A tight query complexity lower bound for phase estimation under circuit depth constraint

Yu Tong

February 22, 2021

We consider the following problem

Problem 1. *Given a black box oracle acting on a single qubit*

$$O_\theta = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix},$$

determine whether $\theta = \theta_0$ *or* $\theta = \theta_1$*, where* $|\theta_1 - \theta_0| = \epsilon$*.*

We can of course solve this problem using the quantum phase estimation (QPE) algorithm with query complexity $\mathcal{O}(\epsilon^{-1})$. Ref. [1] shows this upper bound is tight. However the QPE algorithm requires querying the oracle for a total of $\mathcal{O}(\epsilon^{-1})$ times in a coherent manner. This might not be possible with near-term devices where the circuit depth is limited by the coherence time.

In this note we consider allowing only D queries to the oracle O_θ before all the qubits are measured. We have the following lower bound

Theorem 2 (Main result)**.** *Consider the scenario in which we are allowed to query* O_θ *in a coherent manner for at most* D *times, and before the* $(D+1)$*-th query all qubits need to be measured. In this scenario, in order to solve Problem 1, we need to use on average* $\Omega(D^{-1}\epsilon^{-2})$ *queries to* O_θ *for some* θ*.*

This lower bound is tight because a modification of Kitaev's algorithm can achieve this scaling. The proof is provided in Section 4.

1 Notations

Throughout this note, for a set of variables $(x_1, x_2, \ldots, x_L)$, we will write $\vec{x}_k = (x_1, x_2, \ldots, x_k)$, $k = 1, 2, \ldots, L$. For a joint distribution P of $(X_1, X_2, \ldots, X_L)$, we will write P_{X_k} for the marginal distribution of X_k, and $P_{X_k|\vec{X}_{k-1}}$ for the conditional distribution of X_k on $(X_1, X_2, \ldots, X_{k-1})$. To be more specific, $P_{X_k|\vec{X}_{k-1}}(x_k|\vec{x}_{k-1})$ is the probability of $X_k = x_k$ conditional on $X_1 = x_1, X_2 = x_2, \ldots, X_{k-1} = x_{k-1}$. For convenience we define $P_{X_k|\vec{X}_0} = P_{X_k}$.

We write $|\vec{0}\rangle = |0\rangle\,|0\rangle \cdots |0\rangle$.

2 The Hellinger distance

Definition 3 (The Hellinger distance)**.** *The Hellinger distance between two distributions P and Q is defined as*

$$H(P,Q) = \sqrt{\frac{1}{2}\sum_x \left(\sqrt{P(x)} - \sqrt{Q(x)}\right)^2}.$$

As we can see this is reminiscent of the 2-norm (Euclidean) distance between two quantum states. In fact we have the following relation

Lemma 4. *Let $|\phi\rangle = \sum_x \phi_x |x\rangle$ and $|\psi\rangle = \sum_x \psi_x |x\rangle$. Let P and Q be the distributions obtained by measuring $|\phi\rangle$ and $|\psi\rangle$ respectively in the computational basis, i.e. $P(x) = |\phi_x|^2$ and $Q(x) = |\psi_x|^2$. Then*

$$H(P,Q) \leq \frac{1}{\sqrt{2}}\| \, |\phi\rangle - |\psi\rangle \, \|.$$

The following lemma will be our main tool to bound the information we get in multiple sequential experiments.

Lemma 5. *Let P and Q be two distributions of the random variables (X, Y). We have*

$$H^2(P,Q) \leq H^2(P_X, Q_X) + \mathbb{E}_{x\sim\frac{1}{2}(P_X+Q_X)}\left[H^2\left(P_{Y|X}(\cdot|x), Q_{Y|X}(\cdot|x)\right)\right]$$

Proof. The proof is mainly a modification of the proof of Theorem 2.1 in Ref. [3]. We only prove for $L = 2$ and the rest follows by induction.

$$\begin{aligned} H^2(P,Q) &= 1 - \sum_{x,y}\sqrt{P(x,y)Q(x,y)} \\ &= 1 - \sum_{x,y}\sqrt{P_X(x)P_{Y|X}(y|x)Q_X(x)Q_{Y|X}(y|x)} \\ &= 1 - \sum_{x,y}\frac{P_X(x)+Q_X(x)}{2}\sqrt{P_{Y|X}(y|x)Q_{Y|X}(y|x)} \\ &+ \sum_{x,y}\left(\frac{P_X(x)+Q_X(x)}{2} - \sqrt{P_X(x)Q_X(x)}\right) \\ &\times \sqrt{P_{Y|X}(y|x)Q_{Y|X}(y|x)}. \end{aligned}$$

We have

$$\sum_y \sqrt{P_{Y|X}(y|x)Q_{Y|X}(y|x)} = 1 - H^2\left(P_{Y|X}(\cdot|x), Q_{Y|X}(\cdot|x)\right),$$

Therefore

$$\begin{aligned}H^2(P,Q) &= \mathbb{E}_{x\sim\frac{1}{2}(P_X+Q_X)} H^2\left(P_{X_2|X}(\cdot|x), Q_{X_2|X}(\cdot|x)\right) \\ &+ \frac{1}{2}\sum_x \left(\sqrt{P_X(x)} - \sqrt{Q_X(x)}\right)^2 \\ &\times \sum_y \sqrt{P_{X_2|X}(y|x) Q_{X_2|X}(y|x)}.\end{aligned}$$

Since by the Cauchy-Schwarz inequality

$$\sum_y \sqrt{P_{X_2|X}(y|x) Q_{X_2|X}(y|x)} \leq 1,$$

we have

$$H^2(P,Q) \leq \mathbb{E}_{x\sim\frac{1}{2}(P_X+Q_X)} H^2\left(P_{X_2|X}(\cdot|x), Q_{X_2|X}(\cdot|x)\right) + H^2\left(P_X, Q_X\right).$$

□

By inductively applying this lemma, we have the following corollary

Corollary 6. *Let P and Q be two distributions of the random variables $(X_1, X_2, \ldots, X_L)$, we have*

$$H^2(P,Q) \leq \mathbb{E}_{\vec{x}_L\sim\frac{1}{2}(P+Q)} \left[\sum_{l=1}^{L} H^2\left(P_{\cdot|\vec{X}_{l-1}}(\cdot|\vec{x}_{l-1}), Q_{\cdot|\vec{X}_{l-1}}(\cdot|\vec{x}_{l-1})\right)\right]$$

Proof. We simply use Lemma 5 repeatedly.

$$\begin{aligned}H^2(P,Q) &\leq H^2(P_{\vec{X}_{l-1}}, Q_{\vec{X}_{l-1}}) \\ &+ \mathbb{E}_{\vec{x}_{L-1}\sim\frac{1}{2}(P_{\vec{X}_{L-1}}+Q_{\vec{X}_{L-1}})}\left[H^2\left(P_{X_L|\vec{X}_{L-1}}(\cdot|\vec{x}_{L-1}), Q_{X_L|\vec{X}_{L-1}}(\cdot|\vec{x}_{L-1})\right)\right] \\ &\leq H^2(P_{\vec{X}_{l-2}}, Q_{\vec{X}_{l-2}}) \\ &+ \mathbb{E}_{\vec{x}_{L-2}\sim\frac{1}{2}(P_{\vec{X}_{L-2}}+Q_{\vec{X}_{L-2}})}\left[H^2\left(P_{X_{L-1}|\vec{X}_{L-2}}(\cdot|\vec{x}_{L-2}), Q_{X_{L-1}|\vec{X}_{L-2}}(\cdot|\vec{x}_{L-2})\right)\right] \\ &+ \mathbb{E}_{\vec{x}_{L-1}\sim\frac{1}{2}(P_{\vec{X}_{L-1}}+Q_{\vec{X}_{L-1}})}\left[H^2\left(P_{X_L|\vec{X}_{L-1}}(\cdot|\vec{x}_{L-1}), Q_{X_L|\vec{X}_{L-1}}(\cdot|\vec{x}_{L-1})\right)\right] \\ &\leq \ldots\end{aligned}$$

Here we regard $H^2\left(P_{X_l|\vec{X}_{l-1}}(\cdot|\vec{x}_{l-1}), Q_{X_l|\vec{X}_{l-1}}(\cdot|\vec{x}_{l-1})\right)$ as a random variable that is measurable with respect to the sigma algebra generated by $\vec{X}_{l-1}$. Therefore all expectation evaluations of the form $\mathbb{E}_{\vec{x}_{l-1}\sim\frac{1}{2}(P_{\vec{X}_{l-1}}+Q_{\vec{X}_{l-1}})}$ can be replaced by $\mathbb{E}_{\vec{x}_L\sim\frac{1}{2}(P+Q)}$. Thus we have proved the corollary. □

3 An abstract characterization of phase estimation algorithms

A phase estimation algorithm can be characterized in the following way. We use quantum circuits $U_l(\theta)$, $l = 1, 2, \ldots, L$, each of which queries O_θ for a total of D_l times, where $D_l \leq D$. Starting from $U_1(\theta)$, we prepare the state $U_1(\theta)\,|\vec{0}\rangle$, and measure in the computational basis to obtain a random variable X_1. We then use the measurement result to determine the next quantum circuit $U_2(\theta)$ to be used, and obtain a random variable X_2 by measuring $U_2(\theta)\,|\vec{0}\rangle$. We do this for all the L circuits.

As we can see, the circuit we use depends on measurement results from before. This is the main difficulty in obtaining the lower bound, as otherwise the result in Refs. [2], which shows the quantum counting cannot be non-trivially accelerated through parallelization, already gives a lower bound.

Because of this dependence, we denote the l-th circuit to be used by $U_l(\vec{X}_{l-1};\theta)$. This circuit uses O_θ for a total of $D_l(\vec{X}_{l-1})$ times. Given $X_1 = x_1, X_2 = x_2, \ldots, X_{l-1} = x_{l-1}$, X_l is generated from the conditional distribution $P_{X_l|\vec{X}_{l-1}}(\cdot|\vec{x}_{l-1};\theta)$. At the end of the algorithm, we will have the random variables $(X_1, X_2, \ldots, X_l)$ obeying the distribution $P(\cdot;\theta)$, which is

$$P(\vec{x}_l;\theta) = \prod_{l=1}^{L} P_{X_l|\vec{X}_{l-1}}(x_l|\vec{x}_{l-1};\theta). \tag{1}$$

We then determine whether $\theta = \theta_0$ or $\theta = \theta_1$ through the values of these random variables.

Note that the above framework accounts for repeating the procedure because if we need to repeat then we only need to include the circuits used in repetitions into the circuits $\{U_l(\vec{X}_{l-1};\theta)\}$. The expected total number of queries to O_θ is $\sum_{l=1}^{L} \mathbb{E} D_l(\vec{X}_{l-1})$.

4 Proof of the main result

In this section we prove Theorem 2. In order to be able to distinguish between $\theta = \theta_0$ and $\theta = \theta_1$ with probability at least $2/3$ we need

$$H(P(\cdot;\theta_0), P(\cdot;\theta_1)) = \Omega(1). \tag{2}$$

We first upper bound the Hellinger distance between the pairs of conditional distributions $P_{X_l|\vec{X}_{l-1}}(\cdot|\vec{x}_{l-1};\theta_0)$ and $P_{X_l|\vec{X}_{l-1}}(\cdot|\vec{x}_{l-1};\theta_1)$, and then upper bound the left-hand side of (2) through Lemma 6. By Lemma 4, we have

$$\begin{aligned}
&H\left(P_{X_l|\vec{X}_{l-1}}(\cdot|\vec{x}_{l-1};\theta_0), P_{X_l|\vec{X}_{l-1}}(\cdot|\vec{x}_{l-1};\theta_1)\right) \\
&\leq \frac{1}{\sqrt{2}}\|U_l(\vec{X}_{l-1};\theta_0)\,|\vec{0}\rangle - U_l(\vec{X}_{l-1};\theta_1)\,|\vec{0}\rangle\,\| \\
&\leq \frac{1}{\sqrt{2}}\|U_l(\vec{X}_{l-1};\theta_0) - U_l(\vec{X}_{l-1};\theta_1)\|
\end{aligned}$$

As $U_l(\vec{X}_{l-1};\theta_0)$ and $U_l(\vec{X}_{l-1};\theta_1)$ only differ in O_θ, and $\|O_{\theta_0} - O_{\theta_1}\| \leq \epsilon$, we have

$$\|U_l(\vec{X}_{l-1};\theta_0) - U_l(\vec{X}_{l-1};\theta_1)\| \leq D_l(\vec{X}_{l-1})\epsilon.$$

Therefore

$$H\left(P_{X_l|\vec{X}_{l-1}}(\cdot|\vec{x}_{l-1};\theta_0), P_{X_l|\vec{X}_{l-1}}(\cdot|\vec{x}_{l-1};\theta_1)\right) \leq \frac{1}{\sqrt{2}} D_l(\vec{x}_{l-1})\epsilon.$$

By Corollary 6 we have

$$H^2(P(\cdot;\theta_0), P(\cdot;\theta_1)) \leq \frac{\epsilon^2}{2} \sum_{l=1}^{L} \mathbb{E}_{\vec{x}_L \sim \frac{1}{2}(P(\cdot;\theta_0)+P(\cdot;\theta_1))} \left[D_l^2(\vec{x}_{l-1})\right].$$

Therefore by (2) we have

$$\frac{1}{2}\sum_{l=1}^{L} \left(\mathbb{E}_{\vec{x}_L \sim P(\cdot;\theta_0)} D_l^2(\vec{x}_{l-1}) + \mathbb{E}_{\vec{x}_L \sim P(\cdot;\theta_1)} D_l^2(\vec{x}_{l-1})\right) = \Omega(\epsilon^{-2}).$$

Since $D_l \leq D$,

$$\begin{aligned} & D\sum_{l=1}^{L} \left(\mathbb{E}_{\vec{x}_L \sim P(\cdot;\theta_0)} D_l(\vec{x}_{l-1}) + \mathbb{E}_{\vec{x}_L \sim P(\cdot;\theta_1)} D_l(\vec{x}_{l-1})\right) \\ & \geq \sum_{l=1}^{L} \left(\mathbb{E}_{\vec{x}_L \sim P(\cdot;\theta_0)} D_l^2(\vec{x}_{l-1}) + \mathbb{E}_{\vec{x}_L \sim P(\cdot;\theta_1)} D_l^2(\vec{x}_{l-1})\right) \\ & = \Omega(\epsilon^{-2}). \end{aligned}$$

Therefore the total number of queries satisfies

$$\sum_{l=1}^{L} \left(\mathbb{E}_{\vec{x}_L \sim P(\cdot;\theta_0)} D_l(\vec{x}_{l-1}) + \mathbb{E}_{\vec{x}_L \sim P(\cdot;\theta_1)} D_l(\vec{x}_{l-1})\right) = \Omega(D^{-1}\epsilon^{-2}).$$

This shows that the expected total number of queries for either θ_0 or θ_1 is lower bounded by $\Omega(D^{-1}\epsilon^{-2})$, which proves Theorem 2.

References

[1] A. J. Bessen. Lower bound for quantum phase estimation. *Phys. Rev. A*, 71 (4):042313, 2005. doi:10.1103/physreva.71.042313.

[2] P. Burchard. Lower bounds for parallel quantum counting. *arXiv preprint arXiv:1910.04555*, 2019.

[3] C. Daskalakis and Q. Pan. Square Hellinger subadditivity for bayesian networks and its applications to identity testing. In *Conference on Learning Theory*, pages 697–703. PMLR, 2017.