \approx \sqrt{r/N} \), or \( r \approx N\theta^2 \). Hence, \( \delta r \approx 2\sqrt{rN}\delta \theta \). Since \( \delta \theta = O(1/T) \) and setting \( \delta r \approx 1 \), it follows that \( T = O(\sqrt{rN}) \). Classically, the query complexity is \( O(rN) \), since we would need to determine the values of \( rN \) inputs for which \( X_i = 1 \).